

# Akshay K. Jagadish

Postdoctoral Research Fellow, Princeton AI Lab

📍 Princeton, USA | ✉ akshaykjadish@gmail.com | 🏠 akjadish.github.io | 📷 akjadish | 🐦 akjadish | 🎓 Scholar

## Education

### Postdoctoral Research Fellow

Princeton University  
Natural and Artificial Minds (NAM) Initiative, Princeton AI Lab

Princeton, USA  
2025 - ongoing

### Ph.D. in Computer Science

University of Tübingen · Max Planck Institute for Biological Cybernetics · Helmholtz Munich  
Thesis: Meta-Learning: A Unifying Framework for Testing Theories of Human Learning

Tübingen, Germany  
2021 - 2025

### M.Sc. in Computational Neuroscience

University of Tübingen · Max Planck Institute for Biological Cybernetics  
Thesis: Compositional Reinforcement Learning in Minds and Machines

Tübingen, Germany  
2018 - 2020

### B.Tech. in Electrical and Electronics Engineering

National Institute of Technology Karnataka · Ecole Polytechnique Federale de Lausanne  
Thesis: Structural and Functional Correlates of Personality

Surathkal, India  
2013 - 2017

## Research Experience

### Ph.D. Thesis

University of Tübingen · Max Planck Institute for Biological Cybernetics · Helmholtz Munich  
Computational Principles of Intelligence Lab · Advisors: Dr. Marcel Binz and Prof. Eric Schulz

Tübingen, Germany  
Apr. 2021 - Aug. 2025

#### Natural Minds

- Proposed a new framework for building computational cognitive models termed as “Meta-learned models of cognition” [PDF].
- Demonstrated that resource rational meta-learned inference explains zero-shot compositional reasoning in humans [PDF].
- Illustrated that ecologically rational meta-learned inference explains human category learning [PDF].
- Showed that humans and large language models display symmetric belief updating [PDF].
- Developed bounded ecologically rational meta-learned inference to explore the role of bounded resources and ecological priors in human learning [PDF].
- Demonstrated that human learning and decision making in 15 different experiments can be explained by meta-learning ecological priors from large language models [PDF].

#### Artificial Minds

- Illustrated that inducing anxiety in large language models increases exploration and bias [PDF].
- Demonstrated that narrative of traumatic experiences increases state anxiety in large language models, but using mindfulness-based techniques can help alleviate the same [PDF].
- Used sparse autoencoders to reveal temporal difference learning in large language models [PDF].

#### Automated scientific discovery

- Contributed towards building the first foundation model of human cognition [PDF].
- Generated cognitive models with large language models [PDF].
- Automated scientific minimization of regret to discover cognitive models [PDF].

### Masters Thesis

Max Planck Institute for Biological Cybernetics  
Computational Principles of Intelligence Lab · Advisors: Dr. Marcel Binz and Prof. Eric Schulz

Tübingen, Germany  
Jun. 2020 - Dec. 2020

- Built meta-reinforcement learning agents that can discover and compose latent structures underlying rewards of structured bandit tasks, and further, generalize to novel structures unseen during training.

### Graduate Research Assistant

University of Tübingen  
Sinz Lab · Advisors: Prof. Fabian Sinz and Prof. Edgar Walker

Tübingen, Germany  
Nov. 2018 - Mar. 2021

- Built a factor analysis model on top of a convolutional neural network, which predicts stimulus-based neural activity, to recover non-stimulus-related latent brain states [PDF].
- Developed a novel parameter-efficient readout, called a Gaussian readout, that maps nonlinear features learned by the deep convolutional network to the response of each neuron [PDF].

## Graduate Research Assistant

Max Planck Institute for Biological Cybernetics

Computational Neuroscience Lab · Advisor: Prof. Peter Dayan

- Conducted a literature review on the role of Dopamine in reward-based learning [PDF].
- Analyzed behavior (choices and reaction times) of monkeys, whose dopamine receptors were pharmacologically stimulated, during a rule-based categorization task [PDF].

Tübingen, Germany

Nov. 2019 - Feb. 2020

## AI Researcher

Wadhvani Institute for Artificial Intelligence

AI for Social Impact · Advisor: Dr. Rahul Panicker

- Developed a model based on deep-learning that predicts weight of an object from its images [NDA].

Mumbai, India

May 2018 - Sep. 2018

## Postbaccalaureate Research Assistant

University of Minnesota, Twin-cities

Computational Visual Neuroscience Lab · Advisor: Prof. Kendrick Kay

- Developed a generative model that factors in the bottom-up, stimulus-driven, and top-down, goal-driven attentional state to characterize cortical fMRI responses for various stimuli and attentional loci combinations.

Minnesota, USA

Jul. 2017 - Feb. 2018

## Undergraduate Thesis

Ecole Polytechnique Federale de Lausanne

Medical Image Processing Lab · Advisors: Prof. D. van de Ville and Prof. P. Giannakopoulos

- Investigated the relationship between structural and functional connectivity measures derived from MRI, and Neuroticism Extroversion Openness Personality Inventory-Revised (NEOPI) personality traits [PDF].

Lausanne, Switzerland

Aug. 2016 - May 2017

## Undergraduate Research Assistant

Indian Institute of Science

Computational Tomography Lab · Advisor: Prof. Kasi Rajgopal

- Developed an algorithm, called k-ABC, based on the artificial bee colony algorithm to come up with an optimal variable density sampling scheme for the compressed sensing-based reconstruction of Magnetic Resonance (MR) images [PDF].

Bangalore, India

May 2015 - May 2017

## Selected Publications

---

\* equal contribution, # alphabetical ordering

**Jagadish, A. K.**, Binz, M. & Schulz, E. (2025). Meta-learning ecological priors from large language models captures human learning and decision making. *In Preparation*.

Rmus, M., **Jagadish, A. K.**\*, Mathony, M., & Schulz, E. (2025). Generating Computational Cognitive Models using Large Language Models. *Under review* [PDF]

Binz, M., ..., **Jagadish, A. K.**#, ..., & Schulz, E. (2024). Centaur: a foundation model of human cognition. *Nature* [PDF]

Demircan, C. \*, Saanum, T. \*, **Jagadish, A. K.**, Binz, M., & Schulz, E. (2025). Sparse Autoencoders Reveal Temporal Difference Learning in Large Language Models. *In the 13th International Conference on Learning Representations (ICLR)* [PDF]

**Jagadish, A. K.**, Thalmann, M., Coda-Forno, J., Schulz, E., & Binz, M. (2024). Human-like Category Learning by Injecting Ecological Priors from Large Language Models into Neural Networks. *Proceedings of the 41st International Conference on Machine Learning (ICML)* [PDF]

Schubert, J., **Jagadish, A. K.**, Binz, M., & Schulz, E. (2024). In-context learning agents are asymmetric belief updaters. *Proceedings of the 41st International Conference on Machine Learning (ICML)* [PDF]

**Jagadish, A. K.**, Binz, M., Saanum, T., Wang, J. X., & Schulz, E. (2024). Zero-shot compositional reasoning in a reinforcement learning setting. [PDF]

Ben-Zion, Z., Witte, K. \*, **Jagadish, A. K.**\*, Duek, O., Harpaz-Rotem, I., Khorsandian, M., ... & Spiller, T. R. (2024). "Chat-GPT on the Couch": Assessing and Alleviating State Anxiety in Large Language Models. *npj Digital Medicine* [PDF]

Coda-Forno, J. \*, Witte, K. \*, **Jagadish, A. K.**, Binz, M., Akata, Z., & Schulz, E. (2023). Inducing anxiety in large language models increases exploration and bias. [PDF];

Binz, M., Dasgupta, I., **Jagadish, A. K.**, Botvinick, M., Wang, J.X., & Schulz, E. (2023). Meta-learned models of cognition. *Behavioral and Brain Sciences*, 1-38. [PDF]

Bashiri, M. \*, Walker, E. \*, Lurz, K. K., **Jagadish, A. K.**, Muhammad, T., Ding, Z., ... & Sinz, F. (2021). A flow-based latent state generative model of neural population responses to natural images. *In Advances in Neural Information Processing Systems (NeurIPS)*, 34, 15801-15815. [PDF]

Lurz, K. K., Bashiri, M., Willeke, K. F., **Jagadish, A. K.**, Wang, E., Walker, E. Y., ... & Sinz, F. (2021). Generalization in data-driven models of primary visual cortex. *In International Conference on Learning Representations (ICLR)* [PDF]

## Honors, Awards, & Fellowships

2024	<b>Top Reviewer</b> , Top 10 % reviewer (main track) for Neural Information Processing Systems (NeurIPS)	Canada
2024	<b>ELLIS Winter School on Foundation Models</b> , Top 40 students selected from 607 to present their research	Netherlands
2023	<b>Analytical Connectionism</b> , Top 35 students selected to attend the summer course at the Gatsby Computational Neuroscience Unit in University College London	UK
2020	<b>SMARTSTART Fellowship in Computational Neuroscience</b> , Top 15 students awarded a travel budget of 1000 euros and mentorship from Prof. Peter Dayan and Prof. Fabian Sinz	Germany
2019	<b>Dean's List</b> , Top 3 performers in the summer semester for M.Sc. program in Computational Neuroscience	Germany
2018	<b>Dean's List</b> , Top 3 performers in the winter semester for M.Sc. program in Computational Neuroscience	Germany
2018	<b>Max Planck Society Scholarship</b> , Top 5 students selected to undertake M.Sc in Computational Neuroscience at the University of Tübingen	Germany
2017	<b>Harvard Young Scientist Development Program</b> , Top 25 students selected for training in neuroscience	USA   India
2016	<b>Summer Research Program</b> , Top 20 students funded to conduct research at EPFL	Switzerland
2016	<b>Summer Research Fellowship Program</b> , Top 10 % students funded to conduct research at the Indian Institute of Science	India
2013	<b>Ranked Top 0.1%</b> , Karnataka Common Entrance Test among 150,000 students	India
2013	<b>Ranked Top 0.1%</b> , COMED-K among 200,000 students	India
2011	<b>Most Consistent Performer of the Batch</b> , High school at Presidency School	India

## Press

2025	<b>New York Times</b> , A science-education piece called "Digital Therapists Get Stressed Too, Study Finds" on our article "Chat-GPT on the Couch": Assessing and Alleviating State Anxiety in Large Language Models [URL]	AoE
2025	<b>Fortune Magazine</b> , Interview to discuss our article "Chat-GPT on the Couch": Assessing and Alleviating State Anxiety in Large Language Models [URL]	AoE
2025	<b>ScienceDaily</b> , Accessible take on our article "Chat-GPT on the Couch": Assessing and Alleviating State Anxiety in Large Language Models [URL]	AoE

## Invited talks

2024	<b>Annual Meeting of the Cognitive Science Society (CogSci)</b> , Workshop on "In-context learning in natural and artificial intelligence"	Netherlands
2024	<b>Annual Meeting of the Cognitive Science Society (CogSci)</b> , Workshop on "Compositionality in minds, brains and machines: A unifying goal that cuts across cognitive sciences"	Netherlands
2024	<b>Indian Institute of Science</b> , Seminar talk at the Center for Neuroscience	India
2024	<b>Princeton University</b> , Lab meeting at Computational Cognitive Science Lab	USA
2023	<b>Helmholtz Center in Munich</b> , Joint lab retreat with Explainable Machine Learning Lab	Germany
2023	<b>University of Oxford</b> , Lab meeting at the Human Information Processing Lab	United Kingdom
2022	<b>Max Planck Institute for Human Cognitive and Brain Sciences</b> , Joint lab retreat with Doeller Lab	Germany
2021	<b>Stanford University</b> , Joint lab retreat with Human Information Processing and Causality in Cognition Lab	USA
2019	<b>University of Tübingen</b> , Workshop on "Causality in Neuroscience" at Neuroscience Conference for Young Scientists	Germany

## Organization

2024	<b>Co-organizer</b> , "Connecting Minds and Machines a Foundation Model Approach to Learning" symposium at Helmholtz Pioneer Campus, Munich	Germany
2024	<b>Co-organizer</b> , "In-context learning in natural and artificial intelligence" workshop at CogSci-2024	Netherlands
2023	<b>Co-organizer</b> , Laboratory Retreat of Computational Principles of Intelligence Lab	Germany
2022	<b>Co-organizer</b> , Computation and Cognitive Tübingen Summer School (CaCTüs) aimed specifically at young scientists held back by personal, financial, regional or societal constraints.	Germany
2019	<b>Co-organizer</b> , "Causality in Neuroscience" Workshop at Neuroscience Conference for Young Scientists	Germany
2020	<b>Volunteer</b> , Machine Learning Summer School (MLSS) held at MPI for Intelligent Systems, Tübingen	Germany

## Supervision

---

- 2024 **Co-supervisor (Research Assistant)**, Elif Kara on “Human decision-making in the wild” at the University of Munich *Germany*
- 2023 **Co-supervisor (Master Thesis)**, Johannes Schubert on “Rational Analysis of Optimism Bias” at the University of Tübingen. Converted to a publication at the *International Conference on Machine Learning (ICML)* [PDF] *Germany*

## Teaching

---

- 2023 **Teaching Assistant**, “Computational Cognitive Science” course at the Graduate Training Center for Neuroscience, University of Tübingen *Germany*
- 2022 **Tutor**, Tutorial on “Meta-Reinforcement Learning” at the Max Planck Institute for Biological Cybernetics *Germany*

## Service

---

- 2025- **Reviewer**, Transactions on Machine Learning Research (TMLR) *AoE*
- 2024- **Reviewer**, Neural Information Processing Systems (NeurIPS) *AoE*
- 2023- **Reviewer**, International Conference on Learning Representations (ICLR) *AoE*
- 2022- **Reviewer**, Annual Meeting of the Cognitive Science Society (CogSci) *AoE*
- 2023- **Reviewer**, Cognitive Computational Neuroscience (CCN) *AoE*