

Michael Rawson

Seattle, WA · michael.rawson0@gmail.com · [linkedin.com/in/michael-rawson](https://www.linkedin.com/in/michael-rawson)

Summary

Ph.D. applied mathematician and applied AI scientist with 6+ years building, fine-tuning, and evaluating machine learning systems in production at Amazon, Microsoft, Meta, and PNNL. Hands-on across the full AI lifecycle: LLM and agentic system development (MCP agents, MoE transformers, RAG), foundation model fine-tuning and post-training, reinforcement learning, and rigorous statistical evaluation. Deep background in statistics, experimentation, and uncertainty quantification, applied to enterprise-scale AI in security, logistics, and speech domains. Published researcher (ICML, ICLR, SPIE, JMM) and reviewer for ICML, NeurIPS, and Amazon AMLC.

Core Skills

Generative AI & LLMs: LLM development and evaluation, AI agents (MCP), prompt engineering, context engineering, Retrieval-Augmented Generation (RAG), foundation model fine-tuning and post-training, Mixture-of-Experts transformers, GPT and Phi model families, HuggingFace

ML & Statistics: Statistical modeling and inference, experiment design and A/B testing, causal and Lipschitz-bound analysis, predictive analytics, uncertainty quantification, reinforcement learning (deep Q-learning, bandits), anomaly detection, time series forecasting, model pruning and optimization, drift detection and retraining

Speech & Multimodal: Speaker and speech recognition, adversarial robustness for speech models, multimodal information fusion, hyperspectral and MRI signal processing

Engineering: Python, PyTorch, TensorFlow, C++, CUDA, SQL, Spark, Docker, AWS, SLURM/HPC, MATLAB, data pipelines, production model monitoring

Responsible AI: Adversarial robustness, safety and security evaluation for enterprise AI, model reliability and governance

Professional Experience

AI Applied Scientist — Amazon

Seattle, WA · May 2025 - Present

- Build agentic AI experiences on the Model Context Protocol (MCP), designing tool-use workflows and interaction patterns for long-horizon logistics tasks.
- Develop Mixture-of-Experts transformer models for hierarchical product customs classification supporting global trade, including fine-tuning and post-training to lift task-specific accuracy and enterprise readiness.
- Curate and structure training and evaluation datasets for hierarchical classification, defining benchmarks and quality metrics used to track model improvement.
- Led the Tinc Hackathon team on theft and anomaly detection, placing top 5 nationally.
- Partner with engineering, product, and business stakeholders to define model success metrics and translate evaluation findings into roadmap decisions.

ML Research Engineer — Microsoft

Seattle, WA · April 2024 - May 2025

- Built end-to-end RAG and LLM solutions for enterprise cyber security and safety analysts using OpenAI GPT and frontier Phi models, covering retrieval strategy, prompt and system instruction design, and evaluation.
- Developed automatic prompt differentiation methods for systematic prompt optimization; presented at the Microsoft Security AI Research (MSECAIR) Workshop.
- Designed evaluation approaches for LLM output quality, reliability, and safety in a high-stakes security context, identifying failure modes and closing them through context engineering and retrieval improvements.
- Maintained and retrained production NLP models under data and model drift, monitoring performance telemetry and triggering retraining cycles in PyTorch.
- Communicated technical findings to engineering, product, and executive stakeholders; served as a Microsoft Hackathon judge.

Machine Learning Researcher — Pacific Northwest National Laboratory

Seattle, WA · April 2022 – April 2024

- Built deep learning systems for speaker and speech recognition, including adversarial evaluation to measure and improve model robustness — directly applicable to voice and multimodal AI quality work.
- Developed a knowledge and ontology system mapping semantic distance via a lattice metric, supporting retrieval and structured reasoning over technical corpora.
- Trained, evaluated, tested, and deployed models across LLMs, graph/network transformers, anomaly detection, and time series forecasting on GPU, SLURM, and HPC infrastructure.
- Published first-author and collaborative work at ICML TAG, ICLR Physics4ML, and SPIE on efficient model architectures, neural network pruning, and fast hyperspectral abundance estimation.

Research Assistant — School of Medicine, University of Maryland

College Park, MD · June 2020 – April 2022

- Developed high-dimensional MRI reconstruction methods robust to noisy and corrupted data, combining signal processing, mathematical modeling, optimization, and deep learning.

Machine Learning Scientist Fellow — Army Research Lab

Adelphi, MD · Sept 2020 – Sept 2021

- Designed and implemented a deep reinforcement learning system for multimodal information fusion across heterogeneous sensor sources, with uncertainty quantification to bound decision confidence.

Machine Learning Engineer Intern — Facebook (Meta)

Menlo Park, CA · May 2021 – Aug 2021

- Applied deep Q-learning and deep bandit reinforcement learning to Facebook Marketplace search ranking, improving simulated search performance by 40%.
- Built offline simulation and evaluation infrastructure to compare ranking policies before production exposure. C++, Cython, Python.

Research Assistant — Norbert Wiener Center, University of Maryland

College Park, MD · Aug 2017 – May 2020

- Analyzed deep neural network stability and generalization using Lipschitz bounds; investigated NLP representations with topological data analysis. C++, CUDA, Python.

Software Engineer — Bloomberg LP

New York, NY · Feb 2014 - April 2016

- Modeled financial derivatives data with machine learning over Spark-based data pipelines; built volatility surface approximation and trading order execution/matching systems. C++, SQL, TDD, Agile.

Earlier Roles

- **Intern Research Assistant**, Earth Sciences, NASA GSFC — Greenbelt, MD · June 2017 - Aug 2017
 - **Research Assistant**, Courant Institute, NYU — New York, NY · May 2016 - May 2017
 - **Software Engineer Intern**, Amazon.com — Seattle, WA · May 2013 - Aug 2013
-

Education

Ph.D., Applied Mathematics and Statistics — University of Maryland, College Park · May 2022

M.S., Mathematics — Courant Institute of Mathematical Sciences, New York University · Aug 2017

B.S., Computer Science — Herbert Wertheim College of Engineering, University of Florida · Dec 2013

Selected Publications & Talks

- **Automatic Prompt Differentiation**. Microsoft Security AI Research (MSECAIR) Workshop, Oct 2024.
 - Chung, F., **Rawson, M. G.**, Shen, Z., Xu, M. *The Ricci Curvatures and Fast Algorithms for Random Clustering Graphs*. Joint Mathematics Meetings (JMM), Jan 2024.
 - **Rawson, M.** *Software Architecture: Programming Paradigms with AI Applications*. Hacker Dojo Python Workshop, Nov 2023. DOI 10.13140/RG.2.2.13082.21441.
 - Godfrey, C., **Rawson, M. G.**, Brown, D., Kvinge, H. *Fast Computation of Permutation Equivariant Layers with the Partition Algebra*. ICML TAG 2023; ICLR 2023 Physics4ML.
 - **Rawson, M. G.**, Doster, T., Emerson, T. *Fast, Accurate, High Information Hyperspectral Abundance Methods via Radial Basis Functions*. Proceedings of SPIE Defense + Commercial Sensing, May 2023.
 - **Rawson, M. G.** *Renormalized Sparse Neural Network Pruning*. Sparsity in Neural Networks Workshop, July 2022.
-

Open Source & Community

- **Open source:** PyGraphCurvature, Petri Nets (GitHub)
- **Reviewer:** ICML, NeurIPS, MLADS, TAG, JMM, Amazon CSS/AMLC
- **Organizer:** Joint Mathematics Meetings (JMM); Microsoft Hackathon judge; hackathon mentoring