

## PROFESSIONAL SUMMARY

AI/LLM Specialist with 6 years of hands-on experience designing high-quality training data, crafting advanced prompts, and developing evaluation frameworks for large language models. Skilled in NLP data workflows, domain-specific content engineering, and end-to-end curation pipelines that support scalable machine learning initiatives. Adept at conducting linguistic, semantic, and behavioral audits to strengthen model accuracy and reduce hallucinations. Experienced in building automated validation tools using Python and collaborating closely with ML engineers, data scientists, and product teams to refine model outputs and enhance real-world reliability. Known for delivering precise, well-structured, and technically rigorous content that drives robust AI performance across diverse enterprise applications.

## TECHNICAL SKILLS

**LLM Content Development & Prompt Engineering** – Designing high-quality prompts, training data, and model interaction flows.  
**Dataset Creation & Large-Scale Annotation** – Building domain-specific datasets, annotation guidelines, and scalable data pipelines.  
**NLP Techniques** – Hands-on experience with tokenization, NER, summarization, text classification, and semantic analysis.  
**Conversational AI Dataset Design** – Crafting multi-turn dialogue data, intent/slot structures, and evaluation sets for chat-based AI systems.  
**Python & Java Scripting** – Developing automation scripts for data processing, validation, and quality assurance workflows.  
**AI/ML Data Audits & Quality Evaluation** – Performing linguistic, semantic, and behavioral audits to improve model accuracy and reduce hallucinations.  
**Test Case Design & Evaluation Scenarios** – Creating structured test suites and validation frameworks to assess LLM performance.  
**Technical Documentation & Workflow Design** – Writing clear, scalable documentation for ML pipelines, data processes, and model behaviour guidelines.  
**Cross-functional Collaboration** – Working closely with ML engineers, data scientists, and product teams to refine models and optimize AI development cycles.

## EDUCATION

Master's in MIS from Wilmington University, Delaware. May 2023  
Bachelor's in CS from JNTUH, Hyderabad, India. May 2020

## PROFESSIONAL EXPERIENCE

**Client: Labcorp, TX | Remote**  
**Senior LLM Content Developer** **Dec 2022 – Present**

- Led development of a large-scale, multi-domain dataset for enterprise LLM training and evaluation.
- Created and curated 5,000+ prompts, dialogues, negative samples, and test items across healthcare, finance, retail, logistics, and customer service.
- Designed edge-case scenarios to assess reasoning, contextual understanding, multi-step logic, safety, and factual accuracy.
- Built evaluation templates for tone consistency, semantic accuracy, hallucination detection, and compliance checks.
- Collaborated with ML engineers to analyze model outputs, identify failure patterns, and refine weak areas in the dataset.
- Used Python scripts to automate test generation, clean formatting, enforce token limits, and perform linguistic validation.
- Applied domain-specific constraints including HIPAA-safe phrasing and regulated financial terminology.
- Improved model factual accuracy by 24% and reduced hallucination errors by 17% through targeted dataset

optimization.

- Delivered a reusable evaluation suite adopted as a baseline for multiple model releases.
- Conducted linguistic and semantic audits to ensure accuracy, neutrality, and clarity across all dataset components, significantly improving model stability in real-world use cases.
- Partnered with domain SMEs in healthcare and finance to validate complex terminology, ensuring dataset accuracy and compliance.
- Implemented automated error-flagging scripts to detect formatting inconsistencies, boosting dataset quality and review efficiency.
- Contributed to safety-aligned prompt engineering, reducing harmful or policy-violating outputs in model evaluations.
- Developed scalable annotation guidelines and quality-control checklists that enabled consistent data labeling across distributed teams and reduced rework by over 30%.

**Client: Zomato | Hyderabad, India**

**March 2019 – November 2021**

**Role: Technical Content Writer**

- Developed clear, structured, and SEO-optimized technical content for Zomato's internal knowledge systems, partner documentation, and user-facing guides.
- Created and maintained content taxonomies, category mappings, and structured data templates used for restaurant listings, menu attributes, cuisines, and user intent classification.
- Collaborated with product, data, and engineering teams to ensure accurate metadata, consistent tagging standards, and clean structured content across large datasets.
- Conducted data-driven content audits to improve search relevance, reduce inconsistencies, and streamline content retrieval for the internal CMS.
- Authored high-quality documentation for API workflows, onboarding processes, technical FAQs, and platform feature updates.
- Assisted with early chatbot and support automation content, including intent drafting, response flows, and tone guidelines.
- Standardized content naming conventions, taxonomies, and metadata rules, which improved CMS searchability and internal navigation.
- Created guidelines for restaurant menu digitization, ensuring accurate mapping of cuisines, dietary tags, and attributes.
- Coordinated with regional teams to validate cultural and linguistic differences in menu descriptions and food categories.
- Developed content quality checklists that reduced review turnaround time and improved dataset consistency.
- Supported the localization team by producing content adapted for multiple regions and customer segments.
- Performed quality checks on annotated data, ensuring consistency in labels related to cuisines, delivery attributes, dietary tags, and location metadata.
- Improved content accuracy, reduced duplication, and enhanced knowledge-base navigability, contributing to faster query resolution and improved partner onboarding.
- Worked cross-functionally with UX writers, data analysts, and operations teams to maintain content compliance, linguistic clarity, and data hygiene across the platform.

## Achievements

- Recognized for delivering the highest-quality dataset batch with zero rework.
- Led a cross-team initiative to standardize annotation guidelines across three projects.

## Publication

Cloud Computing Based Efficient Information Retrieval – International Journal of Advanced Science and Technology, Vol. 29, No. 3, 2020.