

Jeffrey Chan

LinkedIn: [jeffrey-chak-him-chan](https://www.linkedin.com/in/jeffrey-chak-him-chan) | Phone: 514-625-6619
Email: chan.jch19@gmail.com | Website: www.jeffreychchan.com
GitHub: github.com/JeffreyCHChan/JeffreyChan

PROFESSIONAL EXPERIENCE

Specialist Data Engineer | January 2024 -

Canadian National Railway

As a Data Engineer, I champion continuous improvement of our data engineering practices, lead initiatives in Databricks observability to optimize compute resources utilization, resolve gaps within our Unity Catalog metadata, and implement disaster recovery solutions that minimize downtime for the critical railway business.

OBSERVABILITY and DISASTER RECOVERY

- Creating an **observability** solution identifying areas for cost reduction, maintaining performance, resiliency of pipelines and highlighting Unity Catalog metadata gaps within **Databricks**.
- Develop a **disaster recovery** codebase for **Databricks** and communicate the execution procedure ensuring continuous critical railway services.

DATA PIPELINING

- Construct pipelines securing and landing data from sources such as **Apache Kafka** into the data lakehouse.
- Lead discussions on best data engineering practices, mentor interns through the development process from requirements gathering and designing architecture to development and deployments into production

Relevant Skills: Azure, Databricks, Apache Spark, Python, Kafka, SQL, Scala, GCP, Informatica

Data Engineer | October 2022 – November 2023

Compass Analytics

As a Data Engineer, I built scalable cloud architecture, deployed it onto AWS, developed data pipelines for consumption primarily using Python, Spark, SQL and Databricks, and create data visualizations using dashboards.

REAL-TIME ANALYTICS

- Developed and optimized **Python** code to capture and assemble data into Data Lakes and Data Warehouses to support operational reporting, data science and self-service analytics
- Create and maintain optimal and scalable data pipeline architecture using **SQL**, **Python**, **Databricks**, **Spark**

DATA INGESTION ENGINE

- Integrated a data ingest system for extracting and storing **25 documents per 30 seconds** into **AWS S3**
- Generated **AWS cloud architecture** designs according to stakeholder plans and requirements

Relevant Skills: Amazon Web Services (AWS), Solution Architecture, Python, Databricks, Spark, Snowflake

Data Engineering Developer | May 2021 - August 2021

Intact Financial Corporation

In this role, I translated business demands into technical requirements for the incoming Snowflake database, ensuring data quality using the Erwin Data Catalog and minimizing costly reconfigurations.

DATA SYSTEM DESIGN

- Documented conceptual, logical, and physical data models using the **Erwin Data Catalog**
- Built source to target mapping documents to describe data transformations for the **Snowflake** system

DATABASE MIGRATION

- Migrated a legacy database to meet new requirements, resulting in monthly **cost savings of over \$15,000**

Relevant Skills: Python, SQL Server databases, Data Analysis, Erwin Data Catalogue, Snowflake

Business Intelligence Developer | January 2021 - April 2021

McGill University

As a developer, I developed SQL scripts to automate maintenance tasks on existing Oracle database servers, eliminating costly and time-consuming manual interventions for large databases.

DATA EXTRACTION AND ANALYTICS

- Extracted requirements, developed, and tested code, **saving over 20 hours monthly** for database administrators
- Automated business reports and **Oracle** database health reports using **PowerBI** and **SSRS**

Relevant Skills: SQL Server databases, Software Development and Analysis

CERTIFICATIONS, DESIGNATIONS & CLEARANCES

Databricks Data Engineer Professional | December 2024

Databricks Inc.

Databricks Data Engineer Associate | December 2024

Databricks Inc.

Databricks Lakehouse Fundamentals | January 2023

Databricks Inc.

AWS Certified Cloud Practitioner | November 2022

Amazon Web Services

EDUCATION

Bachelor of Engineering – Software Engineering C-Edge

Concordia University - Gina Cody School of Engineering, Montreal, QC - Canada

- GPA: 3.57
- Member of the Institute for Co-Operative Education
- Student Leadership: Executive Vice-President, Engineering and Commerce Case Competition

Relevant Courses: *Machine Learning, Big Data Analytics, Artificial Intelligence, Data Structures and Algorithms*

SKILLS & TOOLS

Cloud Services & Tools: Amazon Web Services, Google Cloud Platform (GCP), Databricks, Excel, Word, PowerPoint

Languages: Python, Java, SQL, MongoDB, JavaScript, HTML, CSS, Git

Packages & Libraries: scikit-learn, NumPy, PySpark, pandas, NLTK, TensorFlow, BeautifulSoup, PyTorch

COMMUNITY/PROFESSIONAL ACTIVITIES

Case Competition Judge | September 2022 - Present

Engineering and Commerce Case Competition

- Provide feedback to students on innovative engineering and business solutions to industry challenges
- Review and critique student solutions leveraging expertise in engineering and design

Executive Vice-President | April 2021 - May 2022

Engineering and Commerce Case Competition

- Led a 20-person team to accomplish their mandates
- Handled operational tasks such as setting deadlines and overseeing department progress
- Oversaw a weeklong international engineering and business case competition

PROJECTS

Formula 1 Analytics

- Develop code that streamed real-time car telemetry
- Create dashboards and analysis tools from streamed data

Relevant Skills: *Python, Data Networking, Databricks, AWS*

Capstone

- Developed a data viewing platform for DICOM files improving the workflow of data scientists.
- Improved responsiveness by over **6000% (70 seconds to 1.5 milliseconds)** using Azure Table Storage

Relevant Skills: *Database Design, C#, Backend Development, Azure*

Course Recommendation System

- Completed for Big Data Analytics ingesting data and extracting keywords in Excel containing course descriptions
- Applied recommendation models to provide suggestions based on student ratings and course similarity

Relevant Skills: *PySpark, Data Engineering, Machine Learning, Python, Excel*

ADDITIONAL INFORMATION

English (Native), French (Spoken, Written) | Cantonese (Spoken)