

Russell Akram

Forward Deployed Engineer | Frontier GenAI (Claude, OpenAI) | Agentic Systems & Production AI
Arlington, VA • russellakram77@gmail.com • 571-595-9695 • linkedin.com/in/the-russell • russellakram.com

SUMMARY

Forward deployed engineer and co-founder who turns GenAI ambition into shipped, enterprise-grade software. 8 years building production systems, with 3+ years deploying LLM-powered solutions on the Anthropic Claude API and OpenAI API. I embed with stakeholders, translate ambiguous business problems into high-value use cases, and deliver working software with production-grade practices in testing, logging, versioning, and documentation. My products run in the real world: an AI form-automation tool piloted with the World Bank's risk assessment function, a live IAM governance platform that cuts document-processing effort by 94% and replaces \$150K to \$500K implementations at over 70% lower cost, and a production AI audit tool shipped end to end on AWS. I move with speed and precision in fast-paced, ambiguous delivery environments.

FRONTIER GENAI PLATFORM EXPERIENCE

- **Anthropic Claude API (production, 3+ yrs):** Claude Opus and Sonnet as the reasoning layer across five shipped systems. Tool use, extended thinking, structured output with cited evidence and confidence scoring, multi-step prompt-chain design, prompt libraries, and Claude Code for AI-assisted delivery.
- **OpenAI API (production):** GPT-4o and embeddings powering retrieval and generation in Bernie, with Assistants-style patterns and evaluation iteration.
- **Agentic & human-in-the-loop patterns:** multi-step agent workflows, tool-use orchestration, and human-in-the-loop review controls across FormSync AI, Apex Steward, ApexAudit, and Gnosis. Architecture decisions that balance quality, safety, latency, cost, and model risk through task-based multi-provider routing with cost enforcement.

EXPERIENCE

Co-Founder & Forward Deployed Engineer | RJ Autonomous, Arlington, VA *Sep 2018 – Present*

- **Embedded client delivery:** lead discovery with client stakeholders, translate high-value GenAI use cases into solutions, run working sessions to shape scope, and deliver working software from day one. Own engagements end to end as the primary technical contact across SMB, mid-market, and institutional accounts.
- **FormSync AI, production GenAI piloted with the World Bank's risk assessment function:** built a full-stack AI tool that converts meeting transcripts into pre-filled forms. Claude extracts field values from uploaded .docx and .xlsx forms with cited evidence and confidence scores, a human-review UI lets analysts verify and edit, and completed forms export to PDF, DOCX, or CRM webhook. Built-in zero-config demo mode runs end to end without API keys. Stack: Next.js 14, TypeScript, Supabase, Anthropic SDK, mammoth, SheetJS, pdf-lib.
- **Apex Steward, agentic GenAI platform with measurable enterprise impact:** architected a multi-tenant SaaS platform where Claude and Voyage AI extract roles, applications, and entitlements from vendor documents into a live role matrix and NIST CSF maturity dashboard. Compressed document ingestion by 94% (from roughly 8 hours to 30 minutes per policy), saving about \$750 per document and \$30K per mid-market engagement, and replacing \$150K to \$500K enterprise IGA implementations at over 70% lower total cost of ownership. Human-in-the-loop analyst approval built in. Live, with a partnership conversation underway.
- **ApexAudit, production AI delivered concept to launch:** shipped a two-tier AI audit tool, live at rjautonomous.com/ai-tools. The paid tier runs Python on AWS Lambda calling Claude via the Anthropic API, processing structured client inputs through a multi-step prompt chain to generate custom reports delivered by automated email, with an RDS and S3 data pipeline and Stripe payments. Iterated the prompt chain across 3 evaluation cycles to raise output specificity to a customer-actionable bar.

- **Cardinal Hub, multi-tenant enterprise SaaS:** full-stack conference-center management platform with Marymount University as the first design partner. Subdomain multi-tenancy with Postgres Row-Level Security, Microsoft Entra ID SSO with magic-link auth, a Claude-assisted event intake wizard, an AI report builder with XLSX and PDF export, and Stripe (card and ACH), check, PO, and cost-center payment flows. Stack: Next.js 14, TypeScript, PostgreSQL, Prisma.
- **Gnosis, full-stack agentic AI workspace (sole engineer, ~150 tasks shipped):** Next.js 14, TypeScript, Supabase Postgres with comprehensive RLS across roughly 15 tables and 22 migrations, pgvector retrieval, Voyage embeddings, and a multi-provider model gateway with BYOK and plan-tiered routing. Resolved a complex soft-delete RLS defect via SECURITY DEFINER functions. macOS Electron wrapper in active development.
- **Reusable assets and standards:** codify each engagement into reusable code, prompt libraries, runbooks, and reference architectures. Build with CI/CD, testing, logging, versioning, and documentation as defaults.

Senior IT Analyst | Marymount University, Arlington, VA

Sep 2022 – Present

- **Bernie, production GenAI assistant, lead developer:** built an AI assistant on Claude Opus and OpenAI API serving students, parents, faculty, and live campus tours, with an Azure-hosted RAG pipeline, vector retrieval, and a coordinate-based map layer. Cut hallucination on policy-sensitive queries across 3 retrieval and chunking iterations. Demonstrated to university leadership. Phase 2 adds SSO-integrated agentic ticketing.
- **Stakeholder translation in a regulated environment:** lead working sessions with non-technical sponsors in admissions, faculty, and facilities, translating ambiguous needs into shipped features under FERPA and NSA/DHS CAE-CDE institutional standards.
- **Operational impact:** drove 300+ endpoints to 100% security compliance in 6 months via Intune Autopilot and zero-touch provisioning.

Data Science Researcher (Seasonal) | Netflix

Aug 2022 – Dec 2022

- Built neural prediction models and large-scale data pipelines (Python, pandas, Spark), accelerated dataset processing 2x via parquet, and presented findings and 3D visualizations.

TECHNICAL SKILLS

- **Frontier GenAI:** Anthropic Claude API (Claude Opus and Sonnet, tool use, extended thinking, Claude Code), OpenAI API (GPT-4o, Assistants, embeddings), Groq Llama 3.3, Voyage embeddings, prompt engineering, prompt libraries, evaluation frameworks, LLMops and prompt management
- **Agentic & AI Architecture:** multi-step agents, tool use, human-in-the-loop controls, RAG pipelines, vector databases (pgvector, IVFFlat), multi-provider routing, structured output with confidence scoring
- **Languages:** Python (primary), TypeScript, JavaScript, SQL, Bash, PowerShell
- **Cloud & Platform Services:** AWS (Lambda, EC2, S3, RDS, IAM, VPC, networking), Microsoft Azure, Microsoft Entra ID, serverless, storage, compute, IAM
- **Engineering Practices:** CI/CD, testing, logging, versioning, documentation, REST APIs, microservices and event-driven integration patterns, Postgres and Supabase with RLS, Prisma, Stripe
- **Security & Compliance:** NIST 800-53, NIST CSF, CIS AWS Foundations, IAM least-privilege, FERPA, OAuth, AES-256-GCM, SSO and SAML

CERTIFICATIONS & EDUCATION

- AWS Certified Solutions Architect Associate (Scheduled May 2026) | CompTIA Security+ (March 2026) | AWS Certified Cloud Practitioner (August 2025)

M.S. Cybersecurity (in progress) | Marymount University — NSA/DHS CAE-CDE

Expected 2027

B.S. Cybersecurity (GPA: 3.9) | Marymount University

2018 – 2022

Study Abroad — Post-Quantum Cryptography | University of Oxford

2021