Saptashwa Bandyopadhyay

Backend and distributed-systems engineer · 9+ years · led a team of ~8 · platforms for 10,000+ employees

contact@saptashwab.com | Kolkata, India | linkedin.com/in/saptashwa-bandyopadhyay-ab9a7032 | github.com/saptashwab

KEY ACHIEVEMENTS

- **10,000+ tags a second** on a **150,000**-tag platform: Python ingestion into MongoDB and Azure IoT Hub for one of India's largest steel manufacturers, owned end to end as DRI. Designed the edge-to-cloud architecture and wrote the ingestion.
- **15× faster** semantic search (**30+ s** to under **2 s**), adopted by **90%** of **10,000+** employees across internal tools that were searched separately before.
- **30,000+ user accounts** onboarded through one identity and access module reused by **3** platforms; designed its OAuth token management, session lifecycle and per-application session limits.

EXPERIENCE

Manager, Architecture & DRI Lead — PwC India Apr 2026 – Present

Industrial IoT data platform (steel manufacturer)

- DRI and architecture owner, hands-on in Python.
- Cut end-to-end latency from **8+ s** to under **1.5 s** (**5×**) across the Kepware, Ignition and Azure IoT Hub pipeline.
- Designed dual writes to MongoDB and Azure IoT Hub so data survives the failure of either sink; owns MS SQL schema design and query tuning.
- On-call escalation point; wrote the design documents, runbooks and incident procedures across **3** integrated systems.

Pre-sales and solution architecture

- Leads pre-sales and solution architecture, about 30% of the current role.
- Won **3 of 3** bids since leading pre-sales in September 2025: proposals, solution designs, estimates and client demos.

Tech: Python, Azure IoT Hub, MongoDB, MS SQL, Ignition, Kepware

Senior Associate, Team Lead & Platform Owner — PwC India Apr 2022 – Mar 2026

AI semantic search and insight platform (firm-wide)

- DRI and platform owner, leading a team of **~8** developers.
- Built semantic search on Weaviate and PostgreSQL pgvector, evaluated relevance with a labelled query set, and tuned chunking and hybrid search.
- Self-hosted GitLab CE and Runners and built CI/CD from scratch, after finding the department's instance could not reach the deploy VMs.
- Worked directly with **2** Partners and **3** Directors on requirements and priorities.

Identity and access module (3 platforms)

- Main designer of the module that connects each application to the firm's identity platform.
- Worked closely on implementing RBAC; implemented SSO federation for client organisations and MFA.

Enterprise SaaS (contract management, healthcare)

- Backend engineer and technical lead for code and security reviews, for **5+** enterprise clients.
- Designed multi-tenant PostgreSQL with a schema per client, and moved long-running work onto RabbitMQ and ActiveMQ.
- Resolved **500+** issues as technical lead for code and security reviews; led hardening that cut vulnerabilities by **70%**.

US banking (two banks)

- Two separate US banks: hands-on builder at one, DRI for a platform migration at the other.
- Built a Python pipeline converting SWIFT and Excel payment formats at **99.9%** accuracy across **500+** fields per file.
- Automated validation from **2–3** developers × **2** days to under **10** minutes.
- Led the Jenkins to GitLab CI/CD migration: builds **400%** faster for **20+** engineers; set up and managed the Kafka pipeline.

AWS serverless app (steel pipes manufacturer)

- Worked with the client's managers and shop floor on what to measure, then designed and built the app.

- Built Python Lambda services taking machine events from AWS IoT Core into DynamoDB, with API Gateway endpoints for real-time access.

Tech: Node.js, Python, PostgreSQL (pgvector), Weaviate, RabbitMQ, Kafka, AWS Lambda, GitLab CI/CD, Docker, Terraform

Associate — PwC India Jul 2018 – Mar 2022

- Backend and image-processing projects for enterprise clients.
- Improved licence-plate OCR accuracy by **40%** with a custom-trained KNN model for a smart parking system.
- Built Node.js microservices for a contract lifecycle SaaS serving **5+** clients, with digital signature, DMS and BPM integrations.
- Built a code review tool on SonarQube and Pylint, and document and barcode processing in Python and Flask.

Tech: Python, Node.js, Flask, Cassandra, PostgreSQL

Software Systems Analyst — Open Source Software Solutions Oct 2017 – Jul 2018

CRM data sync platform (HubSpot and Microsoft Dynamics)

- Designed and built a bidirectional HubSpot and Microsoft Dynamics sync engine on a **10-minute** cycle for several enterprise clients, cutting sync time by **65%**.

Tech: PHP, REST APIs

PROJECTS

RoleRadar: LLM job-scoring pipeline (designed and directed; agents wrote most of the code) | saptashwab.com/p/roleradar

- Designed a two-step LLM scoring pipeline and froze the eval protocol before the run: **21 of 21** good jobs caught on **59** untouched postings, **0** ineligible jobs marked eligible.
- Measured inference cost across **9** models: hidden reasoning cost about **42×** more and ran **29×** slower on the same jobs.

Kloister: self-hosted project tracker (designed and directed; agents wrote much of the code under review) | saptashwab.com/p/kloister

- Wrote the spec and release rules and reviewed every milestone: **2** teams at the day job plan their work in it; **8** releases in **18** days, with **80** browser tests on every CI run.

Saltmere: browser city-builder on IPCC sea-level data (designed and directed; agents wrote most of the code) | saptashwab.com/p/saltmere

- Set the game design and the climate model's rules: **410** players in its first **22** days, about **40** minutes a session.

SKILLS

Languages: Python, Node.js, TypeScript, SQL, Dart (Flutter)

Data: PostgreSQL (schema design, multi-tenant, pgvector), MS SQL, MongoDB, DynamoDB, Cassandra, Weaviate

Distributed systems: Kafka, RabbitMQ and ActiveMQ, MQTT, event-driven microservices

Cloud and delivery: AWS (Lambda, IoT Core, DynamoDB, API Gateway), Azure (IoT Hub, DevOps), CI/CD (GitLab, Azure DevOps, Jenkins), Docker, Terraform, Kubernetes (familiarity, actively strengthening)

Security: OAuth and session design, SAML, SSO and MFA, RBAC, security code review

AI and LLM: semantic search (pgvector, Weaviate), LLM evals, LLM cost and latency measurement, coding agents (Codex, Claude Code), MCP servers (exploratory)

EDUCATION

B.Tech, Computer Science and Engineering 2014 – 2018

Indian Institute of Engineering Science and Technology (IIEST), Shibpur

Languages: English, Hindi, Bengali (professional)