

WILLIAM GLAZER-CAVANAGH

438-389-8650 ◇ willcavanagh@hotmail.com ◇ French / English

[Website](#)

[LinkedIn](#)

[GitHub](#)

[Scholar](#)

EDUCATION

University of Montréal - Professional Masters in Machine Learning

Accepted in the MILA program - 3.9 / 4.3 GPA

September 2022 - December 2023

Montreal, Quebec

Polytechnique Montréal - Bachelor in Software Engineering

Artificial Intelligence & Data Science - 3.4 / 4.0 GPA

September 2017 - May 2022

Montreal, Quebec

WORK EXPERIENCE

AbCellera Biologics - Machine Learning Scientist II

May 2023 - Present

Automation of high-throughput screening (HTS) for drug discovery via semi-supervised learning and computer vision

- Developed a semi-supervised pipeline for automatic labeling of new assay classes using **Python** and **PyTorch**. This pipeline outperformed pre-existing models in 80% of the images in out of distribution (OOD) settings as measured by **A/B testing**.
- Engineered parallelized data ingestion processes for image processing with **Dask**. Implemented efficient data structures with **Parquet** and **Zarr** reducing the model training time from 22 Hours to 17 Hours (a 30% reduction) as well as enabling random access.
- Improved throughput of model predictions per second from 400 predictions per second to 1200 predictions per second (a 3X improvement) by modifying state of the art computer vision models in **PyTorch**.
- Collaborated with stakeholders to address errors in internal assays, conducting detailed error analyses and providing specific recommendations for assay optimization.
- Designed reusable abstractions for common ML tasks, streamlining team workflows and fostering collaboration.

Intact Insurance - Data Scientist Intern

May 2022 - August 2022

Optimization and evaluation of call center processes across Canada

- Increased yearly call center sales by an estimated 200K\$ by implementing employee management strategies in production. These strategies were designed and reviewed using our custom discrete event simulator implemented with **Python**, **SimPy**.
- Detected production issues by developing dashboards in **PowerBI** with data ingestion pipelines written in **SQL** and **Python**.

Croesus Finansoft - Research Intern

May 2021 - August 2021

Automatically explain stock price fluctuations from news articles with Natural Language Processing (NLP)

- Identified avenues of research by assessing the state of the art in knowledge injection with knowledge graphs for **NLP** algorithms and **Large Language Models**. Established a research agenda with principal researcher.
- Increased reproducibility and reduced delays by 15% by refactoring and re-designing the algorithm evaluation pipeline with **Argo Workflows** and **Microservice Design Patterns**.

Merck Canada - Data Engineer Intern

January 2021 - May 2021

Development of Merck Canada's internal ETL (extract, transform, load) processes for Business Intelligence

- Produced first insights on GARDASIL-9 HPV vaccine's Canadian market by leading a market opportunity analysis. This project was used to power a canada-wide marketing and awareness campaign against HPV.
- Automated Merck Canada's manual ETL workloads and reduced delays by two days with **AWS Cloud** and **Python**.

PROJECTS & PUBLICATIONS

IVADO - Deep Learning Essentials Contributed to IVADO/MILA online Deep Learning class covering: CNNs, GANs, RNNs

Polytechnique Montreal - MCIS Lab Paper accepted at EMSE: *Towards a Change Taxonomy for ML Systems*

Polytechnique AI Student Club - Founder Kickstarted the club and handled content for the first cohort of 10

Polytechnique Montreal Admission Scholarship Awarded for outstanding academic performance

PROGRAMMING SKILLS

Languages: Python, Bash, C/C++

Machine Learning: PyTorch, Scikit-Learn, NumPy

Data Wrangling: Dask, Pandas, SQL

Others: Flask, Linux, Docker, GIT, LaTeX

Visualization: Plotly, Matplotlib, Seaborn, PowerBI

Books: Designing ML Systems (*ongoing*), Clean Code