

Davi Bastos Costa

University of Chicago, Chicago Human + AI Lab (CHAI Lab), 5730 S. Ellis Ave, Chicago, IL 60637, USA
davicosta@uchicago.edu, (872) 297-9071
Website, GoogleScholar, GitHub, LinkedIn

Physicist bridging theoretical physics and AI

Education

-
- **University of Chicago** Chicago, IL
Ph.D. in Physics, McCormick Fellowship, Sugarman Award for Excellence 2020-2025
Advisor: Clay Córdova
Thesis: Non-Invertible Symmetries in Finite-Group Gauge Theories
 - **University of São Paulo** São Paulo, Brazil
Master in Physics, highlight in Physics in the Brazilian Physics Society 2018-2020
Advisor: Gustavo Burdman
Dissertation: Dark Matter and Neutrino Masses in Twin Higgs Models
 - **University of São Paulo** São Paulo, Brazil
Bachelor of Physics 2014-2018
Scientific initiation advisor: Nestor Caticha
Project: Emerging Collective Properties in NN based Models of Social System
 - **University of São Paulo** São Paulo, Brazil
Bachelor in Geography Candidate 2013-2014
Studies in Sociology and Philosophy of Science

Research Experience

-
- **Schmidt AI in Science Postdoctoral Fellowship, University of Chicago** Chicago, IL
Data Science Institute, CHAI Lab and Knowledge Lab 2026-Present
Mentor: Chenhao Tan and James Evans
Research: AI for Scientific Discovery and AI Safety
 - **Postdoctoral Research Fellow, University of São Paulo** SP, BRA
TELUS Digital Research Hub, Center for AI and Machine Learning 2025-2026
Mentor: Renato Vicente
Research: Emergence in Multi-Agent Systems Based on LLMs
 - **Fermilab, DOE National Lab** Batavia, IL
Fermilab Theory Division Feb-Aug 2019
Advisor: Roni Harnik
Research: Particle Physics Phenomenology and Formal Theory

- **University of São Paulo** São Paulo, Brazil
 The São Paulo Research Foundation Scientific Initiation Fellowship 2016-2017
 Advisor: Nestor Caticha
 Research: Statistical Mechanics and Neural Networks

Professional Experience

- **Independent Contractor, Mercor** Remote
 Computational PhD May 2026-Present
 Physics Expert Oct 2025-Present

Awards and Honors

- Oral Presentation Selection, LXAI Workshop @ ICML 2026 Jul 2026
- Oral Lighting Presentation Selection, AAMAS 2026 Workshop SE May 2026
- Oral Presentation Selection, PersonaNLP Workshop @ NeurIPS 2025 Nov 2025
- University of Chicago nominee for Schmidt Science Fellows May 2025
“Selected as institutional nominee for a highly competitive international fellowship supporting exceptional early-career scientists to pursue interdisciplinary research at world-leading institutions.”
- Nathan Sugarman Award for Excellence in Graduate Research May 2025
“For his original contributions to discrete gauge theory and modern concepts of symmetry in quantum field theory, and his discovery of new examples of electric-magnetic duality.”
- Physical Sciences Division Fellowship Summer 2021
“Your nomination by the Department of Physics and selection by the Physical Sciences Division Fellowship Committee are testimony to your strong academic potential”
- McCormick Fellowship 2020-2021
“Awarded by the Admissions Committee to highly-rated applicants to the Ph.D. program of the Department of Physics.”
- International House Fellow Award 2021-2024
“Fellows are expected to actively participate in the social and cultural life of the House and bring to International House the richness of their cultures and individual interests.”
- Wolfram Staff Pick 2021
“These exceptional posts are curated by Wolfram Community moderators. This is a mark of excellence and a way to show appreciation to our outstanding contributors.”
- Highlight by Brazilian Physics Society 2019
“For finding a new mathematical solution for the $U(1)$ anomaly equations that helps model building in particle physics.”

Publications

-
- | | |
|--|------------|
| 13. Emergent Discrimination and Polarization in Multi-Agent Systems | Sept 2026 |
| Submitted to ICLR 2027 | |
| Authors: Felipe Alves, Davi Bastos Costa, Nestor Caticha and Renato Vicente | |
| 12. Persona-Model Collapse in Emergent Misalignment | May 2026 |
| Oral presentation at ICML 2026 LXAI Workshop, e-Print arXiv:2605.12850 [cs-CL] | |
| Entered Neurips 2026 phase 3 with 1 borderline accept and 2 accepts | |
| Authors: Davi Bastos Costa and Renato Vicente | |
| 11. Moral Susceptibility and Robustness under Persona Role-Play in LLMs | Nov 2025 |
| Oral presentation at NeurIPS 2025 PersonaNLP Workshop, e-Print arXiv:2511.08565 [cs-CL] | |
| Authors: Davi Bastos Costa, Felipe Alves and Renato Vicente | |
| 10. Deceive, Detect, and Disclose: Large Language Models Playing Mini-Mafia | Sept 2025 |
| Oral presentation at AAMAS 2026 Workshop SE, e-Print arXiv:2509.23023 [cs-AI] | |
| Presentation at COLM 2026 Workshop on Agent Behavior | |
| Authors: Davi Bastos Costa and Renato Vicente | |
| 9. Effective field theory for weakly bound two-neutron halo nuclei | Mar 2025 |
| Phys. Rev. C 112, 014001 | |
| Authors: Davi Bastos Costa, Masaru Hongo, Dam Thanh Son | |
| 8. Non-invertible symmetries as condensation defects in finite-group gauge theory | Dec 2024 |
| Submitted to the SciPost journal, e-Print arxiv:2412.16681 [cond-mat.str-el] | |
| Authors: Clay Cordova, Davi Bastos Costa and Po-Shen Hsin | |
| 7. Simons Lectures on Categorical Symmetries | Nov, 2024 |
| e-Print arXiv:2411.09082 [math-ph] | |
| Authors: Davi Bastos Costa, Clay Córdoba, Michele Del Zotto, Dan Freed, et al. | |
| 6. Non-invertible symmetries in finite group gauge theory | Jul 2024 |
| SciPost Phys. 18 (2025) 019 | |
| Authors: Clay Cordova, Davi Bastos Costa and Po-Shen Hsin | |
| 5. Neutrino Masses in the Mirror Twin Higgs with Spontaneous $\mathbb{Z}_2$ Breaking | April 2024 |
| Journal of High Energy Physics 09 (2024) 106 | |
| Authors: Pedro Bittar, Davi Bastos Costa and Gustavo Burdman | |
| 4. Benefits of marriage as a search strategy | Aug 2021 |
| Selected by Wolfram Staff Picks and published online, e-Print arXiv:2108.04885 [econ.TH] | |
| Authors: Davi Bastos Costa (a pandemic pastime) | |
| 3. Anomaly-free $U(1)^m$ extensions of the Standard Model | Jul 2020 |
| Physical Review D 102 (2020) 11, 115006 | |
| Authors: Davi Bastos Costa | |
| 2. Chiral Abelian gauge theories with few fermions | Jan 2020 |
| Physical Review D 101 (2020) 9, 095032 | |
| Authors: Davi Bastos Costa, Bogdan Dobrescu and Patrick Fox | |

1. **General solution to the $U(1)$ anomaly equations** Oct 2019
Physical Review Letters 123 (2019) 15, 151601, highlight in the Brazilian Physics Society
Authors: Davi Bastos Costa, Bogdan Dobrescu and Patrick Fox

Selected Scientific Events

21. COLM 2026, San Francisco	Oct-2026
20. ICML 2026, Seoul	Jul-2026
19. AAMAS 2026, Paphos	May-2026
18. ICLR 2026, Rio de Janeiro	Apr-2026
17. NeurIPS 2025, Mexico City	Dec-2025
16. Strings 2025 Abu Dhabi, NYU Abu Dhabi	Jan-2025
15. 2024 IHES Summer School - Symmetries and Anomalies, IHES	Jun-July 2024
14. Categorical Symmetries in Quantum Field Theory, Edinburgh University	July 2024
13. Paths to Quantum Field Theory 2023, Durham University	Aug 2023
12. Prospects in Theoretical Physics 2023, Princeton University	July 2023
11. TASI 2023: Aspects of Symmetries, University of Colorado Boulder	Jun-July 2023
10. Global Categorical Symmetries School and Conference, Perimeter Institute	June 2022
9. Conference on Early Childhood Development: Economic Theory and Data, UChicago	Nov 2022
8. 12th International Neutrino Summer School 2019, Fermilab	Aug 2019
7. CTEQ School on QCD, Pittsburgh University	July 2019
6. The 2019 Phenomenology Symposium, Pittsburgh University	May 2019
5. 4th Joint Dutch-Brazil School on Theoretical Physics, ICTP-SAIFR	Feb 2019
4. XX Escola de Verão Jorge Swieca de Partículas e Campos, SBF	Jan 2019
3. III IFT-Perimeter-SAIFR Journeys into Theoretical Physics, ICTP-SAIFR	Jul 2018
2. First Joint Trieste/Saifr School on Particle Physics, ICTP-SAIFR	Jun 2018
1. II IFT-Perimeter-SAIFR Journeys into Theoretical Physics, ICTP-SAIFR	Jul 2017