

# Vihaan Gupta

[linkedin.com/in/vgupta06](https://www.linkedin.com/in/vgupta06) | [vihaangupta@college.harvard.edu](mailto:vihaangupta@college.harvard.edu) | +1(857)-829-8855 | [vihaangupta.dev](https://vihaangupta.dev)

## Education

### Harvard University

May 2027

- B.A. in Computer Science with a Joint Concentration in Mathematics and a secondary in Philosophy; GPA: 3.87/4.0.
- Relevant Coursework: Distributed Systems, Systems Programming, Software Engineering with AI, Human-Computer Interaction, Differential Privacy, Data Structures & Algorithmic Theory, Probability, Computer Graphics.
- Activities and Leadership: Teaching Fellow for Data Structures and Algorithms, President of Tech for Social Good.

## Skills

**Programming Languages:** Python, C/C++, Swift, Java, JavaScript/TypeScript, C#, Go, Linux Script, Dart, SQL.

**Frameworks & Tools:** Spring Boot, Kubernetes, Docker, Apache Kafka, Jenkins, React, NumPy, Flutter, Node.js, Git.

## Experience

### MIAX Exchange

Jun - Aug 2025

#### Software Engineering Intern

Princeton, NJ

- Designed and deployed a Kafka-Spring Boot microservice for real-time trade event streaming, achieving sub-millisecond message latency and ensuring end-to-end delivery guarantees under high throughput.
- Implemented idempotent offset tracking and rollback-safe Redis caching to preserve data integrity across failovers, cutting recovery time from 2 minutes to 1 second.
- Integrated kubeLint into Jenkins CI/CD to automatically scan Kubernetes config on pull requests and developed custom linting rules tailored to compliance policies, detecting misconfigurations in 70% of PRs.

### As1 Social

May 2025

#### Founding Software Engineer

Cambridge, MA

- Engineered 10+ features for the MVP of social media app in React Native (frontend) and PostgreSQL (backend).
- Boosted user retention by 30% by devising caption parsing and hashtag ranking for video recommendations, enabling personalised feeds based on preferences, keyword relevance, and trend recency.
- Architected adaptive video streaming using HLS with multi-bitrate transcoding, dynamically adjusting playback quality to users' bandwidth.

### Avalon Infosys

May - July 2024

#### Software Engineering Intern

New Delhi, India

- Created mobile interface for the existing OpenEMIS dashboard, enabling users to locate schools across Nigeria.
- Integrated OpenStreetMap with Flutter to render vector-based school markers, cutting map load times by 40%.
- Built for performance and offline-first usability, enabling 90% of core features to work without connectivity.

## Projects

### Lumi

Nov 2025 - Present

- Built Lumi, a context-aware AI reading and writing assistant using folder-scoped retrieval-augmented generation (RAG), enabling citation-grounded drafting and editing across 10+ document formats.
- Developed a scalable document ingestion and retrieval pipeline (OCR, chunking, embeddings, PostgreSQL + pgvector) capable of serving low-latency (<2s) retrieval-augmented queries at interactive speeds.

### Graphene

Sept - Dec 2025

- Built and launched Graphene, a production research-discovery platform that models academic literature as an interactive knowledge graph, supporting 2.4M+ papers through the arXiv API.
- Optimised graph rendering and querying, supporting interactive exploration of 100-500 nodes per session.

### RL for Inventory Control

Jan 2025

- Implemented tabular Q-learning for a stochastic inventory MDP, learning a near-base-stock policy that significantly outperformed heuristic baselines under uncertain demand patterns.
- Improved RL stability via adaptive learning rates and exploration tuning, scaling performance to longer horizons.