

Andreea Bobu

70 Vassar Street, 31-245 – Cambridge, MA 02142

☎ (617)-417-0993 • ✉ abobu@mit.edu • 🌐 www.mit.edu/~abobu

Research Interests

I develop autonomous agents that learn to do tasks for, with, and around humans. My goal is to ensure that these agents align with people, whether expert designers or novice end users. My work looks at: 1) getting the right data to supervise the training of the robot, whether directly from people or via priors; 2) enabling agents and humans to efficiently and interactively arrive at shared task representations for reliable interaction; 3) quantifying and addressing misalignment caused by different human modeling choices. I ground my work in experiments and user studies with AI systems like assistive robot arms or LLM agents, and draw upon methods from deep learning, mathematical human modeling, inverse reinforcement learning, and Bayesian inference.

Professional Positions

- 2024–present **Boeing Assistant Professor**
Massachusetts Institute of Technology, Department of Aeronautics and Astronautics
- 2023–2024 **Research Scientist**
The AI Institute
- Summer 2021 **Research Intern**
NVIDIA Research, Robotics Group

Education

- 2017–2023 **University of California, Berkeley**
Ph.D. in Electrical Engineering and Computer Sciences
Advisor: Anca Dragan
Thesis: [Aligning Robot Representations with Humans](#)
- 2013–2017 **Massachusetts Institute of Technology**
B.S. in Computer Science and Engineering, Minor in Mathematics
Advisors: Adrian Dalca, Polina Golland, Stefanie Jegelka

Awards and Honors

- 2023 **Emerging Research Award at [I. Conference on Mathematics of Neuroscience & AI](#)**
For the talk on “Aligning Robot and Human Representations”.
- 2022 **[Rising Stars](#) Academic Career Workshop in EECS**
Chosen to participate in an intensive workshop for graduate students and postdocs who are interested in pursuing academic careers in EE, CS, and AI and decision-making.
- 2022 **Robotics: Science and Systems (RSS) [Pioneers](#)**
Selected for workshop bringing together top early career researchers in robotics.
- 2021 **[Apple PhD Scholars in Artificial Intelligence and Machine Learning Fellowship](#)**
Two-year fellowship with an annual stipend of \$45,000 for graduate students in AI/ML.
- 2021 **Best Paper Award Finalist at [ACM/IEEE HRI](#)**
For the paper “Feature Expansive Reward Learning: Rethinking Human Input”.
- 2021 **Best Paper Award Honorable Mention at [IEEE T-RO](#)**
For the paper “Quantifying Hypothesis Space Misspecification in Learning From Human-Robot Demonstrations and Physical Corrections”.
- 2020 **Best Paper Award Winner at [ACM/IEEE HRI](#)**
For the paper “LESS is More: Rethinking Probabilistic Models of Human Behavior”.
- 2020 **Human-Robot Interaction (HRI) [Pioneers](#)**

Chosen to participate in a highly selective workshop seeking to foster creativity, communication, and collaboration across Human-Robot Interaction.

2019 **Cadence Women in Technology Scholarship**

A \$5,000 scholarship for women in EECS demonstrating leadership and a strong academic record.

2016 **Best Paper Award Winner at MICCAI Patch-MI**

For the paper “Patch-Based Discrete Registration of Clinical Brain Images”.

2016 **Google Anita Borg Memorial Scholarship**

A \$10,000 scholarship for women in EECS demonstrating leadership and a strong academic record.

2015–present **Member of Tau Beta Pi (TBP) National Honor Society for Engineering**

Honors society for engineering students with the strongest academic records at their university.

2015–present **Member of Eta Kappa Nu (HKN) National Honor Society for EECS**

Honors society for EECS students with the strongest academic records at their university.

Teaching

Spring 2025	16.S948: Algorithmic Human-Robot Interaction	<i>MIT</i>
	Instructor	
Fall 2024–2025	16.410/16.413: Principles of Autonomy and Decision Making	<i>MIT</i>
	Instructor	
Spring 2021	CS 287H: Algorithmic Human-Robot Interaction	<i>UC Berkeley</i>
	Graduate Student Instructor	
Fall 2019	CS 188: Introduction to Artificial Intelligence	<i>UC Berkeley</i>
	Graduate Student Instructor	
January 2016	6.178: Introduction to Software Engineering in Java	<i>MIT</i>
	Instructor and Lecturer	
2015–2017	6.046: Design and Analysis of Algorithms	<i>MIT</i>
	Tutor	
Spring 2014	6.01: Introduction to Electrical Engineering and Computer Science	<i>MIT</i>
	Student Lab Assistant	

Advising & Mentoring

Current Ph.D. Students

Minyoung Hwang

Current M.S. Students

Ananya Kulshrestha, Audrey Lee, Helena Merker, Jordan Abi Nader

Current Undergraduate Students

George Cao, David Lee, Megan Tian, Min Khant Zaw

Past M.S. Students

Yi Liu (→ ML Research Engineer at Scale AI), Arjun Sripathy (→ Senior ML Scientist at Tesla Autopilot)

Past Undergraduate Students

Regina Wang (→ M.S. at Stanford), David Zhang (→ Codepoint Fellow), Matthew Zurek (→ Ph.D. at UW-Madison), Sampada Deglurkar (→ Ph.D. at UC Berkeley)

Ph.D. Committees

Sean Ye (Georgia Tech), Simon Holk (KTH Royal Institute of Technology), Erdi Sayar (Technische Universitat Munchen), Alex Forsey-Smerek (MIT), Alex Cuellar (MIT).

Outreach

Summer 2024	RoboLaunch	CMU
	Speaker	
	I gave a talk at the CMU RI RoboLaunch Speaker Series, an outreach program for promoting robotics & AI research and education.	
Summer 2019	Girls in Engineering Camp	UC Berkeley
	Lecturer and Mentor	
	I co-organized a Self-Driving Cars workshop, teaching the girls about sensing, planning, and control in autonomous driving, and experimenting with an Evo robot.	
August 2018	AI4ALL	UC Berkeley
	Teaching Assistant	
	I mentored a team of underrepresented high school students as they learned to train a deep reinforcement learning agent in MuJoCo.	
2018–2022	Berkeley Artificial Intelligence Research	UC Berkeley
	Mentor	
	I mentored underrepresented undergraduate students in research and career planning.	
2018–2019	Women in Computer Science and Engineering	UC Berkeley
	Mentor	
	I mentored early-stage female PhD students in career planning and navigating life at UC Berkeley.	
2016	Women in Science and Engineering	MIT
	Mentor	
	I mentored high school girls from the Greater Boston area during monthly sessions designed to introduce them to engineering at MIT.	
2013–2015	Educational Studies Program	MIT
	Lecturer	
	I taught courses on “Water Security in Asia”, “Introduction to Probability”, and “Group Theory” to middle school students in the New England region.	

Professional Activities

Conference Area Chair

2024-2025	CoRL: Conference on Robot Learning
2024-2025	ICLR: International Conference on Learning Representations

Workshops & Seminars Co-organized

2025	Workshop on Generative Modeling Meets Human-Robot Interaction	RSS
2025	Workshop on Human-Centered Robot Learning in the Era of Big Data & Large Models	ICRA
2025	7th Workshop on Long-term Human Motion Prediction	ICRA
2024	Workshop on Task Specification for General-Purpose Intelligent Robots	R:SS
2024	Workshop on Mechanisms for Mapping Human Input to Robots	R:SS
2024	6th Workshop on Long-term Human Motion Prediction	ICRA
2024	6th Workshop on Lifelong Learning and Personalization in Long-Term HRI	HRI
2023	Workshop on Interactive Learning with Implicit Human Feedback	ICML
2022	Workshop on Aligning Robot Representations with Humans	CoRL
2022–2023	Dream/CPAR Seminar	UC Berkeley
2022	2nd Workshop on Social Intelligence in Humans and Robots	R:SS
2021	1st Workshop on Social Intelligence in Humans and Robots	ICRA
2020	Workshop on Advances and Challenges in Imitation Learning for Robotics	R:SS
2020–2021	SemiAutonomous Vehicles Seminar	UC Berkeley

External Reviewer for Workshops, Conferences, Journals, and Grant Panels

Robotics: CoRL, ICRA, R:SS, HRI, IROS, L4DC, RA-L, T-RO, T-MECH, T-HRI

Selected Invited Talks

How Structure Enables Efficient and Human-Aligned Robot Learning		
2024	Workshop on Behavioral Machine Learning	<i>NeurIPS</i>
2024	Global Summit on Open Problems for AI Keynote	<i>AE</i>
Why Robots Aren't Superhuman in Our Human World		
2024	TEDx	<i>MIT</i>
Aligning Robot and Human Representations		
2024	Autonomy Talks	<i>ETH</i>
2024	6.161: Robotics Science & Systems	<i>MIT</i>
2024	16-886: Models & Algorithms for Interactive Robotics	<i>CMU</i>
2023	International Symposium on the Mathematics of Neuroscience	<i>ISMnN</i>
2023	Center for Human-Compatible AI Workshop	<i>CHAI</i>
2023	Stanford Robotics Seminar	<i>Stanford</i>
2023	Department Seminar	<i>MIT, Princeton, Georgia Tech, Cornell, Brown, NYU, UIUC, UCSD</i>
2022	UW Robotics Colloquium	<i>UW</i>
2022	New Trends in Aerospace Seminar Series	<i>MIT</i>
2022	CS 6960: Human-AI Alignment	<i>U of Utah</i>
Inducing Structure in Robot Learning via Human-Guided Representations		
2022	SemiAutonomous Vehicles Seminar	<i>UC Berkeley</i>
2021	Workshop on Aware Learning: How to Benefit from Priors	<i>CDC</i>
2021	Workshop on Human-AI Collaboration in Sequential Decision-Making	<i>ICML</i>
2021	Human And Robot Partners (HARP) Lab Reading Group	<i>CMU</i>
2021	CS287H: Algorithmic Foundations of Human-Robot Interaction	<i>UC Berkeley</i>

Journal Articles

- [J3] **Learning Perceptual Concepts by Bootstrapping from Human Queries**
A. Bobu, C. Paxton, W. Yang, B. Sundaralingam, Y.W. Chao, M. Cakmak, D. Fox.
IEEE Robotics and Automation Letters (RA-L), 2022.
- [J2] **Inducing Structure in Reward Learning via Feature Learning**
A. Bobu, M. Wiggert, C. Tomlin, A. D. Dragan.
The International Journal of Robotics Research (IJRR), 2022.
- [J1] **Quantifying Hypothesis Space Misspecification in Learning from Human-Robot Demonstrations and Physical Corrections**
A. Bobu, A. Bajcsy, J. F. Fisac, S. Deglurkar, A. D. Dragan.
IEEE Transactions on Robotics (T-RO), 2019.
Best paper award honorable mention.

Conference Publications

- [16] **QuickLAP: Quick Language-Action Preference Learning for Autonomous Driving**
J. Abi Nader, D. Lee, **A. Bobu**
(in submission) Conference on Neural Information Processing Systems (NeurIPS), 2025.
- [15] **Context Matters: Learning Generalizable Rewards via Calibrated Features**
A. Forsey-Smerek, J. Shah, **A. Bobu**
(in submission) Conference on Neural Information Processing Systems (NeurIPS), 2025.

- [14] **Goal Inference from Open-Ended Dialog**
R. Ma, J. Qu, **A. Bobu**, D. Hadfield-Menell
(in submission) *Conference on Neural Information Processing Systems (NeurIPS)*, 2025.
- [13] **Learning How Hard to Think: Input-Adaptive Allocation of LM Computation**
M. Damani, I. Shenfeld, A. Peng, **A. Bobu**, J. Andreas
International Conference on Learning Representations (ICLR), 2025.
- [12] **Adaptive Language-Guided Abstraction from Contrastive Explanations**
A. Peng, B. Z. Li, I. Sucholutsky, N. Kumar, J. A. Shah, J. Andreas, **A. Bobu**
Conference on Robot Learning (CoRL), 2024.
- [11] **Preference-Conditioned Language-Guided Abstraction**
A. Peng, **A. Bobu**, B. Z. Li, T. R. Summers, I. Sucholutsky, N. Kumar, T. L. Griffiths, J. A. Shah
ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2024.
- [10] **Aligning Robot and Human Representations**
A. Bobu^{*}, A. Peng^{*}, P. Agrawal, J. A. Shah, and A. D. Dragan.
ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2024.
- [9] **Diagnosing and Repairing Feature Representations Under Distribution Shifts**
I. Lourenço, **A. Bobu**, C. R. Rojas, B. Wahlberg.
IEEE Conference on Decision and Control (CDC), 2023.
- [8] **Diagnosis, Feedback, Adaptation: A Human-in-the-Loop Framework for Test-Time Policy Adaptation**
A. Peng, A. Netanyahu, M. K. Ho, T. Shu, **A. Bobu**, J. A. Shah, P. Agrawal.
International Conference on Machine Learning (ICML), 2023.
- [7] **SIRL: Similarity-based Implicit Representation Learning**
A. Bobu^{*}, Y. Liu^{*}, R. Shah, D. S. Brown, and A. D. Dragan.
ACM/IEEE International Conference on Human Robot Interaction (HRI), 2023.
- [6] **Teaching Robots to Span the Space of Functional Expressive Motion**
A. Sripathy, **A. Bobu**, Z. Li, K. Sreenath, D. S. Brown, and A. D. Dragan.
IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2022.
- [5] **Dynamically Switching Human Prediction Models for Efficient Planning**
A. Sripathy^{*}, **A. Bobu**^{*}, D. S. Brown, A. D. Dragan.
IEEE International Conference on Robotics and Automation (ICRA), 2021.
- [4] **Situational Confidence Assistance for Lifelong Shared Autonomy**
M. Zurek^{*}, **A. Bobu**^{*}, D. S. Brown, A. D. Dragan.
IEEE International Conference on Robotics and Automation (ICRA), 2021.
- [3] **Feature Expansive Reward Learning: Rethinking Human Input**
A. Bobu^{*}, M. Wiggert^{*}, C. Tomlin, A. D. Dragan.
ACM/IEEE International Conference on Human Robot Interaction (HRI), 2021.
Best paper award finalist.
- [2] **LESS is More: Rethinking Probabilistic Models of Human Behavior**
A. Bobu^{*}, D. Scobee^{*}, J. F. Fisac, S. Sastry, A. D. Dragan.
ACM/IEEE International Conference on Human Robot Interaction (HRI), 2020.
Best paper award winner.
- [1] **Learning Under Misspecified Objective Spaces**
A. Bobu, A. Bajcsy, J. F. Fisac, A. D. Dragan.
Conference on Robot Learning (CoRL), 2018.
Invited to special issue.

Workshop Publications

- [W9] **Interpretable Human-in-the-Loop In-Context Preference Learning Via Preference Boundaries**
V. K. Chen, J. Shah, **A. Bobu**
Workshop on Human-in-the-Loop Robot Learning: Teaching, Correcting, and Adapting, RSS 2025.
- [W8] **Masked Inverse Reinforcement Learning for Language Conditioned Reward Learning**
M. Hwang, A. Forsey, **A. Bobu**
Workshop on Human-in-the-Loop Robot Learning: Teaching, Correcting, and Adapting, RSS 2025.
- [W7] **Getting Aligned on Representational Alignment**
I. Sucholutsky, L. Muttenthaler, A. Weller, A. Peng, **A. Bobu**, B. Kim, B. C. Love, E. Grant, I. Groen, J. Achterberg, J. B. Tenenbaum, K. M. Collins, K. L. Hermann, K. Oktar, K. Greff, M. N. Hebart, N. Jacoby, Q. Zhang, R. Marjeh, R. Geirhos, S. Chen, S. Kornblith, S. Rane, T. Konkle, T. P. O’Connell, T. Unterthiner, A. K. Lampinen, K. Muller, M. Toneva, T. L. Griffiths
Workshop on Representational Alignment (Re-Align), ICLR 2024.
- [W6] **Time-Efficient Reward Learning via Visually Assisted Cluster Ranking**
D. Zhang, M. Carroll, **A. Bobu**, A. D. Dragan.
Workshop on Human-in-the-Loop Learning, NeurIPS 2022.
- [W5] **Efficient Robot Teaching by Learning Intermediate Human-Guided Representations**
A. Bobu.
Companion of the Robotics: Science and Systems (RSS), 2022.
- [W4] **Aligning Robot Representations with Humans**
A. Bobu, A. Peng.
Workshop on Collaborative Robots and the Work of the Future, ICRA 2022.
- [W3] **Detecting Hypothesis Space Misspecification in Robot Learning from Human Input**
A. Bobu, A. D. Dragan.
Companion of the ACM/IEEE International Conference on Human-Robot Interaction, 2020.
- [W2] **Adapting to Continuously Shifting Domains**
A. Bobu, E. Tzeng, J. Hoffman, T. Darrell.
Workshop at the International Conference on Learning Representations (ICLR), 2018.
- [W1] **Patch-Based Discrete Registration of Clinical Brain Images**
A. V. Dalca, **A. Bobu**, N. S. Rost, P. Golland.
Patch-based Techniques in Medical Imaging (MICCAI Patch-MI), 2016.
Best paper award winner.

Patents

Concept Training Technique for Machine Learning

A. Bobu, B. Sundaralingam, C. Paxton, M. Cakmak, W. Yang, Y. Chao, D. Fox.
U.S. Patent 17982401.

Press & Media

2025	Meet the Mind: MIT Professor Andreea Bobu	<i>MIT CSAIL</i>
2025	Human Centric Robot Learning Faculty Spotlight	<i>CSAIL Alliances</i>
2024	Robot Conversations Podcast	<i>CSAIL Alliances</i>
2023	A Faster Way to Teach a Robot	<i>MIT News</i>