

AYAN PUTATUNDA

Staff Data Engineer · Data Platforms for AI · Agentic AI Systems
Pleasanton, CA | 510.488.8555 | ayanputatunda87@gmail.com

PROFESSIONAL SUMMARY

Staff Data Engineer with 15+ years architecting petabyte-scale data platforms that power production AI applications. Currently leading enterprise AI-readiness initiatives at Zendesk — designing context governance and Model Context Protocol (MCP) architectures that let AI agents safely consume governed enterprise data, and building agentic frameworks that automate data engineering workflows end to end. Proven record of taking AI-driven systems from concept to production, leading cross-functional teams, and translating ambiguous business problems into scalable engineering solutions. Builds and ships full-stack AI products independently.

CORE SKILLS

GenAI & Agentic AI: Anthropic Claude API, Claude Code, MCP (Model Context Protocol), LangGraph, multi-agent systems, RAG architectures, vector search, Snowflake Cortex (LLM functions, semantic views, Cortex Analyst), LLM evaluation & prompt engineering

Data Platform: Snowflake, dbt, Airflow, Fivetran, Google BigQuery, Spark, ELT/ETL design, streaming & batch pipelines, data modeling (Kimball), semantic layers

Languages & Cloud: Python, SQL, Docker, GitHub Actions / CI-CD, Google Cloud Platform (GCP), Snowflake Data Cloud

Leadership: Technical mentoring, team building & management, stakeholder management, engineering best practices, customer-facing technical discovery

PROFESSIONAL EXPERIENCE

Staff Data Engineer | ZENDESK Nov 2024 – Present

San Francisco, CA

- Designing an enterprise context governance system for the Zendesk Data Platform — curating, governing, and versioning business context (metrics, semantics, lineage, domain knowledge) so AI agents operate on trusted, consistent enterprise data.
- Architecting a context-as-MCP layer that exposes governed data context to LLM-powered agents through Model Context Protocol servers, enabling intelligent, AI-driven workflows across the data stack.
- Pioneered an agentic AI framework using Claude Code and LangGraph, targeting an 80% reduction in manual data pipeline development across multi-repository dbt and ingestion workflows.
- Built Snowflake Cortex semantic layers (semantic views, AI-assisted SQL generation) enabling governed natural-language analytics over petabyte-scale data.
- Architected petabyte-scale data infrastructure with Fivetran, Airflow, and dbt, improving delivery efficiency by 40% while maintaining 99.9% uptime for enterprise-grade data services.
- Established engineering best practices across 5+ teams; mentor engineers on technical decision-making for complex, multi-system, long-term projects.

Staff Data Engineer | ACHIEVE May 2022 – Nov 2024

San Mateo, CA

- Led the FDR Sales Data Transformation on Google Cloud — architected the data model integrating Salesforce CRM into a BigQuery data lake serving 500K+ unified customer profiles.
- Designed robust batch and streaming ETL frameworks on GCP standardizing financial data as ML model inputs for data science teams, reducing data inconsistencies by 60%.
- Partnered with data scientists, analysts, product managers, and business leaders to drive data-driven roadmaps, real-time KPI monitoring, and personalized customer experiences.

Engineering Manager · Principal Data Engineer · Senior Data Engineer | NOODLE.AI Sep 2019 – Apr 2022

Palo Alto / San Francisco, CA

- Architected the end-to-end data pipeline for a first-in-class supply chain AI platform, processing TB-scale customer data with 99.9% uptime as the foundation for ML-driven demand and inventory recommendations.
- Built and scaled the data engineering team from 2 to 8 engineers with 100% retention, promoting 3 engineers to senior roles.

- Designed configurable, customer-agnostic data models and a reusable Airflow framework, cutting customer implementation time by 50% and accelerating pipeline development by 60%.

Team Lead · Senior Developer | COGNIZANT Oct 2010 – Sep 2019

Kolkata, India · Buenos Aires, Argentina · San Francisco, USA

- Designed and maintained multi-channel retail POS data integration solutions (RAYMARK, MSBI, SQL Server) for global retail clients, achieving 99.5% data accuracy with fault-tolerant ETL.
- Built automated intra- and inter-system reconciliation processes, eliminating manual validation and reducing errors by 85%.

AI SIDE PROJECTS — axiomic-ai.com

Personal side-project umbrella for hands-on agentic AI experimentation, spanning product builds (Labs), AI consulting frameworks (Solutions), and tech education (Academy). Each product designed, built, and deployed end to end — full-stack development, LLM integration, and cloud deployment. Link - <https://axiomic-ai.com/>

- **PersonasFlow** — AI-powered life-management platform: intelligent daily routines, persona management, and progress tracking across every role a user plays. Link - <https://personasflow.com/>
- **EstateVision** — rental portfolio management platform for property owners: portfolio tracking, centralized document store, and automated rent-reminder notifications. Link - <https://estatevision.io/>
- **WealthPilot** — AI-driven personal finance system recommending monthly investment amounts and allocations personalized to each user's risk tolerance and financial goals. Link - <https://www.pilotmywealth.com/>

EDUCATION

Master of Computer Science in Data Science — University of Illinois Urbana-Champaign **2018 – 2019**

B.Tech in Computer Science and Engineering — West Bengal Institute of Technology, India **2006 – 2010**

PUBLICATIONS & COMMUNITY

- Technical Reviewer, *Essential PySpark for Scalable Data Analytics* (Packt Publishing, ISBN 978-1-800563094).
- Guest speaker for MS program in Data Science & Artificial Intelligence at University of San Francisco.
- Creator, "Ayan in Tech" — YouTube channel teaching data engineering and applied AI.