

AMEYA RANADE

(413) 479-9346 | ranadeamr@gmail.com | ameyaranade.com | linkedin.com/ameya-ranade

EDUCATION

University of Massachusetts Amherst

Master of Science in Computer Science (GPA: 3.92 / 4.00)

May 2025

Amherst, MA

University of Mumbai (Sardar Patel Institute of Technology)

Bachelor of Technology in Information Technology, Minor in Management (GPA: 3.72 / 4.00)

May 2023

Mumbai, IN

EXPERIENCE

Intellect Design Arena

Software Engineer

Jul 2025 – Present

Piscataway, NJ

- Ship feature improvements and targeted fixes for **Magic Submission**, a multi-tenant document intake platform processing **500+** submissions daily, reducing extraction failure rate from **12%** to **5%** across **9** tenants.
- Maintained **99%+** uptime across distributed Kubernetes and Lambda services using Grafana dashboards and Cloudwatch logs, owning on-call triage, root-cause analysis, and production hotfixes.
- Launch **10+** underwriting workflows end-to-end in Spring Boot for **Xponent**, Insurance Underwriting Workbench, enabling two carrier clients to reduce decision-making time by **60%** and lower operational costs by **10%**.
- Developed an asynchronous audit trail using Spring Boot and AWS SQS, offloading payload diffing to background workers, cutting audit review time by **65%**.

Apple

Graduate Researcher

Feb 2025 – Jul 2025

Amherst, MA

- Spearheaded **ELAPSE** (Efficient Approach to End-to-End Salient Entity Linking), a web-scale system for Safari news indexing, achieving **+24 F1** over prior SOTA while running **3x** faster than LLM baselines.
- Built three ELAPSE variants (flash/base/large) spanning latency-accuracy frontier, extending [ReFinED](#) framework with bi-encoder salience prediction, trained on **100M** document-mention pairs over a **6.2M** entity knowledge base.
- Optimized inference for salience prediction, restricting expensive entity disambiguation (ED) candidate scoring to salient mentions only and cutting per-document ED compute by **75%** relative to baseline.

Autire

Software Engineer

Jun 2024 – Dec 2024

Dallas, TX

- Built **SpecGenie**, an AI platform compressing multi-week document workflows into **under 4 hours** for engineering and construction teams globally, shipped to Fortune 500 energy major.
- Implemented **10+** end-to-end features and core extraction service leveraging AWS, LLMs, LangChain, and Pydantic, achieving **98%** field-level accuracy across **10+** heterogeneous vendor document layouts.
- Engineered transformation engine with Pydantic validation schemas and Celery batch processing, producing enterprise data ready for bulk ingestion into downstream 3D design tools, eliminating **95%** of manual data entry.
- Designed human-in-the-loop review interfaces in React and Material UI integrating **15+** REST endpoints, accelerating approval cycles from days to hours per document batch before 3D-tool ingestion.

LokaVidya Technologies

Software Engineer Intern

Jul 2021 – Jan 2022

Mumbai, India

- Developed **LokaVidya Meet**, a video conferencing platform using Node.js and BigBlueButton APIs to deliver **6+** meeting features including breakout rooms, session recording, and role moderation for secure virtual classrooms.
- Migrated Role-Based Access Control to Redis for session and role caching, reducing access latency by **87%** and offloading Azure SQL queries on subsequent requests.

PROJECTS

Sift | Typescript, Next.js, Python, FastAPI, Uvicorn, Gemini, GCP, Vertex AI, Vercel

- Artwork search for sourcing visual references for literary writing; built multimodal embedding pipeline over Museum artworks using Gemini embeddings and **sub-100ms** semantic search.

Pravas | Typescript, Next.js, Supabase, Mapbox GL, Sarvam AI, Vercel

- Travel journaling app with voice diary recording; built multilingual speech-to-text and translation using **Sarvam AI**, with interactive Mapbox canvas for pinning and reliving trip entries.

TECHNICAL SKILLS

Languages: Python, Java, C/C++, JavaScript, TypeScript, SQL, Bash

Full-Stack & Frameworks: Django, FastAPI, Spring Boot, Node.js, Celery, React, Next.js, Angular, Svelte

Databases & Infra: PostgreSQL, MongoDB, Supabase, Redis, Docker, Kubernetes, AWS, Vercel, Grafana, Datadog

ML Engineering: NumPy, Pandas, Scikit-learn, OpenCV, PyTorch, Hugging Face, MLflow, LangChain, vLLM, Vector DBs