

Emmanuel Noutahi

MACHINE LEARNING SCIENTIST · COMPUTATIONAL BIOLOGIST

☎ (+1) 438-764-3132 | ✉ emmanuel.noutahi@gmail.com | 📱 [maclandrol](#) | 📧 E. Noutahi

Summary

- Machine learning scientist who builds models and tools across computational biology and chemistry. I translate research into actionable program decisions, informed portfolio trade-offs, and deliverables for internal and external stakeholders.
- Expertise in representation learning, generative molecular design, multimodal modeling, and OOD robustness, with strong MLOps + software engineering.
- Proven track record of building and leading large interdisciplinary teams across ML, biology, chemistry, and engineering. I set strategy and roadmaps, hire and mentor, and deliver end-to-end systems adopted internally and used in external collaborations.

Experience

Principal Scientist / Program and Group Lead

VALENCE LABS / RECURSION PHARMACEUTICALS

Montréal, QC

May 2023 – Present

- Lead a ~10-person team of senior, staff, and principal scientists/engineers to build Recursion's *virtual cell* from multi-omics and perturbation data. Set vision and roadmap, define scientific/technical strategy, and contribute hands-on to modeling and integration with Recursion assays and discovery pipelines.
- Conceived and led [LOWE](#), an LLM-orchestrated workflow engine that composes ML models with lab automation over proprietary data. Following a board demo, coordinated a 6-person task force to deliver v1 in 10 weeks for a live JPM 2024 demo; this preceded a [publicly announced collaboration with Bayer](#) and has driven internal adoption of agentic discovery systems.
- Designed and built [SAFE-GPT](#) for constraint-aware molecular design; now adopted as a core tool across three internal programs.
- Founded and lead the **Biomodality Research Unit** (4 FTE + interns); the team's work crystallized the company's Virtual Cell vision and underpins public communications and manuscripts on omics modeling.
- Own group hiring and contribute to broader hub hiring; conduct ~30 interviews/year and serve as final decision-maker (prior to executive approval), resulting in 6 accepted offers to date.
- Named inventor on **7** Recursion patents filed or granted since joining in 2023 (**9** total to date), covering AI for drug discovery, agentic systems, and multimodal modeling.

Head of Platform

VALENCE DISCOVERY

Montréal, QC

Jun 2021 – May 2023

- Built and led an 11-person team across ML, software engineering, computational chemistry, and data science; delivered a generative AI platform optimized for low-data settings and used across internal programs and partner projects.
- Owned platform roadmaps and quarterly OKRs; co-owned compute and pipeline budgets with the Head of Drug Discovery; drove cross-team prioritization and delivery across research and product.
- Led development of **Kernel**, a human-in-the-loop molecule prioritization platform integrating medicinal-chemist feedback to accelerate internal and partner model building and candidate progression.
- Led hiring across the platform group and maintained a strong culture with **100%** retention through the Recursion acquisition.

Senior ML Engineer and Lead, Computational Design

VALENCE DISCOVERY

Montréal, QC

Dec 2018 – Jun 2021

- Founding team member shaping internal IP, research direction, and external scientific positioning.
- Prototyped, built, and productionized core components of Valence's design and prediction stack: molecular representations, multi-objective optimization, property prediction, compound prioritization, and automated SAR extraction from patents.
- Technical lead on pilots and partnerships with Charles River Laboratories, Servier, and Repare Therapeutics. Work with CRL enabled Valence's first multi-program partnership and identified candidates with *nM* potency and desirable ADMET properties.

Co-founder and Head of Engineering

ORCA PROJECT

Montréal, QC

Feb 2018 – Aug 2018

Technical lead and sole engineer for a surgical operating room coordination and management application.

Course Lecturer at University of Montréal

BIOINFORMATICS AND COMPUTER SCIENCE

Montréal, QC

Sep 2014 - May 2018

- Teaching Assistant : weekly tutorials, assignment design and grading, discussion boards, and office hours for **Discrete Mathematics** and **Algorithms in Bioinformatics**.
- Course Lecturer for **Molecular Biology and Bioinformatics** : Linux, Python/Perl/R, biological databases, DNA/RNA sequencing, genome assembly and annotation, macromolecule structure modelling.

Research Assistant

LABORATORY OF COMPUTATIONAL AND THEORETICAL BIOLOGY - UDEM

Montréal, QC

Jan 2014 - Aug 2014

Developed micro/macro evolutionary-aware algorithms for gene tree correction and genome evolutionary history inference.

Research Intern

ZENKLUSEN LAB - UDEM

Montréal, QC

May 2013 - Dec 2013

Developed MATLAB, ImageJ, and Imaris pipelines for cellular compartment segmentation and 2D/3D RNA localization and quantification for smFISH data.

ROBERT CEDERGREN CENTRE - UDEM

May 2012 - Aug 2012

3D modelling and validation of RNA ligase 2 structures in *Diplonema*, a unicellular protist with a fragmented mitochondrial genome.

Education

PhD in Computational Biology

UNIVERSITÉ DE MONTRÉAL (UDEM)

Montréal, QC

Sep 2014 - Nov 2018

Dean's honor list and best thesis award

Thesis on the Inference of Species and Genome Evolutionary Histories

B.S. in Bioinformatics

UNIVERSITÉ DE MONTRÉAL (UDEM)

Montréal, Qc

Sep 2011 - May 2014

Dean's honors list

Awards & Honors

2019	Dean's Honour List (PhD) , Faculty of Medicine - UdeM	Montréal, Qc
2019	Best Thesis Award , Faculty of Graduate and Postdoctoral Studies - UdeM	Montréal, Qc
2015-2017	Tuition Fee Waivers , Faculty of Graduate and Postdoctoral Studies - UdeM	Montréal, Qc
2014-2017	DIRO Excellence Scholarships , Department of Computer Science and Operations Research - UdeM	Montréal, Qc
2014-2017	PhD Fast-Track Scholarship , Faculty of Graduate and Postdoctoral Studies - UDEM	Montréal, Qc
2012-2014	Dean's Honour List (Undergraduate) , Faculty of Arts and Science - UdeM	Montréal, Qc
2011-2015	National Excellence Scholarships , Direction des Bourses et Secours Universitaires du Bénin	Cotonou, Bénin
2010	Best Student Awards , Vallet Foundation	Cotonou, Bénin

Technical Skills

I build machine learning models and software for drug discovery. You can find some of the open-source packages I maintain in the datamol.io ecosystem and on my [GitHub](#) profile.

SKILLS

Programming	Very Proficient : Python, Matlab, Bash/Zsh Well-acquainted : JavaScript, C, C++, Fortran, R, Java
Systems & Databases	Slurm, cloud compute orchestration, deployment and scaling, SQL/NoSQL
Software & Infrastructure	Git, CI/CD, packaging, Conda, Docker, Kubernetes, AWS/GCP, Argo, React, Next.js, Vercel, FastAPI
Scientific Computing	Computational chemistry/biology, algorithm development, ML frameworks (PyTorch, TensorFlow, scikit-learn), NumPy, SciPy, pandas, Dask, Ray, Spark, RDKit
LLM Tooling & Agents	Transformers (Hugging Face), LangChain, retrieval/agents, prompt orchestration and evaluation
Visualization	Matplotlib, Seaborn, Plotly, Dash, D3.js
Deep Learning	Transformers, GNNs, RNNs, CNNs, GFlowNets, diffusion and flow-matching, deep kernel methods, reinforcement learning, Bayesian optimization

LANGUAGES

French (native) - English - Fon - Mina

Service & Leadership

Lecturer - AI in Radiology

McGILL UNIVERSITY

Montréal, QC

2024 - Present

Invited Talks and Keynotes

MILA TEA TALKS, AMLD AFRICA, ASHG, MOML AND VARIOUS FACULTIES AT UDEM

2020 - Present

Program Committee Member

INTERNATIONAL CONFERENCE ON INTELLIGENT SYSTEMS FOR MOLECULAR BIOLOGY (ISMB)

2019 - Present

Peer-review Activity

ACTIVE REVIEWER FOR >10 MACHINE LEARNING AND COMPUTATIONAL BIOLOGY CONFERENCES AND JOURNALS

2015 - Present

Mentorship

SUPERVISED >15 INTERNS FROM THE START OF MY PHD AT UDEM TO CURRENT POST AT RECURSION

2014 - Present

Students Association of UDEM

DOCTORAL STUDENT REPRESENTATIVE

2016 - 2017

Led initiatives to improve the quality of graduate education.

VICE PRESIDENT, COMMUNICATIONS

2014 - 2016

Organized networking events between students, principal investigators, and bioinformatics professionals.
Served as an organizing committee member for the 1st and 2nd bio-hackathons at University of Montréal.

Selected Publications

For a complete list of publications, see my [Google Scholar](#).

- Zhu L. & **Noutahi E.** (2025). Conditional chemical language models are versatile tools in drug discovery. *arXiv preprint arXiv:2507.10273*.
- **Noutahi E.**, Hartford J., Tossou P., Whitfield S., Denton A., Wognum C., Ulicna K., Craig M., Hsu J., Cuccarese M., et al. (2025). Virtual cells : Predict, explain, discover. *arXiv preprint arXiv:2505.14613*.
- Wenkel F., Tu W., Masschelein C., Shirzad H., Eastwood C., Whitfield S. T., Bendidi I., Russell C., Hodgson L., El Mesbahi Y., **Noutahi E.**, Denton A. K. (2025). TxPert : Leveraging biochemical relationships for out-of-distribution transcriptomic perturbation prediction. In revision at *Nature Biotechnology*.
- Bendidi I., Whitfield S., Kenyon-Dean K., Ben Yedder H., El Mesbahi Y., **Noutahi E.**, Denton A. K. (2024). Benchmarking transcriptomics foundation models for perturbation analysis : one PCA still rules them all. *arXiv preprint arXiv:2410.13956*.
- **Noutahi E.**, Gabellini C., Craig M., Lim J.S., & Tossou, P. (2024). Gotta be SAFE : A New Framework for Molecular Design. *Digital Discovery*, doi :10.1039/d4dd00019f.
- Tossou P, Wognum C, Craig M, Mary H & **Noutahi E.** (2023). Real-World Molecular Out-Of-Distribution : Specification and Investigation. *Journal of Chemical Information and Modeling*, doi :10.1021/acs.jcim.3c01774.
- Horwood J. & **Noutahi E.** (2020). Molecular Design in Synthetically Accessible Chemical Space via Deep Reinforcement Learning. *ACS Omega* 5(51), 32984-32994.
- **Noutahi E.**, Calderon V., Blanchette M., El-Mabrouk N & Lang B.F. (2019). Rapid genetic code evolution in Chlorophyta mitochondrial genomes. *Molecular Biology and Evolution*, 36(4), 766-783.
- Rahman S., Zorca C.E., Traboulsi T., **Noutahi E.**, Krause M.R., Mader S., & Zenklusen D. (2017). Single-cell profiling reveals that eRNA accumulation at enhancer-promoter loops is not required to sustain transcription. *Nucleic Acids Research*, 45(6), 3017-3030.
- Sabir J.S., Jansen R.K., Arasappan D., Calderon V., **Noutahi E.**, Zheng C., Park S., Sabir M.J., Baeshen M.N., Hajrah N.H., & Khiyami M.A. (2016). The nuclear genome of *Rhazya stricta* and the evolution of alkaloid diversity in Apocynaceae. *Scientific Reports*, 6, 33782.
- **Noutahi E.**, Semeria M., Lafond M., Seguin J., Boussau B., Guéguen L., El-Mabrouk N., & Tannier E. (2016). Efficient Gene Tree Correction Guided by Genome Evolution. *PLoS ONE* 11(8) : e0159559.