

Deqing Fu

Email: deqingfu@usc.edu **Address:** Ginsburg Hall, USC, Los Angeles, CA 90089
Website: deqingfu.github.io **Office:** GCS 403E

Education	University of Southern California <i>Doctor of Philosophy (Ph.D.) in Computer Science</i> Advisors: Vatsal Sharan & Robin Jia	2022 - Present
	University of Chicago <i>Master of Science (S.M.) in Statistics</i>	2020 - 2022
	University of Chicago <i>Bachelor of Science (S.B.) in Mathematics</i> with Honors <i>Bachelor of Science (S.B.) in Computer Science</i> with Honors <i>Bachelor of Arts (A.B.) in Statistics</i>	2016 - 2020
Experience	Simons Institute at UC Berkeley <i>Visiting Graduate Student</i> Program: Modern Paradigms in Generalization	Fall 2024
	Meta Llama Team <i>AI Research Scientist Intern</i> Host: Lawrence Chen; Peers: Tong Xiao, Rui Wang, and Pengchuan Zhang	May – Dec., 2024
	Google <i>Software Engineering Intern</i> Teams: Image Search; Knowledge Engine; Bard (now Gemini) Multilingual	Summer 2021, 2022, 2023
	Summer Student Researcher , Argonne National Laboratory Hosts: Paul Hovland & Sri Hari Krishna Narayanan	Summer 2017
All Publications	VisualLens: Personalization through Task-Agnostic Visual History Wang Bill Zhu, <u>Deqing Fu</u> , Kai Sun, Yi Lu, Zhaojiang Lin, Seungwhan Moon, Kanika Narang, Mustafa Canim, Yue Liu, Anuj Kumar, Xin Luna Dong	NeurIPS 2025
	TLDR: Token-Level Detective Reward Model for Large Vision Language Models <u>Deqing Fu</u> , Tong Xiao, Rui Wang, Wang Zhu, Pengchuan Zhang, Guan Pang, Robin Jia, Lawrence Chen	ICLR 2025
	Transformers Learn Low Sensitivity Functions: Investigations and Implications	

* denotes
equal contribution

	Bhavya Vasudeva*, <u>Deqing Fu</u> *, Tianyi Zhou, Elliot Kau, You-Qi Huang, Vatsal Sharan	ICLR 2025
	DeLLMa: Decision Making Under Uncertainty with Large Language Models Ollie Liu*, <u>Deqing Fu</u> *, Dani Yogatama, Willie Neiswanger <i>ICLR 2025 Spotlight (Top 5.1%)</i>	ICLR 2025
	DreamSync: Aligning Text-to-Image Generation with Image Understanding Feedback Jiao Sun*, <u>Deqing Fu</u> *, Yushi Hu*, Su Wang, Royi Rassin, Da-Cheng Juan, Dana Alon, Charles Herrmann, Sjoerd van Steenkiste, Ranjay Krishna, Cyrus Rashtchian	NAACL 2025
	Transformers Learn to Achieve Second-Order Convergence Rates for In-Context Linear Regression <u>Deqing Fu</u> , Tian-Qi Chen, Robin Jia, Vatsal Sharan <i>SoCalNLP Symposium 2023 Best Paper Award</i>	NeurIPS 2024
	Pre-trained Large Language Models Use Fourier Features to Compute Addition Tianyi Zhou, <u>Deqing Fu</u> , Vatsal Sharan, Robin Jia	NeurIPS 2024
	IsoBench: Benchmarking Foundation Models on Isomorphic Representations <u>Deqing Fu</u> *, Ruohao Guo*, Ghazal Khalighinejad*, Ollie Liu*, Bhuwan Dhingra, Dani Yogatama, Robin Jia, Willie Neiswanger <i>Huggingface 🗝️: [Dataset]</i>	COLM 2024
	SCENE: Self-Labelled Counterfactuals for Extrapolating to Negative Examples <u>Deqing Fu</u> , Ameya Godbole, Robin Jia	EMNLP 2023
	Harnessing the Conditioning Sensorium for Improved Image Translation Cooper Nederhood, Nicholas Kolkin, <u>Deqing Fu</u> , Jason Salavon	ICCV 2021
Pre-prints	Are LLM Decisions Faithful to Verbal Confidence? Jiawei Wang, Yanfei Zhou, Siddartha Devic, Deqing Fu	
	When Do Transformers Learn Heuristics for Graph Connectivity? Qilin Ye*, <u>Deqing Fu</u> *, Robin Jia, Vatsal Sharan	
	Resa: Transparent Reasoning Models via SAEs Shangshang Wang, Julian Asilis, Ömer Faruk Akgül, Enes Burak Bilgin, Ollie Liu, <u>Deqing Fu</u> , Willie Neiswanger	

Technical Report: Qwen3 Cross-layer Transcoders

Jakob Hansen, Deqing Fu, Shangshang Wang, John Carlsson

Zebra-CoT: A Dataset for Interleaved Vision Language Reasoning

Ang Li*, Charles L. Wang*, Deqing Fu*, Kaiyu Yue*, Zikui Cai*, Wang Bill Zhu*, Ollie Liu*, Peng Guo*, Willie Neiswanger, Furong Huang, Tom Goldstein, Micah Goldblum

Huggingface 🍌: [*Dataset*] [*Models*]

Textual Steering Vectors Can Improve Visual Understanding in Multimodal Large Language Models

Woody Haosheng Gan*, Deqing Fu*, Julian Asilis*, Ollie Liu*, Dani Yogatama, Vatsal Sharan, Robin Jia, Willie Neiswanger

FoNE: Precise Single-Token Number Embeddings via Fourier Features

Tianyi Zhou, Deqing Fu, Mahdi Soltanolkotabi, Robin Jia, Vatsal Sharan

Awards

NeurIPS Top Reviewer

NeurIPS, 2025

Viterbi Best Research Assistant Award (*one per department*)

USC, 2025

USC Capital One Center (CREDIF) PhD Fellow

USC, 2025

SoCalNLP Symposium 2023 Best Paper Award

SoCalNLP, 2023

Provost's Fellowship

USC, 2022

Susanne H. Rudolph Scholarship

UChicago, 2020-2021

Liew Family College Research Fellows Fund

UChicago, 2020

Dean's List

UChicago, 2016-2020

Grants

VESSL AI Academia Program, 1,200 NVIDIA A100 Hours (\approx \$2,500 value)

PI: Deqing Fu

August 2025

Title: *Emergence of Fourier Features in Large Language Model Pretraining*

NSF ACCESS Allocation, 750,000 Credits (\approx \$75,000 value)

PI: Deqing Fu, Co-PI: Robin Jia

August 2025 - August 2026

Title: *Robust Mathematical Reasoning with Fourier Language Models*

Invited Talks

Algorithmic Perspectives on Understanding Transformers

FlaNN Talk Seminar

Aug 25, 2025

Closing the Modality Gap:

Benchmarking and Improving Visual Understanding in Multimodal LLMs

Stanford NLP Seminar

May 22, 2025

Algorithmic Perspectives on Understanding Transformers

	UChicago/TTIC NLP Seminar	May 08, 2025
	<i>Tutorial on the interplay between Theory and Interpretability</i>	
	USC ML/NLP Joint Seminar	Apr 25, 2025
	<i>Algorithmic Perspectives on Understanding Transformers</i>	
	Duke NLP Seminar	Feb 12, 2025
	<i>Guest Lecture on Interpretability and AI Safety</i>	
	DSO 575: Driving Business Transformation with GenAI	Oct 24, 2024
	<i>Simplicity Bias orf Transformers to Learn Low Sensitivity Functions</i>	
	USC NLP Lunch	Feb 22, 2024
	<i>Aligning Text-to-Image Generation with Image Understanding Feedback</i>	
	USC NLP Lunch	Oct 19, 2023
	<i>Transformers Learn to Achieve Second-Order Convergence for In-Context Linear Regression</i>	
	USC NLP Lunch	Mar 29, 2023
Services	Reviewer for conferences in	
	• Machine Learning : ICLR (2025, 2026), NeurIPS (2024, 2025), ICML 2025;	
	• Natural Language Processing : COLM 2024, EMNLP (2023, 2024), ACL (2024, 2025);	
	• Computer Vision : CVPR (2023, 2024, 2025), ECCV 2022, ICCV 2025	
	Organizer of USC ML/NLP Joint Reading Group	Jan 2025 – Present
	Organizer of USC AI Safety Club	Jan 2024 – Jan 2025
Students Supervised	Master's and Undergraduate Students	
	Zitong Cynthia Huang (USC UG)	Sep 2025 – Present
	Gustavo Carvalho (USC MS)	Feb 2024 – Present
	Qilin Ye (USC UG \rightarrow Duke MS)	Sep 2023 – Present
	Woody Haosheng Gan (USC UG)	Sep 2023 – Present
	You-Qi Huang (USC UG)	Sep 2023 – Mar 2024
	Elliot Kau (USC MS)	Sep 2023 – Mar 2024
	Tian-Qi Chen (USC MS)	Mar 2023 – Sep 2023
Teaching Experience	Teaching Assistant , CSCI 467: Introduction to Machine Learning	
	University of Southern California, Los Angeles, CA	Spring 2025
	Teaching Assistant , CSCI 566: Deep Learning and its Applications	
	University of Southern California, Los Angeles, CA	Spring 2023
	Teaching Assistant , BUSN 41301: Statistical Insight	

Booth School of Business, University of Chicago, Chicago, IL Fall 2021

Course Assistant, CMSC 31230: Fundamentals of Deep Learning
University of Chicago, Chicago, IL Fall 2021

Course Assistant, MATH 20300: Analysis in \mathbb{R}^n - I
University of Chicago, Chicago, IL Spring 2018

Skills

C/C++, Python, \LaTeX , MATLAB, R, Java, CUDA
PyTorch, TensorFlow, Keras, JAX, HuggingFace