Akshay K. Jagadish

Postdoctoral Research Fellow, Princeton AI Lab

Princeton, USA | akshaykjagadish@gmail.com | akjagadish.github.io | akjagadish | akjagadish | Scholar

Education

Research Fellow — *Princeton, USA*
Princeton AI Lab — *2025 - ongoing*
Independent researcher part of the Natural and Artificial Minds Program

Ph.D. in Computer Science — *Tübingen, Germany*
University of Tübingen · Max Planck Institute for Biological Cybernetics · Helmholtz Munich — *2021 - 2025*
Thesis: Meta-Learning: A Unifying Framework for Testing Theories of Human Learning

M.Sc. in Computational Neuroscience — *Tübingen, Germany*
University of Tübingen · Max Planck Institute for Biological Cybernetics — *2018 - 2020*
Thesis: Compositional Reinforcement Learning in Minds and Machines

B.Tech. in Electrical and Electronics Engineering — *Surathkal, India*
National Institute of Technology Karnataka · Ecole Polytechnique Federale de Lausanne — *2013 - 2017*
Thesis: Structural and Functional Correlates of Personality

Research Experience

Ph.D. Thesis — *Tübingen, Germany*
University of Tübingen · Max Planck Institute for Biological Cybernetics · Helmholtz Munich — *Apr. 2021 - Aug. 2025*
Computational Principles of Intelligence Lab · Advisors: Dr. Marcel Binz and Prof. Eric Schulz

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 mdoels [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 mdels [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 — *Tübingen, Germany*
Max Planck Institute for Biological Cybernetics — *Jun. 2020 - Dec. 2020*
Computational Principles of Intelligence Lab · Advisors: Dr. Marcel Binz and Prof. Eric Schulz

- 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 — *Tübingen, Germany*
University of Tübingen — *Nov. 2018 - Mar. 2021*
Sinz Lab · Advisors: Prof. Fabian Sinz and Prof. Edgar Walker

- 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 *Tübingen, Germany*

Max Planck Institute for Biological Cybernetics *Nov. 2019 - Feb. 2020*

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].

AI Researcher *Mumbai, India*

Wadhwani Institute for Artificial Intelligence *May 2018 - Sep. 2018*

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].

Undergraduate Thesis *Lausanne, Switzerland*

Ecole Polytechnique Federale de Lausanne *Aug. 2016 - May 2017*

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].

Undergraduate Research Assistant *Bangalore, India*

Indian Institute of Science *May 2015 - May 2017*

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].

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. *Under Review at PNAS.*

Rmus, M., **Jagadish, A. K.***, Mathony, M., & Schulz, E. (2025). Generating Computational Cognitive Models using Large Language Models. *In Advances in Neural Information Processing Systems (NeurIPS)* [PDF]

Binz, M., **Jagadish, A. K.**, Rmus, M., & Schulz, E. (2025). Automated scientific minimization of regret. *In the AI4Science Workshop at the Advances in Neural Information Processing Systems (NeurIPS)* [PDF]

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

Demirican, 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]

Rodriguez, C.*, **Jagadish, A. K.***, Meskaldji, D. E., Haller, S., Herrmann, F., Van De Ville, D., & Giannakopoulos, P. (2019). Structural correlates of personality dimensions in healthy aging and MCI. *Frontiers in psychology, 9, 2652.* [PDF]

Honors, Awards, & Fellowships

2025 **Top Reviewer**, Top 10 % reviewer (main track) for Neural Information Processing Systems (NeurIPS) *USA*

2025 **Natural and Artificial Mind Fellowship**, Top 5 candidates with a fellowship to conduct independent research at the Princeton AI Lab *USA*

2025 **Ph.D. with Magna cum Laude**, for exceptional research conducted during doctoral studies *Germany*

2024 **Top Reviewer**, Top 10 % reviewer (main track) for Neural Information Processing Systems (NeurIPS) *Canada*

2024 **ELLIS Winter School on Foundation Models**, Top 40 students selected from 607 to present their research *Netherlands*

2023 **Analytical Connectionism**, Top 35 students selected to attend the summer course at the Gatsby Computational Neuroscience Unit in University College London *UK*

2020 **SMARTSTART Fellowship in Computational Neuroscience**, Top 15 students awarded a travel budget of 1000 euros and mentorship from Prof. Peter Dayan and Prof. Fabian Sinz *Germany*

2019 **Dean's List**, Top 3 performers in the summer semester for M.Sc. program in Computational Neuroscience *Germany*

2018 **Dean's List**, Top 3 performers in the winter semester for M.Sc. program in Computational Neuroscience *Germany*

2018 **Max Planck Society Scholarship**, Top 5 students selected to undertake M.Sc in Computational Neuroscience at the University of Tübingen *Germany*

2017 **Harvard Young Scientist Development Program**, Top 25 students selected for training in neuroscience *USA | India*

2016 **Summer Research Program**, Top 20 students funded to conduct research at EPFL *Switzerland*

2016 **Summer Research Fellowship Program**, Top 10 % students funded to conduct research at the Indian Institute of Science *India*

2013 **Ranked Top 0.1%**, Karnataka Common Entrance Test among 150,000 students *India*

2013 **Ranked Top 0.1%**, COMED-K among 200,000 students *India*

2011 **Most Consistent Performer of the Batch**, High school at Presidency School *India*

Press

2025 **New York Times**, A science-education piece called *"Scientists Use A.I. to Mimic the Mind, Warts and All"* on our article A foundation model to predict and capture human cognition [URL] *AoE*

2025 **New York Times**, 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 **Fortune Magazine**, Interview to discuss our article "Chat-GPT on the Couch": Assessing and Alleviating State Anxiety in Large Language Models [URL] *AoE*

2025 **ScienceDaily**, Accessible take on our article "Chat-GPT on the Couch": Assessing and Alleviating State Anxiety in Large Language Models [URL] *AoE*

2025 **Tagesspiegel**, An accessible piece in German called *"How does ChatGPT "think"?: A chatbot goes to a psychologist"* on our article "Chat-GPT on the Couch": Assessing and Alleviating State Anxiety in Large Language Models [URL] *AoE*

Supervision

2025 **Doctoral Student**, Liyi Zhang's doctoral research project on "Distilling inductive reasoning into LLMs" at the Princeton University *USA*

2025 **Doctoral Student**, Solim Legris's doctoral research project on "Neuro-symbolic models for ARC benchmark" at the Princeton University *USA*

2025 **Doctoral Student**, Younes Strittmatter's doctoral research project on "Tools for automated scientific discovery" at the Princeton University *USA*

2025 **Master Student**, Daniel Braga's master research project on "Automated representational structure discovery" at the Princeton University *USA*

2024 **Master Student**, Elif Kara's master research project on "Human decision-making in the wild" at the University of Münich *Germany*

2023 **Master Student**, Johannes Schubert's master thesis 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*

Invited talks

2025	**Cornell-CUNY-UC Davis**, Cross-lab meeting part of an NSF grant on ”Collective behavior in smart environments”.	*USA*
2025	**New York University**, NYUConcat talk series in the department of psychology at NYU	*USA*
2025	**Princeton University**, Natural and Artificial Minds monthly talk series in the the Princeton AI Lab	*USA*
2025	**Princeton University**, Workshop on Automated Discovery of Mind and Brain	*USA*
2025	**University of Osnabrück**, Lab meeting at the Laboratory for Automated Scientific Discovery of Mind and Brain	*Germany*
2024	**Annual Meeting of the Cognitive Science Society (CogSci)**, Workshop on “In-context learning in natural and artificial intelligence”	*Netherlands*
2024	**Annual Meeting of the Cognitive Science Society (CogSci)**, Workshop on “Compositionality in minds, brains and machines: A unifying goal that cuts across cognitive sciences”	*Netherlands*
2024	**Indian Institute of Science**, Seminar talk at the Center for Neuroscience	*India*
2024	**Princeton University**, Lab meeting at Computational Cognitive Science Lab	*USA*
2023	**Helmholtz Münich**, Joint lab retreat with Explainable Machine Learning Lab	*Germany*
2023	**University of Oxford**, Lab meeting at the Human Information Processing Lab	*United Kingdom*
2022	**Max Planck Institute for Human Cognitive and Brain Sciences**, Joint lab retreat with Doeller Lab	*Germany*
2021	**Stanford University**, Joint lab retreat with Human Information Processing and Causality in Cognition Lab	*USA*
2019	**University of Tübingen**, Workshop on “Causality in Neuroscience” at Neuroscience Conference for Young Scientists	*Germany*

Organization

2024	**Co-organizer**, “Connecting Minds and Machines a Foundation Model Approach to Learning” symposium at Helmholtz Pioneer Campus, Munich	*Germany*
2024	**Co-organizer**, “In-context learning in natural and artificial intelligence” workshop at CogSci-2024	*Netherlands*
2023	**Co-organizer**, Laboratory Retreat of Computational Principles of Intelligence Lab	*Germany*
2022	**Co-organizer**, Computation and Cognitive Tübingen Summer School (CaCTüS) aimed specifically at young scientists held back by personal, financial, regional or societal constraints.	*Germany*
2019	**Co-organizer**, “Causality in Neuroscience” Workshop at Neuroscience Conference for Young Scientists	*Germany*
2020	**Volunteer**, Machine Learning Summer School (MLSS) held at MPI for Intelligent Systems, Tübingen	*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*
2025-	**Reviewer**, International Conference on Machine Learning (ICML)	*AoE*
2022-	**Reviewer**, Annual Meeting of the Cognitive Science Society (CogSci)	*AoE*
2023-	**Reviewer**, Cognitive Computational Neuroscience (CCN)	*AoE*