

Sourav Banerjee

Houston, TX | sourav@utexas.edu | (832) 973-1168 | souravbanerjee.com | linkedin.com/in/sourav-ban

SUMMARY

Data Governance Lead and Data Engineer with 2+ years driving enterprise-scale data strategy at a Fortune 500 manufacturer. Proven track record leading cross-functional data initiatives, architecting MDM programs, and establishing data governance practices with measurable business outcomes.

EXPERIENCE

Data Governance Lead

Daikin

July 2024 - Present, Houston, TX

- Led the Product Master Data Management (MDM) program to establish a single source of truth across 12 business domains, cutting data acquisition time by 70% and improving consistency by 80%; designed data mapping frameworks and automated workflows in Semarchy, Snowflake, and AWS as part of a broader SAP transition, reducing manual user input time by 50%.
- Eliminated \$100K+ in projected spend and hundreds of anticipated manual-classification hours by architecting an LLM-integrated product classification system with Snowflake Cortex, delivering a validated product hierarchy as a critical asset within the MDM program.
- Championed company-wide data governance policy across 10 business units. Established data ownership frameworks, stewardship models, and metadata classification standards that standardized how the business receives and handles data.
- Deployed Atlan as the central data governance platform to 280 users, standardizing metadata practices across Snowflake, Tableau, and SQL Server with integrations into Okta, ServiceNow, ThoughtSpot, and SAP; making enterprise data discoverable and trustworthy at scale.
- Directed the implementation of a natural-language analytics platform using ThoughtSpot, enabling self-service business intelligence for upper management across 3 executive dashboards; reduced reliance on ad hoc data requests.
- Built robust ETL pipelines using Snowflake, AWS Glue, and Matillion to extract data from AS400 and other legacy systems, reducing latency and enabling seamless connectivity with reporting platforms including Tableau.

SKILLS

Data & Cloud: Snowflake, AWS (Glue, Lambda, SQS), Atlan, Semarchy, Matillion, Tableau, SQL Server, ThoughtSpot, SAP, Git, Firebase, Figma

Programming: Python, SQL, Java, React.js, Kotlin, Swift, Dart / Flutter

Governance & MDM: Data Governance, Master Data Management, Metadata Management, Data Stewardship, ETL / ELT

Languages: Native English, Intermediate Japanese, Conversational Bengali, Basic Hindi, Basic Spanish

EDUCATION

Bachelor of Science, Computer Science

Minor in Japanese • The University of Texas at Austin • Austin, TX • 2024 • 3.82

Relevant Coursework: Machine Learning, Artificial Intelligence, Human-Computer Interaction, Operating Systems, Software Engineering, Networks, Mobile Development (iOS & Android), Game Programming Paradigms

PROJECTS

Gakuji | *React Native, Python, Supabase, Netlify, Railway, Git*

- Desktop application for learning Japanese vocabulary and Kanji through song lyrics.
- Built frontend in React Native with a Python FastAPI backend, connected to a PostgreSQL database on Supabase.
- Originally deployed using AWS Lambda, SQS, and Netlify, converted to a packaged desktop application with electron.

Chaat | *React, Figma, Lovable*

- Spotify for food: a discovery platform helping consumers find dishes that match their tastes and explore local culinary scenes.
- Operated as UI/UX design lead and front-end developer: prototyped wireframes in Figma, conducted A/B testing with end users, and led development through two implementation cycles (Flutter → React).
- Managed the complete development lifecycle of the Chaat application, ensuring seamless progression from user requirements and prototyping to final feature delivery.

HONORS & AWARDS

- Certificate of Global Engagement, University of Texas at Austin, Tokyo, Japan (2023)
- Outstanding Japanese Language Student Award, 2023–2024