

Piyush Datta

Website: <https://www.piyushdatta.com> |
GitHub: <https://github.com/PiyushDatta> |
Email: piyushdattaca@gmail.com |

Work Experience

- Meta** San Francisco Bay Area, CA
Software Engineer 2025 – Present
 - Core engineer on Meta Ads Infrastructure, developing low-latency C/C++ services powering global ad delivery.
 - Designed large-scale ML calibration systems that increased revenue by ~\$10M/year and reduced fleet CPU usage by 2–3% across services handling millions of RPS.
- Amazon Web Services** San Francisco Bay Area, CA
Software Engineer 2020 – 2025
 - Database kernel (MySQL/Postgres) engineer on the AWS Aurora Database Engine team using C/C++.
 - Core engineer on the new *Aurora Limitless* product, presented at 2024 re:Invent (<https://aws.amazon.com/about-aws/whats-new/2023/11/amazon-aurora-limitless-database/>).
 - Led initiatives to create a sharded distributed storage engine capable of handling more than 1M TPS.
- Givex** Toronto, ON
Software Engineer 2019 – 2020
 - Backend development using Python, JavaScript, C/C++, PostgreSQL, and Unix.
 - Implemented gift card, point-of-sale, and other transaction-based solutions for Fortune 500 companies.

Contributions & Projects

- 1st Place – Google Technology Prize, Hack for Social Impact 2025** <https://devpost.com/software/doorwai>
Python, React Native, Node/Express, Google Cloud
 - Created a fully functional cross-platform app that acts as an AI assistant for youth facing homelessness.
 - Enabled users to chat with the AI assistant via text or speech; the assistant could also place phone calls to agencies on the user's behalf without using the user's phone number.
- 1st Place – Novice Champion, LLM Chess Hackathon (StrongCompute)** <https://github.com/PiyushDatta/chess-hackathon>
Python, PyTorch, TensorBoard
 - Trained a 3M-parameter vision-LLM agent from scratch using 8xH100s, FlashAttention, FSDP, GELU, and AMP.
 - Won Novice Champion by having my LLM agent defeat other LLM agents in 1-on-1 chess matches (100-move max).
- Amazon Patent (Pending) – AWS Aurora AI/RL-Based Database Configuration System**
 - Lead designer and inventor of a database configuration system that automatically predicts optimal database configurations based on customer workloads using AI reinforcement learning techniques.
- Contributor to TensorFlow – Google Brain Open-Source Project** <https://www.tensorflow.org>
C/C++, Python
 - Contributed to an open-source project enabling developers to create and train ML models.

Skills

Technologies/Concepts: C/C++, Rust, Python, JavaScript/TypeScript, Distributed Systems, IPC, 2PC, ML/AI, LLM, Reinforcement Learning, Consensus Algorithms, NewSQL/Distributed SQL

Education

University of Toronto, St. George Campus Toronto, ON