Julian Yocum

juliany@berkeley.edu | (980) 579-0864 | 2455 Prospect St. Berkeley, CA, 94704

EDUCATION

Massachusetts Institute to Technology **Cambridge, MA**

B.S. in Physics and B.S. in Artificial Intelligence and Decision Making — GPA 4.8 *Fall 2019 – Spring 2023*

Masters of Engineering in Computer Science and Electrical Engineering — GPA 4.8 *Fall 2023 – Spring 2024*

Coursework: Machine Learning, Deep Learning, Reinforcement Learning, Large Language Models, Computer Vision, Algorithms, Abstract Algebra, Topology, Quantum Physics III, Quantum Computation, Relativity

UC Berkeley **Berkeley, CA**

PhD in Computer Science *Fall 2024 – Present*

Coursework: LLM Agents, Natural Language Processing, Physics-inspired Machine Learning, Information Theory

RESEARCH

Center for Human-Compatible Artificial Intelligence (CHAI) **Berkeley, CA**

PhD Student (advised by Dr. Stuart Russell) *Fall 2024 – Present*

— Developed theory and empirical support for a novel form of topological representation coined “feature fields”.

— Found emergence of world models in neural networks with statistical structure, e.g. HMMs, Ising model.

— Advised undergraduate interns on two world model projects, one leading to a workshop paper.

MIT Algorithmic Alignment Group **Cambridge, MA**

Master’s Student (advised by Dr. Dylan Hadfield-Menell) *June 2023 – May 2024*

— Submitted a workshop paper extending previous results from contracting in multi-agent reinforcement learning into the multi-agent generative agent domain.

— Reviewer for Foundation Models for Decision Making 2023 NeurIPS Workshop.

— Developed a new pipeline extending previous workshop results to improve negotiation with contract rollouts.

Center for Human-Compatible Artificial Intelligence (CHAI) **Berkeley, CA**

Research Intern (mentored by Dr. Justin Svegliato under Dr. Stuart Russell) *June 2023 – Aug. 2023*

— Created a novel multi-agent contracting framework for generative agents in Minecraft.

— Constructed a benchmark of social dilemmas agents to facilitate the study of multi-agent coordination.

— Demonstrated that generative agents can generate and respond to contracts in a way that overcomes social dilemmas, leading to better outcomes for all agents.

Quantum Applications Team at IonQ **College Park, MD**

Research Intern (mentored by Dr. Daiwei Zhu) *June. 2022 – Aug. 2022*

— Designed novel patent-pending optimization techniques to make quantum circuits more efficient by training reinforcement learning agents based on graph neural network architectures.

— Developed a circuit transformation library to serve as the environment for RL agents.

MIT Research Laboratory of Electronics **Cambridge, MA**

Research Intern (under Dr. William Oliver) *Jan. 2022 – May 2022*

— Studied resonance properties of superconducting microwave resonators.

— Developed analysis software to do qubit initialization for Bosonic error correcting codes.

MIT Laboratory for Nuclear Science **Cambridge, MA**

Research Intern (under Dr. Lindley Winslow) *Feb. 2020 – May 2021*

— Published a paper on a novel multi-objective optimization technique for reconstructing particle trajectories for the Cryogenic Underground Observatory for Rare Events (CUORE) .

— Designed a Monte Carlo for the CUORE underground ground detector to simulate particle events.

— Conducted statistical data analysis in Python and parallelized algorithms for a 50-node compute cluster.

INDUSTRY & ACADEMIC SERVICE

6.4110 Representation, Reasoning, and Inference in AI (Prof. Leslie Kaelbling) **Cambridge, MA**
Teaching Assistant *Sep. 2023 – May 2024*

— Conducted weekly office hours to provide students with personalized assistance on course material.
— Lead problem sessions to reinforce key concepts and demonstrate problem-solving techniques.
— Collaborated closely with Prof. Leslie Kaelbling to ensure course materials are effectively communicated and to gather feedback for continuous course improvement.

AI Bill Legislation Tracker (https://ai-bill-scraper.web.app) **Berkeley, CA**
Developer *Sep. 2023 – May 2024*

— Developed a legislation tracker for policy-makers to follow bills related to artificial intelligence in the US.
— Built infrastructure for cloud-based database of bills with automatic updating.
— Implemented automated ChatGPT categorization and analysis of bills.
— Collaborated closely with policy researchers at CHAI to ensure the tool is as helpful as possible.

Global Teaching Labs **Bremen, Germany**
Teacher *January 2023*

— Designed and taught a curriculum on quantum computing to ~30 German high school students.
— Collaborated with a student at Jacobs University to ensure the curriculum was accessible and engaging.

MIT AI Alignment **Cambridge, MA**
Communications Director *Sep. 2020 – May 2023*

— Mentors three undergraduate members with career, research, academic, and personal advice.
— Organizes public talks featuring prominent figures in the AI research field.
— Designs student surveys and strategizes about meetings with administration.
— Facilitated weekly reading groups over two semesters, fostering collaborative discussions on latest research.

Pozen Fellow – Entrepreneurship Intern at Tracflo **New York, NY**
Full Stack Developer *June 2021 – Aug 2021*

— Designed and developed a dashboard in React and NodeJS for a startup as the centerpiece product.
— Interviewed clients to collect feedback and plan new features

PUBLICATIONS

J. Yocum, C. Allen, B. Olshausen, and S. Russell, "Neural manifold geometry encodes feature fields," in *Proc. NeurIPS 2025 Workshop on Symmetry and Geometry in Neural Representations*, 2025.

E. Altuzar, **J. Yocum**, and C. Allen, "Transformers represent causal abstractions," in *Proc. NeurIPS 2025 Workshop on Symmetry and Geometry in Neural Representations*, 2025.

J. Yocum, P. J. K. Christoffersen, M. Damani, J. Svegliato, D. Hadfield-Menell, and S. Russell, "Mitigating generative agent social dilemmas," in *Proc. NeurIPS 2023 Workshop on Foundation Models for Decision Making*, 2023. https://social-dilemmas.github.io/

Yocum, J., Mayer, D., Ouellet, J. L., & Winslow, L. (2022). Muon Track Reconstruction in a Segmented Bolometric Array Using Multi-Objective Optimization. *JINST* **17** P07004.

TALKS & PRESENTATIONS

"Realization". Artificiality Summit. 2025
"Towards Strong Emergence: Non-Computability and Infinity". The Science of Consciousness Conference 2025
"Neural Manifold Geometry Encodes Feature Fields". CHAI Workshop 2025
"Mitigating Generative Agent Social Dilemmas". NeurIPS 2023 FMDM Workshop 2023
"Characterizing Track-like Events in CUORE with Multi-Objective Optimization". APS April Meeting 2021
"Searching for Magnetic Monopoles in CUORE". MIT PRISM Conference 2020

SKILLS

Python (matplotlib, numpy, Qiskit, pytorch), C, C++, C#, ReactJS, NodeJS, SQL, HTML & CSS, Linux, Bash