Leilani H. Gilpin

Engineering 2 Office 347B
1156 High Street
Santa Cruz, CA 95060
☎ +1 (415) 937 1806
✉ lgilpin@ucsc.edu
🖥 lgilpin.com
🐦 leilanigilpin

Last updated: November 6, 2024

Research Interests

The theories and methodologies towards monitoring, designing, and augmenting machines that can **explain** themselves for diagnosis, accountability, and liability.

Employment

2021-present	**UC Santa Cruz**, *Assistant Professor of Computer Science and Engineering (CSE)* *Faculty affiliate in the Science and Justice Research Center (SJRC).*
2020-2021	**Sony AI**, *Research Scientist II*
2013-2015	**Palo Alto Research Center (PARC)**, *Member of Technical Staff*

Education

2015–2020	**Ph.D., Electrical Eng. and Computer Science**, *MIT* Dissertation: Anomaly Detection through Explanations Advisors: Gerald Jay Sussman and Lalana S. Kagal
2011–2013	**M.S., Computational and Mathematical Engineering**, *Stanford University*
2006–2011	**B.S., Computer Science and B.S., Mathematics**, *UC San Diego (UCSD)* Highest Honors in Computer Science, Honors in Mathematics, Music Minor

Honors and Awards

2024	Air Force Office of Scientific Research Young Investigator Program (AFOSR YIP)
2022	AAAI New Faculty Highlight
2020	Rising Stars in EECS
2020	ACM FAccT Travel Award
2018	Nokia Bell Labs Prize Finalist
	Finalist for prize that recognizes research that "changes the game" in the field of information and communications technologies by a factor of 10.
2018	AAAI Doctoral Consortium Travel Award
2017	Nokia Bell Labs Prize Semi-finalist

2016	USENIX Security Student Travel Award
2016-2020	MIT University Center for Exemplary Mentoring (UCEM) Sloan Scholar
2015	MIT ODGE Diversity Fellowship
2011-2013	National Science Foundation (NSF) Graduate Research Fellowship
2013	Stanford SSB Health IT Competition 1st Place
2011	Stanford School of Engineering Fellowship
2011	Yahoo! HackU All Stars Finalist
2011	Yahoo! HackU First Place
2011	Yahoo! Excellence Award
2010	CRA Outstanding Undergraduate Researcher Honorable Mention
2009-present	Member of Tau Beta Pi and Eta Kappa Nu
2010	Tau Beta Pi Scholarship
2009	Gary C. Reynolds Memorial Scholarship
2009	BAE Scholarship Finalist

Publications

According to Google scholar: 3,615 citations, h-index: 10 , i10-index: 11.

Refereed Journal Publications

[1] Adam Amos-Binks, Dustin Dannenhauer, and Leilani H Gilpin. "The anticipatory paradigm". In: *AI magazine* 44.2 (2023-06). ISSN: 0738-4602.

[2] Sharan Subramanian and Leilani Gilpin. "Convolutional Neural Network Model for Diabetic Retinopathy Feature Extraction and Classification". In: *Journal of Student Research* 13.1 (Feb. 2024). DOI: 10.47611/jsrhs.v13i1.6093. URL: https://www.jsr.org/hs/index.php/path/article/view/6093.

[3] Ryan Jenkins, Kristian Hammond, Sarah Loehr, and Leilani Gilpin. "Separating Facts and Evaluation: Motivation, Account, and Learnings from a Novel Approach to Evaluating the Human Impacts of Machine Learning". In: *AI and Society*. 2022.

[4] Peter R Wurman, Samuel Barrett, Kenta Kawamoto, James MacGlashan, Kaushik Subramanian, Thomas J Walsh, Roberto Capobianco, Alisa Devlic, Franziska Eckert, Florian Fuchs, Leilani Gilpin, et al. "Outracing champion Gran Turismo drivers with deep reinforcement learning". In: *Nature* 602.7896 (2022), pp. 223–228.

[5] Michal Araszkiewicz, Ilaria Angela Amantea, Saurabh Chakravarty, Robert van Doesburg, Maria Dymitruk, Marie Garin, Leilani Gilpin, Daphne Odekerken, and Seyedeh Sajedeh Salehi. "ICAIL Doctoral Consortium, Montreal 2019". In: *Artif. Intell. Law* 28.2 (2020), pp. 267–280. DOI: 10.1007/s10506-020-09267-z. URL: https://doi.org/10.1007/s10506-020-09267-z.

[6] Ioana Baldini, Clark Barrett, Antonio Chella, Carlos Cinelli, David Gamez, Leilani Gilpin, Knut Hinkelmann, Dylan Holmes, Takashi Kido, Murat Kocaoglu, et al. "Reports of the AAAI 2019 Spring Symposium Series". In: *AI Magazine* 40.3 (2019), pp. 59–66.

[7] Leilani H. Gilpin, Jamie C. Macbeth, and Evelyn Florentine. "Monitoring Scene Understanders with Conceptual Primitive Decomposition and Commonsense Knowledge". In: *Advances in Cognitive Systems* 6 (2018).

[8] Ayesha Bose, Leilani Gilpin, Jamin Agosti, and Quinn Dang. "The Veicl Act: Safety and Security for Modern Vehicles". In: *Willamette L. Rev.* 53 (2016), p. 137.

[9] Juan Liu, Eric Bier, Aaron Wilson, John Alexis Guerra-Gomez, Tomonori Honda, Kumar Sricharan, Leilani Gilpin, and Daniel Davies. "Graph analysis for detecting fraud, waste, and abuse in healthcare data". In: *AI Magazine* 37.2 (2016), pp. 33–46.

Peer Reviewed Conference Publications

[10] Li Liu, Diji Yang, Sijia Zhong, Kalyana Suma Sree Tholeti, Lei Ding, Yi Zhang, and Leilani H. Gilpin. "Right this way: Can VLMs Guide Us to See More to Answer Questions?" In: *To appear in NeurIps 2024*. 2024.

[11] Priyesh Vakharia, Abigail Kufeldt, Max Meyers, Ian Lane, and Leilani H Gilpin. "ProSLM: A Prolog Synergized Language Model for explainable Domain Specific Knowledge Based Question Answering". In: *International Conference on Neural-Symbolic Learning and Reasoning*. Springer. 2024, pp. 291–304.

[12] Rui-Jie Zhu, Ziqing Wang, Leilani Gilpin, and Jason K. Eshraghian. "Autonomous Driving with Spiking Neural Networks". In: *To appear in NeurIps 2024*. 2024. URL: https://arxiv.org/abs/2405.19687.

[13] Diego Ortiz Barbosa, Leilani H. Gilpin, and Alvaro A. Cardenas. "Semi-Automated Synthesis of Driving Rules". In: *ISOC Symposium on Vehicle Security and Privacy (VehicleSec 2023)* (2023).

[14] Biagio La Rosa, Leilani H. Gilpin, and Roberto Capobianco. "Towards a fuller understanding of neurons with Clustered Compositional Explanations". In: *Thirty-seventh Conference on Neural Information Processing Systems*. 2023. URL: https://openreview.net/forum?id=51PLYhMFWz.

[15] Adam Amos-Binks, Dustin Dannenhauer, and Leilani H Gilpin. "Anticipatory Thinking Challenges in Open Worlds: Risk Management". In: *AAAI Spring Symposium on Designing Artificial Intelligence for Open Worlds*. 2022.

[16] Gregory Falco and Leilani H Gilpin. "A stress testing framework for autonomous system verification and validation (v&v)". In: *2021 IEEE International Conference on Autonomous Systems (ICAS)*. IEEE. 2021, pp. 1–5.

[17] Leilani H. Gilpin, Vishnu Penubarthi, and Lalana Kagal. "Explaining Multimodal Errors in Autonomous Vehicles". In: *2021 IEEE 8th International Conference on Data Science and Advanced Analytics (DSAA)*. 2021, pp. 1–10. DOI: 10.1109/DSAA53316.2021.9564178.

[18] Leilani H. Gilpin. "System-wide Monitoring for Anomaly Detection". In: *Advances in Cognitive Systems* (2020).

[19] Jason R. Wilson, Leilani H. Gilpin, and Irina Rabkina. "A Knowledge Driven Approach to Adaptive Assistance Using Preference Reasoning and Explanation". In: *arXiv preprint arXiv:2012.02904* (2020).

[20] Leilani H. Gilpin. "Explaining Possible Futures for Robust Autonomous Decision-Making". In: *Proceedings of the AAAI Fall Symposium on Anticipatory Thinking* (2019).

[21] Leilani H. Gilpin, David Bau, Ben Z Yuan, Ayesha Bajwa, Michael Specter, and Lalana Kagal. "Explaining explanations: An overview of interpretability of machine learning". In: *2018 IEEE 5th International Conference on data science and advanced analytics (DSAA)*. IEEE. 2018, pp. 80–89.

[22] Leilani H. Gilpin, Danielle M. Olson, and Tarfah Alrashed. "Perception of Speaker Personality Traits Using Speech Signals". In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM. 2018, LBW514.

[23] Leilani H. Gilpin, Cagri Zaman, Danielle Olson, and Ben Z Yuan. "Reasonable perception: Connecting vision and language systems for validating scene descriptions". In: *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM. 2018, pp. 115–116.

[24] Leilani H. Gilpin and Ben Ze Yuan. "Getting Up to Speed on Vehicle Intelligence". In: *AAAI Spring Symposium Series*. 2017.

[25] Leilani Gilpin, Laurent Ciarletta, Yannick Presse, Vincent Chevrier, and Virginie Galtier. "Co-simulation solutions using AA4MM-FMI applied to smart space heating models". In: *Seventh International Conference on Simulation Tools and Techniques*. 2014.

Peer Reviewed Short Conference Publications and Abstracts

[26] Jonathan Wellington Morris, Vishrut Shah, Alex Besanceney, Daksh Shah, and Leilani H. Gilpin. *Slug Mobile: Test-Bench for RL Testing*. 2024. arXiv: 2409.10532 [cs.RO]. URL: https://arxiv.org/abs/2409.10532.

[27] Leilani H Gilpin. "Accountability layers: explaining complex system failures by parts". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 37. 13. 2023, pp. 15439–15439.

[28] Leilani H. Gilpin. "Anticipatory Thinking: A Testing and Representation Challenge for Self-Driving Cars". In: *2021 55th Annual Conference on Information Sciences and Systems (CISS)*. 2021, pp. 1–2. DOI: 10.1109/CISS50987.2021.9400212.

[29] Leilani H. Gilpin and Lalana Kagal. "An Adaptable Self-Monitoring Framework for Opaque Machines". In: *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*. International Foundation for Autonomous Agents and Multiagent Systems. 2019, pp. 1982–1984.

[30] Leilani H. Gilpin. "Reasonableness Monitors". In: *Thirty-Second AAAI Conference on Artificial Intelligence*. 2018.

Peer Reviewed Workshop Publications

[31] Shengjie Xu and Leilani H. Gilpin. "DANGER: A Framework for Generating Dangerous Scenes for Testing Robustness of Machine Learning". In: *Neurips Workshop on Progress and Challenges in Building Trustworthy Embodied AI* (2022).

[32] Leilani H. Gilpin. "Reconciling System-wide Errors with Symbolic Explanations". In: *Proceedings of the IJCAI Workshop on AI for Anomaly Detection* (2020).

[33] Leilani H. Gilpin. "Monitoring Opaque Learning Systems". In: *ICLR 2019 Debugging ML Models Workshop* (2019).

[34] Leilani H. Gilpin, Tianye Chen, and Lalana Kagal. "Learning From Explanations for Robust Autonomous Driving." In: *ICML Workshop on AI for Autonomous Driving*. 2019.

[35] Leilani H. Gilpin, Cecilia Testart, Nathaniel Fruchte, and Julius Adebayo. "Explaining Explanations to Society". In: *arXiv preprint arXiv:1901.06560* (2018).

Other Publications and Pre-Prints

[36] Shiyuan Huang, Siddarth Mamidanna, Shreedhar Jangam, Yilun Zhou, and Leilani H. Gilpin. *Can Large Language Models Explain Themselves? A Study of LLM-Generated Self-Explanations*. 2023. arXiv: 2310.11207 [cs.CL].

[37] Shreyan Mitra and Leilani Gilpin. *A novel post-hoc explanation comparison metric and applications*. 2023. arXiv: 2311.10811 [cs.LG].

[38] Shreyan Mitra and Leilani Gilpin. *The XAISuite framework and the implications of explanatory system dissonance*. 2023. arXiv: 2304.08499 [cs.LG].

[39] Sharan Subramanian and Leilani H. Gilpin. *Convolutional Neural Network Model for Diabetic Retinopathy Feature Extraction and Classification*. 2023. arXiv: 2310.10806 [eess.IV].

[40] Karianne Bergen and Leilani Gilpin. *Negative News No More: Classifying News Article Headlines*. Tech. rep. 11. 2012, p. 2017.

[41] Leilani Gilpin. *The Impact of Topology and Communication Models on Connectivity in Networks*. 2011.

[42] Leilani Hendrina Gilpin. "Anomaly detection through explanations". PhD thesis. Massachusetts Institute of Technology, 2020.

Invited Talks and Panels

Talks	**NeuroSymbolic Approaches for Trustworthy, Explainable Complex Systems**	
	Alan Turing Institute NeSy AI group (Remote)	*2024*
	Responsible Explainability for Complex Systems	
	adeptID Lunch & Learn	*2024*
	Explaining Self-Driving Cars for Accountable Autonomy	
	NeurIps Workshop on XAI in Action	*2023*
	Explaining and Generating Dangerous Scenarios for Testing AVs	
	Workshop on Scenario Generation for Testing AVs (SceGen)	*2023*
	Kansas State University	*2023*
	Accountability Layers	
	Stanford University: Computing & Society Seminar	*2024*

	Brown University	*2023*
	UCSC Applied Mathematics Seminar	*2023*
	ML Fairness at Illinois University	*2023*
	University of Birmingham (joint with R. Marinescu)	*2022*
	University College London	*2022*

Explaining Errors in Autonomous Driving

	JHU Department of Civil & Systems Engineering	*2021*

Anomaly Detection through Explanations

	55th Annual Conference on Information Sciences and Systems	*2021*
	FUZZ-IEEE	*2021*
	Cornell Computing and Information Science	*2021*
	Utah School of Computing	*2021*
	USC Computer Science	*2021*
	UC Santa Cruz CSE	*2021*
	Institute of Assured Autonomy at JHU	*2020*
	31st International Workshop on Principles of Diagnosis (DX)	*2020*
	USC - Information Science Institute	*2020*
	ICML Workshop on Monitoring ML Systems	*2020*
	Sony AI	*2020*
	Northeastern University - Experiential AI	*2020*
	Northwestern University - Computer Science	*2020*
	Salesforce Research (canceled due to Covid-19)	*2020*
	Rochester Data Science Consortium (postponed due to Covid-19)	*2020*
	Idexx (postponed due to Covid-19)	*2020*

Explaining Explanations in AI

	Statistics Canada (STATCAN)	*2021*
	TU Berlin	*2020*
	U. of Cambridge Distributed Info. and Automation Lab (DIAL) XAI Day	*2020*
	Stanford University - Knowledge Graphs Seminar	*2020*

The Car Can Explain!

	UC San Diego - Halicioglu Data Science Institute	*2019*
	CSAIL-Toyota Meeting	*2018*
	MIT Museum	*2018*
	Columbia Law School: Software Freedom Law Center	*2018*
	CSAIL-Toyota Meeting	*2017*
Panelist	UCSC Open Source Symposium on AI and Open Source	*2024*

	ICLR Workshop on AI for Earth and Space Sciences	*2022*
	ESIP Winter Meeting visionary panel on knowledge graphs	*2021*
	Northwestern ML Impact Initiative: Intelligibility, Fairness, Transparency	*2021*
	Columbia Law School: Software Freedom Law Center	*2019*
Invite-only	Bay Area Alignment Workshop	*2024*
	Dagstuhl Seminar	*2024*
	Trustworthiness and Responsibility in AI - Causality, Learning, and Verification	
	8th Annual CHAI Workshop (Asilomar)	*2024*
Outreach	Rotatary Club of Salinas	*2024*
	BayHacks	*2024*
	UCSC Slugs and Steins	*2023*
	UCSC AI Club: XAI Talk	*2022*
	St. Olaf Hackathon	*2023*
	Baskin Day	*2023*

Grants

PI **Air Force Office of Scientific Research Young Investor Program (AFOSR YIP)**. *Frame-Based Monitoring to Detect and Explain Multimodal Autonomous System Errors*. March 1, 2024-February 28, 2027. Award Amount: $450,000.00.

Underwriters Laboratory (subaward under Northwestern University). *Adversarial Examples to Test Explanation Robustness*. March 1, 2022-April 1, 2023. Award Amount: $125,477.00.

UCSC Office of Research Seed Funding for Early Stage Initiatives. *Robustifying Machine Learning for Safe and Secure Autonomous Vehicles*. March 15, 2022 - March 31, 2023. Alvaro Cardenas (Co-PI), Daniel Fremont (Co-PI), and Cihang Xie (Co-PI). Award Amount: $40,000.00.

Underwriters Laboratory (subaward under Northwestern University). *Explaining Hallucinations in LLMS with Commonsense Monitoring*. March 1, 2024-April 1, 2025. Award Amount: $150,000.00.

Co-PI **Department of Transportation (DoT) National Center for Transportation Cybersecurity and Resiliency**. *Reducing Transportation Cybersecurity Risks*. 2022-2026. UC Santa Cruz Team: Alvaro Cardenas (Co-PI), Daniel Fremont (Co-PI), and Cihang Xie (Co-PI).

Learning Lab's Grand Challenge: Building Critical Mass for Data Science. *Building Data Science Communities for Improving Student Success*. 2023-2026. Team: Pedro Morales-Almazan(PI), Marcella Gomez (Co-PI), Judith Canner (Co-PI), Leilani Gilpin (Co-PI), Sangwon Hyun (Co-PI). Award Amount: $345,000.00

.

Gifts

Siemens CITRIS Research Gift 2023-2024. Gift Amount: $50,000

UCSC Committee on Research (COR) Faculty Allowance (CFA). 2022, 2023. Gift Amount: $2,000 per year.

Tutorials

AAAI '23	Trustworthy and Responsible AI [tutorial website]	2023
	Co-lead with Yilun Zhou, Jieyu Zhao, Harsha Nori, and Besmira Nushi.	
US2TS '22	Knowledge-based commonsense reasoning and explainability	2022
	Co-lead with Filip Ilievski of USC ISI.	

Selected Press

- Awe and trepidation as AI comes for smartphones: [Yahoo! News]
- The Tricky Ethics of AI in the lab: [Chemical and Engineering News (CEN)]
- Celebrating Baskin Engineers on International Women in Engineering Day: [UCSC News]
- New data science project seeks to span disciplines and create community: [UCSC News].
- UC Santa Cruz engineers join major transportation cybersecurity project: [UCSC News].
- Santa Cruz County's potential vulnerability to a cyber attack: [KION 46 article]
- Researchers Develop DANGER Framework for Stress Testing. [Research Spotlight]
- Robustifying AVs Project Chosen for Seed Funding. [UCSC News] .
- Mob.ly App Makes Driving Safer by Changing How Drivers Navigate. [AI Pulse Report] .
- MIT CSAIL Student Spotlight. [Student Profile]
- MIT student lead AI and Ethics Reading Group. [MIT News].
- MIT Internet Policy Research Initiative (IPRI) [Student profile].

Teaching

Lead Instructor

UCSC	Advanced ML (CSE 290C) Graduate	*Spring 2024*
	Artificial Intelligence (CSE 240) Graduate	*Winter 2023, Winter 2024, Fall 2024*
	Responsible Data Science (CSE 246)	*Fall 2022*
	Artificial Intelligence (CSE 140)	*Winter 2021, Spring 2022/2023, Fall 2023*
MIT	Artificial Intelligence and Global Risks	*IAP 2018*

Developed, taught, managed a new course on the risks of AI from a global perspective. [course webpage].

Stanford	SMASH Institute - Calculus	*Summer 2015*

Planned and lead weekly lectures to teach a semester-long calculus class over the summer.

Guest Lectures

UCSC	Crown College plenary	*2024*
	Navigating the Ethical Landscape of AI a panel with Sri Rao	
	Artificial Intelligence: CSE140	*Spring 2024*
	Delivered two lectures on Adversarial Search.	
	Advanced Computer Security: CSE233	*Winter 2024*
	Delivered two lectures on (deep) Reinforcement Learning.	
	Machine Learning (Graduate): CSE 242	*Fall 2022*
	Machine Learning: CSE 142	*Fall 2022, Fall 2023*
	Projects in Artificial Intelligence: CSE 247	*Spring 2022*
	Research & Teaching in CSE: CSE 200	*Winter 2022, Fall 2022, Fall 2023*
UConn	Projects in Artificial Intelligence: 5095 - IETs	*Fall 2022*
USC	DSCI-558 class on Common Sense and Explainability	*Fall 2022*
Emmanuel	Practical Machine Learning (Invited talk on GANs and XAI)	*Spring 2022/2023*
Northwestern	CS 496: AI Perspectives (Invited talk on XAI)	*Spring 2021/2022/2023*
Stanford	CS 520: Knowledge Graphs (Invited talk on XAI)	*Spring 2020*
MIT	6.905/6.945: Large-scale Symbolic Systems	*Spring 2019*
	6.S978: Privacy Legislation in Practice: Law and Technology	*Spring 2017*

Teaching Assistant

MIT	6.905/6.945: Large-scale Symbolic Systems	*Spring 2019*
Stanford	CS 348A: Geometric Modeling (PhD Level Course)	*Spring 2013*
UCSD	COGS 5A (beginning java)	
	CSE 8A/8B (beginning java)	
	CSE 5A (beginning C)	
	CSE 21 (discrete mathematics)	
	CSE 100 (Advanced Data Structures)	
	CSE 101 (Algorithms)	

Mentoring

PhD students (primary supervisor)

	Oliver Chang	*2022-present*
	Li Liu	*2023-present*
	Jack Fox Keen	*2023-present*
	Peiyu (Olivia) Wang	*2023-present*
	Shiyuan Huang	*2024-present*
	Priyesh Vakharia	*2024-present*
	Creighton Glasscock	*2024-present (incoming)*

Postdoctoral Researchers (primary supervisor)

	Biagio La Rosa	*2024-present*

Visiting PhD students (primary supervisor)

Sapienza U.	Biagio La Rosa	*2022-2023*

PhD student committee service (other supervisor)

Def. Committee	Connor Priyor	*2024*
	Eliana Stefani	*2024*
	Jayanth Yetukuri	*2024*
	Zhaowei Zhu	*2023*
Adv. Committee	Héctor Carrión	*2024*
	Sebastián Castro	*2024*
	Zonglin Di	*2024*
	Changmao Li	*2023*
	Michael Briden	*2022*
	Anuj Kaul	*2022*

Master's Students

Thesis Advisor	Sijia Zhong	*2023-present*
	Sanjana Sudhakar Patil	*2023-2024*
	Co-advised with Kate Ringland, Prof. of Computational Media	
	Shengjie Xu	*2021-2023*
	Sainaya Brid	*2022-2023*
	Sanjana Sudhakar Patil	*2023-2024*
	Co-advised with Kate Ringland, Prof. of Computational Media	
Thesis Reader	Ex Taranenko	*2024*
Project Advisor	Batuhan Salih	*2024-present*
	Aaja Ouellette	*2024-present*
	Ziyuan Wang	*2024-present*
	Rithesh Kumar	*2024-present*
	Akashaleena Sarkar	*2024-present*

Zhonghui Li	*2024-present*
Vinh Le	*2024-present*
Pratik Doshi	*2024-present*
Kajal Patil	*2023-present*
Karthik Bhatt	*2024-present*
William Self	*2024-present*
Nistha Kumar	*2024-present*
Vidhata Singh	*2024-present*
Rami Wilson	*2024-present*
Khin (Sabrina) Yone	*2024-present*
Anish Kumar	*2023-present*
Gene Park	*2024-present*
Richard Ho	*2024*
Ryan Leung	*2024*
Michael Soebroto	*2024*
Yanwen Xu	*2024*
Jimmy Franknedy	*2023-2024*
Rohan Ghosalkar	*2023-2024*
Shiyuan Huang	*2023-2024*
Ameyaa Biwalkar	*2022-2024*
Dhanishtha Patil	*2023-2024*
Harshini Venkataraman	*2023-2024*
Anuj Kamat	*2023-2024*
Olivia Anastassov	*2022-present*
On leave since June 2023	
Kalyana Suma Sree Tholeti	*2023-2024*
Arka Pal	*2023*
Maithili Luktuke	*2023*
Supriya Paulose	*2023*
Ryan Ahari	*2023*
Shayan Salsabilian	*2023*
Joseph Silberman	*2023*
Sukhveer Karlcut	*2023*
Vidyasri Ravi	*2022-2023*
John Yu	*2022-2023*
Rohit Das	*2022-2023*

	Smruthi Pobbathi	*2022-2023*
	Rangasri Chakravarthy	*2022-2023*
	Jacob Low	*2022-2023*
	Yash Chhabria	*2022-2023*
	Keely Wheelan	*2022-2023*
	Zhihang Zhou	*2022*
Project Reader	Nafisa Hussain	*2024*
	Jay Mehta	*2024*
	Valentina Tang	*2024*
	Zichao Li	*2024*
	Daniel Sabo	*2024*
	Najmeh Mashhadi	*2023*
	Golam Md Muktadir	*2023*
	Srilekha Vutukuru	*2023*
	Aaron Monajjemi	*2023*
	Christopher Liu	*2023*
	Alexander Miller	*2023*
	Sreevani Suvarna	*2023*
	Rahul Bharadwaj	*2022*
	Alec Siegel	*2022*
	Ian Hardy	*2023*
	Jinghao Shen	*2023*

Undergraduate Students (UCSC)

Primary Advisor	Rishika Srinivas	*2022-present*
	Arnav Kartikeya	*2022-present*
	Shreedhar Jangam	*2023-present*
	2024 Dean's Award Winner.	
	Siddarth Mamidanna	*2023-present*
	2024 Dean's Award Winner.	
	Neha Suresh	*2024-present*
	Rajat Maheshwari	*2024-present*
	Anshika Agarwal	*2024-present*
	Selam Mitike	*2024-present*
	Daniel Yang	*2024-present*
	Ashwin Pazhoor	*2024-present*
	Bill Zhang	*2024-present*

Jihang Li	*2024-present*
Anish Talluri	*2024-present*
Ian Yam	*2024-present*
Thomas Wang	*2024-present*
Umair Rizwan	*2024-present*
Sri Seethammagari	*2024-present*
Samiyah Shaikh	*2024-present*
Macklin Reeve-Wilson	*2024-present*
Sebastian Woo	*2024-present*
Sam Cheng	*2023-present*
Ish Khandelwal	*2023-present*
Tommy Chen	*2022-present*
Raj Singh	*2023-2024*
Coen Adler	*2021-2024*
Bachelor's thesis student. 2024 Chancellor's Award Winner.	
Batuhan Salih	*2022-2024*
Nick Wang	*2022-2024*
Tong Jia	*2023-2024*
Roger Li	*2023-2024*
Camden Beard	*2023-2024*
Daniel Hsieh	*2023-2024*
Max Myers	*2023-2024*
Oliver Lokhandwala	*2023-2024*
Koret Scholarship Recipient.	
Fariha Lateef	*2023-2024*
Andrew Susanto	*2023-2024*
Brandon Shapiro	*2023-2024*
Daniel Sarni	*2024*
Vaibhav Honakere	*2024*
Erick Hernandez	*2024*
Janvi Rochlani	*2024*
Sebastian Osorio	*2024*
Sahil Gupta	*2022-2023*
Bachelor's thesis student.	
Alex Feghhi	*2023*
Michael McAllister	*2023*

	Suneet Katrekar	*2022-2023*
	Owen Shi	*2022-2023*
	Anish Kumar	*2023*
	Tanisha Khemka	*2022*
	Ariel Kamen	*2022*
	Lan Mi	*2022*
	Radhika Gadre	*2022*
	Aylin Akkus (visiting student)	*2022*
	Christopher Oey	*2022*
	Aaja Ouellette	*2022*
	Isaac Plotkin	*2022*
Reading Committee	Nishanth Jayram	*2022*

MIT Thesis Students (12+ month fulltime student)

MEng	Tianye Chen	*2018-2019*
	Co-advised with Lalana Kagal. Co-authored paper on rule-learning [34].	
SuperUROP	Evelyn Florentine	*2017-2018*
	Co-authored journal paper on monitoring opaque learning systems [7].	
	Zoe Lu	*2017-2018*

MIT Research Project Students (6+ month semester course)

UROP	Shayda Moezzi	*Fall 2020-2021*
	Vishnu S Penubarthi	*Fall 2019-2021*
	Co-authored paper on multimodal explanations [17].	
	Marla E. Odell	*Spring 2019*
	Elizabeth Han	*Spring 2019*
	Obada Alkhatib	*IAP/Spring 2018*
	Michal Reda	*IAP/Spring 2018*
	Ishan Pakuwal	*IAP/Spring 2018*
UAP	Matthew Kalinowski	*Spring 2017*

University Service

University Service	Senate Committee on Rules, Jurisdiction and Elections (RJE)	*2023-present*
Dept. Service	PhD Visit Days/Open House Organizer	*2023*
	MS/PhD Recruitment Committee	*2023-present*
	NLP MS Capstone Project Mentor	*2023-present*

	Graduate Admissions Committee	*2022-present*
	Security and Privacy Faculty Search Committee	*2023-present*
	LSOE Faculty Search Committee	*2023-present*
	LSOE Faculty Search Committee	*2022-2023*
	Grad AI Day Panel Organizer	*2024*
Committee	External Member: HCI Associate Prof. Search (UCSC CM Dept.)	*2022-2023*

Professional Activities

Editorial Board	Journal of Neurosymbolic Artificial Intelligence	*2022-present*
Guest Editor	Special Issue on Neurosymbolic AI for Cyberphysical systems	*2023-present*
	AI Magazine Special Issue on Anticipatory Thinking	*2022-2023*
Program Chair	Advances in Cognitive Systems (ACS) [link]	*2022*
Organizer	AAAI workshop on Explainable AI for Deep RL: [link]	*2024*
	NeurIPS workshop on Building Trustworthy Embodied AI: [link]	*2022*
	AAAI Fall Symposium on Anticipatory Thinking [link]	*2020/2021*
	NeurIps Workshop on XAI for debugging and diagnosis [link].	*2021*
	ACS Workshop on Story Enabled Intelligence. [link].	*2019*
	AAAI Spring Symposium 2019: Story-Enabled Intelligence. [link].	*2019*
	MIT Machine Learning Interpretability Reading Group	*2018-2020*
	MIT AI and Ethics Reading Group. [link].	*2018-2020*
	MIT IPRI Privacy, Security and Policy (PSP) Meeting	*2018-2019*
	MIT Path of Professorship Workshop	*2018*
	MIT EECS Visit Days and Orientation	*2016*
Advisory Board	Center for Advancing Safety of Machine Intelligence (CASMI)	*2022-present*
	Governance Advisory Committee. [CASMI Leadership]	
	Bias Advisory Board (EU)	*2024-present*
	[Bias Advisor Board]	
Local Chair	NeuroSymbolic AI Conference (NeSy)	*2025*
	Advances in Cognitive Systems	*2019*
Senior PC	AAAI	*2022-present*
PC	Advances in Cognitive Systems (ACS)	*2022-present*
	AIC	*2023*
	Facct	*2023*
	Scheme Workshop	*2023*
	Semantic Web Journal issue on Commonsense Knowledge and Reasoning	*2021*
	AAAI	*2021-2022*

	ACS Workshop on Goal Reasoning	*2021*
	AAMAS	*2021*
	DX-2020	*2020*
Grant Reviewer	AFOSR Trust and Influence program	*2023*
	NSF CISE FRR Panel Reviewer	*2022*
Paper Reviewer	The Web Conf	*2023*
	AAAI SRRAI track	*2023*
	EAAI	*2023*
	NeurIPS Workshop	*2023*
	AI4HRI	*2022*
	Information Issue on Foundations and Challenges of Interpretable ML	*2021*
	HRI	*2020-present*
	ACS Workshop on Goal Reasoning	*2021*
	Artificial Intelligence Review	*2020*
	IEEE Transactions on Cybernetics	*2019*
	NeurIPS	*2019*
	AAAI Spring Symposium	*2019*
	Slovak-Israeli Scientific Research Program	*2018*
	MIT MITES	*2018*
	HRI Late Breaking Reports (LBR)	*2018*
	AAAI (Guest Reviewer)	*2015*
Student Rep.	Stanford ICME	*2011-2013*
	MIT EECS Visiting Committee	*2017*
	Met with the EECS Visiting Committee and gave a personal perspective on the EECS Department, student life, and diversity.	
	MIT Grad Rat	*2017-2019*
Volunteer	UCSD Alumni Board	*2015-2019*

Outreach

UCSC Fall Open House Faculty Panel	*2024*
Invited Talk at Los Altos High AI Club	*2024*
Invited Plenary Speaker-BayHacks	*2024*
Rotary Club of Salinas talk	*2024*
Rotaract Club of Santa Cruz County: Judging Panel	*2024*
UCSC Women in Science Society Panel	*2024*
UCSC INSPIRE AI Panel	*2024*

ACM Club at UCSC: "What Do Profs Do All Day?"	2023
SC Works AI Horizons Part 2 Speaker	2023
UCSC ACM/GraceHacks Keynote Speaker	2023
UCSD CSE Ph.D. Industrial Job Search Panel	2023
Tri-Valley Hacks Judge	2023
Santa Cruz County Science & Engineering Fair Judge	2022

Advocacy

I am passionate about AI and policy, AI and ethics, and informing legislation pertaining to liability for Autonomous Vehicles (AVs).

| UC congressional briefing on artificial intelligence | 2023 |

Joint with Ken Goldberg (UC Berkeley), William Wang (UCSB), David Danks (UCSD), and Brandie Nonnecke (UC Berkeley).

Diversity, Equity and Inclusion

| | UCSC MESA Faculty Forum steering committee | 2022-present |

Joint with Imelda Meza, Academic Advisor/Program Coordinator and Marcella Gomez, Associate Professor, AM

	Almanac of Technology and Advancement (ATA) Faculty Sponsor	2024-present
	UCSC ACM-W Board Faculty Sponsor	2022-present
Mentor	Women in ML (WiML) PhD Mentor	2022-present
	Black in AI	2020
	MIT EECS GAAP	2020

The Graduate Application Assistance Program (GAAP) is a student-run, volunteer-based program which provides assistance to EECS PhD applicants from underrepresented groups, including students from groups underrepresented in STEM and students with non-traditional academic backgrounds.

| | Xerox ABI Mentoring Program | 2014-2015 |