Gagan N Bangaragiri

Bangalore, KA | gagan.bangaragiri@gmail.com | +91 9611706653 | linkedin | portfolio

Work Experience

AI Engineer, Genpact – Bangalore, KA June 2026 – Present

- Built AgenticDocExtractor, a 7-pass pipeline for digital and scanned invoice bundles. It uses PyMuPDF and Marker for digital PDFs, with Claude VLM and PaddleOCR for scanned PDFs.
- Reduced LLM cost by 98%, from $2.65 to $0.05 per 10-page document, by running deterministic extraction before LLM fallback and caching identical model calls.
- Implemented confidence-based routing with 95%+ auto-approval, an LLM judge for 60–95% confidence, and human review below 60%. Added versioned vendor schemas, PII-safe audit events, and a FastAPI and React review system backed by PostgreSQL.

Intern, Genpact – Bangalore, KA January 2026 – May 2026

- Developed internal GenAI systems for model evaluation and ETL migration with FastAPI, LangGraph, Azure AI Foundry, Databricks, PostgreSQL, Redis, and Docker.
- Built reusable evaluation and migration services with automated validation, deployment checks, failure diagnosis, and support for enterprise data workflows.

Summer Intern '25, Data Science, Genpact – Bangalore, KA May 2025 – July 2025

- Developed two GenAI agents with LangChain and the Gemini API for risk model validation. The agents reduced manual validation time by 25% and improved issue logging accuracy.
- Built an AI-powered VS Code extension with 10+ features for code explanation, refactoring, debugging, and other developer workflows.
- Designed Figma prototypes and Streamlit applications for a Risk Model Validation UI and an Employee Document Chatbot.

Projects

Checkpoint: AI-Powered Developer Onboarding Tool Personal Project

- Built and published a pip-installable CLI that creates repository architecture documentation, commit summaries, PR reports, and per-developer catch-up briefings through GitHub Actions.
- Built a LangGraph code-analysis pipeline with AST-based Mermaid diagrams. LiteLLM supports OpenAI, Anthropic, Mistral, Google Gemini, Azure, and Ollama through one interface.
- Public repository: github.com/BurntDosa/Checkpoint

Muso: Multiplayer AI Development Workspace Personal Project – In Development

- Building a browser-based AI workspace where developers share threads, branches, artifacts, files, tools, skills, approvals, and project context.
- Developing a model-agnostic execution layer for shared and personal provider connections, concurrent agent runs, persistent memory, and human approval controls.
- Implementing GitHub App access, OpenBao secret storage, artifact versioning, proposed code diffs, branch review, merge workflows, and auditable agent actions.

EvalHarness: GenAI Evaluation Microservice Internship Project

- Built a FastAPI GenAI evaluation service with 25+ REST endpoints for calibration, risk-weighted scoring, platform adapters, and drift detection with configurable alert rules.
- Implemented RAG metrics for faithfulness, context precision and recall, MRR, NDCG, MAP, noise sensitivity, and counterfactual robustness. Added HHEM hallucination scoring, Microsoft Presidio PII detection, Celery, and Redis.

ETL Migration Toolkit: Informatica & Alteryx to Databricks Internship Project

- Built a multi-agent pipeline that converts Informatica PowerCenter and Alteryx exports into Databricks notebooks for source DDL, target DDL, and ETL logic.

- Added direct Databricks deployment, Unity Catalog support, synthetic test data, LLM-based validation, and a Judge agent that diagnoses and patches deployment failures.

Education

Vellore Institute of Technology, Vellore, TN — August 2026

- B.Tech, Computer Science and Engineering

Greenwood High International School, Bangalore, KA — July 2022

- X (ICSE) & XII (ISC)

Skills

Languages: Python, TypeScript, JavaScript, SQL, Java, C++

AI & LLM: LangChain, LangGraph, LiteLLM, Atomic Agents, Anthropic Claude, Google Gemini API, Azure AI Foundry, AWS Bedrock

ML & Vision: PaddleOCR, OpenCV, YOLOv8, SAM2, spaCy, Scikit-learn, Microsoft Presidio, HHEM-2.1-Open

Frameworks & Libraries: FastAPI, React, Streamlit, Gradio, Svelte, Pydantic, PyMuPDF, OpenPyXL, Celery, GitPython

Cloud & Infrastructure: GCP (Cloud Run, Cloud SQL, GCS, Cloud Build, Cloud Logging, Document AI), Azure (Document Intelligence, AI Foundry, OpenAI, Content Understanding), AWS Bedrock, Databricks, Docker, PostgreSQL, Redis, SQLite, OpenBao

Developer Tools: Git, GitHub Actions, Alembic, Pytest, uv, Figma

Languages

- **English**: Fluent
- **Kannada**: Native
- **French**: Conversational
- **Hindi**: Conversational

Certifications

- Supervised Machine Learning: Regression and Classification – DeepLearning.AI
- Foundations of User Experience (UX) Design – Google Career Certificates
- Start the UX Design Process: Empathise, Define, and Ideate – Google
- Build Wireframes and Low-Fidelity Prototypes – Google