

# JOSEPH D VIVIANO

joseph@viviano.ca | [viviano.ca](http://viviano.ca) | [GScholar](https://GScholar) | [twitter.com/josephdviviano](https://twitter.com/josephdviviano) | [github.com/josephdviviano](https://github.com/josephdviviano)

## EXPERIENCE

### ML Research Scientist

2021–Now

[Deep Genomics](#)

- Self-supervised learning (auto-regressive and masked-language modelling) for biological sequence representations.
- Multi-modal methods facilitating biological sequence-to-sequence models to generalize to novel cell types.
- Anomaly detection pipeline for our drug screening platform.

### Scientific Mentor

2021–Now

[Creative Destruction Lab Montréal](#)

- Performed technical assessments of startups and provided feedback for a business and investor focused audience.
- Technical direction for startups in the NLP, CV, and ML for Bio spaces.

### Applied Research Scientist

2020–2021

[Mila Technology Transfer](#), [Mila Québec AI Institute](#)

- Research on medical applications, ethics of language models, and modular deep learning.
- Deep-learning driven equity momentum trading ([CDPQ](#)) and robust representation learning for digital forgery detection ([Jumio](#)).

### Fellow

2021

[PhD to VC, Fifty Years](#)

- I sourced, diligenced, and deep-tech pre-seed/seed stage companies, performed due diligence on them, & helped source technical talent.

### Research Intern

2020

[Google](#), [Smart Ad Relevance System Predicted Click Through Rate Team](#)

- Uncertainty estimation research for predicted click through rates.
- Produced research pipeline for future search ad pCTR research.

### Research Intern

2019

[Imagia Cybernetics](#)

- Developed deep learning method for cancer localization without segmentations, and multimodal classification combining images with clinical notes.

### Research Methods Specialist

2014–17

[Kimmel TIGRlab](#), [Centre for Addiction and Mental Health](#)

- Biomarkers for vulnerable [schizophrenia](#) and [Alzheimer's](#) patients.
- Designed, built, and managed (team of 5) a [data management platform](#) 22-node compute cluster, and [QA tools](#) used by team of 20.

## PROJECTS

### What's In the Box?

Undesirable content

identifier for the Common Crawl.

### TorchXRayVision

Open source library for

chest XRay datasets and models.

## RESEARCH & COMMUNITY

**Bio & Health:** Genomics, biomarkers for precision medicine.

**Deep Learning:** Representation learning for robust and explainable models, multimodal learning, attention mechanisms.

**Teaching:** Stats and ML courses taught at CAMH, [McGill BrainHack](#), & [MAIN](#).

**Paper Reviews:** [ICLR](#), [NeurIPS Pre-registration in ML Workshop](#), [Biological Psychiatry](#), [MIDL](#), & [PLOS ONE](#)

## EDUCATION

### MSc. Computer Science, Machine Learning Specialization

2018–20

[Mila](#), Université de Montréal, Montréal, QC

- Research on whether saliency methods can be relied on to [diagnose generalization failures](#) of deep learning models.
- Coursework on Reinforcement Learning, Deep Learning, NLP, & CS Fundamentals.

### MSc. Biology, Neuroscience Specialization, With Distinction

2011–13

[Schneider Lab](#), York University, Toronto, ON

- Research using Magnetic Resonance Imaging of the [human visual input](#) and [feedback](#) systems.

### BSc. Psychology, Hons.

2005–09

Queen's University, Toronto, ON

## TECHNOLOGIES

### Python

Proficient

Pytorch, tensorflow, numpy, scipy, pandas, sklearn.

### Linux Administration

Intermediate

Linux Foundation certified system administrator.

### R, C, Java, SQL, MATLAB

Familiar