AKASH KUMAR JHA

Generative AI Engineer | LLM Engineer | RAG Developer

Noida, Uttar Pradesh, India | +91 7643067271 | akashkumarjha067@gmail.com

LinkedIn | GitHub | Portfolio | Live AI Project

PROFESSIONAL SUMMARY

Entry-level Generative AI Engineer and Python developer with hands-on experience designing, building, and deploying **Large Language Model (LLM)** applications and **Retrieval-Augmented Generation (RAG)** systems. Developed a production-deployed multilingual AI document assistant supporting **20 languages**, semantic search, vector embeddings, contextual question answering, page-level citations, and voice interaction. Skilled in **Python, LangChain, OpenAI API, Hugging Face, Prompt Engineering, NLP, Vector Databases, and Machine Learning**. Experienced in document processing, embedding pipelines, REST APIs, SQL, and full-stack AI application development.

TECHNICAL SKILLS

Generative AI & LLMs: Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), Prompt Engineering, LangChain, OpenAI API, Hugging Face, Transformers, LLM Application Development

RAG & Information Retrieval: Text Embeddings, Vector Databases, Semantic Search, Vector Similarity Search, Document Chunking, PDF Text Extraction, Context Retrieval, Grounded Question Answering, Source Citations

Programming & Development: Python, SQL, JavaScript, TypeScript, C++, REST APIs, Next.js, Git, GitHub

NLP & Machine Learning: Natural Language Processing, Text Preprocessing, Tokenization, Multilingual NLP, Speech-to-Text, Text-to-Speech, Scikit-learn, TensorFlow, Classification, Regression, Clustering, Feature Engineering, Model Evaluation

Data Science & Tools: Pandas, NumPy, MySQL, PostgreSQL, MongoDB, Power BI, Matplotlib, Seaborn, Jupyter Notebook, Google Colab, Vercel, Pygame

TECHNICAL PROJECTS

Akarika – Multilingual RAG Document Assistant *Generative AI | Python | RAG*

Technologies: Python, LLMs, RAG, Embeddings, Vector Database, Semantic Search, Next.js

- Designed, built, and deployed an end-to-end **Retrieval-Augmented Generation (RAG)** application enabling users to upload PDF documents and interact using natural-language questions.
- Implemented an LLM retrieval pipeline covering PDF text extraction, document chunking, embedding generation, vector storage, and semantic similarity search for relevant context retrieval.
- Developed context-aware question answering with **page-level source citations**, improving response traceability and supporting verification against uploaded documents.
- Integrated **20-language support**, Speech-to-Text, Text-to-Speech, authentication, persistent chat history, and a responsive interface; deployed the application on Vercel.

Live: rag-pdf-nine.vercel.app

Quick Chat – Real-Time Messaging Platform *Next.js | JavaScript | WebSockets*

- Developed and deployed a full-stack real-time messaging application supporting individual/group chats, typing indicators, read receipts, reactions, and voice notes.
- Implemented WebSocket-based communication, voice/video calling, privacy controls, optional two-factor authentication, and Progressive Web App (PWA) support.

Live: quickchat.akarikatechnologies.me

Snake & Ladder – Player vs Computer Game *Python | Pygame | OOP*

- Built an interactive Snake & Ladder game using **Python, Pygame, and Object-Oriented Programming (OOP)**, featuring custom graphical interfaces and player-vs-computer gameplay.
- Implemented dice-rolling mechanics, turn management, snake/ladder movement logic, game-state handling, and interactive gameplay functionality.

Live: snake-ladder-game-by-akarika.vercel.app

PROFESSIONAL EXPERIENCE

Data Science & Analytics Intern **Jul 2025 – Oct 2025**

Zidio Development | Remote

- Cleaned, transformed, and preprocessed real-world datasets using **Python, Pandas, NumPy, and SQL**, handling missing values, outliers, and inconsistent data.
- Conducted Exploratory Data Analysis (EDA), statistical analysis, and feature engineering to identify data patterns and actionable insights.
- Developed predictive machine learning models using Scikit-learn, applied standard evaluation metrics, and created dashboards using Power BI, Matplotlib, and Seaborn.

EDUCATION

Bachelor of Computer Applications (BCA) **2024 – 2027**

Dr. Babasaheb Bhimrao Ambedkar Bihar University, Muzaffarpur | *Pursuing, 6th Semester*

Senior Secondary (Class XII) – CBSE **2022 – 2024**

Best Model Public School, Hajipur, Bihar **87%**

CERTIFICATIONS

- **Generative AI** – Tutedude, Aug 2026 (Certificate ID: TD-AKAS-GA-1050)
- **SQL for Data Analysis** – LinkedIn Learning, Mar 2026
- **Basics of Python (CR641)** – Cambridge International, UK