# Allen Thomas

(217) 298-6572 | allenthomasdev@gmail.com | linkedin.com/in/allenthomasdev | github.com/AllenThomasDev

## Education

**University of Illinois at Urbana-Champaign** — Urbana, IL
*Master of Computer Science, GPA: 3.85/4.0* — *August 2024 – May 2026*

**University of Pune** — Pune, India
*Bachelor of Engineering in Computer Science, GPA: 8.49/10.0* — *August 2017 – May 2021*

## Experience

**Founding Engineer** — Remote
*Cerebion* — *May 2025 – Present*

- Architected a desktop code review platform with React TypeScript, FastAPI with Pydantic validation, analyzing large enterprise codebases for vulnerable security patterns.
- Engineered a hybrid repair engine combining Semgrep's deterministic static analysis with LLM reasoning, automating fixes for complex vulnerabilities where standard linters fail.
- Built a fault-tolerant patching system with atomic Git ops + automated rollbacks, ensuring 100% repository integrity during autonomous code repair.

**Software Engineer** — Pune, India
*Helpshift* — *June 2022 – June 2024*

- Raised analytics uptime from **99.0% to 99.99%** and **cut $250,000/yr** by migrating analytics infrastructure to AWS.
- Re-architected analytics database from HBase to Redshift OLAP warehouse, refactoring APIs and query patterns to enable 10× traffic growth for 200+ customers with zero downtime.
- Eliminated stream processing bottlenecks affecting real-time analytics by migrating legacy Storm infrastructure to Flink, **reducing analytics event latency by 35%** for 40K+ support agents.
- Preserved **350+ TB of historical data** during migration, maintaining 6+ years of customer analytics access.
- Established Airflow standards and documentation across 5+ teams, saving 15 developer hours weekly.
- Reduced ad-hoc engineering data requests by 40% by implementing Metabase self-service analytics platform.
- **Mentored 10+ hires on coding practices**, reducing time to first release by 35% compared to previous year.

## Projects

**Control Vector-Based LLM Steering** | *Python, PyTorch, Transformers* — *March 2025*

- Implemented Activation Engineering to steer behavior by extracting steering vectors via PCA on contrastive prompt pairs.
- Built an automated evaluation pipeline to stress-test model coherence at varying control strengths, ensuring structural integrity of JSON outputs while altering persona.

**Distributed Stream Processing Platform** | *Golang, Distributed Systems* — *September 2024 – November 2024*

- Implemented SWIM failure detection and consistent hashing to manage dynamic node churn, **achieving sub-3s convergence** for cluster membership updates.
- Designed a **custom distributed file system** (HyDFS) with chain replication, ensuring linearizability for concurrent appends across 10+ nodes.
- Implemented a stream processing engine (RainStorm) with **exactly-once semantics**, utilizing distributed write-ahead logs to track tuple lineage and handle worker failures.
- Engineered autoscaling resource manager that dynamically provisioned worker nodes based on throughput watermarks, optimizing cluster utilization under varying load.

## Awards & Research

**Open Philanthropy Grant Recipient** | *AI Safety, Mechanistic Interpretability* — *September 2025 – Present*

- Awarded competitive funding to investigate **mechanistic causes of alignment-faking** in Large Language Models.
- Scope includes identifying specific patterns that trigger deceptive behavior during chain-of-thought reasoning.

**Multi-Agent Reinforcement Learning: Hide and Seek** | *Reinforcement Learning* — *2021*

- Researched and implemented multiple reinforcement learning algorithms for multi-agent systems, developing a novel hide-and-seek simulation environment inspired by OpenAI research on Multi-agent Autocurricula.

## Technical Skills

**Languages**: Python, Go, Clojure, JavaScript
**Frameworks**: React, FastAPI, Flask, PyTorch, Transformers, React Native
**Infra/Devops**: AWS, Docker, Kubernetes, Airflow, Apache Flink, CI/CD
**Databases & Storage**: PostgreSQL, MySQL, MongoDB, Apache HBase, Redis, Kafka, Redshift, Distributed Storage
**Architecture & Systems**: Microservices, Distributed Systems, Stream Processing, RESTful APIs, Event-Driven Architecture