

# Alan Reyes

314-478-1897 | [alanreyes6747@gmail.com](mailto:alanreyes6747@gmail.com) | [github.com/alanreyes24](https://github.com/alanreyes24)

## EDUCATION

---

### Washington University in St. Louis

*Bachelor of Science in Computer Science | GPA: 3.58*

St. Louis, MO

*Aug. 2024 – May 2028 (Expected)*

## EXPERIENCE

---

### Software Engineering Intern, AI R&D

*Enterprise Mobility*

May 2026 – July 2026

*St. Louis, MO*

- Designed a new AI-driven technology request feature for company devices by architecting an end-to-end RAG pipeline, streamlining enterprise workflows and optimizing data retrieval logic.
- Modernized legacy Java services by migrating them to Spring Boot in collaboration with Application Security, reinforcing enterprise security standards and reducing technical debt.

### Software Engineering Intern

*Nomad eCommerce*

Oct. 2025 – Present

*St. Louis, MO*

- Developed and optimized over 30 client web applications across the full stack using React and SQL, designing efficient database schemas to boost query performance.
- Partnered directly with stakeholders to translate business requirements into technical specifications, managing feature implementation, debugging, and deployment.

### AI / LLM Research Assistant

*Washington University in St. Louis*

Jan. 2025 – Aug. 2025

*St. Louis, MO*

- Engineered a Retrieval-Augmented Generation (RAG) pipeline over a 50,000+ document dataset, dynamically injecting the top 5–10 most relevant policy entries into the LLM context window.
- Deployed models using Hugging Face Spaces and implemented light fine-tuning alongside post-inference processing to enforce structured output generation for researchers.

## PROJECTS

---

### The Tech Digest | *Python, AWS (Lambda, Bedrock, S3, SAM), LangGraph, Pinecone, FastAPI, Next.js*

- Architected a serverless RAG agent (AWS Bedrock, LangGraph, FastAPI) with scale-to-zero infrastructure, enabling live streaming of reasoning steps to a Next.js UI via Server-Sent Events.
- Engineered a self-extending memory system that cross-corroborates web sources before citing, persisting verified data to a Pinecone vector index and AWS S3 to ensure highly accurate, hallucination-free repeat answers.
- Developed a staleness-aware retrieval system that escalates outdated queries to live web search, supported by an automated scraping pipeline covering 14+ RSS feeds with local Llama 3.1 scoring.

### BetterTyping | *JavaScript, React, Vite, Node.js, Express.js, MongoDB, Railway*

- Developed a full-stack React and Node.js typing test application, scaling the platform to support 5,000+ completed sessions across 150+ active users.
- Implemented secure authentication, session management, and interactive performance visualizations, allowing users to track speed and accuracy trends over time.

## TECHNICAL SKILLS

---

**Languages:** Python, JavaScript, Java, C/C++, SQL, HTML/CSS, Swift

**Frameworks/AI:** React, Next.js, Node.js, Express.js, LangChain, LangGraph, LlamaIndex

**Libraries:** Django, Flask, Pandas, NumPy

**Databases:** MongoDB, PostgreSQL, SQLite, ChromaDB

**Tools:** Git, Docker, Postman, Figma, AWS, CI/CD