

# Basanta Rai

Kathmandu, Nepal | [developer.basanta@gmail.com](mailto:developer.basanta@gmail.com) | +977 9849070118 | [basantarai.com.np](http://basantarai.com.np)

[linkedin.com/in/iambasantarai](https://linkedin.com/in/iambasantarai) | [github.com/iambasantarai](https://github.com/iambasantarai)

## Summary

---

Backend engineer with two years of experience, specializing in building scalable APIs, real-time data pipelines and data-intensive applications. Experienced in designing REST APIs, microservices, and RAG-based AI integration. Passionate about system design, performance optimization, and building robust developer-first infrastructure.

## Experience

---

**Associate Software Engineer**, AMNIL Technologies Pvt. Ltd., Nepal May 2023 – Present

- Collaborated with AI engineers to integrate backend services with large language models (LLMs) and RAG pipelines, enhancing AI-driven features
- Implemented real-time AI response streaming using Server-Sent Events (SSE), improving user interaction and system reactivity
- Implemented a domain and URL-based inclusion filter for AI-driven web queries, increasing control over external search context
- Developed a custom file management system on top of AWS S3, simplifying secure document access and ingestion for backend AI workflows
- Integrated BullMQ and Redis to offload heavy tasks to worker threads, enhancing application throughput and system responsiveness
- Developed a webhook broker and monitoring service to receive, route, and track real-time events, eliminating polling and improving system observability
- Engineered scalable REST APIs and relational schemas to manage complex 3D model and metadata workflows in digital construction systems
- Built dynamic CMS modules to control chatbot layout and content for banking clients, enabling flexible intent-driven customer interactions

## Education

---

**Bachelor of Science in Computer Science and Information Technology [TU]** 2019 - 2024

*Prime College*, Kathmandu

## Projects

---

- **panoramer - Panoramic Image Stitching Web Application** View in GitHub

A final year project designed to **simulate panorama generation** by stitching multiple images taken from the same point. It utilizes **SIFT**, **RANSAC**, and **Levenberg-Marquardt** algorithms for precise image alignment and blending, ensuring seamless and high-quality results.

– **Technologies Used:** Python, Flask, openCV, ReactJS, Tailwind CSS

- **Prime IT Club Website** View in GitHub

The official website of **Prime IT Club**, built as a CMS-driven platform to streamline the management of blog posts, events, and members. Provides an intuitive interface for administrators to create, edit, and publish content efficiently.

– **Role:** Led the development, collaborating with other contributors to ensure a scalable and maintainable system

– **Technologies Used:** Laravel, Tailwind CSS, Alpine.js, MySQL

## Technologies

---

**Programming Languages:** JavaScript, Typescript, Go, PHP, Bash, Lua, Python

**Libraries/Frameworks:** Node.js, Express.js, NestJS, Langchain, React, Alpine.js, Laravel

**Databases:** MySQL, PostgreSQL, SQLite, Redis, ChromaDB

**Tools/Technologies/Platforms:** Linux, Vi/Vim/Neovim, Tmux, Docker, Git, GitHub, GitHub Actions, GitLab, Jenkins, Portainer, REST API, Open API, OpenAI, RabbitMQ, AWS, S3, Web Sockets, SSE