Marium Ashraf

Phone: (437) 220-1426 • E-mail: mariumashraf.MA@gmail.com • [LinkedIn](#) • [Website](#) • [Github](#)

EDUCATION

B.Eng in Computer Engineering (Major in Software Engineering)**,** Toronto Metropolitan University   Graduating **April 2027**

TECHNICAL SKILLS

Languages: Python, Java, JavaScript/TypeScript

AI & LLMs: GPT-4o, LLM evaluation, RAG, embeddings, semantic analysis, MLflow

Distributed Systems: Docker, microservices, message queues (RabbitMQ), observability, fault tolerance

APIs & SDKs: REST, gRPC, Slack API, SDK design patterns, versioning

Databases: PostgreSQL, MongoDB, BigQuery, vector databases (ChromaDB)

Cloud & DevOps: AWS, GCP (Pub/Sub, Airflow), CI/CD, GitHub Actions, Linux

Data & ML Systems: ETL/ELT, analytics pipelines, feature engineering, usage telemetry

Testing & QA: PyTest, Jest, automated test suites, validation pipelines

RELEVANT EXPERIENCE

Operations & AI Intern, Lyrata   *July 2025 – Present*

- Designed and deployed AI assisted tooling to parse, validate, and surface operational knowledge from structured and unstructured sources (Python, Flask, PostgreSQL, ChromaDB, GPT-4o), reducing SOP lookup latency by ~90% and enabling consistent, auditable system behavior.

- Built a production-grade internal "SDK-like" interface (LyrataGPT) that delivers cited, procedure-accurate responses to technicians, improving developer-style ergonomics for non-technical users and cutting onboarding ramp time by ~50%.

- Applied Vector Institute DaRMoD methodologies to produce ML systems, implementing reproducible evaluation pipelines (regex vs RAG vs hybrid extraction), MLflow experiment tracking, containerized services, and fault-tolerant routing under real-world conditions.

- Engineered distributed automations across Slack APIs, Google Sheets, and PostgreSQL, implementing schema versioning, idempotent handlers, validation layers, and recovery logging; centralized heterogeneous data sources into a unified reporting and analysis stack, improving data accuracy by ~40%.

Prompt & AI Engineer, Outlier   *Nov 2022 – Jul 2025*

- Developed microservice LLM evaluation pipelines with automated rubric scoring, anomaly detection, and usage analytics, scaling to daily inferences while improving model quality by ~35%.

- Built SQL-backed observability and analytics systems with real-time dashboards and automated alerts, reducing debugging cycles by ~60% and improving reliability of production deployments.

- Led architectural refactors of distributed inference infrastructure, optimizing container orchestration, caching strategies, and load balancing to improve latency and system stability under heavy traffic.

- Translated business and product requirements into technical specifications for cross-functional teams, introducing standardized evaluation and release workflows that accelerated feature delivery by ~30%.

Web Developer Intern, While She is True   Jan 2022 – Dec 2023

- Built production full-stack features using React and backend APIs, including authentication flows, SQL data models, and Stripe integrations, ensuring secure and reliable client–server communication.

- Worked against clearly defined API contracts, contributing to stable interface boundaries between frontend and backend systems and reducing integration errors during feature iteration.

- Implemented automated testing and CI/CD pipelines (Jest, React Testing Library, GitHub Actions) to support predictable releases, minimize regressions, and improve system reliability over time.