

MICHAEL LEI

michael.lei@duke.edu | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

EDUCATION

Duke University

Durham, NC

B.S. Electrical and Computer Engineering, B.S. Computer Science, GPA: 4.0/4.0

May 2028

- Relevant Coursework: Software Design, Algorithms, Databases, OS, Networking, Machine Learning, Computer Architecture

EXPERIENCE

Forewarned

May 2026 – Present

Software Engineer

Durham, NC

- Architected a threat-intelligence engine for a **90-university** honeypot network, distilling **~1M Elasticsearch logs/day** into a **deduplicated Postgres entity graph** that collapses redundant sessions so each distinct threat is enriched once instead of per session.
- Built a **STIX 2.1 projection layer** with a custom **bounded graph-traversal** algorithm that powers exports and interactive threat-graph visualization, plus **TAXII 2.1 servers** with dynamic collections partners pull straight into their existing SIEMs.
- Designed fleet-wide **zero-day and campaign detection** pairing cross-campus **prevalence signals** with retrieval-augmented triage against known vulnerabilities to flag **candidate zero-days** and emerging campaigns.

MealMarkOS

Jan 2026 – Present

Co-founder & Engineer

Durham, NC

- Launched a multi-tenant B2B2C wellness platform piloting at **10+ universities**, where each school's branded instance serves live campus dining menus, meal and training logging, and a universal AI assistant.
- Designed the assistant's tool-loading architecture, organizing **100+ MCP tools** into namespaces like meals and gym, loading only the active tab's tools, with a `load_tools` meta-tool to pull others mid-turn, cutting schema payload **up to 80%**.
- Provisioned the multi-tenant AWS infrastructure on Terraform-managed ECS Fargate, isolating each university in its own Aurora Postgres schema and cutting new-school onboarding to **one CLI transaction**.

VSee

May 2025 – Present

Software Engineering Intern

Sunnyvale, CA

- Trained a medical-coding model over **8,000+ ICD codes**, deployed in a dockerized clinical AI service that cut physician charting time **~93%**.
- Authored a TypeScript Elasticsearch MCP server that lets non-technical employees query usage and telemetry data in natural language.
- Built an AI code-review service triggered by Bitbucket webhooks that posts line-cited issues directly on pull requests, running on **100+ production PRs across 3 repos**.

Chilkoti Group

Summer 2025

Machine Learning Intern

Remote

- Automated OpenCV preprocessing and circle-detection parameter tuning, cutting manual tuning from **20 to 2 minutes per slide** for cancer-screening ML training data.
- Productionized the lab's Jupyter notebook CV prototype into a deployed React annotation tool, letting researchers label slides interactively without touching code.

PROJECTS

PublicCVE | *Python, FastAPI, TypeScript, Postgres, pgvector, Voyage embeddings* | [Live](#)

- Engineered a vulnerability platform that ingests and normalizes **330K+ CVEs** from multi-source threat feeds into one Postgres corpus, enriched in real time from security news, advisories, and exploit activity.
- Developed a **hybrid search** that fuses semantic and keyword results with **reciprocal rank fusion**, letting analysts and API clients retrieve CVEs in plain language; ranked over Voyage embeddings in pgvector.
- Matched live attack payloads to the right CVE using **HyDE RAG**, constraining the model to a fixed schema so it can never invent an ID, and exposed it to Claude and Cursor through a **custom TypeScript MCP server**.

PyDeid | *Python, FastAPI, React, HuggingFace, Presidio* | [GitHub](#) · [Demo](#)

- Engineered a clinical PHI de-identification platform that merges regex, ML, and LLM detectors into one ensemble to catch PHI any single method would miss, plus a fine-tuning toolkit to train and drop in custom token-classification detectors.
- Structured it as one FastAPI backend with dual-scoped API keys driving two React apps, an admin pipeline builder and a production review workspace, shipped as a self-hosted Docker image.

TECHNICAL SKILLS

Languages: Python, TypeScript / JavaScript, Java, Swift, C++, SQL

AI / ML: LLM fine-tuning, RAG, MCP, AI agents, semantic search, PyTorch, HuggingFace, CoreML, NER

Backend: FastAPI, React, Node.js, REST APIs, WebSockets, Microservices, PostgreSQL, Elasticsearch

Cloud & DevOps: AWS, Docker, Kubernetes, Terraform, GitHub Actions, CI/CD, Prometheus, Linux, Git