Harkanwalpreet Singh

harkanwal.p.singh@gmail.com | LinkedIn | GitHub | +91-9779238824

Experience

Senior Software Engineer (L3)

Twilio   *Sept 2024 – Present*

Governance Team
- Implemented and launched **Consent in Engage (Public Beta)**, enabling compliance with user data preferences for Engage Audiences.
- Re-architected the **Deletion Orchestration Reporting Subsystem**, eliminating data loss incidents and improving reliability.
- Optimized Redis usage in the **Privacy System**, removing performance bottlenecks and ensuring **100% detection** of PII classifications.

Compute(Batch) Team
- Implemented key components of the **Batch-to-Realtime Compute Migration** project, projected to reduce company **COGS by ~$6M annually**; owned design of control plane RPC changes and Conductor workflow orchestration.
- Optimized SQL queries in **Batch Compute**, reducing DPU costs by up to 80% and query execution times by up to 95% (15 hours → 45 minutes).
- Designed a resumable S3 offset reader for **Realtime backfills**, enabling continuation from failure points and saving days of reprocessing during interruptions.

Senior Software Engineer (L3)

Bazaarvoice   *Aug 2022 – Aug 2024*

Social Commerce Team
- Designed and implemented a **GDPR-compliant Deletion Orchestrator**, leveraging AWS Step Functions and DynamoDB, ensuring secure data deletion across multiple downstream systems.
- Developed **cache invalidation system** (Redis Hash + Cloudflare APIs) enabling instant media updates, unlocking partnerships with news publishers requiring near-real-time content delivery.
- Built a **media transcription & translation pipeline** (AWS MediaConvert + message queues + OpenAI Whisper API) producing WebVTT captions at scale.
- **Led development of Generative AI Captioning** from POC to production, integrating OpenAI APIs and evolving prompt strategies; adopted by multiple enterprise clients.
- Reduced **storage costs by 20%** by implementing S3 retention policies and a Content Tracking Service to auto-delete unused media/metadata.

Software Engineer II (L2)

Morgan Stanley   *Jan 2021 – Aug 2022*

Data Distribution Team
- Designed and delivered **data distribution API** for financial reporting with Redis caching, ETL for PII masking, and scalable data link APIs.

- Automated **Teradata schema deployments** (tables, views, stored procs) with Jenkins-integrated plugins, reducing manual effort and errors.
- **Led migration of legacy ETL (Informatica → in-house tool)**: recruited and mentored contractors, defined migration strategy, and helped in building architectural roadmap.
- Replaced **CRON-based jobs with Kafka-based pipelines**, reducing downstream data lag by **90%**.

Software Engineer (L1)

Morgan Stanley   *Apr 2019 – Dec 2020*
ETL team
- Built **Teradata query simulator** with ElasticSearch + Scala pipeline, enabling pre-deployment performance testing.
- Enhanced internal ETL tool (Python & Perl) used by multiple teams for high-volume data processing.
- Automated migration from **Perforce → Git**, preserving commit history for legacy projects.

Software Analyst (L1)

Capgemini   *Sep 2018 – Mar 2019*

- Automated **HR reporting dashboards** by integrating Tableau with Python (TabPy).
- Completed 5-month **industry-level Core Java training**.

Skills

Languages: Python, Go, Scala, Java, JavaScript/TypeScript, C++, Bash, SQL
Frameworks: Spring Boot, Node.js, React, FastAPI, Finatra, Conductor, Airflow, Spark, Hadoop
Cloud & DevOps: AWS (EC2, ECS, Lambda, S3, DynamoDB, Step Functions, Kinesis, CloudFront, CloudWatch, RDS), Kubernetes, Docker, Terraform, Cloudflare, Jenkins, GitHub Actions
Data & Streaming: Kafka, Redis, Elasticsearch, Trino, Snowflake, Databricks, Athena
Observability: Datadog, Grafana

Education

Indian Institute of Technology (IIT) BHU, Varanasi   *Jul 2014 – May 2018*

Recognitions

Twilio - Operational Spotlight (Q2 2025)
Bazaarvoice - Global Engineering Recognition, Global Gong Award
Morgan Stanley - Highest Rated Employee (two consecutive years)