

JACQUES THIBODEAU

thibo.jacques@gmail.com
[GitHub Profile](#)

[LessWrong Profile](#)
Tel: (506) 530-4970

ML Eng. Skills: Python, PyTorch, SQL, Docker, Wandb, FastAPI, Streamlit, and more. (Full stack DL).

WORK EXPERIENCE

2/2023 – Present **Independent Researcher**

- Working on accelerating alignment research with LLMs; private work, but can be shared upon request ([hackathon](#) I co-organized). [Evaluations](#) for stable alignment ([github](#)).

5/2022 – 1/2023 **SERI MATS: AI Alignment Scholar**

- Model editing and causal tracing with GPT-2/GPT-J. [Upcoming post](#).
- Collaborating with others on a sequence regarding [Simulators](#).
- Spent time reading tons of papers and alignment forum posts. Facilitated [AGISF 201](#).
- [Mentoring people](#) in technical alignment and governance via the [AISM&M program](#).

1/2022 – 5/2022 **AI Safety Camp**

- Created [dataset](#) of AI alignment texts. Wrote most of the code; managed contractors.
- Published [paper](#) on arXiv and Alignment Forum [post](#) based on dataset analysis.

2/2020 – 1/2022 **Canada Energy Regulator, Data Science Team: Data Scientist (NLP)**

- Worked collaboratively with the team to **extract tables and figures from PDFs**, and **helped build a data search tool**. I wrote [code](#) and helped with UX. [Video](#). [Search tool](#).
- Took initiative to teach and apply **code reviews** and Huggingface in projects.

6/2019 – 12/2019 **Government of New Brunswick, Health: Data Analyst & Strategist**

- Improving the mental health and addictions sector with data analysis and efficiency gains.

7/2019 – 10/2019 **Montreal AI Ethics Institute: Research Intern**

- Conducted research on the state of AI Ethics.

6/2018 – 6/2019 **Government of New Brunswick, Executive Council Office:
Innovation Strategist & Facilitator**

- Founding member of Social Innovation Lab. Worked on increasing effective policy innovation and acted as liaison or mentor for stakeholders at all levels of government.

7/2017 – 9/2017 **Université de Moncton: Deep Learning Researcher**

- Completed a literature review for using deep RL for CV. Used GCP to train ML models.

2013 – 2015 **Université de Moncton: Physics Researcher**

- (1/2015 - 9/2015) Invented thin film that controls the polarization of light. *Published paper*.
- (5/2014 - 9/2014) Fabricated films to be used to protect satellites. *Published paper*.
- (5/2013 - 9/2013) Used optical systems to fabricate and test a thin film. *Published paper*.

SELECTED PERSONAL PROJECTS

- [A Case Study on Weakly Supervised Learning](#): Chosen as a **top project** in the [Full Stack Deep Learning](#) course of over 300 professionals. Deployed models on **Azure**.

EDUCATION

9/2014 – 5/2018 **Université de Moncton, M.Sc., Physics,**
9/2010 – 4/2014 **Université de Moncton, B.Sc., Physics**