

BASANT BHATTARAI

SENIOR DATA & AI ENGINEER / PLATFORMS, DATABASES & AGENTIC SYSTEMS

Kathmandu, Nepal | basantbhattarai.com.np | technobasant9@gmail.com | linkedin.com/in/technobasant | github.com/technobasant

PROFILE

Senior data and AI engineer with 9 years in software engineering, six of them owning production data platforms. Open-source first — Kafka, Spark, Iceberg, Trino, dbt and Airflow on Kubernetes — across ingestion, transformation and analytical serving, plus the databases underneath: modeling, replication, backup and recovery, upgrades. Designs and runs the agent workflows on that same stack — LangGraph, Google ADK and MCP tool servers — so model output lands as typed, reviewable rows rather than a transcript. Track record: 40% lower infrastructure cost, 50% faster processing, 60% faster analytics delivery.

CORE STRENGTHS

Platform ownership Consumer-facing reliability, governance, recovery, and operational readiness.	Data engineering Batch and streaming pipelines, orchestration, modeling, and change-safe processing.
Database lifecycle Modeling, replication, failover, backup and PITR, upgrades, and query performance.	Agentic systems LangGraph, Google ADK and MCP tool servers, evaluation, and typed persistence.
Technical leadership Design review, mentoring, runbooks, and making system ownership transferable.	Distributed work Long-term collaboration across European and US time zones from Kathmandu.

EXPERIENCE

Senior Data Engineer | UXCam

Feb 2024 - present

- Own reliability and governance for a platform handling 15 TB+ and 100M+ events daily: freshness and latency objectives per dataset, 99.9% uptime, incident-ready runbooks, reviewable data contracts, and a named owner for every published table.
- Design and operate the production agent workflows — LangGraph and Google ADK orchestration, MCP tool servers, retrieval and evaluation — against the existing platform, so model output lands as typed, provenanced rows rather than a transcript: analytics delivery 60% faster with 75% less analyst effort, review and fallback paths kept.
- Lead architecture and operational review across ingestion, transformation, storage, analytical serving, observability and recovery, without treating any single tool as the architecture.
- Run capacity and cost work as a design input rather than a monthly report, using storage layout, partitioning and compaction cadence as the primary levers.
- Mentor 5 engineers through pairing, code review and written design feedback; team delivery time down 30%, with the emphasis on making system ownership transferable.

Data Engineer | UXCam

Feb 2020 - Feb 2024

- Built and operated batch and streaming pipelines, then took on the data models, orchestration, storage and analytical serving paths around them.
- Rebuilt the core ETL layer in Spark and PySpark with Airflow orchestration, cutting processing time for 5 TB+/day by 50% and unblocking real-time product dashboards.
- Made change and recovery safe by treating schema evolution, replay, late-arriving data and backfills as designed interfaces with idempotent loads and explicit watermarks, not as emergency procedures.

-
- Tuned Trino, ClickHouse, Citus and TimescaleDB serving 100M+ queries/day for a 50% p95 latency reduction through partitioning, clustering and model redesign.
 - Led the Kubernetes migration of the data infrastructure: 40% lower infrastructure cost, 99.9% uptime, and autoscaling that held under peak load.
 - Worked across European and US time zones to turn product questions into maintainable datasets.

Project Leader | SVCET, India

2019 - 2020

- Led 4 developers building an intelligent dialogue system on an NLP and deep-learning stack, reaching the client's acceptance threshold for intent recognition and delivering 2 weeks ahead of schedule.
- Coordinated requirements, planning and review across academic and industry stakeholders, and introduced the code-review and version-control discipline the department kept afterwards.

Software Developer | SV Technology, India

2017 - 2019

- Built Python and FastAPI backend services carrying 10K+ users at 85% test coverage.
- Designed and maintained PostgreSQL schemas, stored procedures and reporting paths, and set up CI with Jenkins alongside the early pipeline work that moved me into data engineering.

CAPABILITY MAP

Platform ownership

Python and SQL; Kafka ingestion; Spark/PySpark processing; Airflow orchestration; Iceberg, dbt and lakehouse modeling; Kubernetes, Docker and Terraform delivery; contracts, lineage, cost and recovery.

Database lifecycle

PostgreSQL, ClickHouse, Trino, Citus, TimescaleDB, Redis and MongoDB; transactional and analytical modeling, partitioning, replication, backup, restore, high availability and query performance.

Agentic systems

LangGraph, LangChain, Google ADK, MCP servers, Pydantic, Milvus and MLflow; bounded tools, retrieval, evaluation before rollout, typed output, provenance, human review and honest fallback paths.

Technical leadership

Architecture and operational review, mentoring, incident-ready runbooks, service and freshness objectives, observability with Prometheus and Grafana, and distributed delivery across time zones.

EDUCATION AND LANGUAGES

Bachelor of Technology in Computer Science and Engineering

JNTUA College of Engineering, Anantapur, India | 2015 - 2019

English: C1 | Nepali: native

Selected work, writing and case studies: basantbhattarai.com.np/work/