# Stephen Guo

Toronto, Ontario, Canada     stephenguo14@gmail.com
+1 (437) 231-1263     Website     GitHub     LinkedIn

## EDUCATION

**University of Toronto**                                                                                          **Toronto, ON**
Honours Bachelor of Science, Computer Science                                    September 2019 - April 2024
cGPA: 3.57 / 4.0

## EXPERIENCE

**Ridges**                                                                                                              **Toronto, ON**
Backend Software Engineer                                                                    September 2025 - Present

- Developed and maintained a **Python FastAPI**, **PostgreSQL**, and **Next.js** platform enabling **1200+** users to build AI agents, benchmark model performance, and participate in ranking systems, resulting in a state of the art agent solving **96%** of Polyglot Python problems
- Took cross-functional ownership in a high-pressure startup environment by balancing backend development, SRE tasks, support tickets, and infra work, accelerating feature delivery by **30%** and reducing dependency on specialized roles
- Maintained production reliability for a high-uptime AI platform by SSHing into machines, analyzing logs, deploying live patches and applying targeted caching strategies, resulting in **99% uptime**
- Reduced cloud infrastructure costs by implementing EC2 autoscaling policies, automatically shutting down idle development environments, and migrating workloads off Vercel to self-managed infrastructure, resulting in **45%** lower monthly cloud spend and **$60k+** annualized savings while maintaining system performance
- Prototyped a **Kubernetes** deployment platform using **ArgoCD**, **Helm**, **Kustomize**, and **OpenTofu**, creating a launch-ready infrastructure foundation that was later deprioritized due to shifting product priorities

**Leap Softwares**                                                                                              **Toronto, ON**
Software Engineer                                                                           October 2024 - September 2025

- Built secure, type-safe **GraphQL** endpoints and automated PDF/Excel reporting tools for a financial services management system, using **TypeScript**, **NestJS**, **PostgreSQL**, **TypeORM**, and **Redis**, ensuring reliable data delivery and improved financial decision-making
- Introduced **Azure** Monitor alerts and health checks integrated with Slack via **Terraform**, **Rust**, and websockets, cutting service downtime by **95%** and reducing on-call load for engineers
- Responded directly to customer support tickets, troubleshooting production issues and coordinating fixes across teammates, improving average resolution time from 1 hour to under **5 minutes** and ensuring a consistent positive customer experience.
- Migrated data from a legacy financial application with a non-normalized schema and redundant records to a modern platform, designing scripts that cleaned, and transformed **100,000+** rows with zero downtime or data loss
- Optimized core reporting features by analyzing **PostgreSQL** query plans and adding targeted indexes, cutting average response times from 2 seconds to under **400 milliseconds**

**BlueCat Networks**                                                                                          **Toronto, ON**
Software Engineer Co-op                                                                           May 2021 - April 2022

- Deployed and configured **Kubernetes** clusters for our data platform on **AWS** using **Helm**, **Terraform**, **Terragrunt** and **GitLab CI**, to automate and scale production deployments, reducing errors and accelerating release cycles for Fortune 500 companies serving **100 million+** people worldwide
- Developed a **Python** query microservice using **SQL**, **Docker**, and **Apache Airflow**, to automate daily report generation from terabytes of DNS and advertising data, cutting report time to under **5 minutes**
- Wrote Kubernetes horizontal and vertical pod autoscalers to dynamically adjust resources, increasing max traffic by over **9 times** and improving service stability during spikes
- Built REST APIs, integration tests, and end-to-end tests for a **Java** codebase with **Gradle** and **Jenkins**, ensuring higher code quality and seamless deployment in CI/CD pipelines
- Configured **Prometheus** Alertmanager and **Grafana** dashboards for production monitoring, reducing alert response times by **85%** through improved error tracking and timely notifications
- Merged 7 pull requests for highly requested features for the open source Trino Helm chart on GitHub, enhancing database security and scalability features based on user feedback, leading to broader adoption and improved reliability

## TECHNICAL SKILLS

**Languages:** JavaScript/TypeScript, Python, Java, Rust, Go, Bash/Shell, SQL
**Libraries/Frameworks:** React, Next.js, GraphQL, NestJS, TypeORM, Tailwind CSS, Jenkins, Jest, Svelte
**Other:** Kubernetes, Helm, Terraform, AWS, Azure, GitLab CI, Docker, Unix/Linux, Git, Vim