# Jay Darshan Vakil

jdvakil@gmail.com | jdvakil.github.io | github.com/jdvakil | linkedin.com/in/jdvakil | Jay Vakil

## Research Interests

Robot Learning | Mobile Manipulation | Computer Vision | Embodied AI

## Education

**Doctor of Philosophy in Computer Science,** *University of Colorado* | Boulder, CO *Aug 2024 - Present*

- **Advisors:** Alessandro Roncone, Nikolaus Correll, and Christoffer Heckman
- **Courses**: Advanced Robotics, Transformers based AI in robotics

**Bachelor of Science in Electrical Engineering,** *University of Washington* | Seattle, WA *Sept 2018 - Mar 2022*

- Minor in Mathematics and Computer Science
- **Awards:** Dean's List
- **Courses:** Object-Oriented programming with Java, Software Engineering, Data structures and algorithm, Linear/discrete signal processing, Microprocessor system design, AC/DC circuit analysis, Digital circuit analysis, Electrodynamics, Transistors and amplifiers, Electrical testing, Biomedical instrumentation.

## Research Experience

**Facebook AI Research,** *Robotics Research Engineer* | Pittsburgh, PA and Fremont, CA *Apr 2022 - July 2024*

- **Research Focus:** Visual imitation learning for robot manipulation, Offline reinforcement learning, Mobile manipulation, Visual language models, Instance exploration and navigation
- Designed and constructed a distributed robotic arm cluster specifically tailored for large-scale experimentation and data collection. This cluster served as the hardware backbone for numerous impactful research projects.
- Researched a generalized universal agent proficient in solving challenging manipulation tasks within a kitchen-top environment. Our approach efficiently addressed over 38 tasks exhibiting 12 skills. By incorporating world priors from existing foundational models, we augmented the dataset, enhancing the agent's ability to generalize across different scenes. Outperformed prior methods by over 40%.
- Curated and released RoboSet, a scalable multi-task real-world robotic dataset for generalized offline learning.
- Conducted a comprehensive empirical investigation on leveraging pre-trained visual representations for training real-world task execution policies, uncovering insights about the transferability from simulation to real-world environments, achieving groundbreaking results in indoor navigation, and showcasing the advantages of data augmentation and fine-tuning for real-world performance improvement. Evaluated the performance of five different PVRs across two policy-learning paradigms (visual imitation and reinforcement learning) and on three distinct robotic platforms.
- Developed a mobile manipulation agent capable of autonomously planning and exploring environments, while simultaneously recording instances of objects of interest in a 3D map. Utilizing a visual-language model and 3D perception backbone, the agent can accurately identify and localize objects and their corresponding instances based on open-vocabulary requests. Subsequently, the robot can navigate to the specified object, pick it up, and then transport it to a target location for placement.

## Publications

**Dynamics-Compliant Trajectory Diffusion for Super-Nominal Payload Manipulation**

Anuj Pasricha, Joewie Koh, **Jay Vakil**, Alessandro Roncone

- Payload conditioned diffusion model that denoises joint angles, velocities, and torques to generate dynamic motions at 3x nominal payload.
- Accepted at **CoRL 2025**
- Website Paper

**Towards Open-World Mobile Manipulation in Homes: Lessons from the Neurips 2023 HomeRobot Open Vocabulary Mobile Manipulation Challenge**

Sriram Yenamandra*, Arun Ramachandran*, Mukul Khanna*, Karmesh Yadav*, **Jay Vakil***, Andrew Melnik, Michael Büttner, Leon Harz, Lyon Brown, Gora Chand Nandi, Arjun PS, Gaurav Kumar Yadav, Rahul Kala, Robert Haschke, Yang Luo, Jinxin Zhu, Yansen Han, Bingyi Lu, Xuan Gu, Qinyuan Liu, Yaping Zhao, Qiting Ye, Chenxiao Dou, Yansong Chua, Volodymyr Kuzma, Vladyslav Humennyy, Ruslan Partsey, Jonathan Francis, Devendra Singh Chaplot, Gunjan Chhablani, Alexander Clegg, Theophile Gervet, Vidhi Jain, Ram Ramrakhya, Andrew Szot, Austin Wang, Tsung-Yen Yang, Aaron Edsinger, Charlie Kemp, Binit Shah, Zsolt Kira, Dhruv Batra, Roozbeh Mottaghi, Yonatan Bisk, Chris Paxton

- Open Vocabulary Mobile Manipulation as a key benchmark task for robotics: finding any object in a novel environment and placing it on any receptacle surface within that environment.
- Website Paper

**Open X-Embodiment: Robotic Learning Datasets and RT-X Models**

Open X-Embodiment Collaboration

- The largest robot learning dataset with 1 million+ real-world robot trajectories over 22 robot embodiments.
- Accepted at **International Conference on Robotics and Automation (ICRA) 2024 (Best Conference Paper Award)**
- Website Paper

**OK-Robot: What Really Matters in Integrating Open-Knowledge Models for Robotics**

Peiqi Liu*, Yaswanth Orru*, **Jay Vakil**, Chris Paxton, Nur Muhammad "Mahi" Shafiullah†, Lerrel Pinto†

- A zero-shot integrated framework for open vocabulary mobile manipulation tasks by leveraging visual-language models and navigation primitives.
- Accepted at **RSS 2024**, **Robotic Tasks and How to Specify Them? (RSS 2024 workshop)**, **Workshop on Semantic Reasoning and Goal Understanding in Robotics (RSS 2024)**
- Website Paper

**RoboHive: A Unified Framework for Robot Learning**

Vikash Kumar, Rutav Shah, Gaoyue Zhou, Vincent Moens, Vittorio Caggiano, **Jay Vakil**, Abhishek Gupta, and Aravind Rajeswaran

- A comprehensive platform for Robot Learning and Embodied AI research, offering unified task interfaces, diverse environments, and expert benchmarks for algorithmic research in various domains.
- Accepted at **NeurIPS 2023**
- Website Paper

**RoboAgent: Towards Sample Efficient Robot Manipulation with Semantic Augmentations and Action Chunking**

Homanga Bharadhwaj*, **Jay Vakil***, Mohit Sharma*, Abhinav Gupta, Shubham Tulsiani, and Vikash Kumar

- An efficient end-to-end pipeline for training universal agents with multi-task manipulation skills using semantic augmentations and compact action representations, achieving superior performance and adaptability in a low-data regime.
- Accepted at **ICRA 2024**, **OOD Workshop at CoRL 2023 Outstanding Presentation Award at Robot Learning Workshop, NeurIPS 2023**
- Website Paper

**What do we learn from a large-scale study of pre-trained visual representations in sim and real environments?**

Sneha Silwal*, Karmesh Yadav*, Tingfan Wu*, **Jay Vakil***, Arjun Majumdar*, Sergio Arnaud*, Claire Chen, Vincent-Pierre Berges, Dhruv Batra, Aravind Rajeswaran, Mrinal Kalakrishnan, Franziska Meier, and Oleksandr Maksymets

- Extensive research on using pre-trained visual representations (PVRs) for real-world task execution, involving multiple PVRs, policy-learning methods, and robots. Identified key insights, including the correlation between simulation and real-world performance and the success of PVRs in achieving zero-shot indoor navigation.
- Accepted at **ICRA 2024**, and **MoMa workshop ICRA 2024**
- Website Paper

**Where are we in the search for an Artificial Visual Cortex for Embodied Intelligence?**

Arjun Majumdar, Karmesh Yadav, Sergio Arnaud, Yecheng Jason Ma, Claire Chen, Sneha Silwal, Aryan Jain, Vincent-Pierre Berges, Tingfan Wu, **Jay Vakil**, Pieter Abbeel, Jitendra Malik, Dhruv Batra, Yixin Lin, Oleksandr Maksymets, Aravind Rajeswaran, Franziska Meier

- Conducted the largest study on pre-trained visual representations (PVRs) for Embodied AI, creating CORTEXBENCH with 17 tasks. Demonstrated that data scaling doesn't universally improve performance, but task-specific adaptation of VC-1 models achieved competitive or superior results on all benchmarks.
- Accepted at **NeurIPS 2023**
- Website Paper

**SLAP: Spatial-Language Attention Policies**

Priyam Parashar, Vidhi Jain, Xiaohan Zhang, **Jay Vakil**, Sam Powers, and Chris Paxton

- Address challenges in language-guided mobile manipulation, achieving an 80% success rate in real-world tasks with a single model, and a 30% improvement over prior work in challenging conditions with limited data, demonstrating robust execution of open-vocabulary task plans.
- Accepted at **CoRL 2023**
- Website Paper

**Master Controller For High Energy Electron Source Part II**

**Jay Vakil**, Esayas Abera, Cyrus Safi, Wayne Kimura

- Designed a master controller system for a high-energy electron source for particle-accelerator based on diamond wafer technology. System consisted of complex remote controllers and monitoring systems.
- Undergraduate capstone research project
- Presentation

## Teaching Experience

**University of Colorado,** *Teaching assistant* | Boulder, CO *August 2024 - Present*

- Introduction to Artificial Intelligence (CSCI 3202)
- Computer systems (CSCI 2400)

## Select media coverage

- **Carnegie Mellon University** Parenting a 3-Year-Old Robot
- **TechCrunch** Human toddlers are inspiring new approaches to robot learning
- **IEEE Spectrum** Video Friday: A 3-Year-Old Robot Your weekly selection of awesome robot videos
- **Hackaday** ROBOAGENT GETS ITS MT-ACT TOGETHER
- **TechBrief** Meet RoboAgent: Enabling Robots to Acquire Manipulation Abilities
- **Communications of the ACM** Parenting a Three-Year-Old Robot
- **Venture Beat** Meta's OK-Robot performs zero-shot pick-and-drop in unseen environments
- **MIT Technology Review** This robot can tidy a room without any help
- **Fox News** Stop loading the dishwasher; this robot aims to do all the cleanup for you
- **AI at Meta** Robots that learn from videos of human activities and simulated interactions
- **Rocking Robots** 'RoboAgent' Modeled After Human Infants' Learning Techniques

## Academic service

### Reviewer

- IEEE Robotics and Automation Letters (RA-L
- 2025 Robotics: Science and Systems, RobotEvaluation Workshop (RSS 2025)
- 2025 International Conference on Machine Learning, Building Physically Plausible World Models (ICML 2025)
- 2025 IEEE International Conference on Robotics and Automation (ICRA 2025)
- 2024 IEEE International Conference on Robotics and Automation (ICRA 2024)
- 2024 Conference on Robot Learning (CORL 2024)

## Skills

| | |
|---|---|
| **Programming** | Python, C/C++, C#, Java, CUDA, Matlab, Bash, LaTeX, Vim, Linux, Tensorflow, Pytorch, Docker, OpenCV, AWS S3, Git |
| **Robotics/Hardware** | ROS 1/2, MuJoCo, Arduino, Raspberry PI, Sensor Interfacing, Signal processing, Digital/Analog circuit design and simulation, Microprocessors, Embedded systems, Circuit testing, Transistor-level design, CAD, Electrical circuit testing, Isaac Sim, Isaac Lab |