Sawyer Czupka

+1 (757) 818-0742 — sawyerczupka@gmail.com — linkedin.com/in/sawyerczupka

SKILLS

Programming & ML Technologies
Python, TypeScript, JavaScript, Java, C++, Rust, PyTorch, Scikit-Learn, Pandas, Hugging Face, React, Vue, FastAPI, Django, SQLAlchemy, PostgreSQL, TimescaleDB, MongoDB, Qdrant, LlamaIndex, DeepEval, vLLM, Azure OpenAI, Git

DevOps & Infrastructure
Docker, Kubernetes, Terraform, Packer, ArgoCD, GitHub Actions, CI/CD, Prometheus, Grafana, Traefik, MLflow, Azure, AWS, AWS Lambda, Linux, Nginx, Let's Encrypt

EDUCATION

College of William & Mary *Sept. 2021 – May 2025*
BS in Computer Science & Data Science, Cum Laude & Dean's List GPA: 3.6

EXPERIENCE

Software Engineering Intern, Luminexis AI / Threat Tec *Aug. 2025 – Dec. 2025*

- Built full-stack React/FastAPI application with JWT authentication, role-based access control, admin dashboards, and ROI tracking tools for identifying AI automation opportunities.
- Engineered large language model (LLM) powered document analysis pipeline using Azure OpenAI and Document Intelligence to extract requirements, risk factors, and action items from 140+ contract PDFs, reducing manual review time from hours to minutes per document.
- Deployed production retrieval-augmented generation (RAG) chatbot with Qdrant vector database and LlamaIndex achieving 70-80% on RAGAS evaluation metrics for domain-specific Q&A across configurable knowledge bases.
- Provisioned 4-environment Azure infrastructure (Dev, Staging, Demo, Production) using Terraform with shared modules, managing Container Apps, PostgreSQL, Key Vault, AI Foundry, and private networking; implemented CI/CD with GitHub Actions.

Machine Learning & Software Engineer, Teamculture.ai / L10.tech *Jan. 2024 – Sept. 2024*

- Developed serverless RAG evaluation system using AWS Lambda, API Gateway, and DeepEval to automate performance monitoring, reducing evaluation runtime from manual testing to automated pipeline execution.
- Built Vue/FastAPI interface for creating evaluations and curating golden example datasets, enabling systematic quality measurement and iterative improvement of AI services at scale.

Machine Learning Technical Lead, GeoLab @ William & Mary *Jan. 2023 – May 2025*

- Led cross-functional ML integration across frontend, backend, and ML subteams for SCOPE research platform, establishing technical requirements and coordinating implementation of RAG capabilities.
- Architected microservice-based RAG system on Kubernetes using Qdrant, vLLM, LlamaIndex, and FastAPI; built custom document processing pipeline with Vision-Language Models (Qwen-2.5-VL) for extracting text and structure from complex PDF/DOCX documents, reducing document parsing errors by over 50% compared to typical tools.

Machine Learning & Software Engineer Intern, The World Bank & GEF *May 2023 – Sept. 2023*

- Built RAG-based system automating analysis of 24,000+ GEF project documents for large-scale socioeconomic impact evaluation, enabling data-driven funding allocation and strategic planning decisions for the first time.

PERSONAL PROJECTS

LapEvo: iRacing Telemetry Tracker & AI Racing Coach — `https://lapevo.com` *Jan. 2025 – Present*

- Architected real-time telemetry pipeline ingesting iRacing simulator data at 60Hz using Python/FastAPI microservices with TimescaleDB hypertables for time-series storage and event-driven architecture with thread-safe asynchronous processing.
- Built full-stack monorepo with React 19/TypeScript dashboard, auto-generated OpenAPI client via Orval, and Docker Compose orchestration; implemented pytest with testcontainers for isolated integration testing, enabling CI/CD without live game dependencies.

Personal Homelab Infrastructure *July 2023 – Present*

- Maintain production-grade homelab with 6 VMs and 116-pod Kubernetes cluster on Proxmox, hosting services including Vaultwarden, Technitium DNS, Jellyfin, and Gitea with Traefik reverse proxy and automated Let's Encrypt TLS certificates.
- Leverage infrastructure-as-code with Packer, Terraform, Kubernetes manifests, and ArgoCD for GitOps-based continuous deployment, enabling rapid experimentation while maintaining consistent, reproducible environments.