

Aadarsh Shrestha

Lalitpur, Nepal • +977 9803522629 • aadarsh@blackfire.com.np • aadarshshrestha1@gmail.com
linkedin.com/in/aadarsh-shrestha • github.com/Aadarshao • blackfire.com.np

Summary

AI and data engineer, and co-founder of BlackFire AI, which builds AI for Nepali and the low-resource languages of Nepal. Co-built Aawazz, a Nepali and English text-to-speech and voice-cloning product, and lead Tipni, meeting notes for Nepali, English and code-switched meetings. Before BlackFire, built ERP data pipelines and newsroom AI tools at Kantipur Media Group and MediaStack AI. Comfortable across the stack: data pipelines, speech and LLM workflows, backend APIs and React frontends. Available for selected freelance and contract work.

Technical Skills

AI & Speech: Text-to-speech and voice cloning, speech recognition, speaker diarization, LLM pipelines (Gemini, OpenRouter), transcript-grounded generation and claim verification, fine-tuning open models

Low-Resource Languages: Devanagari text processing, grapheme-to-phoneme conversion, phonetic coverage and recording-script design, Nepali-English code-switching

Data Engineering & Orchestration: Apache Airflow, Astronomer Cosmos, dbt-core, Apache Spark, ETL/ELT pipelines, medallion architecture (Bronze/Silver/Gold), data quality (QA/QC), PII pseudonymization

Backend & APIs: Python, Node.js (Express, Fastify, Hono), REST and WebSocket APIs, Prisma

Frontend: TypeScript, React, Next.js, Vite, Tailwind CSS, TanStack Query

Databases: PostgreSQL, MySQL, Oracle (source systems), SQL

Cloud & Tools: Docker, GitHub Actions, AWS, Cloudflare (R2, Workers), Vercel, Linux, Git & GitHub

Professional Experience

Co-founder, Chairperson & CEO

BlackFire AI (Blackfire A.I. Pvt. Ltd.), Lalitpur, Nepal

Jul 2026 – Present

- Co-founded BlackFire with Anup Adhikari to build AI for Nepali and the low-resource languages of Nepal; incorporated on 22 July 2026.
- Decide what the company builds and in what order, set pricing in NPR, and handle partnerships and startup programmes.
- BlackFire joined NVIDIA Inception in August 2026 and received startup credits from AWS Activate and Cloudflare for Startups.
- Still build: co-built Aawazz, lead Tipni and work on LipiSub (see Projects); delivered client websites for Root.ed and SHODH Nepal.

Data & AI Engineer

MediaStack AI (M&S Labs), Kathmandu, Nepal

Jun 2026 – Sep 2026

- **Cutaway** (frontend developer): built the React frontend for an AI clipping service that turns long-form video into short clips, with Nepali transcription, face-tracked reframing and Devanagari or Roman-Nepali captions (Fastify API, Python worker, Postgres, Cloudflare R2).
- **Newsflow** (full-stack developer): built a web editor for composing a newspaper's print edition, with a TypeScript layout engine and a PDF service (Hono, Playwright, Ghostscript) that renders print-ready PDF/X.
- **KMG Analytics Dashboard** (developer and data analyst): developed the group's analytics dashboard and analysed the data behind it.

Data Engineer

Kantipur NEO, Kantipur Media Group, Kathmandu, Nepal

Apr 2026 – Jun 2026

- Consolidated a company-wide Oracle ERP into PostgreSQL and rebuilt it as a medallion (Bronze/Silver/Gold) pipeline in dbt and Apache Airflow.
- Delivered AI-ready data to an in-house corporate AI for decision support.
- Pseudonymised personal data before the business layer, in line with Nepal's Individual Privacy Act 2075.

Data Engineering Intern

eXtensoData (F1Soft Group), Lalitpur, Nepal

Aug 2025 – Jan 2026

- Completed a structured internship covering core data engineering tools and workflows: ETL pipelines and Airflow DAGs with Spark, SQL and Linux.
- Practised monitoring, debugging and validating data pipelines, including formal QA/QC test plans.

Projects

Aawazz – Nepali and English Text-to-Speech

Co-creator & Engineer – BlackFire

- Built with Anup Adhikari from May 2026: text-to-speech, voice cloning from a 30-second sample and video dubbing, live on web and Android.
- Designed around Nepali phonetics from the start (conjuncts, independent vowels, sandhi) rather than adapted from an English model.
- 20+ studio voices, Nepali and English mixed in one sentence, and a WebSocket streaming API with under 300 ms to first audio.

Tipni – AI Meeting Notes

Creator & Lead – BlackFire (public beta)

- Meeting notes for Nepali, English and code-switched meetings; every note links to the transcript quote behind it, and a verification pass drops claims the transcript does not support.
- On-device recording with verified, resumable chunk uploads; consent-based speaker recognition; daily and total caps on AI spend; audited admin console.
- React, TypeScript, TanStack Query, Express, Prisma/MySQL (also the job queue), S3/R2 and Playwright tests, deployed on AWS.

LipiSub – AI Subtitles for Nepali Video

Developer – BlackFire

- Code-switched transcription, an AI correction pass, word-level timing, a live editor, and export as burned-in subtitles or SRT, VTT and ASS files; the browser extracts the audio track, so uploads stay small (React, TypeScript, Vite, Tailwind).

Newari Text-to-Speech Recording Script

Research – BlackFire

- Phased Python pipeline (Wikipedia and Wiktionary acquisition, language filtering, grapheme-to-phoneme, phonetic inventory, set-cover selection) producing a 507-line script (~49 minutes) covering 100% of reachable phonemes, 99.75% of diphones and 99.63% of conjuncts, verified by 12 automated acceptance checks.

Root.ed and SHODH Nepal Websites

Client Work – BlackFire

- Root.ed (sole developer): Next.js 16, React 19 and Vite on Cloudflare Workers, with a Three.js hero scene, low-power fallback and keyboard-accessible navigation.
- SHODH Nepal: React, TypeScript, Vite and Tailwind, with per-route SEO, GitHub Actions quality checks and browser regression tests.

Airflow File Processing System and Spark ETL Pipeline

Internship Projects – eXtensoData

- Event-driven Airflow DAG that compresses files uploaded to a shared Google Drive folder and emails their metadata; ran locally in Docker, with expected versus actual behaviour documented for QA.
- Spark ETL job that ingests CSV data, replaces missing values with the mode, splits datetime fields and loads consolidated output for analysis.

Volunteering & Community

Volunteer, TEDxIIMS

Kathmandu – Nov 2024

- Co-hosted the audience Q&A on stage at TEDxIIMS, an officially licensed TEDx event on the theme “Inspiring Ideas, Igniting Change”.

Aawazz Demo, Cursor Kathmandu

Lalitpur – Aug 2026

- Demoed Aawazz with Serena Shakya at Cursor Kathmandu, a community meetup for AI builders.

Relevant Coursework & Training

- Big Data Technologies (Hadoop ecosystem, Spark batch and streaming concepts); Apache Airflow (DAG design, scheduling, event-driven pipelines)
- Linux Systems, Version Control with Git, SQL to Python (Udemy), Pandas

Education

Bachelor of Computer Science (Hons)

Taylor's College, studied at IIMS College, Kathmandu

2022 – 2026

High School

Little Angels' School, Lalitpur

Jul 2019 – Jul 2021

Certifications

- SQL (Advanced) and SQL (Intermediate) – HackerRank, 2025
- AWS Certification (Foundational) – Amazon Web Services, 2022