

Nikhil Pavan Kanaka

 [npkanaka.github.io](https://github.com/npkanaka)  nikhilpavan.kanaka@gmail.com  github.com/npkanaka  +1 (813)705-2796

EDUCATION

Ohio State University

MS by thesis in Computer Science (NLP)

Jan 2023 - Dec 2024

GRIET

Bachelor of Technology in Computer Science and Engineering

Oct 2020

INDUSTRIAL EXPERIENCE

Enterprise Data Analytics - The Ohio State University

AI Research Intern

Jan 2024 – Dec 2024

- Developed a GenAI-powered educational assistant using RAG pipelines with LangChain, Pinecone, and Llama to help students query course materials. Stored data on Azure Blob Storage, automated updates with Event Grid, and built a Gradio app on Azure App Service.
- Led a data migration project to optimize preprocessing for large-scale NLP datasets, reducing ETL time from 6 hours to under 56 minutes using Airflow MWAA and fine-tuned connectorX.
- Built an automated lecture transcription system using OpenAI's Whisper model, enhancing accessibility and learning experience within the university's learning management system.

Wexner Medical Center

Machine Learning Engineer Intern

May 2023 – Aug 2023

- Built an NLP-driven data validation system by integrating LLMs with Great Expectations, featuring human-in-the-loop processes and domain-specific fine-tuning to automate data quality checks. The system generates dynamic data validation rules from natural language using GPT-3.
- Developed an NLP pipeline to analyze 250K+ patient feedback using BERT (sentiment), BERTopic (topic modeling), and medspaCy (keyword extraction) to generate actionable healthcare insights.

Ninjacart

Senior Software Engineer

Jun 2022 – Dec 2022

- Led an 8-member backend development team for ninjacart.com's B2B platform, driving API integrations with key partners using Java, Spring Boot, Node.js, MongoDB, Kafka, and Apicurio Registry. Achieved a 28% increase in transaction volume within two weeks of implementation.
- Automated API schema generation and validation pipelines using T5 and OpenAPI Validator.

Head Digital Works

Software Development Engineer

Feb 2020 - Jun 2022

- Led the design, feature development, and deployment of backend systems for cricket.com and a23.com, supporting over 80M users, utilizing Java, Spring Boot, AWS EC2, ECS, DynamoDB.
- Engineered a fraud detection system by combining anomaly detection with a rule-based system for monitoring player behavior, resulting in an 18% increase in player retention. Integrated NLP models to detect suspicious communication with retraining pipelines for evolving fraud patterns.

ACADEMIC PROJECTS

Comparative Investigation of Grokking: A Study on Reasoning, Mentor: Dr. Sachin Kumar, OSU
Understanding and Improving LLM reasoning by investigating parametric memory. [\[info\]](#)

911AIPredict, Mentor: Dr. Rajiv Ramnath, OSU

AI assistant for 911 services, integrating call transcription, fine-tuned LLMs, and vector database to extract critical details, suggest health conditions, and enhance dispatcher efficiency. *(under review)*

Voicequity, Mentor: Dr. Srinivasan Parthasarathy, OSU

Advanced ASR fairness by creating a novel metamorphic testing framework to expose and address bias, driving robust and inclusive voice recognition solutions for global demographics. [\[report\]](#)

Long-Answer Question, *Mentor: Dr. Chris Brew, OSU*

Developed a pipeline using LongFormer attention and generation models to create long-answer questions from academic texts, achieved 42.3% semantic accuracy through evaluation. [\[report\]](#)

HiDL, *Mentor: Dr. DK Panda, OSU*

- Engineering Java Bindings for MVAPICH and advancing GPU-accelerated workflows within Nvidia's Spark-Rapids, GDS, and DALI to set new benchmarks in HPC+AI. [\[info\]](#)
- Investigating GPU offloading in Nvidia Rapids. [\[report\]](#)

LogiChain, *OSU*

Developed a system integrating LLMs with symbolic solvers to enhance reasoning by translating natural language into solver-executable code, executing it with symbolic solvers, and feeding results back to the LLM, achieving up to 30% performance gain over LLM-only approaches. [\[code\]](#)

Benchmarking of File Systems, *Mentor: Dr. Ting Zhu, OSU*

Evaluation and optimization of file systems performance to enhance applicability. [\[report\]](#)

ProdReview, *Amazon Research*

Developed an e-commerce review analysis platform incorporating sentiment analysis and a novel graph-based summarization technique, enabling actionable, product-specific insights. [\[info\]](#)

SERVICE

- Founding member of the Advanced Academic Center at GRIET, for academic research. [\[link\]](#)
- Technical Head for Spirals, the Language and Literature club of GRIET. [\[link\]](#)
- Tech mentor at Sahaya Society, fostering skill development for the less privileged. [\[link\]](#)

TALKS

Domain-Specific Fine-Tuning <i>Advancing Analytics in Medical Research through NLP -at- Wexner Medical Center</i>	2023
Unlocking Meaning <i>NLP's Role in Shaping the Future of Intelligent Systems -at- Ninjacart</i>	2022
Quantum Perspectives on Intelligence <i>How quantum reshapes our understanding of intelligence, new possibilities for AI -at- GRIET</i>	2020
Bridging Minds and Machines <i>Psychological Impact of AI Evolution -at- GRIET</i>	2020