

PATRICK LAVICTOIRE

patricklavic@gmail.com

<https://www.patricklavictoire.com>

<https://www.linkedin.com/in/patricklavictoire>

Expertise in mathematics, machine learning, management, and AI alignment (agent foundations).

EXPERIENCE

Machine Learning Manager, Upstart 2022 - 2023

- I managed the team that generates loan terms advertised on partner sites (like Credit Karma, LendingTree, NerdWallet, etc) by predicting the terms an applicant would get on-site (using only the features available to Upstart via the partner site).

Machine Learning Engineer / ML Engineering Manager, Apple 2020 - 2022

- I worked on (and later managed) the team tasked with protecting the integrity of Apple's knowledge graph as new data is ingested, filtering out mistakes and falsehoods before the data is integrated into the knowledge graph.

Research / Data Scientist, Lyft 2017 - 2020

- I built an end-to-end machine-learned prototype for the task of location prediction given noisy sensor data. I also developed and updated ML models to catch fraudulent behavior on the platform, identifying a previously untracked vector of fraud.

Research Fellow, Machine Intelligence Research Institute 2015 - 2017

- I coauthored the *Alignment for Advanced Machine Learning Systems* paper, after having coauthored the *Program Equilibrium in the Prisoner's Dilemma via Löb's Theorem* paper prior to officially joining MIRI. I also facilitated a series of AI alignment workshops for early-career researchers, many of whom are now significant figures in the AI alignment field.

Principal Engineer, Quixey 2013 - 2014

- Quixey was an app search startup. I dramatically improved the accuracy of the search relevance model while simplifying its architecture, and successfully pushed for the company's first responsible data collection policy.

Visiting Assistant Professor, University of Wisconsin–Madison 2010 - 2013

- This was a three-year postdoctoral position in pure mathematics. I published research in harmonic analysis, maximal inequalities, and ergodic theory, and I taught 8 courses at the undergraduate level.

EDUCATION

PhD (Mathematics), University of California at Berkeley 2005-2010

- Advisor: Michael Christ

BA (Mathematics), University of Chicago 2001-2005

SELECTED PUBLICATIONS AND GRANTS

Alignment for Advanced Machine Learning Systems

- *Ethics of artificial intelligence (Vol. 8, 2016)*
With Jessica Taylor, Eliezer Yudkowsky, and Andrew Critch.

Program Equilibrium in the Prisoner's Dilemma via Löb's Theorem

- *Workshops at the Twenty-Eighth AAAI Conference on Artificial Intelligence (2014)*
With Benja Fallenstein, Eliezer Yudkowsky, Mihaly Barasz, Paul Christiano, and Marcello Herreshoff.

Multivariable averaging on sparse sets

- *Transactions of the AMS 366 (2014), 2975-3025*
With Andrew Parrish and Joseph Rosenblatt.

National Science Foundation, Analysis Grant DMS-1201314

2012-2013

Pointwise Convergence for Subsequences of Weighted Averages

- *Colloquium Mathematicum 124 (2011), 157-168*

Universally L^1 -Bad Arithmetic Sequences

- *Journal d'Analyse Mathématique 113 (2011, no. 1), 241-263*

An L^1 Ergodic Theorem for Sparse Random Subsequences

- *Mathematical Research Letters 16 (2009), no. 5, 849-859.*