

SHARIQUE KHATRI

(814) 769-0678 | sharique.khatri@gmail.com

linkedin.com/in/sharique-khatri | shariquekhatri.com | github.com/sharique2004

EDUCATION

The Pennsylvania State University

University Park, PA

Bachelor of Engineering in Computer Science, Minor in Cybersecurity, GPA: 3.60

May 2026

- Dean's List: Fall 2023, Spring 2024, Fall 2024, Spring 2025, Fall 2025, Spring 2026
- Coursework: Data Structures & Algorithms, Operating Systems, Computer Architecture, Database Systems, Compilers, Distributed Systems, Computer Vision, Cybersecurity

SKILLS

AI & Agents: Agent Architectures, Multi Model Routing, Tool Use Orchestration, MCP Servers, RAG, Evaluation

AI Tooling: Claude Code, Cursor, LangChain, LangGraph, Ollama, Whisper, Mistral, DeepSeek, Qwen2 VL

Backend & Data: Python, TypeScript, JavaScript, C, C# (.NET), SQL, REST APIs, PostgreSQL, MongoDB, Pinecone, FAISS

Frontend & Infra: React, Next.js, Tailwind, SwiftUI, Electron, Flask, Node.js, AWS, GCP, Vercel, Docker, CI/CD, Linux

PUBLICATIONS

OmniSch: A Multimodal PCB Schematic Benchmark

Under Review, ECCV 2026

- Contributed to the large scale human annotation pipeline and schematic entity labeling used to benchmark large multimodal models on visual grounding, diagram to graph reasoning, and tool augmented agentic visual search. [arXiv:2604.00270 \[cs.CV\]](https://arxiv.org/abs/2604.00270).

EXPERIENCE

WellX AI

Dubai, UAE (Remote)

Software Engineering Intern

Dec 2025 to Mar 2026

- Built production **RAG** services serving **10,000+** users, integrating hosted and open weight **LLMs** with wearable data through **Python, FastAPI, REST APIs**, embeddings, and vector search.
- Designed backend ingestion endpoints across **3** wearable sources (**Fitbit, Apple Watch, WHOOP**), translating device data into structured contracts for downstream AI features used by internal engineering teams.
- Delivered AI features across **4** lifecycle stages (prompt design, offline evaluation, deployment, monitoring), reporting to the CTO and partnering with **2** teams (research, operations) on feedback loops and quality checks.

Fourth Square

Texas, USA (Remote)

Software Engineering Intern

May 2024 to Aug 2024

- Implemented a **LangChain** document question answering and content generation system handling **1,000+** production queries, integrating **OpenAI** and **Cohere** with chunking, caching, and prompt chains.
- Delivered **11 user stories** across backend services and AI tooling, tuning model outputs against researcher feedback and creating lightweight evaluations to track output quality across iterations.

Office of Research Information Systems, Penn State

State College, PA

DevOps Intern

May 2025 to Aug 2025

- Modernized legacy research administration systems by migrating backend services to **C# .NET**, writing parameterized **SQL**, and delivering **11** backend enhancements across internal research workflows.
- Established **CI/CD** pipelines on **Azure DevOps**, containerized services with **Docker**, and standardized deployments across **Linux** infrastructure, reducing deployment time by **40%**.

FEATURED PROJECTS

Bibi: Local First Agent Harness with Multi Model Routing

shariquekhatri.com/bibi | May 2026

- Engineered a **Python, Flask**, and **Electron** agent harness that routes each task across **4** models (**Whisper, Mistral, DeepSeek, Claude** via **Ollama**, plus **Claude**) by signal, cost, and latency through a custom agent loop.
- Wrote task specific scorers that rank candidate outputs from the **4** backends per request, showing that orchestration logic, not model size, drives practical agent reliability.
- Shipped the harness as an open source **Electron** desktop app with a runnable demo, adding human in the loop approval control over every action the agent executes.

Gmail Recruitment Pipeline: LLM Workflow Automation

Apr 2026

- Automated job search tracking with a **Python** agent that parses the full **Gmail** inbox through an **LLM**, writing structured pipeline state to **PostgreSQL** with retry logic and rate limit handling.
- Added a date scoped parsing mode over the **Gmail API** to classify a single day of messages on demand, replacing manual application tracking with a queryable system of record.

AI Travel Planner: Agentic Planner with Eval Harness

Feb 2026

- Created an agentic planner using **LangChain** orchestration and **RAG** over **50+** destination corpora with **MongoDB Vector Search**, wiring Skyscanner and Weather tool calls into the agent loop.
- Developed a Ragas style offline evaluation harness scoring **2** metrics (retrieval faithfulness, answer relevance) across the **50+** corpora, surfacing hallucination cases before they reached output.

Aman UAE: Verified News Platform with MCP Integrations

dubai-news-tracker.vercel.app | Oct 2025 to Jan 2026

- Launched a **Next.js** and **TypeScript** platform built end to end with **Claude Code**, using **MCP** server integrations for data aggregation and **LLM** source verification, reaching **26,000+** LinkedIn impressions at launch.
- Configured Git based continuous deployment on **Vercel** so pushes ship to production automatically, keeping a live real time news feed serving real traffic.