

# CHARLES GAO

Montreal, QC, Canada • charlesgao.c@gmail.com • 647-762-2488 • github.com/cgwelcome

## SKILLS

---

- **Languages:** Scala, Python, SQL, Bash
- **Technologies:** Apache Airflow, Apache Spark, PySpark, Apache Iceberg, dbt, ElasticSearch, BigQuery, GCP, PostgreSQL, pandas, Docker, Kubernetes

## EXPERIENCE

---

### Data Developer

May 2022 to Present

*Clinia, Montreal*

- Integrated a scalable data platform for health data to enable time travel, and schema evolution using Scala, Apache Spark, and Apache Iceberg
- Deployed data jobs on Apache Airflow on Kubernetes
- Indexed address data on ElasticSearch to normalize and geocode raw address
- Used functional programming principles with Scala Cats library to enhance type-safety in data pipeline

### Data Acquisition Engineer

Jan. 2021 to Apr. 2022

*Clinia, Montreal*

- Created an ETL infrastructure to scrape, clean, and load from different healthcare data sources using Apache Airflow, Python, and GCP
- Implemented a record linkage solution to improve data quality using dbt, and BigQuery
- Extracted, and cleaned health data to serve a data provider for ML model training

### Backend Developer

Sept. 2019 to May 2020

*Viafly Technologies, New York*

- Created a point-of-sale e-commerce platform with Ruby on Rails, Heroku, React, and PostgreSQL integrated with Shopify, and Square API
- Implemented the core e-commerce data model for OLTP database platform in PostgreSQL

### ETL Developer

May 2018 to Mar. 2019

*Energy Profiles Limited, Toronto*

- Maintained ETL pipelines consolidating energy data sources into a Data Warehouse using Python
- Implemented Stored Procedures in T-SQL for monthly energy reporting
- Synchronized temperature streaming data gathering from a network of sensors with BACnet Protocol

## EDUCATION

---

### University of Waterloo

Sept. 2015 to Apr. 2020

B. Math., Computer Science Major and Combinatorics & Optimization Major

## PROJECTS

---

### JOOS1W Compiler (CS 444)

Jan. 2020 to Apr. 2020

- Created a compiler for JOOS1W, a large subset of Java 1.3 to i386 assembly language in C++17

### Trains Microkernel (CS 452)

Sept. 2019 to Dec. 2019

- Created an ARM bare-metal real-time operating system (RTOS) in C
- Implemented context switching, hardware interrupt handler, RS-232 serial port device driver for processes to time-share CPU time