

David Turturean

davidct@mit.edu | linkedin.com/in/daviddturturean | theroadmostcalculated.substack.com

Education

Massachusetts Institute of Technology (MIT)	May 2026 (expected)
B.Sc. in Physics, with a concentration in Astrophysics (Course 8-Flex) and Artificial Intelligence & Decision-Making (Course 6-4)	Cambridge, MA, USA

Selected Coursework

Artificial Intelligence & Computing:

6.C57 Optimization Methods (Graduate-level); 6.7960 Deep Learning (Graduate-level); 6.4610 Natural Language Processing; 6.4110 Representation, Inference and Reasoning in AI; 15.S04 Generative AI Lab; 6.S058 Introduction to Computer Vision; 6.3900 Introduction to Machine Learning; 6.3800 Introduction to Inference; 6.1010 Fundamentals of Programming in Python

Mathematics:

18.112 Functions of a Complex Variable (Complex Analysis); 18.03 Differential Equations; 18.700 Linear Algebra (Proof-based); 18.901 Introduction to Topology

Physics:

8.962 General Relativity (Graduate-level); PHYS 212 Cosmology (cross-registration at Harvard University); 8.13 Experimental Physics I; 8.07 Electromagnetism II; 8.05 Quantum Physics II; 8.044 Statistical Physics I

Research Experience

SERC Projects, 2025–2026: Aligning AI with Human What? September 2025 – Present

Group Leader: Dr. Jakob Stenseke MIT CSAIL

- Member of SERC scholar group on conceptual foundations of value alignment across philosophy, social science, and technical alignment methods
- Developing an auto-autoethnography protocol where LLMs act as reflexive ethnographers to compare models on value conflicts (e.g., safety vs. autonomy), role/power dynamics (deference vs. coercion), and epistemic authority (overconfident guidance), producing a failure-mode catalog beyond benchmarks

Cross-Model Preference Transfer for AI Alignment June 2025 – Present

Faculty Supervisor: Prof. Dylan Hadfield-Menell

Direct Supervisor: Rachel Ma MIT Algorithmic Alignment Group

- Leveraged shared preference geometry to transfer steering vectors across LLM families (Gemma, Llama-3, Mistral/Qwen) using orthogonal alignment maps and behavioral validation
- Expanded analysis across model sizes and layers; alignment strengthens with scale and is bottlenecked by the smaller model in each pair, with within-family transfer stronger than cross-family

- Tested the *Platonic Representation Hypothesis* for steering vectors: absolute preference poles (positive/negative) align robustly across models while contrastive differences are less stable, supporting shared geometry; preparing ICML 2026 submission

AI Risk Repository & AI Risk Index Research

February 2025 – Present

Faculty Supervisor: Dr. Neil Thompson

Direct Supervisor: Dr. Peter Slattery

MIT FutureTech Initiative, MIT CSAIL

- Contributing to database of global AI risk experts across 20+ risk subdomains for the MIT AI Risk Index
- Developed conversational AI system for 1,612 categorized AI risks from 65 papers, currently in alpha testing under “Chatbot” section at project website
- Implemented advanced RAG pipeline with ChromaDB achieving 85%+ relevance accuracy through hybrid retrieval, reducing research time by 75% via automated intent classification

Gamma-Ray Burst (GRB) Science with Machine Learning

June 2024 – Jan 2025 (full-time)

Feb 2025 – Present (part-time)

Faculty Supervisor: Dr. Marshall Bautz

Direct Supervisors: Dr. Benjamin Schneider, Dr. Javier Viaña

MIT Kavli Institute

- Built a real-data pipeline to estimate GRB distance (redshift) from afterglow spectra with supporting X-ray features, with calibrated uncertainty
- Diagnosed mean-prediction collapse on imbalanced data and redesigned normalization and sampling to improve calibration and reduce extreme errors
- Preparing ICML 2026 submission; continued part-time since Feb 2025

Analysis of Metal-Poor Stars' Spectroscopic Data

September 2023 – December 2023

Faculty Supervisor: Prof. Anna Frebel

MIT Kavli Institute

- Identified chemical chronometers and nucleosynthesis patterns in ancient stars to trace early universe conditions
- Processed and analyzed spectroscopic data using Python-based astrophysics tools and pipelines
- Contributed to understanding of r-process element formation in the earliest stellar populations

Enhancing Black Hole Imaging Algorithms

May 2023 – July 2023

Research Supervisor: Dr. Kazunori Akiyama

MIT Haystack Observatory

- Developed image-processing algorithms for the Event Horizon Telescope
- Optimized Julia-based high-performance computing implementations
- Benchmarked real and synthetic data to validate algorithm improvements

Publications & Presentations

• Cross-Model Preference Geometry & Steering Vector Transfer (PRH)

- Presented “*The Geometry of Preference: Transferable Steerability Across Language Model Families*” at New England Mechanistic Interpretability Workshop (Aug. 2025, Northeastern University), demonstrating 85% behavioral preservation in cross-model transfer
- Preparing ICML 2026 submission on the *Platonic Representation Hypothesis* for steering vectors:

using PCA + orthogonal rotations to transfer directions across model families, revealing shared preference geometry strongest for positive/negative poles, with transfer quality improving with model scale

- **Gamma-Ray Burst Redshift Prediction**

- Presented “*Residual Neural Networks for GRB Afterglow Spectra*” at MIT Kavli Institute Undergraduate Research Forum (Aug. 2024)
- Preparing ICML 2026 submission on calibrated GRB redshift estimation from afterglow spectra; built a real-data pipeline and showed the model initially defaulted to mid-range predictions on scarce, noisy data, then redesigned preprocessing and calibration to improve reliability and reduce catastrophic outliers

- **Center for AI Policy’s Congressional Exhibition on Advanced AI (24 Feb. 2025)**

- Selected via a three-step selection process to present an AI-driven voice emulation system to congressional staffers at the Rayburn House Office Building (Washington, DC) [Link]
- Engaged policymakers on the risks of low-latency neural voice technologies deployed en masse in business operations and mass communications

- **Invited Talks**

- “*AI Before the 2025 Milestone: From Origins to Revolution*” (Bridging Gaps RenAIssance #1, online, Sep. 30, 2025; presented in English)
 - * Bridging Gaps: Romania-based pro bono mentoring and coaching community for IT leaders and managers
- “*AI pe înțelesul tuturor*” (“*AI Explained for Everyone*”) (Organizatia Studentilor Economisti (OSE), FSEGA sala 118, Dec. 2, 2025; presented in Romanian)
 - * OSE: Economics Students’ Organization at UBB FSEGA focused on student development and professional opportunities

- **Republica.ro — Contributor (Long-Form Essays)**

- “*Un nou sofism: alibi-ul sintetic*” (“*A New Sophism: The Synthetic Alibi*”) (Nov. 2025) — on synthetic media, epistemic doubt, and political manipulation [Link]
- “*Ultimul presedinte muritor?*” (“*The Last Mortal President?*”) (Nov. 2025) — published as inaugural long Substack post (24.134 Experiential Ethics final project) [Link]

- **The Road Most Calculated** (Launched August 2025) — Bilingual Substack exploring AI safety, longevity, and societal implications of artificial general intelligence

Applied AI Projects

Text-to-Data Pipeline System for NASDAQ

Spring 2025

15.S04 Generative AI Lab, MIT Sloan – Team of 4

- Collaborated with 3 teammates to develop agentic AI system using LangGraph and RAG architecture, enabling natural language data transformations for enterprise operations
 - Implemented multi-agent orchestration with DeepSeek V3 for intelligent validation, timezone conversion, and table joining operations
 - Created modular pipeline projected to reduce 5,000+ hours of manual engineering effort annually through automated code generation and self-correcting workflows
-

Professional Education & Development

AI Education & Safety Engagements

IAIFI PhD Summer School

August 4–8, 2025

- Selected as one of the few undergraduate attendees for intensive week-long program on AI applications in fundamental physics at Harvard University
- Engaged in advanced lectures on reinforcement learning, geometric deep learning, physics-motivated optimization, and domain adaptation

Y Combinator AI Startup School

June 16–17, 2025

- Selected as one of 2,500 attendees for competitive AI conference in San Francisco with speakers including leaders from OpenAI, Microsoft, Anthropic, and DeepMind
- Participated in technical sessions on frontier AI research, safety considerations, and practical applications in industry and startups

MAIA AI Technical Workshop

April 11–13, 2025

- Participated in technical workshop with researchers from Anthropic, OpenAI, Redwood Research, UT Austin and more on interpretability approaches and alignment auditing
- Engaged in discussions ranging from mechanistic interpretability to safety challenges in state-of-the-art AI systems

ARENA (Alignment Research Engineer Accelerator) IAP Workshop

January 2025

- Completed a workshop on AI safety, mechanistic interpretability, and Reinforcement Learning, hosted by MIT AI Alignment
- Developed skills to evaluate deep learning models, including transformers and RL agents

AI Safety Fundamentals Reading Group

September 2024 – December 2024

- Participated in a seven-week technical reading group covering neural network interpretability and AI risk, hosted by MIT AI Alignment
- Engaged in weekly paper readings and discussions on safety considerations and misgeneralizations

MIT AI Alignment (MAIA) Member

December 2024 – Present

- MAIA Executive Committee (Fall 2025)

Leadership Development

MIT Leadership Capstone Program

January 2025 – May 2025

- Completed semester-long leadership program as one of 10 selected seniors, hosted by MIT SOLE Office
- Participated in weekly interactive case studies, biweekly structured mentoring, and team projects focused on strategic decision-making

MIT Rising Leaders Program

September 2023 – December 2023

- Selected among a cohort of 6 students for a semester-long program exploring various leadership theories and models

- Engaged in weekly discussions to develop personal leadership competencies, strategic thinking, effective communication

Society of Physics Students (SPS) September 2023 – Present

- Served as Social Chair (2023–24), Career Chair (2024–25), and Career Committee Member (2025–Present)
 - Coordinated professional development panels and networking events involving MIT Physics alumni
 - Led the creation of the SPS Career Handbook, compiling experiences from students and alumni
 - Organized student–faculty dinners and departmental socials; rekindled digital engagement through a dedicated Discord server active for 2+ years
-

Additional Leadership & Activities

MIT Orientation Leadership 2022 – 2024

- 2024: Orientation Leader for MIT Class of 2028 Orientation
- 2023: Paid International Orientation Coordinator for MIT International Student Orientation
- 2022: MIT International Orientation Mentor

Theta Delta Chi (TDC) Fraternity 2021 – Present

Positions: Member Intake Chair & Academic Chair (2022), Vice-President (2023), Treasurer (2024), Alumni Relations Chair (2025)

- Managed fraternity finances and budget planning of \$100,000+ as Treasurer
- Coordinated internal operations as Vice-President, overseeing 15+ officers
- Represented the MIT chapter at the International Convention
- Led recruitment and new member education initiatives, increasing chapter membership from 21 to 50 across a calendar year

MIT Interfraternity Council (IFC) 2022

Executive Board Member

- Spearheaded an initiative to boost financial transparency for prospective fraternity members by publishing aggregated cost data for all MIT chapters in the Recruitment Booklet
 - Organized leadership development workshops for fraternity chapter officers
-

Teaching & Mentoring Experience

MIT Associate Advisor September 2022 – Present

- Guided 7 first-year students (advisees of Dr. Silvina Hanono-Wachman, of MIT Computer Science and Artificial Intelligence Laboratory) annually in academic development, course selection, and MIT integration
- Provided ongoing mentorship and connected students with MIT resources

Physics Mentoring Program – Lead Mentor (Paid) September 2023 – Spring 2025

- Mentored 3–4 students per semester in core Physics courses
- Led Community of Practice sessions for new mentors

Astronomy Mentor at Rubik School (Paid)

September 2020 – February 2021

- Taught two weekly online astronomy courses for students aged 6–13
- Developed engaging educational materials and structured course content

Mathematics, Physics & Astronomy Tutor at TeachU (Volunteer)

May – July 2019;
May – July 2020

- Volunteered over 200 hours per year tutoring students individually for the Romanian Baccalaureate in Mathematics and Physics
 - All tutored students achieved scores of 85%+ in tutored subjects in the Romanian Baccalaureate
-

Awards & Honors

- **1st Place, IonQ Challenge at MIT iQuHack 2025** – Awarded to team “Quack Quack” for designing a novel quantum algorithm leveraging variational quantum imaginary time evolution to address a maximum-cut-like graph problem
 - **Amy Vojta Award in Membership Recruitment/Intake (2023)** – Sole award presented by the Northeast Greek Leadership Association for outstanding recruitment and member intake initiatives in Greek Life across 57 campuses
 - **Cornelius Vanderbilt Scholarship (2021)** – Full-ride, merit-based scholarship awarded by Vanderbilt University to less than 1% of incoming first-year class
 - **Gold Medal, Global e-Competition in Astronomy and Astrophysics (2020)** – Awarded during the COVID period at the virtual equivalent of the International Olympiad in Astronomy and Astrophysics (competition was held by same organizing committee)
 - **Gold Medal and Absolute Winner of the Planetarium Round, International Olympiad on Astronomy and Astrophysics (2019, Hungary)**
 - **Qualifier, International Earth Science Olympiad Selection Camp (2019, Romania)**
 - **Silver Medal, International Olympiad on Astronomy and Astrophysics (2018, China)**
 - **Citizen of the Centenary of Suceava (2018)** – One of 28 citizens honored for significant achievements across various fields
 - **Gold Medal, International Astronomy Olympiad (2018, Sri Lanka)**
-

Skills

Programming: Python and PyTorch, C++, Julia, L^AT_EX

Machine Learning: Neural Networks (Convolutional, Residual), Supervised Learning, Ensemble Learning, Reinforcement Learning, Model Evaluation & Calibration

Research: Physics-based Modeling, Data-driven Approaches in AI, Algorithm Development, Photometric

and Spectroscopic Data Analysis

Languages: Romanian (Native), English (Fluent), Spanish (Intermediate)