

Adithya S Kolavi

[GitHub](#) | [LinkedIn](#) | [adithyaskolavi.com](#) | [Google Scholar](#) | [Blog](#)

Location: Bengaluru

Email: adithyaskolavi@gmail.com

TECHNICAL SKILLS

- **AI|ML|DL:** PyTorch, Transformers, PEFT, Bitsandbytes, Ray, OpenEnv, Harbor, TRL, SkyRL
- **Python Libraries:** FastAPI, Flask, OpenCV, BeautifulSoup, Selenium, Pandas, Poetry, Langchain, LlamaIndex
- **Web Frameworks:** React.js, Next.js, Express, Node.js, Vue.js, Bootstrap, Tailwind
- **Cloud:** Azure, Azure Machine Learning, AWS, AWS SageMaker, Docker, Kubernetes(AKS), Modal
- **Big Data:** Databricks, Azure Data Factory, Apache Spark, Hadoop, Kafka
- **Databases:** MongoDB, PostgreSQL, Firebase, Redis, MySQL, Supabase, Pinecone, FAISS, Qdrant, ChromaDb, LanceDB
- **Languages:** HTML, CSS, JavaScript, TypeScript, Python, C/C++, SQL

EXPERIENCE

Hugging Face

Research Engineer

Hybrid

April 2026 to Present

- **Built infrastructure to scale RL environments** across **data generation, training, and evaluation**, including **FineEnvs**, an open ecosystem of **cross-domain RL environments, post-trained models, training recipes, and reproducible artifacts** spanning coding, computer use, simulation, multilingual, biology, and more.
- **Built multi-harness RL training from scratch** around **OpenEnv**, enabling the same environments to run across multiple training and evaluation stacks; integrated **TRL, Harbor, and OpenReward**, documented in the **Multi-Harness RL** project.
- **Created and scaled Repo2RL**, which has grown into a **go-to open-source library for synthetic coding data generation**, unifying multiple repository-to-task generation approaches into a common pipeline for **reproducible RL environments and coding tasks at scale**.
- Authored the **Ultimate Guide to Reinforcement Learning Environments**, a comprehensive survey and reproduction of modern RL environment frameworks, reaching **25k+ reads** and **150k+ impressions**.

CognitiveLab

Founder , Lead AI Researcher

Bengaluru

May 2023 to present

- Leading an open-source AI research lab that received the **LLaMA Impact Grant** a six-figure, non-dilutive award by **Meta** for our **Nayana** initiative, building a multimodal vision-language model covering 22 languages and driving breakthroughs in OCR and cross-lingual reasoning Specializing in **synthetic data generation pipelines** for large-scale fine-tuning; created one of the **largest multilingual, multimodal document datasets** using synthetic data **2TB+**.
- Pioneered **India's first Kannada-English bilingual LLM — Ambari-7B**, achieving GPT-3-like performance for cross-lingual tasks and an 85% improvement in Kannada tokenization efficiency; also developed **Indic Eval / Leaderboard**, the first open-source benchmark suite for Indic LLMs with language-specific evaluation metrics.
- Developed open-source projects with ~10,000 GitHub stars, actively used by thousands of developers worldwide.

Microsoft Research

Research Fellow

Bengaluru

August 2025 to March 2026

- Built a **real-time voice agent for conversational interaction with video**, enabling users to talk directly with video content through low-latency speech, vision, and LLM pipelines; the work was **patented** and won the **Microsoft Global Hackathon** among **10k+ participants**.
- Developed **agentic memory at enterprise scale** across the Microsoft stack, designing shared long-term memory infrastructure, retrieval and memory evaluation systems, and real-time LLM pipelines; achieved **40% lower inference latency** and **90%+ retrieval accuracy**, with the work resulting in a **second patent** and technology transfer to **M365 Copilot**.

Apple

ML Intern

Bengaluru

January 2025 to July 2025

- Built an internal **LLM-powered knowledge graph generation and retrieval platform**, enabling the creation of **dense, well-connected knowledge graphs** from unstructured enterprise data with **over 100k+ entities**.
- Designed an **agentic retrieval pipeline** that dynamically queries and refines knowledge nodes, reducing hallucinations by **45%** and improving factual consistency across multi-hop reasoning tasks.
- Leveraged LLMs for **automated relation extraction and schema induction**, achieving a **3x increase in graph completeness** and significantly enhancing retrieval accuracy across Apple's internal intelligence systems.

TurboML

AI Developer

Bengaluru

June 2024 to November 2024

- Architected and implemented a comprehensive **LLMOps observability platform**, enabling real-time monitoring and evaluation of large language models. Developed custom instrumentation for tracking inference latency, token usage, and model performance.
- Built an advanced **distributed tracing system** for AI workflows, reducing debugging time by 60% and enabling granular analysis of model behavior across multiple deployment environments.
- Pioneered automated evaluation frameworks integrating multiple metrics (ROUGE, BLEU, RAG Metrics), enabling continuous assessment of model updates. Reduced manual evaluation effort by 75% while maintaining robust quality standards.

RESEARCH

NetraEmbed - SoTA Multimodal Multilingual Document Retrieval

Transformers, ColPali, Vision-Language Models, MTEB

- Built **state-of-the-art cross-lingual and cross-modal document retrieval system** with both dense (**NetraEmbed**) and multi-vector (**ColNetraEmbed**) variants, achieving **0.716 NDCG@5**, a **152% improvement** over existing baselines, enabling queries in one language to retrieve documents in any of **22 supported languages** while preserving visual elements.
- Developed **NayanaIR Benchmark**, a **first-of-its-kind multilingual multimodal evaluation suite** with **28,000+ document images** and **5,400+ queries** across 23 datasets.

Indic Eval/Leaderboard (Open Source)

spaCy, NLTK, Transformers, SkyPilot, Azure

- Developed a first-of-its-kind evaluation framework for Indic Large Language Models, addressing the current English-centric bias in performance measurement across tasks like machine translation and cross-lingual question answering.
- The framework, compatible with various datasets, integrates TGI to speed up evaluation by 45%, offering metrics like accuracy, F1 score, perplexity, ROUGE, and BLEU. It also supports multi-cloud environments for easy execution.

PUBLICATIONS

• **M3DR: Towards Universal Multilingual Multimodal Document Retrieval**

AAAI 2026 | Workshop on New Frontiers in Information Retrieval | **Accepted**

• **TRACE: Transparent Reasoning and Attribution Chains for Extended Multimodal Contexts**

NeurIPS 2025 | Workshop on Efficient Reasoning | **Accepted**

• **Nayana: A Foundation for Document-Centric Vision-Language Models via Multimodal, and Multilingual Data Synthesis**

ICCV 2025 | Workshop on Computer Vision for Developing Countries (CV4DC) | Proceedings Track | **Accepted**

• **Nayana - A Unified Foundation Model for Multilingual, Multimodal, and Multitask Intelligence**

LlamaCon 2025 | Winner of 2024 LLaMA Impact Grant from Meta | **Accepted**

• **ViViD - Vision Language model for Unified Visual Understanding of Documents**

CVPR 2025 | Emergent Visual Abilities and Limits of Foundation Models (EVAL-FoMo 2025) | **Accepted**

• **Nayana OCR: A Scalable Framework for Document OCR in Low-Resource Languages**

NAACL 2025 | Language Models for Underserved Communities | **Accepted**

• **CAPTAIN: Continuous Automated Planning Through Autonomous Internet Navigation**

AAAI 2025 | Large Language Models for Planning (LM4Plan) | **Accepted**

• **A Multi-Agent Approach for Iterative Refinement in Visual Content Generation**

AAAI 2025 | Advancing LLM-Based Multi-Agent Collaboration | **Accepted**

PROJECTS

OmniParse(6K Github Stars)

Python, FastAPI, Transformers, Redis Queue, Containerization, K8s

- Open Source Data Ingestion and Processing platform, achieving 85% accuracy in structured data extraction from diverse data source like Documents, Media, Web for GenAI applications
- Integrated cutting-edge models for OCR, captioning, and transcription, improving data quality for AI applications by 90%
- Optimized for T4 GPUs deployment, reducing infrastructure costs by 40% compared to cloud-based alternatives.

GitVizz (Open Source)

Python, TypeScript, Langraph

- An open-source tool for developers to intuitively explore codebases through **interactive dependency graphs** visualizing files, modules, and commit relationships in a clear, navigable structure. **Accepted into Vercel's Open Source Sponsorship Program.**
- Designed a powerful **graph-centric UI** enabling fast search, filtering, and navigation, making it effortless to understand complex repositories at a glance.

AI Engineering Academy (Open Source)

Education, Community Building, AI Engineering

- Founded and operate **open-source learning platform** with **10,000+ monthly visitors** and **1.6K GitHub stars**, providing structured roadmaps for AI engineering covering Prompt Engineering, RAG systems, LLM fine-tuning, Agents, and MLOps.
- Built a thriving community-driven platform with comprehensive curriculum spanning fundamentals to advanced concepts, bridging the gap between theoretical understanding and production-ready AI systems through freely accessible resources.

EXTRA CURRICULAR/AWARDS

Google Developers Student Club Lead

University Lead

- Led the Google Developers Student Club, fostering a collaborative environment for technology enthusiasts at PES University.
- Organized events during Hacktoberfest, including hands-on Git and GitHub workshops, to promote open source contributions.
- Facilitated industry collaborations, guest lectures, and networking opportunities for club members.

Samarpana, Shunya

Technical Head

- Directed technical teams for Samarpana (a fundraising marathon for families of martyrs) and Shunya (Math club) events.
- Led website development projects to enhance registration awareness and streamline participant engagement.

Hackathons

Participant

- Won National Level Hackathons in Generative-AI - GenAI-Rush and Kodikon3
- Participated in more than **25 Hackathons** in the span of 2 years (adithyask.com) with a 90% finalist selection rate.

EDUCATION

- BTech, Computer Science, PES University