



Prasad Thipparthi

India

Education

B.Tech in Computer Science and Engineering, NIT Puducherry, IN, 2017

Advanced Certification in Generative AI in Data Science - AI , Indian Institute of Technology, Guwahati, IN

Digital Credentials

IBM Generative & Agentic AI Data Scientist - 2026

IBM Generative & Agentic AI Developer - 2026

AWS Partner: Generative AI Technical - Training Badge - 2025

AWS Partner: Agentic AI Essentials - Training Badge - 2025

IBM Generative & Agentic AI Foundation - 2024

Containers & Kubernetes Essentials - 2021

Languages

English - Fluent

Telugu - Fluent

Hindi - Good

Tamil - Good

Profile

Experienced **Data Scientist** with a strong track record of delivering **AI-driven solutions** in the **Information Technology and Services industry**. Specialized in designing, developing, and deploying **Large Language Model (LLM) applications**, fine-tuning models, and implementing **Retrieval-Augmented Generation (RAG) pipelines** to address complex business challenges.

Proficient in **Python**, **Natural Language Processing (NLP)**, **Machine Learning**, and **Generative AI**. Hands-on experience with LLM frameworks such as **LangChain** and **Hugging Face**, embedding models, vector databases including **ChromaDB**, **FAISS**, and **Pinecone**, and building **hybrid AI systems** combining structured data querying with LLM intelligence.

Experienced in developing scalable AI solutions using **Docker**, **Kubernetes**, **AWS (ECS, EKS)**, **IBM Cloud**, and **IBM Watson AI services**. Skilled in **end-to-end AI application development, deployment, and optimization** for enterprise environments.

Holds a **B.Tech in Computer Science** from the **National Institute of Technology Puducherry**.

AI Trainer & Technical Speaker:

Conducted hands-on workshops on **Agentic AI** and **Generative AI** for IBM practitioners during ICA Day events across Hyderabad, enabling participants to build AI-powered applications. Recognized for contributions to **AI training programs, Agentic AI exercises, and technical enablement sessions**.

Also serves as an **AI Trainer at a reputed international AI training institute**.

Key skills

Generative AI & LLMs

- Prompt Engineering, Retrieval-Augmented Generation (RAG), Vector Search, Langchain.
- Vector Databases: Chroma, FAISS, Milvus, Pinecone
- Agentic AI Solutions, Langflow, Orchestration of Agents, MCP Servers , Strands Agents
- Gen AI Models: LLaMA-70B, LLaMA3, GPT, CodeLLaMA, AWS Bedrock Agents, WatsonX , ICA.

Natural Language Processing (NLP) & Analytics

- Text & Data Analytics, Complex Document Extractions(docling).
- Watson Knowledge Studio, Watson Assistant

Machine Learning & AI Frameworks

- End-to-end ML model development, training, and deployment
- Flask /Fast API Web Framework, Streamlit UI for AI/ML applications



Publications

Complex Table Variant Analysis
with AI System Inputs, 2026

Article on Data Science - The
Future, 2022

Cloud & Infrastructure

- IBM Cloud, AWS Bedrock
- Docker Containers, OpenShift, Kubernetes

Databases

- NoSQL: Cloudant, Cassandra, SQLite
- Graph & Time-Series Databases: Neo4j, InfluxDB

DevOps & Packaging

- Containerization, CI/CD pipelines, Cloud-native deployments

Key courses and training

Advanced Generative AI Certification from IIT Guwahathi.

LangChain for LLM Application Development, Fine tuning LLMs from
DeepLearning.ai

Generative AI / LangChain / OpenAI Foundational Models. Prompt Engineering
Techniques.

Work experience

Data Scientist-Artificial Intelligence

IBM India Pvt Ltd, India | Aug 2017 - Present

Worked as Data Scientist - Gen AI Developer, Led the Cognitive Team for the CP Editor Application for BOXCO Client. Utilized Watson Knowledge Studio for action capturing from contract agreement statements and conducted text processing/analysis (NLP) for document sets. Managed the application stack, including packaging (Dockerization), deployments (CI/CD), API development with Flask, and DevOps implementation. Improved performance for document processing and implemented schedulers and thread pools. Styled generated documents using pdfminer, PyPDF2, and diff-match libraries.

Assignment history

Senior Data Scientist

HUBBELL INC, India | Dec 2025 - Present

Project name: Hubbell Project – Product Intelligence & RAG Solution

Worked on AI-driven use cases for **product categorization** and **cross-referencing** of Hubbell products. Designed and implemented a **Retrieval-Augmented Generation (RAG)** solution using **LlamaIndex** to efficiently process and retrieve information from large-scale product datasets. Improved data discovery, contextual search, and product mapping accuracy by leveraging vector-based retrieval and generative AI capabilities.

Contribution

- Developed AI-powered solutions for **product categorization** and **product cross-referencing**.
- Designed and implemented a **RAG (Retrieval-Augmented Generation)** framework using **LlamaIndex** for large-scale Hubbell product data.
- Built data ingestion, indexing, and retrieval pipelines to enable efficient semantic search and contextual information retrieval.



- Improved product mapping accuracy and streamlined access to product information through vector-based search and AI-driven recommendations.

Lead Data Scientist

BT GROUP PLC | Mar 2025 - Dec 2025

Project name: IBM Project Stark - BT

The project focused on designing, developing, and deploying Agentic AI solutions to streamline business analyst (BA) and solution architect (SA) workflows. It aimed to demonstrate the potential of Generative AI and Agentic AI in automating documentation, requirement gathering, impact analysis, and decision support. The solution was implemented on both the ICA Platform and AWS Bedrock Agentic Platform, with scalable ingestion pipelines and reusable agentic components for enterprise-wide adoption.

Contribution

Proof of Concept (POC): Designed and implemented a POC for building intelligent assistants, including prompt engineering flows and BA user story generation, showcasing the applicability of Agentic AI in automating requirement gathering.

Agentic AI Implementation (Post-POC): Extended the POC into a production-ready solution by developing and deploying Agentic Orchestration Flows in AWS Bedrock, tailored for BA use cases (user story generation) and SA use cases (impact analysis, decision support).

Framework & Pipeline Development: Contributed to the Ingestion Pipeline design, integrated with the ICA4AA framework, enabling scalable handling of documents and data. Built reusable Agentic AI components to accelerate adoption and serve as a foundation for future enterprise automation.

Data Scientist - Gen AI

IBM - Gen AI HCS, India | Oct 2023 - Feb 2025

Project name: IBM - Gen AI Auto Augmented Quality Engineering (AAQE)

The IBM Gen AI Auto Augmented Quality Engineering (AAQE) project utilizes advanced Generative AI, Prompt Engineering, and Advance RAG Techniques to automate and enhance the generation of test cases and code snippets.

Contribution

Worked as Gen AI Developer : LLM Fine tuning, Prompt Engineering, chunking and advanced RAG Techniques, LangChain framework.

Data Scientist - AI

IBM Internal, India | Jan 2021 - Sep 2023

Project name: IBM Cloud Garage Method - Mainframe Application Modernizer (MAM).

IBM Garage Method for Cloud Toolkit is a suite of integrated tools and assets with a scalable, flexible architecture supported by an analytics engine and automation that helps our clients accelerate their transformation journey to a multi-cloud environment.

Contribution



Building the MAM asset backend API's / Integration of Analytics, Ops Analysis, and Other AI Services.

Data Scientist-AI

BOXCO-DIABOS, India | Sep 2018 - Nov 2020

Project name: DIABOS Cognitive Digitization

To build a Cognitive CP Application which is capable of applying the Automatic changes to charterer party after reading Recap. The world's first cognitive enterprise solution in shipping industry capable to extracting the information from documents and updating in the required format.

Contribution

Led the Cognitive Team for the ChartererParty Editor Application for BOXCO Client. Utilized Watson Knowledge Studio for action capturing from contract agreement statements and conducted text processing/analysis (NLP) for document sets.

Data Scientist - Cognitive Computing

IBM FCDD, India | Jan 2018 - Aug 2018

Project name: Cognitive KYC

Developed a minimum viable product for Cognitive KYC, automating KYC checks including Adverse Press, Sanctions, Politically Exposed Persons lists, and basic identity verification for entities, integrating cognitive capabilities to enhance the client's system.

Contribution

Worked for DB2 Database Management for the Application.

Additional information

Conducted workshops on Ethical Hacking Countermeasures (1 day), Real-time Data Science/AI Projects (2 days), Python for Data Science (3 days), and Generative AI - LLM Applications (3 days).