

Siddharth Reddy

CONTACT sgr@berkeley.edu
<https://siddharth.io>

EDUCATION **University of California, Berkeley**
Ph.D. in Computer Science, 2017 - Present, GPA: 3.91
Cornell University
B.S. in Computer Science, Graduated 2017

AWARDS NVIDIA Graduate Research Fellowship (2019-2020)
INFORMS Undergraduate Operations Research Prize (2017)

INDUSTRY INTERNSHIPS **DeepMind**
Ph.D. Intern **May 2019 - August 2019**
Collaborated with Jan Leike and the AI safety team on an algorithm for training reinforcement learning agents from human feedback in the presence of unknown unsafe states. Details at <https://sites.google.com/berkeley.edu/request>.

Knewton
Data Science Intern **May 2014 - October 2015**
Collaborated with Kevin Wilson and Chaitanya Ekanadham on mathematical models for profiling student behavior and recommending personalized content in online courses. Details at <https://siddharth.io/lentil>.

RESEARCH EXPERIENCE **Berkeley Artificial Intelligence Research Lab**
Graduate Student Researcher **July 2017 - Present**
Collaborating with Anca Dragan and Sergey Levine on machine learning algorithms for assistive human-computer interfaces. Details at <https://siddharth.io/files/statement.pdf>.

Machine Learning Research Group, Cornell University
Undergraduate Researcher **September 2014 - May 2017**
Collaborated with Thorsten Joachims, Siddhartha Banerjee, and Igor Labutov on a mathematical framework for spaced repetition that combines ideas from queueing theory and psychology to optimize review scheduling in flashcard software. Details at <https://siddharth.io/leitnerq>.

PUBLICATIONS Jensen Gao, Siddharth Reddy, Glen Berseth, Nicholas Hardy, Nikhilesh Natraj, Karunesh Ganguly, Anca D. Dragan, Sergey Levine, *X2T: Training an X-to-Text Typing Interface with Online Learning from User Feedback*, International Conference on Learning Representations (ICLR), 2021.

Siddharth Reddy, Sergey Levine, Anca D. Dragan, *Assisted Perception: Optimizing Observations to Communicate State*, Conference on Robot Learning (CoRL), 2020.

Siddharth Reddy, Anca D. Dragan, Sergey Levine, Shane Legg, Jan Leike, *Learning Human Objectives by Evaluating Hypothetical Behavior*, International Conference on Machine Learning (ICML), 2020.

Siddharth Reddy, Anca D. Dragan, Sergey Levine, *SQIL: Imitation Learning via Reinforcement Learning with Sparse Rewards*, International Conference on Learning Representations (ICLR), 2020.

Gokul Swamy, Siddharth Reddy, Sergey Levine, Anca D. Dragan, *Scaled Autonomy: Enabling Human Operators to Control Robot Fleets*, IEEE Conference on Robotics and Automation (ICRA), 2020.

Siddharth Reddy, Anca D. Dragan, Sergey Levine, *Where Do You Think You're Going?: Inferring Beliefs about Dynamics from Behavior*, Neural Information Processing Systems (NeurIPS), 2018.

Siddharth Reddy, Anca D. Dragan, Sergey Levine, *Shared Autonomy via Deep Reinforcement Learning*, Robotics: Science and Systems (RSS), 2018.

Siddharth Reddy, Igor Labutov, Siddhartha Banerjee, Thorsten Joachims, *Unbounded Human Learning: Optimal Scheduling for Spaced Repetition*, ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2016.