Gabriel Sarch

Princeton Artificial Intelligence
41 William Street
Princeton, NJ 08540, USA

Email: gabrielsarch@princeton.edu
Homepage: gabesarch.me
Google Scholar: Gabriel Sarch

Academic Appointments

Postdoctoral Research Fellow Sep 2025–Present
Princeton Language and Intelligence, Princeton University – Princeton, NJ

Education

Ph.D. in Machine Learning and Neural Computation (Joint Degree) 2020–2025
Carnegie Mellon University – Pittsburgh, PA
Advisors: Katerina Fragkiadaki and Michael J. Tarr.

M.S. in Machine Learning Research 2023
Carnegie Mellon University – Pittsburgh, PA

B.S. in Biomedical Engineering, *magna cum laude* 2016–2020
University of Rochester – Rochester, NY

Research Interests

machine learning; computer vision; reinforcement learning; cognitive science; perception, reasoning, & memory

Publications

[*]Equal contribution

Preprints & Under Review

[P1] **Gabriel Sarch**[*], Declan Campbell[*], Nicholas Budny[*], Kia Ghods, Katerina Fragkiadaki, Jonathan D. Cohen, Brenden Lake, and Taylor Webb. Learning Grounded Reasoning from Human Gaze. *arXiv:2611.23209*. *In submission*.

[P2] **Gabriel Sarch**, Haoyang Wu, Yipeng Zhang, Katerina Fragkiadaki, Danqi Chen, Zhuang Liu. Learning to Persist in Visual Exploration. *In submission*.

[P3] Xueyang Yu, Yifan Jing, Taiming Lu, Linrong Cai, Zhipeng Cai, Yunyang Xiong, **Gabriel Herbert Sarch**, and Zhuang Liu. Can Image Generation Benefit Visual Reasoning? *In submission*.

[P4] Advait Gupta, Sachin Konan, Tom Yin, **Gabriel Sarch**[*], and Zhuang Liu[*]. AgentVQA: A Unified Benchmark for Agentic Visual Understanding. *arXiv:2610.31494*. *In submission*.

[P5] Boya Zeng, Tianze Luo, Shu Pu, Jucheng Shen, Taiming Lu, **Gabriel Sarch**, and Zhuang Liu. i1: A Simple and Fully Open Recipe for Strong Text-to-Image Models. *arXiv:2606.11289*. *In submission*.

[P6] Yida Yin, Harish Krishnakumar, Chung Peng Lee, Boya Zeng, Wenhao Chai, Shengbang Tong, Wenhu Chen, Hu Xu, Xingyu Fu, **Gabriel Herbert Sarch**, Aleksandra Korolova, and Zhuang Liu. WorldBench: A Challenging and Visually Diverse Multimodal Reasoning Benchmark. *arXiv:2606.06538*. *In submission*.

Conference & Journal Publications

[1] **Gabriel Herbert Sarch**[*], Linrong Cai[*], Qunzhong Wang, Haoyang Wu, Danqi Chen, and Zhuang Liu. Vero: An Open RL Recipe for General Visual Reasoning. *European Conference on Computer Vision (ECCV), 2026*. **Oral at ECCV 2026 UniWorld and CVPR 2026 DataMFM workshops. 300k+ all-time downloads**.

[2] Chengshuai Shi[*], Wenzhe Li[*], Xinran Liang[*], Yizhou Lu, Wenjia Yang, Ruirong Feng, Seth Karten, Ziran Yang, Zihan Ding, **Gabriel Herbert Sarch**, Danqi Chen, Karthik Narasimhan, and Chi Jin. Odysseus: Scaling VLMs to 100+ Turn Decision-Making in Games via Reinforcement Learning. *Conference on Language Modeling (COLM), 2026*.

[3] **Gabriel Herbert Sarch**, Snigdha Saha, Naitik Khandelwal, Ayush Jain, Michael J. Tarr, Aviral Kumar, and Katerina Fragkiadaki. Grounded Reinforcement Learning for Visual Reasoning. *Advances in Neural Information Processing Systems (NeurIPS), 2025*.

[4] Jacob Yeung, Andrew F. Luo, **Gabriel Herbert Sarch**, Margaret M. Henderson, Deva Ramanan, and Michael J. Tarr. Reanimating Images using Neural Representations of Dynamic Stimuli. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2025*. **Oral (Top 1%)**.

[5] Sahithya Ravi, **Gabriel Herbert Sarch**, Vibhav Vineet, Andrew D. Wilson, and Balasaravanan Thoravi Kumaravel. Out of Sight, Not Out of Context? Egocentric Spatial Reasoning in VLMs Across Disjoint Frames. *Conference on Empirical Methods in Natural Language Processing (EMNLP), 2025*.

[6] **Gabriel Herbert Sarch**, Balasaravanan Thoravi Kumaravel, Sahithya Ravi, Vibhav Vineet, and Andrew D. Wilson. Grounding Task Assistance with Multimodal Cues from a Single Demonstration. *Findings of the Association for Computational Linguistics: ACL, 2025*.

[7] **Gabriel Herbert Sarch**, Lawrence Jang, Michael J. Tarr, William W. Cohen, Kenneth Marino, and Katerina Fragkiadaki. VLM Agents Generate Their Own Memories: Distilling Experience into Embodied Programs of Thought. *Advances in Neural Information Processing Systems (NeurIPS), 2024*. **Spotlight (Top 2%)**.

[8] Ayush Jain, Pushkal Katara, Nikolaos Gkanatsios, Adam W. Harley, **Gabriel Herbert Sarch**, Kriti Aggarwal, Vishrav Chaudhary, and Katerina Fragkiadaki. ODIN: A Single Model for 2D and 3D Segmentation. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024*. **Highlight (Top 3%)**.

[9] **Gabriel Herbert Sarch**, Yue Wu, Michael J. Tarr, and Katerina Fragkiadaki. Open-Ended Instructable Embodied Agents with Memory-Augmented Large Language Models. *Findings of the Association for Computational Linguistics: EMNLP, 2023*.

[10] **Gabriel Herbert Sarch**, Michael J. Tarr, Katerina Fragkiadaki, and Leila Wehbe. Brain Dissection: fMRI-trained Networks Reveal Spatial Selectivity in the Processing of Natural Images. *Advances in Neural Information Processing Systems (NeurIPS), 2023*.

[11] **Gabriel Herbert Sarch**, Hsiao-Yu Fish Tung, Aria Y. Wang, Jacob S. Prince, and Michael J. Tarr. 3D View Prediction Models of the Dorsal Visual Stream. *Conference on Cognitive Computational Neuroscience (CCN), 2023*.

[12] Jacob L. Yates, Shanna H. Coop, **Gabriel Herbert Sarch**, Ruei-Jr Wu, Daniel A. Butts, Michele Rucci, and Jude F. Mitchell. Detailed Characterization of Neural Selectivity in Free Viewing Primates. *Nature Communications, 2023*.

[13] **Gabriel Herbert Sarch**, Zhaoyuan Fang, Adam W. Harley, Paul Schydlo, Michael J. Tarr, Saurabh Gupta, and Katerina Fragkiadaki. TIDEE: Tidying Up Novel Rooms using Visuo-Semantic Commonsense Priors. *European Conference on Computer Vision (ECCV), 2022*. **Winner, CVPR 2023 Embodied AI Rearrangement Challenge**.

[14] Zhaoyuan Fang*, Ayush Jain*, **Gabriel Herbert Sarch***, Adam W. Harley, and Katerina Fragkiadaki. Move to See Better: Self-Improving Embodied Object Detection. *British Machine Vision Conference (BMVC), 2021*.

Honors, Fellowships & Awards

National Science Foundation Graduate Research Fellowship (NSF-GRFP) 2020–2025
Supported three years of graduate research (top 15%).

NeurIPS Spotlight Paper 2024
Selected among the top 2% of NeurIPS 2024 submissions.

Outstanding Reviewer at CVPR 2026
Review quality placed among the top 5% of reviewers

Winner, Embodied AI Workshop Rearrangement Challenge at CVPR 2023
Winning model submission of object rearrangement robotics challenge at CVPR 2023.

Runner-Up, Amazon Alexa Prize SimBot Embodied Dialogue Challenge 2022
Runner-up among 10 university teams building multimodal agents.

Biomedical Engineering Award for Academic Excellence 2020
Awarded to the highest GPA student in the University of Rochester BME graduating class.

University of Rochester Center for Visual Science Research Fellowship 2019
One-year grant for research in visual neuroscience at the University of Rochester.

National Institutes of Health Ruth L. Kirschstein National Research Service Award 2018
Supported a fully funded summer of research at the National Institutes of Health.

Selected Research & Industry Experience

Technical Staff, AI Aug 2024–Dec 2024
Yutori – San Francisco, CA
Built distributed training infrastructure and data pipelines for multimodal language models supporting autonomous web agents. Founders: Devi Parikh, Dhruv Batra, Abhishek Das.

Research Intern May 2024–Aug 2024
Microsoft Research
Led development of a real-time assistance system that uses gaze, speech, and hand gestures to personalize task guidance. Mentors: Andy Wilson, Bala Kumaravel, and Vibhav Vineet.

Research Assistant Aug 2018–Aug 2020
Active Vision Laboratory, University of Rochester
Investigated neural correlates of eye movements and visual attention in primates. PI: Jude Mitchell.

Research Intern May 2018–Aug 2018
Neurobiology Laboratory, National Institutes of Health
Studied computational models of the pre-Bötzinger respiratory generator. PI: Jeffrey Smith.

Invited & Selected Talks

Vero: An Open RL Recipe for General Visual Reasoning 2026
ECCV 2026 Workshop on Universal Representations for Perception, Reasoning, and World Modeling. **Oral**.

An Open Reinforcement Learning Recipe for General Visual Reasoning 2026
CVPR 2026 Workshop on Data for Multimodal Foundation Models. **Oral**.

Grounded Reinforcement Learning for Visual Reasoning 2025
Princeton Language and Intelligence Seminar, Princeton University.

Task Planning with Large Language Models 2024
Search-Based Planning Laboratory, Carnegie Mellon University.

Open-Ended Embodied Agents with Memory-Augmented Large Language Models 2024
CMU Catalyst LLM Agents Seminar, Carnegie Mellon University.

Spatial Processing During Natural Viewing: Insights from Neural Network Modeling 2023
brAIn Seminar, Carnegie Mellon University.

Teaching Experience

Guest Lecturer, Introduction to Machine Learning (COS 324) Spring 2026
Princeton University.

Guest Lecturer, Frontiers in Deep Learning (COS 597K) Fall 2025
Princeton University.

Teaching Assistant, Deep Reinforcement Learning (10-403) Spring 2023
Carnegie Mellon University.

Teaching Assistant and Lecturer, Biologically Intelligent Exploration (85-435) Spring 2023
Carnegie Mellon University.

Research Mentoring

Graduate Research Mentoring

Boya Zeng	Research Mentor	Ph.D., Princeton University
Yida Yin	Research Mentor	Ph.D., Princeton University
Linrong Cai	Research Mentor	M.S.E., Princeton University
Lawrence Jang	Research Mentor	M.S., Carnegie Mellon University → Ph.D., Carnegie Mellon University
Aditya Kumar	Research Mentor	M.S., Carnegie Mellon University → Ph.D., Carnegie Mellon University
Snigdha Saha	Research Mentor	M.S., Carnegie Mellon University → Software Engineer, Google
Advait Gupta	Research Mentor	M.S., University of Maryland → Ph.D., UC Riverside
Naitik Khandelwal	Research Mentor	M.S., Carnegie Mellon University → Software Engineer, PlayStation

Sahil Somani	Research Mentor	M.S., Carnegie Mellon University → Software Engineer, Google
Raghav Kapoor	Research Mentor	M.S., Carnegie Mellon University → Software Engineer, Adobe

Undergraduate Research Mentoring

Emily Luo	Senior Thesis Co-Advisor	B.S., Princeton University
Yen-Ru Chen	Research Mentor	B.S., University of California, San Diego
Haoyang Wu	Research Mentor	B.S., University of Michigan
Joshua Wolford	Research Mentor	B.S., Princeton University
Yipeng Zhang	Research Mentor	B.S., Xi'an Jiaotong University
Yifan Liu	Research Mentor	B.S., Tsinghua University → Ph.D., Stanford University

Professional Service & Leadership

AI Office Hours 2026–Present
Host free, weekly online Q&A on AI, research, and careers, open to all backgrounds.

Conference and Workshop Organization 2026
Co-Organizer, *CogVL: Cognitive Foundations for Multimodal Models*, CVPR.

Conference Area Chair 2026–2027
ICLR (2027).

Conference Reviewing 2023–2026
ICLR (2024–2026); CVPR (2023–2026); NeurIPS (2023–2026); ICML (2024–2026); AISTATS (2024–2025, Program Committee 2025); ECCV (2025–2026); ACL (2024–2025).

Departmental Service 2024–2025
Member, CMU Machine Learning Department Ph.D. Peers Committee; supported incoming Ph.D. students.

Outreach
Fireside-chat interviewer for Bobyard founder Michael Ding, Princeton AI Lab. 2026

Volunteer with UrbanPromise Trenton's afterschool program, leading academic activities for local youth. 2026–Present