

Sarbagya Gho Shrestha

sarbagyaghoshrestha@gmail.com | +977 9823833365 | Bhaktapur, Nepal

[linkedin.com/in/sarbagyashrestha](https://www.linkedin.com/in/sarbagyashrestha) | github.com/sarbagyag | <https://sarbagyaghoshrestha.com.np/>

Engineer who builds with intent. Drawn to technology that connects people to services, automates the meaningful, and scales what matters, whether that's governance infrastructure serving thousands of citizens or a startup's first intelligent system. Comfortable at any layer of the stack, from embedded hardware to distributed systems. Always learning, always shipping.

WORK EXPERIENCE

Mandala Foods — Fruit upcycling startup

Jan 2026 – Present

Full-Stack AI Engineer | Kathmandu, Nepal (Remote)

- Sole technical architect for AI and automation infrastructure. Designed and deployed **Maya**, an internal RAG knowledge chatbot, and a public-facing website chatbot, both on loosely coupled microservices.
- Architected and shipped **Maya Genie v1**, a multi-channel commerce platform (WhatsApp, Viber, Messenger) with a Go orchestration gateway, LangChain RAG cognitive core, function calling, and Khalti/eSewa payments; submitted to the *AI by HER: Global Impact Challenge*. Migrated its vector store from Supabase to self-hosted pgvector and built a UI for knowledge-base updates.
- **Internal Chatbot**: Vikunja task-tracking system integrated with the chatbot for Mandala Foods, a hybrid enterprise assistant that runs Text-to-SQL over normalised xls/csv data and pgvector document search in parallel, with embedding-based schema retrieval over a Postgres catalog (Python, LlamaIndex, Gemini embeddings).
- Prototyping on-device defect detection for apples and kiwis on STM32 nodes: fine-tuned lightweight vision models via transfer learning on open agricultural datasets extended with custom-collected samples, then applied INT8 quantization to fit SRAM/Flash limits.
- Engineered a Go service syncing Vikunja task attachments (XLSX, CSV, DOCX) to Google Drive to feed the n8n RAG embedding pipeline, and a Vikunja-backed reporting tool that generates AI narrative summaries straight from the database.
- Run the production VPS (Ubuntu 24.04, Docker, Traefik, Let's Encrypt) hosting n8n, Open WebUI, PostgreSQL/pgvector, PocketBase, Redis and a Claude API proxy that cut LLM spend; hardened it with SSH key auth, fail2ban, UFW and Docker port audits. Version-controlled n8n workflows, added backup scripts, and set up Google Workspace mail (DNS, DKIM, DMARC).

• Tech: Python, Go, LlamaIndex, LangChain, n8n, Docker, Traefik, PostgreSQL/pgvector, Redis, Gemini Embeddings, Anthropic Claude API, STM32, TinyML

Ninja Infosys (formerly ISRC) — Government Technology & Civic Systems

Jan 2025 – Apr 2026

Full-Stack & IoT Engineer | Kathmandu, Nepal

- Co-architected **eShashan**, a federated, multi-tenant Nepali GovTech platform (organizational-unit tenancy) on Go microservices, GraphQL Apollo Federation, gRPC, Kafka, Redis and Kubernetes, with ORY Kratos/Oathkeeper authentication and Module Federation micro-frontends.
- Built the iCMS site-builder pipeline for eShashan: a Go build-runner consuming Kafka build events through Redis-locked job queues, compiling layouts, optimizing assets and publishing to S3, plus a backend-for-frontend that renders tenant pages.
- Led technical architecture of Digital e-Palika (50+ municipalities) and the Integrated Content Management System (iCMS) powering 100+ government websites on a distributed, event-driven microservices architecture.
- Deployed digital profile systems for 10+ municipalities including Pokhara Metropolitan City in 4 months with a 2-person team, serving 100,000+ citizens; built a Survey Management System with a custom ODK Fetcher processing 100,000+ responses with live analytics.
- Built ESP32/STM32 field telemetry: GNSS over UART/NMEA, I2C and SPI for location anti-spoofing, offline-first logging to flash (NVS/LittleFS) for zero-connectivity surveys, streamed via MQTT/HTTP into PostGIS.
- Managed CI/CD and containerised deployments (Docker, AWS EC2/S3, MinIO, Jenkins) and trained 300+ local officials and enumerators in digital tools.

• Tech: Go, GraphQL Federation, gRPC, Kafka, Kubernetes, NestJS, Next.js, PostgreSQL/PostGIS, Redis, MinIO, ORY Kratos/Oathkeeper, ESP32, STM32, Docker, CI/CD

Ibtidaa Softwares Pvt. Ltd.

Oct 2023 – Nov 2024

Full Stack Developer & Tech Lead | Nepal (Hybrid)

- Developed the "Bidesh" App/Website using Next.js for RP Srijan Way To Success.
- Built an Event Management and Visualization Tool using Django and JavaScript for Shramik Sanjal with Chart.js integration.
- Tech: Next.js, NestJS, PostgreSQL, MongoDB, Django REST Framework, Redux.js, Jenkins

PointZero

Jan 2022 – Sep 2023

Frontend Developer | Nepal (Hybrid)

- Maintained an e-commerce website and developed a web application for SSoChe, Pulchowk Campus.
- Tech: JavaScript, React, MongoDB, Node.js, HTML5, CSS

LEADERSHIP

IEEE Pulchowk Student Branch

Apr 2023 – Jun 2024

Chair & Creative Director | Lalitpur, Nepal

- Led the IEEE Pulchowk Student Branch as Chair, organizing technical events, workshops, and coordinating student activities.
- Managed visual communications and promotional content for the branch as Creative Director.

KEY PROJECTS

8-Bit Computer From Scratch: A Breadboard-Based CPU

Jul 2024 – Dec 2024

6th Semester Minor Project, IOE Pulchowk Campus | with Rohit Joshi, Yugal Pariyar

- Engineered a fully functional 8-bit von Neumann computer on solderless breadboards using discrete 74-series TTL logic ICs, with no microcontrollers or off-the-shelf CPUs; built a custom 128-bit RAM module when ICs were unavailable.
- Hand-wired the variable clock, program counter, registers, RAM, 4-bit/8-bit ALU, instruction register, shared 8-bit bus and hardwired control logic; implemented a microcoded fetch-decode-execute cycle with LED observability on every bus and register.
- [Report](#) | [LinkedIn post](#)

ghoLang: A Custom Bytecode Virtual Machine & Interpreter in C

Jun 2026 – Present

Programming language built from scratch

- Hand-written single-pass Pratt parser compiling source directly to bytecode; stack-based VM with call frames, constant pools, string interning and an instruction dispatch loop.
- Mark-and-sweep tri-color garbage collector with self-adjusting heap growth; Lua-style closures with heap-migrating upvalues; classes with single inheritance; string interpolation, hash maps and modules.
- [GitHub](#)

zorLang: AST Tree-Walk Interpreter & DSL in Java

Apr 2026 – Jul 2026

Programming language built from scratch

- Recursive descent parser building a typed AST evaluated with the Visitor pattern; hierarchical environments for lexical scoping, closures, control flow and an object-oriented class model with initializers, plus lexical and runtime error reporting.
- [GitHub](#) | [Docs](#)

Maya Genie: Bilingual RAG Chatbot on Go and Python Microservices

Feb 2026 – Present

AI / production system | Mandala Foods

- Go WebSocket gateway and orchestrator decoupled through Redis Streams and Pub/Sub, with a Python FastAPI service for retrieval and generation behind them; only the gateway is internet-facing and chat sessions live in Redis.
- LangChain retrieval over self-hosted pgvector: 512-character chunks, Gemini embeddings with separate document and query task types, top-4 retrieval with source citations, English and Nepali answers grounded in the retrieved context.
- Provider-agnostic LLM layer (Anthropic, OpenAI, Gemini or a self-hosted endpoint) and an admin ingestion flow with document tracking, running in production on a Docker and Traefik VPS.
- [Mandala Foods](#)

Nepali Text To Speech

2024

Deep Learning Project | 8th Semester Major Project, IOE Pulchowk Campus

- Nepali text-to-speech platform using Tacotron, synthesizing natural-sounding speech to improve accessibility for Nepali-speaking audiences.
- [Report](#)

eShashan GovTech Platform

2025 – 2026

Federated multi-tenant government platform | Ninja Infosys

- Auth, site-builder/build-runner, media and organization services across a Go + GraphQL Federation stack; one of two architects of the platform.
- [Live demo](#)

Government & Civic Platforms

2025

Ninja Infosys | Digital Profiles, Survey Management System, ICMS

- Digital profile systems for 10+ municipalities including Pokhara Metropolitan City, serving 100,000+ citizens (pokhara.digprofile.com); Survey Management System with custom ODK Fetcher processing 100,000+ responses (kerabari-admin.vercel.app).
- Integrated Content Management System powering 100+ government websites nationwide on microservices with Kafka and Kubernetes (fansep.moald.gov.np).

zorbaOS: Multithreaded 32-bit x86 Operating System Kernel

Jul 2026 – Present

Operating system built from scratch | Assembly, C

- Custom x86 bootloader in assembly (real mode to 32-bit protected mode with custom GDT/IDT); paging-based virtual memory and a malloc/free heap allocator.
- Linux-inspired VFS with a FAT16 driver; multitasking scheduler and ELF loader running user-space programs across Ring 0 / Ring 3; keyboard and PIC interrupt drivers with an interactive kernel shell.

• [GitHub](#)

STM32-Based Self-Balancing Robot

Aug 2026 – Sep 2026

Embedded firmware | STM32 (ARM Cortex-M), Embedded C

- Bare-metal C firmware using HAL and Low-Layer drivers for peripheral configuration, interrupts and precise timing; MPU-6050 IMU over I2C with a real-time complementary filter for tilt estimation.
- Hardware-timer PWM motor control (TIM1), UART over USB-CDC telemetry; integrating an L298N H-bridge and tuning a discrete PID loop to close the balancing feedback (in progress).

• [GitHub](#) | [Documentation](#)

TDA2030 2.1-Channel Active Amplifier & Analog Signal Processing

Sep 2026

Analog design | KiCad

- Schematic, layout and hand-soldering of a KiCad PCB with three TDA2030 stages, split +/- supply from a centre-tapped transformer and a shared heatsink; 2nd-order Sallen-Key Butterworth crossover (120 Hz) splitting satellite and subwoofer outputs; modelled equivalent digital IIR biquads via bilinear transform.

• [Documentation](#)

ESP32 Gesture-Controlled Robot Car

2026

Embedded systems | ESP32, ESP-NOW, FreeRTOS

- Handheld MPU6050 tilt remote streaming over ESP-NOW to a 4WD car unit on a custom chassis, powered by 2×18650 cells with onboard charging.

EDUCATION

Tribhuvan University, IOE, Pulchowk Campus

May 2021 – May 2025

B.E. in Electronics, Communication and Information Engineering

GPA: 3.37/4 (74.64%)

SOS Hermann Gmeiner School

2019 – 2021

+2 Science (Physics and Mathematics)

Grade: A+

SOS Hermann Gmeiner School

2019

SEE

Grade: A+

TECHNICAL SKILLS

Languages: Go, TypeScript/JavaScript, Python, C/C++, SQL, Shell Scripting

Frontend: Next.js, React.js, Redux.js, Module Federation, Tailwind CSS

Mobile: React Native, Flutter

Backend: NestJS, GraphQL (Apollo Federation), gRPC, tRPC, Express.js, Django, Node.js, Microservices, ORY Kratos/Oathkeeper

AI / Data: RAG (LangChain, LlamaIndex), pgvector, Text-to-SQL, TinyML / INT8 quantization, n8n, Tacatron, OpenRefine, Pytorch, Numpy, Pandas

Databases: PostgreSQL/PostGIS, MongoDB, Redis

DevOps & Cloud: Docker, Kubernetes, Traefik, Kafka, Jenkins, AWS (EC2, S3, Lambda), CI/CD, Nginx, Tailscale

Embedded & Hardware: ESP32 (ESP-IDF, ESP-NOW), STM32, FreeRTOS, UART/I2C/SPI, KiCad PCB design, Digital Logic

Other: ODK Collect/Central, Machine Learning, Music Production