Jack Geraghty

Professional Profile

Research Engineer with a **PhD specializing in the robustness and evaluation of multimodal systems** under degraded or missing inputs. I possess a deep, first-principles understanding of machine learning fundamentals, from mathematical foundations to implementing custom architectures in PyTorch. My **experience spans high-scale production environments** at Amazon, maintaining 99.999% uptime, to **research collaborations on beyond-5G networks and real-time motorway control systems**. I combine academic rigor with low-level systems proficiency in Rust and Python to build reliable, high-performance ML pipelines.

Education and Research

PhD in Computer Science | University College Dublin 2021 — 2025

- **Thesis:** *Learning to Associate: Handling Missing Modalities in Multimodal Systems.*
- Developed **lightweight reconstruction** and stress-testing methods to analyze model behavior under degraded, missing, or manipulated inputs.
- Researched **post-training techniques to generate missing data**, enhancing model robustness and adaptability in decentralized and federated environments.

B.Sc. in Computer Science | University College Dublin 2016 — 2021

- First Class Honours; Dissertation: *Analysing the Energy Consumption of Websites.*

Industry and Research Experience

Software Engineer | Amazon 2019 — 2020

- Contributed to **large-scale observability systems** processing billions of data points per AWS region per hour to ensure high reliability for AWS products.
- Integrated a **statistical anomaly detection** method into production for **proactive identification of anomalous behavior on devices**.
- Optimized system performance through deferred execution methods and performance tuning in latency-sensitive distributed environments.

Applied ML Research Collaborator | InterDigital 2021 — 2022

- Investigated the intersection of multimodal ML and **beyond 5G networks**, focusing on robustness issues caused by degraded inputs and distribution shifts in intelligent transport.
- **Researched the transition of ML solutions** from controlled environments to real-world, high-stakes network applications.

ML Research Collaborator | Roughan & O'Donovan Engineering 2022 — 2024

- Applied ML and statistical methods to **predict traffic flow breakdowns**, integrating results into live control systems **on Ireland's busiest motorway**.
- **Investigated failure modes and representation stability** in operational prediction models using real-time sensor data and network telemetry.

Technical Skills

- **ML Fundamentals:** Deep understanding of backpropagation, attention mechanisms, and optimization; expertise in model robustness, multimodal machine learning, federated learning, model calibration, and uncertainty analysis.

- **Engineering:** Proficient in Python, Rust with strong knowledge of C, and C++; experienced in performance engineering, SIMD optimization, and building deterministic data pipelines.
- **ML Tooling:** PyTorch, Candle and MLOps; experience with multi-GPU training and HPC orchestration (SLURM).

Selected Projects

Multimodal Stress-Testing & Behavioural Analysis Toolkit

- Built a modular PyTorch framework for **systematically stressing models under degraded inputs**, implementing reconstruction-based probes to analyze representation stability.
- **Supports both centralized and decentralized** multimodal learning under missing modality inputs.

AudioSamples: A high-performance audio processing toolkit

- Developed a comprehensive **Rust-native audio processing ecosystem** providing type-safe, high-performance sample representations, conversions, statistical analysis and DSP primitives with intrinsic metadata handling — addressing common pain points in audio workflows by preserving sample rate, channel layout and format throughout processing.
- Extended the core toolkit with modular crates for **efficient audio file I/O** in Rust (audio_samples_io), including benchmarks against established libraries, **and seamless Python bindings** (audio_samples_python) to expose powerful audio generation, analysis, editing and resampling capabilities to the Python ecosystem.
- Enabled **cross-language interoperability and reproducible research workflows** through PyO3-based bindings that embed audio properties directly into Python objects, simplifying common tasks like signal generation, format conversion, multi-channel manipulation and high-performance I/O.

Selected Publications

- **Learning to Associate: Multimodal Inference with Fully Missing Modalities** (2025)
  Geraghty, J., Hines, A., & Golpayegani, F. *ACM TIST*. Lightweight post-training models for reconstructing missing modalities in large-scale ML systems.
- **Understanding the Relevancy of Modality Information in Multimodal ML** (2023)
  Geraghty, J., Hines, A., & Golpayegani, F. *ECAI*. Empirical analysis of modality contribution and inference robustness.
- **Intelligent Shared Mobility Systems: A Survey on Whole System Design Requirements, Challenges and Future Direction**
  Golpayegani, F., et al. *IEEE Access*. A systems-level survey of shared mobility systems, analysing how AI, optimisation, and autonomous decision-making techniques jointly address interconnected challenges across multi-modal transportation as a unified whole.
- **AQP: An Open, Modular Python Platform for Objective Speech and Audio Quality Metrics** (2022)
  Geraghty, J., Li, J., Ragano, A., & Hines, A. *ACM MMSys*. Modular evaluation platform for audio quality used by multiple research groups.

Teaching & Awards

- **Teaching Excellence Awards (2021–2023):** Recognized for delivering lectures on algorithms and concurrent programming to 100+ students.
- **John Kelly Memorial Award (2021):** Awarded for the highest academic performance in the graduating class.

References available upon request.