

Jarod Bloch

Pittsburgh, PA • 484-550-5079 • jarodb@andrew.cmu.edu • [linkedin.com/in/jarodbloch](https://www.linkedin.com/in/jarodbloch) • jarodblo.ch

Education

Carnegie Mellon University, School of Computer Science

May 2027

Bachelor of Science, Artificial Intelligence

Pittsburgh, PA

GPA: 3.3, Dean's List: Fall 2024 (High Honors), Fall 2025 (High Honors)

Awards: 3rd Place, CASI × Gray Swan AI Jailbreaking Hackathon (Fall 2025); 6th Place, top solo vs. teams of 2–3 (Spring 2026)

Member, **Plaid Parliament of Pwning (PPP)**, multi-time DEF CON CTF champion team, consistently world top-ranked

Coursework: ML for Cybersecurity (grad), Computer Security (15-330), Autonomous Agents, Natural Language Processing (grad)

Publications

Cheng, Y. F.*, **Bloch, J.***, et al. "Auditorily Embodied Conversational Agents: Effects of Spatialization and Situated Audio Cues on Presence and Social Perception." *Published in Proceedings of ACM CHI 2026*. <https://doi.org/10.1145/3772318.3791794>

*(Co-first author; co-presented oral talk at ACM CHI '26)

Experience

Connected Vehicle Software Engineering Intern | Python, LangGraph, RAG, BigQuery, Pandas May 2026 - August 2026
Ford Motor Company Dearborn, MI

- Building an agentic framework (Python, LangGraph) that orchestrates Ford's internal LLM platform (FordLLM) to predict severity classifications for vehicle diagnostic trouble codes (DTCs), helping the triage team prioritize EV issues.
- Designing the RAG pipeline (diagnostic documentation and live EV telemetry via BigQuery, Pandas) so the agent reasons from objective evidence rather than prior labels, establishing a consistent ground-truth for a previously subjective process.

Research Assistant | Python, Docker, vLLM May 2025 - May 2026
Carnegie Mellon University Pittsburgh, PA

- Conducted adversarial ML research on LLM-based vulnerability detection under Prof. Lujo Bauer (CyLab), evaluating recall and failure modes against 100+ semantics-preserving JavaScript perturbations.
- Optimized a containerized vulnerability benchmarking framework, increasing self-hosted LLM inference throughput by 8.6x.

Teaching Assistant | C, Google Apps Script May 2025 - August 2025
Carnegie Mellon University Pittsburgh, PA

- Served as head of code reviews for 220+ primarily graduate-level students in Computer Systems (15-213/15-503), designing rubrics and coordinating sessions to evaluate C-based memory allocators, proxy servers, and shell programs.
- Built an automated TA management pipeline in Google Apps Script (roster imports, tracking, statistics) whose auto-generated mailing lists replaced manual lookup, saving the 13-person staff ~3 hrs/week and adopted as infrastructure for future semesters.

Projects

Repo-Level Type Inference for Security | Python, CodeQL, NetworkX, vLLM, Slurm February 2026 - May 2026

- Architected a language-agnostic neuro-symbolic pipeline (CodeQL + LLMs) that recovers types across repositories exceeding frontier-model context limits, injecting runtime isinstance guards to harden dynamically-typed code against type-confusion bugs.
- Reduced cross-file AI type errors by 37% by mapping dependency DAGs to inject explicit type signals into context.
- Optimized annotation quality via large-scale Slurm GPU ablation studies on 50 real-world repositories across Qwen3 model sizes (0.6B–30B), evaluating seeded root nodes, model types (Coder vs. Instruct), and context augmentation strategies.

Indirect Prompt Injection Portfolio Site | HTML, CSS, JavaScript February 2026

- Designed and deployed a multi-vector indirect prompt injection demonstration exploiting DOM text extraction, image alt-text, and metadata pipelines using invisible Unicode obfuscation to test tokenizer resilience (OWASP LLM01).
- Deployed an injection payload that caused 4 leading LLMs (Claude, ChatGPT, Gemini, Copilot) to execute benign hidden instructions in the page to educate end-users when asked to summarize or analyze it, highlighting systemic weaknesses.

Discord E2EE Protocol Reverse Engineering | JavaScript, WebSocket, UDP, MLS March 2026

- Restored full voice functionality to a deprecated Node.js user client after Discord's mandatory DAVE E2EE migration broke unofficial audio support, delivering stable real-time audio in live group-DM calls.
- Reverse-engineered the undocumented voice-gateway changes from live WebSocket and UDP traffic and implemented DAVE end-to-end (MLS group key exchange, binary frame encryption, and session lifecycle management).

Skills

AI Security & Red Teaming: Adversarial ML, Prompt Injection, RAG Poisoning, MITRE ATLAS, OWASP LLM Top 10, garak, PyRIT

Offensive Security & AppSec: Reverse Engineering, Binary Exploitation, Penetration Testing, SAST/CodeQL, Threat Modeling

Machine Learning & AI: PyTorch, scikit-learn, Hugging Face, vLLM, LoRA/PEFT, DPO, RAG, Agentic Systems, Anomaly Detection

Programming Languages: Python, C, C++, Java, JavaScript/TypeScript, C#, SML, SQL

Infrastructure & MLOps: Docker, Slurm/HPC, AWS, Oracle Cloud, Weights & Biases (W&B), FastAPI, Linux/Bash, Git