

# AMNISH SINGH ARORA

Toronto, ON | 647-885-9860 | [amnishsingh04@gmail.com](mailto:amnishsingh04@gmail.com) | [LinkedIn](#) | [GitHub](#) | [Blog](#)

---

Software Developer with a strong background in **AI** and **Web Development**, and experience in large enterprises and fast-paced startups.

## Profile

---

- Love exploring new technologies and contributing to **open-source** projects.
- Sharing insights about my work through **blog posts** at [Dev.to](#).

## Experience

---

### Software Engineer – [DeepStructure](#)

Jul 2024 – Present

- Joined as the 6<sup>th</sup> employee at a [VC-backed](#) tech startup in Silicon Valley, building a cutting-edge AI infrastructure platform.
- Architected and built core components of our in-house implementation of OpenAI's **Assistants API**, like **streaming** powered by **NATS**, retries with **assertions**, **fallback models** with concurrently racing requests, and robust **rate-limiting** using a TTL cache, complemented by a comprehensive **smoke test suite** setup in **CI**.
- Rewrote the **structured completions** module in the SDK to add support for non-OpenAI **LLM providers** like Google, Anthropic, Sambanova, Groq, and Amazon Bedrock, while reorganizing the code to follow best practices inspired by open-source projects like [Vercel's AI SDK](#).
- Worked on improving the performance and accuracy of structured completions in a **customer-facing** project by exploring various **prompt engineering** techniques, ultimately concluding that non-nested output schemas reduce **LLM confusion** and improve generation speeds.
- Built several **Slack bot** reference apps on top of DeepStructure to [dogfood](#) platform capabilities, implemented **dynamic self-improving** responses leveraging user reactions, and a mechanism to refresh access tokens for deployed apps automatically.
- Built an **LLM evaluation** React **dashboard** enabling side-by-side model comparisons using the "LLM as a Judge" approach; also implemented a concept of **versioned Assistant families**, allowing users to mark a favorite for consistent routing within each family.

### Frontend Developer – [Ontario Treasury Board Secretariat](#)

Sept 2022 – Jul 2024

- Contributed to the **Angular** frontend of OnTrack, a large enterprise application integrating multiple legacy systems for the Cabinet Office and Treasury Board, within an **Agile** team.
- Played a key role in a **major feature redesign**, successfully implementing **26 User Stories** within a 2-week sprint.
- Developed domain-specific features such as submitting and applying **suggestions** for Key Activities and Milestones tied to the Cabinet Office's election commitments and generating document reports from **aggregated data** across multiple critical applications.
- Started as a co-op student and earned four consecutive contract renewals due to **demonstrated reliability** and **high performance**.
- Worked hard to make web applications and generated documents **AODA** compliant, ensuring accessibility for users with disabilities.

## Education

---

**Advanced Diploma:** Computer Programming and Analysis  
**Seneca College of Applied Arts and Technology** – Toronto

**Sept 2021-April 2024**

- Named to the President's Honor List in all semesters.
- Graduated with a perfect GPA of 4.0 and high honors.
- Selected high-impact electives in Cloud Computing, Open Source, and iOS development.

## Projects

---

### ChatCraft:

- [3rd largest contributor](#) on a developer-oriented Open-Source **AI chatbot** with **168** GitHub stars, with cool features like a tool's editor, speech capabilities, image processing, and more.
- Integrated **Text-to-Speech** capabilities like streaming audio messages in real-time, the ability to pick different voices, and downloading messages as audio with **concurrent** API calls, improving **download speeds** by **500%**, and allowing downloads of [infinitely long messages](#).
- Implemented experimental features like [Web Handlers](#), allowing users to call their **backend functions** within ChatCraft, effectively allowing users to plug and play their microservices.

URL: <https://github.com/tarasglek/chatcraft.org>

### Dine-A-Night (College Capstone Project):

- A highly customizable restaurant **reservation** application, allowing restaurant owners to sign up and build **visual layouts** using a drag-and-drop interface, customers to make reservations with seats of their choice, and admins to **monitor** upcoming and historical reservations.
- Built on top of a complex distributed [architecture](#), with **Angular** on the Frontend, **Express.js** for API and business logic, **MongoDB** for structured data storage, **Firebase** for authentication and blob data storage, and **Google Cloud** for Maps and Address Autocomplete features.
- Comes with a comprehensive [User Manual](#) and [YouTube Demos](#) for onboarding new users and providing relevant information to interested investors.

URL: <https://dine-a-night.web.app/home>

## Technical Skills

---

- **AI Engineering:** Working knowledge of prompt engineering techniques like **RAG**, **Few-shot prompting**, etc., and consuming/extending services offered by different AI providers.
- **Backend Development:** Capable of creating **secure** and reliable REST APIs and server-side rendered apps with Node frameworks like **Express.js** and **Hono**, and some experience in **ASP.NET**.
- **Frontend Development:** Capable of building highly interactive web apps leveraging modern Frontend frameworks like **Angular** and **React**.
- **Cloud Computing:** Deep understanding of Cloud Fundamentals, **Docker**, and experience in leveraging a wide range of services provided by major cloud providers like **AWS**, **Azure**, and **Google Cloud**, including authentication, storage, compute, and LLM infrastructure.
- **RDBMS Development:** Proficient in designing normalized databases with a good understanding of **ACID** properties and leveraging powerful features of different flavors of **SQL** like, **Postgres**, **Oracle**, and **MySQL**.
- **NoSQL Databases:** Familiar with tradeoffs between different types of NoSQL databases and experience with using Document type databases like **MongoDB**.
- **Object-Oriented Programming:** Strong grasp of both object-oriented and procedural paradigms with experience in **C/C++**, **Python**, [Java](#), **C#**, and **Swift**.