Somil Arora

linkedin.com/in/somil855 | github.com/Somilstar | somilarora55@gmail.com | +1(647) 612-5182 | Toronto,ON

TECHNICAL SKILLS

Languages: Java, Python, TypeScript, JavaScript, C/C++, Go, SQL, Shell/Bash, HTML5, CSS3

Backend & APIs: Spring Boot, Node.js, Express.js, gRPC, Protocol Buffers (proto3), REST API design, Microservices architecture, Distributed & event-driven systems

Frontend: React.js, Next.js, Tailwind CSS, responsive UI development

Cloud & DevOps: AWS (EC2, S3, RDS, EKS, IAM, CloudWatch), Azure (AKS, ARM, VMs, Managed Instances), GCP (GKE), Docker, Kubernetes, Jenkins, GitHub Actions, CI/CD pipelines, Gradle, Maven, Git/GitHub

AI & Machine Learning: LLM API integration (OpenAI, Anthropic Claude), Retrieval-Augmented Generation (RAG), vector databases, LangChain, prompt engineering, TensorFlow.js, AI-assisted development (GitHub Copilot, Claude Code, Cursor)

Databases & Testing: PostgreSQL, MySQL, MongoDB, JUnit, Postman, Swagger, TestRail, SQL-based data validation

EDUCATION

York University **Toronto, ON**
Bsc (Hons) Computer Science *September 2022 - April 2027*

- Courses: Data Structures and Algorithms, Design and Analysis of Algorithms, Operating systems, Introduction to AI and Logic programming, Database Management, Advanced Software Engineering, Object Oriented Programming.

WORK EXPERIENCE

IBM **Markham, ON**
Software Developer Intern *May 2025 - August 2026*

- Delivered end-to-end **AWS EKS** parking support across **3 microservices**, extending Turbonomic's managed-Kubernetes cost-optimization coverage and rolling out safely to production behind phased feature flags.
- Designed and shipped a full-stack Parking Mode capability spanning new **Protobuf/gRPC** schemas, service and repository layers, and 2 new **REST** endpoints, giving enterprise customers programmatic control over workload termination vs. suspension.
- Re-architected **AWS probe** discovery for **cost-optimization** deployments by selectively skipping **8 non-essential entity** types, cutting **unnecessary API calls** and reducing pipeline data volume in production.
- Built a customer-facing Azure resource-lock safety system integrating the **Azure Resource Manager** API across **5+ entity** types, resolving a live enterprise support escalation and preventing accidental parking of protected resources.
- Root-caused and fixed a silent data-correctness bug that was dropping database entities without error, and traced a customer-reported **Azure API** throttling incident to its source through production log analysis.
- Merged **39+ pull requests** and **closed 44 Jira tickets** across **8 repositories and 3 cloud platforms** over 16 months, coordinating feature delivery with backend, cloud-probe, frontend, and QE teams.

Gao Tek Inc **Remote**
Tech support Intern *October 2023 - January 2024*

- Improved site performance through backend query and server-side logic optimization, cutting page load time by 25% and reducing ongoing maintenance effort by 30%.
- Authored internal technical documentation that shortened onboarding time by 20% and reduced average support-ticket resolution time.
- Partnered with developers and support staff in a fully remote environment to triage incoming tickets, prioritize recurring issues, and communicate resolutions to non-technical stakeholders.

PROJECTS

Smart Event Processing and Analytics Platform
Developed a real-time, event-driven analytics platform for processing and analyzing high-throughput data streams across various industries.

- Architected a real-time, event-driven analytics platform for ingesting, processing, and visualizing high-throughput data streams, built on **Java** and **Spring Boot** with **Apache Kafka** as the event backbone.
- Designed the **Kafka** topic and partitioning strategy to support parallel consumption across consumer groups, enabling horizontal scaling of processing workers without downstream reprocessing.
- Built a real-time **KPI** dashboard in **React.js** integrated with **AWS QuickSight**, surfacing streaming metrics to end users with minimum refresh latency.
- Integrated real-time alerting and notification workflows via **AWS SNS/SQS**, decoupling alert delivery from the processing pipeline and improving operational response time by **50%**.