+1 (917) 250-2368,
+919985971997
Linkedin: siddsuresh97

Siddharth Suresh

sidd.suresh97@gmail.com
siddharth.suresh@wisc.edu
Github:siddsuresh97
https://sidsuresh.com

Employment

Applied Scientist Intern — **Amazon AGI** — **June 2024 – September 2024**

- Building better recommendation systems using LLMs.
- Alignment using SFT
- LLM as a Judge

Graduate Research Assistant — **University of Wisconsin-Madison** — **Fall 2021-present**

Knowledge and Concepts lab - Advised by Dr. Timothy T Rogers (Fall 2022 - present)

- Worked on contrasting the differences in conceptual structure of humans and Large Language models and also aligning them with Humans. Fine-tuned LLMs using PEFT techniques like LORA. Two papers were accepted at the ICLR 2023 Tiny Paper Track. Long paper accepted at the main track EMNLP 2023. Poster presented at Cogsci 2023 in Sydney.
- Comparing differences in FT vs ICL for compositional tasks in LLMs. Exploring how LLM hidden representations capture task vectors.
- Mechanistic interpretability in LLMs by doing causal interventions in attention-heads.
- Worked on using LLM agents to simulate human opinion- dynamic (wisdom of partisan crowds). Two papers accepted at ICLR 2024 workshop on LLM agents.
- Working on building human-aligned vision models by leveraging human semantic knowledge. Trained more than 1000 model instances using remote distributed training with different objectives - vision-contrastive, vision-language contrastive (CLIP), categorisation.Preliminary work from this project was accepted as a talk at VSS 2023. Paper accepted at ICLR 2024 workshop on Representational alignment.

Research Specialist — **University of Wisconsin-Madison** — **September 2019 – May 2021**

- Found the presence of ensemble representations like average size, and average color in DNNs trained for object recognition. Presented the work as a talk at VSS 2021 and NMA 2021.

Research Intern — **Brown University** — **Summer 2018**

Serre Lab - Advised by Dr. Thomas Serre

- Compared the performance of various action recognition models on an in-house data set behavioral phenotyping of mice.
- Compiled a dataset that was used to analyse the neural spikes from the V4 neurons in non-human primates.
- Analysed electrophysiological data from the monkey brain to find the receptive field of V4 neurons.

Education

Madison, WI — **University of Wisonsin-Madison** — **Fall 2021 – Fall 2026**

- Ph.D. in Cognitive Science, Expected Fall 2026.
- MS. in Computer Science Engineering, Expected Fall 2024.
- Graduate Coursework: Computer Vision; Learning-based Image Synthesis; Advanced Natural Language Processing; Learning based methods in Computer Vision; Computational Network Biology; Machine Learning;Optimisation; Learning from language; Language Learning in Infants

Bengaluru, India — **PES University** — **Fall 2015 – Fall 2019**

- B.Tech in Computer Science Engineering
- Undergraduate Coursework: Machine Learning; Robotics; Natural Language Processing; Artificial Intelligence; Operating Systems; Databases; Algorithms; Programming Languages; Comp. Architecture

Publications

- **AI-enhanced semantic feature norms for 786 concepts** (2025). **Suresh, S.**, Mukherjee, K., Giallanza, T., Yu, X., Patil, M., Cohen, J. D., Rogers, T. T. (Under review at CogSci and ICLR Bi-align workshop).
- **Bridging the Creativity Understanding Gap: Small-Scale Human Alignment Enables Expert-Level Humor Ranking in LLMs** (2025). Zhou, K. L., Chen, J., **Suresh, S.**, Rogers, T., Jain, L., Nowak, R., Mankoff, R., &. Zhang, J. (Under review).
- **Conceptual structure coheres in human cognition but not in large language models** (2023). **Suresh, S.**, Mukherjee, K., Yu,X., Huang,W., Padua,L., & Rogers, T. T. Accepted to the main track at EMNLP 2023.
- **Humor in AI: Massive Scale Crowd-Sourced Preferences and Benchmarks for Cartoon Captioning** (2024). Zhang, J., Jain, L., Guo, Y., Chen, J., Zhou, K. L., **Suresh, S.**, Wagenmaker, A., Sievert, S., Rogers, T., Jamieson, K., Mankoff, R., & Nowak, R. *Advances in Neural Information Processing Systems* (Spotlight Poster).
- **Probing LLM World Model: Enhancing Guesstimation with Wisdom of Crowds Decoding** (2024). Chuang, Y.-S., Harlalka, N., Narendran, S., Cheung, A., Gao, S., **Suresh, S.**, Hu, J., & Rogers, T. T. Advances in Neural Information Processing Systems 2024 Workshop on Behavioral Machine Learning
- **Categories vs Semantic Features: What shapes the similarities people discern in photographs of objects?** (2024). **Suresh, S.**, Mukherjee, K., Huang, W., & Rogers, T. T. Workshop on Representational Alignment ICLR 2024.
- **The role of shared labels and experiences in representational alignment** (2024). $Mukherjee, K.^*$, **Suresh**, **S.***, Yu,X., & Lupyan, G. Workshop on Representational Alignment ICLR 2024.
- **The Wisdom of Partisan Crowds: Comparing Collective Intelligence in Humans and LLM-based Agents** (2024). Chuang, Y., **Suresh, S.**, Harlalka, N., Goyal, A., Hawkins, R., Yang, S., Shah, D., Hu, J., & Rogers, T.T. ICLR 2024 workshop on LLM Agents
- **Simulating Opinion Dynamics with Networks of LLM-based Agents** (2024). Chuang, Y., Goyal, A., Harlalka, N., **Suresh, S.**, Hawkins, R., Yang, S., Shah, D., Hu, J., & Rogers, T.T. ICLR 2024 workshop on LLM Agents
- **Simulating Opinion Dynamics with Networks of LLM-based Agents** (2024). Chuang, Y., Goyal, A., Harlalka, N., **Suresh, S.**, Hawkins, R., Yang, S., Shah, D., Hu, J., & Rogers, T.T. NAACL 2024.
- **Semantic Feature Verification in FLAN-T5** (2023). **Suresh, S.**, Mukherjee, K., & Rogers, T. T. The First Tiny Papers Track at ICLR 2023.
- **Human-machine cooperation for semantic feature listing.** (2023). Mukherjee, K., **Suresh, S.**, & Rogers, T. T. The First Tiny Papers Track at ICLR 2023.
- **Evaluating LLM Agent Group Dynamics against Human Group Dynamics: A Case Study on Wisdom of Partisan Crowds**(2023). Chuang, Y., **Suresh, S.**, Harlalka, N., Goyal, A., Hawkins, R., Yang, S., Shah, D., Hu, J., & Rogers, T.T.
- **Simulating Opinion Dynamics with Networks of LLM-based Agents.**(2023) Chuang, Y., Goyal, A., Harlalka, N., **Suresh, S.**, Hawkins, R., Yang, S., Shah, D., Hu, J., & Rogers, T.T.
- **Learning interactions to boost human creativity with bandits and GPT-4** (2023). Vartanian, A., Sun, X., Chuang, Y.-S., **Suresh, S.**, Zhu, J., & Rogers, T.T. CogSci 2024
- **GrCluster: a score function to model hierarchy in knowledge graph embeddings** (2020). Ranganathan V., **Suresh S.**, Mathur Y., Subramanyam N, Barbosa D. SAC '20: Proceedings of the 35th Annual ACM Symposium on Applied Computing.
- **Tracking Large Group Synchronous Movement through Space** (in prep). Marji, M., Jackson, J. C., **Suresh, S.**, Miller, K. J., Huang, A., Liu, A., Andresen, C., Pompey, C., & Niedenthal, P.
- **Visual ensemble representations emerge in Deep Neural Networks trained for object recognition** (in prep). **Suresh, S.** & Ward, E.

Conference Presentations

- **Beyond Object recognition: The Role of Visual-Semantic Representations in Understanding the Ventral Visual Stream Suresh, S.**, Mukherjee, K., Huang, W., & Rogers, T. T. Poster presentation at the 2024 CNS Annual meeting in Toronto, Canada.

- **Semantic Feature Verification in FLAN-T5** (2023). **Suresh, S.**, Mukherjee, K., & Rogers, T. T. Poster presented at The First Tiny Papers Track at ICLR 2023.
- **Behavioral estimates of conceptual structure are robust across tasks in humans but not large language models** (2023). **Suresh, S.**, Mukherjee, K., Padua,L., & Rogers, T. T. Poster presented at Cogsci 2023 in Sydney, Australia.
- **Can deep convolutional networks explain the semantic structure that humans see in photographs?** (2023). **Suresh, S.**, Mukherjee, K.,& Rogers, T. T. Talk presented at Vision Science Society conference 2023 in St.Pete's beach, Florida.
- **GrCluster: a score function to model hierarchy in knowledge graph embeddings** (2020). Ranganathan V., **Suresh S.**, Mathur Y., Subramanyam N, Barbosa D. Talk presented online at SAC '20: Proceedings of the 35th Annual ACM Symposium on Applied Computing.
- **Visual memory for causal and coincidental events** (in prep). **Suresh, S.** & Ward, E.Poster presented at Vision Science Society 2022.
- **Visual ensemble representations emerge in Deep Neural Networks trained for object recognition**. **Suresh, S.** & Ward, E.Talk presented at Neuromatch 3.0 Conference 2020.

University and Department Colloquia

- Conceptual structure coheres in human cognition but not large language models. Invited talk at PDP seminar at **Princeton University** - March 2024.
- Guest lecture on using version control while colaborating on research projects. AI + Society seminar (**Wisconsin Institute of Discovery**)- Nov 2023
- Talk at the International Conference on Semantic Cognition at the **Medical Research Council, Cognition and Brain Science Unit, Cambridge, UK.**- Sep 2023
- Guest lecture on using Computer Vision models in Cognitive Science. Class - Environment and tools for Large-Scale Behavioral data science. - June 2023
- Tutorial on using Deep Learning Models to run psychology experiments. Post-baccalaureate Research Education Program (PREP) tutorial series(**University of Wisconsin, Department of Psychology**). - June 2023
- Tutorial on using Large Language Models. AI + Society seminar (**Wisconsin Institute of Discovery**)- February 2023
- Visual memory for causal and coincidental events. First-year Project Symposium (**University of Wisconsin, Department of Psychology**) - Nov 2022
- Visual ensemble representations emerge in Deep Neural Networks trained for object recognition. McPherson Eye Research Institute Seminar Series (**University of Wisconsin, McPherson Eye Research Institute**) - April 2022
- Tutorial on using Deep Learning Models to run psychology experiments. Post-baccalaureate Research Education Program (PREP) tutorial series(**University of Wisconsin, Department of Psychology**). - June 2021
- Visual ensemble representations emerge in Deep Neural Networks trained for object recognition. HAMLET (**University of Wisconsin, Department of Psychology**) - Oct 2020

Relevant projects

- **New yorker cartoon generator using Diffusion Models** - Fine-tuned a diffusion model to generate new yorker style cartoons given an image description.
- **Simulating human opinion dynamics using using agent-based models** Using LLMs to simulate human opinion- dynamic (wisdom of partisan crowds) using human-like agents using langchain.

Awards

- **Integrate** Fellow (NSF-NRT, Work at the intersection of Human-Robot Interaction and Psychology)
- Departmental travel award of 1500 $ to present research at CogSci 2023 in Sydney, Australia
- Menzies award of $500 to conduct research using Large Language Models.
- Departmental travel award of 1500 $ to present research at EMNLP 2023 in Singapore

Teaching Experience

- Programming for behavioral data science (Fall 2023) - Led code along for about 30 students and also held office hours to help them out with becoming better programmers and scientific thinkers.
- **Lead TA, Neuromatch Academy, Deep Learning (Summer 2021):** Taught a three-week-long summer course about Deep Learning and was responsible to lead 5 TAs.

Scientific Service

- Production Editor, Neuromatch Academy - 2021
- Neuromatch 3.0 conference - Volunteered to be a backend host at the NMC 3.0 conference for 20 hours., Was responsible for starting, recording and live-streaming each session., Contributed to the post-production of individual talks after the conference.
- Served as the reviewer for ICLR 2024 workshop on representational alignment.

Languages and Technologies

- Python; Go; Pytorch; Keras (Tensorflow backend); Pandas; scikit-learn; NumPy; Pandas; OpenCV; Langchain; accelerate; Google Cloud Platform; AWS; Git; CLI; Linux; Scipy; Matplotlib; Tensorflow

Relevant coursework

- Artificial Intelligence, Machine Learning, Natural Language Processing, Digital Image Processing, Autonomous Mobile Robotics, Big Data, Design and Analyses of Psychological Experiments, Learning from Language, Learning based image synthesis, Advanced Natural language processing, Learning based methods for computer vision, Computational Network Biology, Modelling user interaction for AI applications.

Mentorship Experience

- Mentored an undergrad on a research project (Hugh Yu) which resulted in a publication at a top AI conference and is now a Masters Student in Computer Science at Brown University.
- Mentored an undergrad on a research project (Agam Goyal) which resulted in a publication at a top AI conference and is now a PhD student in Computer Science at UIUC.
- Mentored an undergrad on a research project (Alex Huang) which resulted in a publication at a top AI conference and is now a Machine Learning Engineer at Fetch.
- Mentored a 3rd-year undergraduate student (Ruohong Wang) to secure a 3-month research internship at a systems neuroscience lab Harvard University.
- Mentored a 3rd-year college student (Samarth Mathur) to secure a 9-month research internship at a systems lab at University of Toronto.
- Mentored a student (Akash Nagraj)to secure a 1-year RAship in Deep Learning (Computer Vision) at Brown University.
- Mentored a student (Gowtham Ramesh) to secure a 1-year RAship in Deep Learning (Natural Language Processing) at IIT Madras.