

# Gabriel Mousa

Seattle, Washington | (954) 279-3664 | gab.01@hotmail.com | <https://www.linkedin.com/in/GabeMousa> |  
<https://github.com/gabe-mousa/> | <https://gabe-mousa.github.io/gabemousa-website/>

## Summary

---

Software Engineer with distributed systems expertise seeking AI Research Engineer, ML, or Backend-SWE roles. Built open-source AI safety evaluation tools implementing SOTA interpretability and alignment research. Experienced in scaling infrastructure to millions of users and automating complex ML pipelines.

## Experience

---

### Software Engineer II

*Microsoft — Seattle, Washington | October 2023 – Present*

- Architected distributed systems to account for large-scale token distribution for authentication and authorization of over 5M+ machines per month. This included complex request routing and scaling solutions to manage load balancing traffic from 4 Azure services across 9 regions with custom caching and networking solutions to support specific partner scenarios.
- Synchronized logs and metrics for service performance data at billions of requests/month scale and across 7 Azure services improving end-to-end traffic flow tracking. Thus bringing incident resolution time down by 60% on average and allowing for improvements to service quality by +20% and up to +60% depending on the service, over the period of 6 months.
- 0 to 1 product development for infrastructure buildout automation which leverages MCP Client and Servers, CI/CD pipelines, CosmosDB thread tracking, Ev2, and Kubernetes workers. Automating infrastructure rollout to bring new region rollout times down from ~2 months to ~2 weeks per region.
- Designed custom horizontal autoscaling for message processing when customer load exceeds infrastructure capabilities with automated software deployments to handle scaling of 1M+ customers per region. This included custom networking and token solutions to enable resource allocation for over 9 active regions, relieving scaling concerns as service grew at over 100% YoY.

### Software Engineer

*Dignitas Technologies — Orlando, Florida | November 2022 – September 2023*

- Developed SDKs for C, C++, Java, and Python for users to interact with Dignitas simulation software. Working low-level on compilers and building systems to support 6 releases across 3 Linux distros and Windows.
- Introduced API integrations for 15+ endpoints to handle customer traffic with rate limiting and auth/authz.

### Software Engineering Intern

*Microsoft — Redmond, Washington | May 2022 – August 2022*

- Working as a part of Azure Containers Instances team to build custom resource tracking for child resources handling 20,000+ resources per month.
- Implementation of custom resource tracking across 3 Microsoft Azure services.

## Skills

---

**Languages:** Python, Typescript, Golang, C++, C#/.NET, Java

**Frameworks and Packages:** TensorFlow, Scikit-learn, XGBoost, Semantic Kernel, Pandas, NumPy, FastMCP, Python Notebooks

**Infrastructure:** Kubernetes, Docker, Azure, Azure DevOps, GCP, AWS, Ev2

**Research Areas:** Mechanistic Interpretability, CoT Faithfulness, Sycophancy Evaluation, AI Alignment

## Projects

---

### Apolien — AI Safety Evaluation Library for Python:

<https://github.com/gabe-mousa/Apolien> — Actively maintain AI Safety Evaluation Library implementing SOTA alignment research methodologies:

- Implemented post-hoc Chain-of-Thought faithfulness analysis based on Anthropic's attribution graph methodology and circuit tracing research for mechanistic interpretability evaluation.
- Built sycophancy evaluation techniques informed by Anthropic's reward tampering research, detecting specification gaming and deceptive patterns in model outputs.
- Benchmarked against frontier models including GPT-5.1 Family, Claude Family, and Llama family.
- Published technical blog posts on LinkedIn documenting research findings and implementation approaches.

## Education

---

### University of Central Florida — Orlando, FL

*B.S. in Computer Science*

**Relevant Coursework:** Machine Learning, Matrices and Linear Algebra, Advanced Statistics