

BEVIN BENET

Applied AI Engineer | Agentic AI | LLM Systems | RAG

+91 9074332091 | bevinbenet@gmail.com | linkedin.com/in/bevin-benet | github.com/bevinbenet

SUMMARY

Applied AI Engineer and hands-on CTO (4+ years) building end-to-end agentic AI systems with LangGraph, FastAPI, RAG, vector search, LLM evals, LangSmith, and CI/CD; first-author researcher in multimodal deep learning.

TECHNICAL SKILLS

AI & LLMs: NLP, RAG, AI agents, function calling, tool use, structured outputs, vector search, guardrails

Frameworks: LangGraph, LangChain, CrewAI, MCP, OpenAI SDK, Instructor, Pydantic

LLMOps & Testing: LangSmith (tracing & evals), LLM-as-judge, pytest, GitHub Actions CI/CD

ML & Deep Learning: PyTorch, TensorFlow, Keras, computer vision, LSTM, LoRA/QLoRA, model evaluation

Backend & Web: Python, FastAPI, JavaScript, React, Next.js, Node.js, Strapi, REST APIs

Data & Infra: PostgreSQL (pgvector), MySQL, MongoDB, Docker, AWS (ECS/Fargate scripting), Nginx, Linux, Git

EXPERIENCE

Chief Technology Officer (Hands-On) | Kirel Creations Pvt. Ltd., Kochi 2024 - Present

- Cut engineering headcount by **50%+** with AI-assisted development: delivered the **IAM platform** (Next.js/Strapi; subscriptions, Razorpay, role-based access) in under 6 months with 3 engineers (typically 6–7), and **Findflo** with 1–2 (typically 4–6). Sole engineer maintaining IAM.
- Designed and built a **RAG website chatbot framework** with guardrail-gated agentic workflows (e.g., a literature-review assistant that withholds output until clarifying questions are answered).
- Lead web engineering (architecture, Nginx tuning, delivery); recruited, trained, and code-reviewed 2–3 interns.
- Owned client relationships across **15+ engagements**; personally built **10+ production websites**.

Web Developer / Intern | Kirel Creations Pvt. Ltd., Kochi 2022 - 2024

- Built production WordPress sites, then moved to full-stack work (React, Next.js, Node.js, Strapi) with clients.

PROJECTS

Super Sales Agent – Agentic AI B2B Lead Generation | Kirel Creations

- Built for internal lead generation: a 6-node LangGraph pipeline turning a plain-language target into qualified, scored B2B leads with typed Instructor + Pydantic outputs (Python, FastAPI, LangGraph, pgvector, LangSmith).
- Hybrid scoring (deterministic code + grounded LLM judgements); compliance enforced in code and schema.
- LLM evaluation harness (golden dataset, LLM-as-judge) and LangSmith tracing; pgvector semantic search.
- Dockerized FastAPI, mocked pytest suite, GitHub Actions CI; authored AWS ECS/Fargate deploy scripts.

Entropy-Gated Tri-Stream Human Action Recognition | IISER Thiruvananthapuram 2026

- Fused RGB, skeleton, and optical-flow streams with entropy-gated routing: **83.12%** mean accuracy on JHMDB and **59–84%** lower GFLOPs with minimal accuracy trade-off. Published in *Smart Health* (Elsevier).

Personal Agentic Chat Application | RAG, Tool Use

- Chat app combining RAG, external tools, APIs, and conversational context for multi-step tasks.

EDUCATION

Integrated M.Sc. Computer Science (Data Science) | Sacred Heart College, MG University 2021 - 2026

CGPA: 4.72/5 | Gold Medalist, Batch Topper

PUBLICATIONS

- Entropy Gated Tri-Stream Normalized Fusion Architecture for HAR - *Smart Health* (Elsevier), 2026
- Similarity Score Estimation of Kata Movement Using LSTM and DTW - ICAECT, 2023
- Privacy Preserving Machine Learning: A Study on the Methods and Trends - VIGYAAN, 2022-23

CERTIFICATION

Certification: The Complete Agentic AI Engineering Course - Udemy, 2025