Kaiyu Yang

Postdoctoral Scholar @ Caltech
☎ 734-389-9696
✉ kaiyuy@caltech.edu
https://yangky11.github.io

ACADEMIC APPOINTMENTS

California Institute of Technology — Pasadena, CA
Computing, Data, and Society Postdoctoral Fellow — 9/2022 – Present

>   Advisor: Anima Anandkumar

EDUCATION

Princeton University — Princeton, NJ
Ph.D. in Computer Science — 7/2022

>   Advisor: Jia Deng
>   Dissertation: "Neurosymbolic Machine Learning for Reasoning"
>   Committee: Danqi Chen, Jia Deng, Mayur Naik, Karthik Narasimhan, Olga Russakovsky

University of Michigan — Ann Arbor, MI
M.S. in Computer Science and Engineering — 8/2018

Tsinghua University — Beijing, China
B.Eng. in Computer Science — 7/2016
B.S. in Mathematics and Applied Mathematics — 7/2016

RESEARCH INTERESTS

AI · Machine Learning · Neuro-symbolic Reasoning · Automated Theorem Proving

PUBLICATIONS

Preprint — **LeanDojo: Theorem Proving with Retrieval-Augmented Language Models**

Kaiyu Yang, Aidan Swope, Alex Gu, Rahul Chalamala, Peiyang Song, Shixing Yu, Saad Godil, Ryan Prenger, and Anima Anandkumar.
Under review at NeurIPS (Datasets and Benchmarks Track), 2023

CVPR 2023 — **Infinite Photorealistic Worlds using Procedural Generation**

A Raistrick, L Lipson, Z Ma, L Mei, M Wang, Y Zuo, K Kayan, H Wen, B Han, Y Wang, A Newell, H Law, A Goyal, K Yang, and J Deng.
Conference on Computer Vision and Pattern Recognition (CVPR), 2023

TMLR 2023 — **Learning Symbolic Rules for Reasoning in Quasi-Natural Language**

Kaiyu Yang and Jia Deng.
Transactions on Machine Learning Research (TMLR), 2023

EMNLP 2022 — **Generating Natural Language Proofs with Verifier-Guided Search**

Kaiyu Yang, Jia Deng, and Danqi Chen.
Conference on Empirical Methods in Natural Language Processing (EMNLP), 2022, *Oral*

ICML 2022	**A Study of Face Obfuscation in ImageNet**
	Kaiyu Yang, Jacqueline Yau, Li Fei-Fei, Jia Deng, and Olga Russakovsky.
	International Conference on Machine Learning (ICML), 2022

NeurIPS 2020	**Strongly Incremental Constituency Parsing with Graph Neural Networks**
	Kaiyu Yang and Jia Deng.
	Neural Information Processing Systems (NeurIPS), 2020

NeurIPS 2020	**Rel3D: A Minimally Contrastive Benchmark for Grounding Spatial Relations in 3D**
	Ankit Goyal, Kaiyu Yang, Dawei Yang, and Jia Deng.
	Neural Information Processing Systems (NeurIPS), 2020, *Spotlight*

FAT* 2020	**Towards Fairer Datasets: Filtering and Balancing the Distribution of the People Subtree in the ImageNet Hierarchy**
	Kaiyu Yang, Klint Qinami, Li Fei-Fei, Jia Deng, and Olga Russakovsky.
	Conference on Fairness, Accountability, and Transparency (FAT)*, 2020

ICML 2019	**Learning to Prove Theorems via Interacting with Proof Assistants**
	Kaiyu Yang and Jia Deng.
	International Conference on Machine Learning (ICML), 2019

ICCV 2019	**SpatialSense: An Adversarially Crowdsourced Benchmark for Spatial Relation Recognition**
	Kaiyu Yang, Olga Russakovsky, and Jia Deng.
	International Conference on Computer Vision (ICCV), 2019

ECCV 2016	**Stacked Hourglass Networks for Human Pose Estimation**
	Alejandro Newell, Kaiyu Yang, and Jia Deng.
	European Conference on Computer Vision (ECCV), 2016

AWARDS AND GRANTS

Neurosymbolic AI for Autonomy 2023
Co-authored proposal awarded by Caltech's Center for Autonomous Systems and Technologies

Siebel Scholar 2022
42 computer science graduate students awarded annually from selected institutions worldwide

Outstanding Reviewer 2020, 2021
Top 20% at the conference on Computer Vision and Pattern Recognition (CVPR)

Google Cloud Research Credits 2019
Google Cloud Platform

ICML Travel Award 2019
International Conference on Machine Learning (ICML)

SEAS Travel Grant 2019
School of Engineering and Applied Science (SEAS), Princeton University

Outstanding Teaching Assistant Award 2015, 2016
Tsinghua University

MEDIA

Exploring the Tradeoff Between Privacy and Algorithm Performance 2022
Princeton Insights

Researchers Devise Approach to Reduce Biases in Computer Vision Data Sets 2020
Princeton Engineering News

AI Is Biased. Here's How Scientists Are Trying to Fix It 2019
Wired

TALKS

Theorem Proving via Machine Learning

      Lean for the Curious Mathematician   9/2023

LeanDojo: Theorem Proving with Retrieval-Augmented Language Models

      Rutgers University   Host: Alex Kontorovich, 7/2023

      Hoskinson Center for Formal Mathematics, CMU   Host: Jeremy Avigad, 5/2023

      Conference on Artificial Intelligence and Theorem Proving (AITP)   9/2023

      Stanford Software Research Lunch   10/2023

Neurosymbolic Reasoning, From Formal Logic to Natural Language

      University of California, Los Angeles   Host: Guy Van den Broeck, 2/2023

      University of California, Santa Barbara   Host: Lei Li, 11/2022

      University of Southern California   Host: Xiang Ren, 10/2022

Teaching Machines to Reason Symbolically

      OpenAI   3/2022

      Google   Host: Denny Zhou, 2/2022

      University of Pennsylvania   Host: Mayur Naik, 2/2022

      NSF "Understanding the World Through Code" Program   Host: Swarat Chaudhuri, 1/2022

      Caltech   Host: Anima Anandkumar, 1/2022

Generating Natural Language Proofs with Verifier-Guided Search

      N2Formal Group, Google   Host: Markus Rabe, 7/2022

A Study of Face Obfuscation in ImageNet

      International Conference on Machine Learning (ICML)   7/2022

      NeurIPS Workshop on "ImageNet: Past, Present, and Future"   12/2021

      CVPR Workshop on "Learning from Limited and Imperfect Data (L2ID)"   6/2021

Learning Symbolic Rules for Reasoning in Quasi-Natural Language

      Princeton NLP Group   7/2021

Towards Fairer Datasets: Filtering and Balancing the Distribution of the People Subtree in the ImageNet Hierarchy

      Conference on Fairness, Accountability, and Transparency (FAT*)   1/2020

Learning to Prove Theorems via Interacting with Proof Assistants

      Princeton Programming Languages Group   10/2019

      International Conference on Machine Learning (ICML)   6/2019

RESEARCH MENTORING

Peiyang Song   2022 – Present
Undergraduate student @ UCSB

Rahul Chalamala   2022 – Present
Undergraduate student @ Caltech

Shixing Yu   2022 – 2023
Master's student @ UT Austin → Ph.D. student @ Cornell

Gene Chou   2021
Undergraduate student @ Princeton → Ph.D. student @ Cornell

Jacqueline Yau   2019 – 2020
Master's student @ Stanford → Machine Learning Engineer @ Apple

TEACHING EXPERIENCE

COS484/584: Natural Language Processing 2021/2 – 2021/5
Teaching assistant, Department of Computer Science, Princeton University
Data Structures and Algorithms 2013/8 – 2016/7
Head teaching assistant, Department of Computer Science and Technology, Tsinghua University

SERVICE

Organizer

The 3rd Workshop on Mathematical Reasoning and AI @ NeurIPS 2023

Tutorial on Machine Learning for Theorem Proving @ NeurIPS 2023

Reviewer

International Conference on Machine Learning (ICML)

Neural Information Processing Systems (NeurIPS)

International Conference on Learning Representations (ICLR)

Journal of Machine Learning Research (JMLR)

IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)

ACM Transactions on Programming Languages and Systems (TOPLAS)

Artificial Intelligence to Assist Mathematical Reasoning: Proceedings of a Workshop

Computer Vision and Pattern Recognition (CVPR)

International Conference on Computer Vision (ICCV)

European Conference on Computer Vision (ECCV)

Nature Human Behaviour

Volunteer

Neural Information Processing Systems (NeurIPS)

Session Chair

Caltech SURF Seminar Day

Committee Member

Caltech CMS Graduate Admission Committee

REFERENCES

Anima Anandkumar
Bren Professor
Computing + Mathematical Sciences
California Institute of Technology
Pasadena, CA 91125
✉ anima@caltech.edu

Danqi Chen
Assistant Professor
Department of Computer Science
Princeton University
Princeton, NJ 08544
✉ danqic@cs.princeton.edu

Jia Deng
Associate Professor
Department of Computer Science
Princeton University
Princeton, NJ 08544
✉ jiadeng@princeton.edu

Olga Russakovsky
Assistant Professor
Department of Computer Science
Princeton University
Princeton, NJ 08544
✉ olgarus@cs.princeton.edu