John Gimba

+1 (765) 430-3607 | LinkedIn | johngimba15@gmail.com | U.S. Citizen

EDUCATION

Purdue University West Lafayette, IN
Bachelor of Science in Computer Science Expected May 2027
Honors & Awards: Inspirit AI Scholar; Dean's List; multiple Best Student in Computer Science honors
Relevant Coursework: Object-Oriented Programming, Abstract Data Structures and Algorithms, Discrete Math, Linear Algebra, Introduction to Artificial Intelligence, Combinatorics, Machine Learning, Computer Networking
Organizations: AWS Cloud Club, Google Student Developer Club, National Society of Black Engineers

TECHNICAL SKILLS

Programming: Python, TypeScript, JavaScript, Java, C++, C, SQL, HTML/CSS
AI & Agents: PyTorch, Hugging Face Transformers, RAG, Tool Use, Model Routing, Evaluation
Frameworks & APIs: React, Node.js, GraphQL, REST APIs, Docker, Kubernetes
Cloud & Systems: AWS, DynamoDB, Fargate, Linux, Git, State Machines, Caching, Observability

EXPERIENCE

Amazon Web Services Herndon, VA
Software Developer Intern *May 2026 - August 2026*
Enterprise certificate customers needed one reliable interface to coordinate multiple stateful Certificate Authorities across issuance workflows.

- Developed a Java service and REST APIs that let customers manage Certificate Authority pools through one endpoint. Owned the backend path from data modeling through AWS Fargate deployment and console integration.
- Defined active and disabled states plus safe draining behavior. Added routing and failover for zero-downtime rotation. RDE and end-to-end tests verified recovery paths and state transitions.

Amazon Web Services Herndon, VA
Software Developer Intern *May 2025 - August 2025*
Certificate customers needed stable references while the cloud resources behind their workflows changed.

- Implemented Java APIs and DynamoDB-backed alias resolution so customers could use persistent human-readable identifiers. The mapping layer remained stable as underlying resources changed.
- Integrated aliases with certificate issuance and revocation across new and existing APIs. Wrote unit and integration tests for data consistency, invalid references, error handling, and backward compatibility.

Forterra Clarksburg, MD
Software Engineering Intern *May 2024 - August 2024*
An autonomous platform needed synchronized visual capture and interpretable sensor signals for downstream perception.

- Developed Raspberry Pi 5 software that coordinated eye-width cameras with precision motors and encoders. Encoder feedback synchronized complete 360-degree stereoscopic capture on a General Dynamics robotic platform.
- Created MATLAB acquisition and visualization workflows for a loop-antenna and ADC sensor system. Signal-processing filters improved electric-vehicle sensing accuracy by 35% and reduced noise by 20% against the raw-signal baseline.

PROJECTS

Agentic Workflow Orchestrator and Monitoring Console *2026*

- Multi-step agents needed predictable behavior when tools failed or required human approval. Engineered an asynchronous Python state machine that coordinated parallel tool calls and persisted workflow state.
- Defined typed tool schemas and policy controls. Added retries, timeouts, idempotency, and checkpoint recovery. A React and TypeScript console used a Node.js GraphQL API to inspect runs, approve actions, and replay structured traces.
- Added trace-based evaluations for task success and tool-call accuracy. Failure labels and latency measurements made unsuccessful runs easier to reproduce and debug.

Retrieval-Augmented Agent Memory and Evaluation Suite *2026*

- Long-running agents needed relevant context without resending complete histories. Developed a Python memory service that combined vector and keyword retrieval with metadata filters and checkpointed conversation state.
- Created replayable evaluations for retrieval precision and context recall. Separate checks measured groundedness and answer consistency.
- Added chunk deduplication and cache invalidation. Replayed prior runs after prompt or model changes to catch retrieval and generation regressions.

LLM Inference Gateway and Evaluation Harness *2025*

- Agent services needed one reliable way to route requests across model backends. Developed a Python API gateway with streaming responses, caching, rate limits, and policy-based routing.
- Added request batching, prioritization, and circuit breakers. Reproducible load tests measured time to first token, p50/p99 latency, cache hits, and recovery from backend failures.

Facial Expressions to Emojis *2024*

- Emoji selection needed a reliable way to translate facial expressions into emotion labels. Cleaned more than 1,500 images and used an OpenCV Haar Cascade classifier to detect faces.
- Trained and compared four supervised CNNs with different hyperparameters. The strongest model reached 83% accuracy and mapped predicted emotions to matching Apple emojis.

LEADERSHIP & PROFESSIONAL DEVELOPMENT

Management Leadership for Tomorrow (MLT) - Career Prep Fellow *2024 - Present*

- Selected for a competitive fellowship. Completed training in structured problem solving and behavioral interviewing. Strengthened professional networking.