

Han Fang |

✉ hanfang.info@gmail.com • 🌐 hanfang.info
Github Linkedin Google Scholar

EXPERIENCE

Meta

AI Research Scientist, Meta Super Intelligence Labs

Jan 2025 - Present

- My research focuses on teaching models to reason, plan, and use tools effectively. I work at the frontier of agentic capabilities through RL, with a particular interest in emerging forms of long-term memory.
- Lead efforts to improve agentic tool-use capabilities, including zero-shot generalization to unseen domains and long-horizon tasks. My work spans the full end-to-end stack, including pre- & mid-training / SFT, synthetic data generation, agentic RL (recipe and environments), & agentic evaluation.
- Drive step-change in model quality across many competitive agentic benchmarks, leading to more robust JSON adherence over multi-step multi-turn, long-horizon execution, and native MCP support.
- Built RL environments (in both simulation and human-in-the-loop fashion) to improve the model's agentic capabilities: created complex tasks, code-based graders, model-based graders, and rubric-based scoring.

Meta

Senior Staff Research Scientist / TLM, Meta AI

Jan 2023 - Dec 2024

- Led post-training of Llama 2 and 3 in production, [debut Meta AI](#), and grew it to 1 billion MAU. Talk at [Meta's Connect Conference 2023](#). Drove the integrated training runs, core capabilities, tool use and orchestration, and data flywheel.
- Ran the core alignment work stream, drove model integration runs (SFT, RL, Eval). Developed Meta AI's online RL framework with [Mixture of Judges](#), improving reasoning, instructions following, safety, refusals.
- Created the Orchestrator system and the Planner model for enabling [tool use in Meta AI](#), enabling real-time information with search, image generation, code execution, and photo editing with images ([Blog](#)).
- Led the core capabilities work streams, including [instructions following](#), [factuality](#), tool use, and [multilinguality](#) (enabled the roll-out to [12 languages and 40+ countries](#)).
- Developed Meta AI's data flywheel for [Reinforcement Learning from User Feedback](#), including sampling algorithms, user preference alignment, and flywheel annotation.
- Led development of the search tool for LLM grounding. Drove complex negotiations with Microsoft and closed contract alignment on the [Bing-Meta Search API partnership](#).

Facebook (now Meta)

Staff Research Scientist / TLM, Facebook AI

June 2020 - Jan 2023

- Led the CTO's special project to [advance AI for better multi-modal multi-lingual content understanding](#).
- Led the development [Meta's first Few-shot Learner](#) which can work in 100+ languages, learns from images & text, and detects new forms of violations without human in the loop.
- Led the development of linear complexity Transformer architecture, Linformer, to enable scaling up GPU inference and analysis of billions of contents a day ([Blog](#)).
- Drove step-function improvement in topline metrics in IG & FB. Developed efficient Transformer based models in recs system and multi-domain user modeling. ([Blog](#)).

Facebook (now Meta)

Senior Research Scientist

Seattle, WA

July 2017 - June 2020

- Developed [Reinforcement Learning based model optimizer](#) to E2E optimize models for downstream tasks.
- Identified the [AI gaps around multilingual and multimodal capabilities to correctly detect violation](#).
- Built [Tetris Planner](#), FB's first global storage rebalancing algorithm for PiB-scale data placement.

University of Washington

Industry Advisory Board

Seattle, WA

Aug 2019–Dec 2024

- Serves on the industry advisory board of the Data Science graduate program to improve its curriculum.

Facebook	Menlo Park, CA
<i>Data Scientist Intern (Rating: Rock-star, top 3%)</i>	<i>Jun 2016 - Aug 2016</i>
○ Developed ML models to predict billion-scale user engagement and found effective infrastructure strategies.	
○ Developed an optimization framework for geo-spatial data on air field recommendation in 30 countries.	
Cold Spring Harbor Laboratory (CSHL)	Cold Spring Harbor, NY
<i>Research Assistant (PhD student in residence)</i>	<i>Aug 2014 - Jun 2017</i>
○ PhD thesis: Graphical and machine learning algorithms for large-scale genomics data (2000+ citations).	
○ Developed the best practice for detecting mutations in large-scale sequencing data (in Nature Protocols)	
○ Built a negative binomial mixture model to estimate and infer the genome properties (1000+ citations).	
Cold Spring Harbor Laboratory	Cold Spring Harbor, NY
<i>Computational Science Developer</i>	<i>June 2013 - June 2014</i>
○ Co-developed Scalpel, a SoTA graph assembly algorithm to detect mutations (in Nature Methods).	

EDUCATION

Stony Brook University	Stony Brook, NY
<i>PhD in Applied Mathematics, President's Award to Distinguished Doctoral Students</i>	<i>2014–2017</i>
Johns Hopkins University	Baltimore, MD
<i>Visiting graduate student in Computer Science - Advisor moved to JHU</i>	<i>2016–2017</i>
Stony Brook University	Stony Brook, NY
<i>M.S. in Applied Mathematics and Statistics</i>	<i>2011–2013</i>
Sun Yat-sen University	Guangzhou, China
<i>B.S. in Optical Physics, Cum laude</i>	<i>2007–2011</i>

AWARDS

- 2017 President's Award to Distinguished Doctoral Students @ SBU (5 out of 449)
- 2017 The Woo-Jong Kim Dissertation Award @ AMS (1 out of 34)
- 2017 Excellence in Research Award @ AMS (5 out of 34)
- 2016 Research Access Project @ SBU
- 2016 Department Travel Grant @ AMS
- 2015 Reviewers' Choice @ The American Society of Human Genetics (Top 10%)
- 2015 Summer Institute in Statistics for Big Data Scholarship @ University of Washington
- 2014 Research Assistant Fellowship @ CSHL
- 2013 Research Access Project @ SBU
- 2013 Department Travel Grant @ AMS
- 2010 Outstanding Student Scholarship @ SYSU (Top 10%)

PUBLICATIONS

[Google Scholar](#)

In Industry:

- Generalized Parallel Scaling with Interdependent Generations *Harry Dong, David Brandfonbrener, Eryk Helenowski, Yun He, Mrinal Kumar, Han Fang, Yuejie Chi, Karthik Abinav Sankararaman* [ICML](#) (2026)
- Dual-Weighted Reinforcement Learning for Generative Preference Modeling *Shengyu Feng, Yun He, Shuang Ma, Beibin Li, Yuanhao Xiong, Songlin Li, Karishma Mandyam, Julian Katz-Samuels, Shengjie Bi, Licheng Yu, Hejia Zhang, Karthik Abinav Sankararaman, Han Fang, Riham Mansour, Yiming Yang, Manaal Faruqui [arXiv](#) (2025)*
- Boosting LLM Reasoning via Spontaneous Self-Correction. *Xutong Zhao, Tengyu Xu, Xuewei Wang, Zhengxing Chen, Di Jin, Liang Tan, Zishun Yu, Zhuokai Zhao, Yun He, Sinong Wang, Han Fang, Sarath Chandar, Chen Zhu* . [COLM](#) (2025)
- Reinforcement Learning from User Feedback. *Eric Han, Jun Chen, Karthik Abinav Sankararaman, Xiaoliang Peng, Tengyu Xu, Eryk Helenowski, Kaiyan Peng, Mrinal Kumar, Sinong Wang, Han Fang, Arya Talebzadeh* . [arXiv](#) (2025)
- Learning Auxiliary Tasks Improves Reference-Free Hallucination Detection in Open-Domain Long-Form

Generation. *Chengwei Qin, Wenxuan Zhou, Karthik Abinav Sankararaman, Nanshu Wang, Tengyu Xu, Alexander Radovic, Eryk Helenowski, Arya Talebzadeh, Aditya Tayade, Sinong Wang, Shafiq Joty, Han Fang, Hao Ma* · [ACL](#) (2025)

- Think Smarter not Harder: Adaptive Reasoning with Inference Aware Optimization. *Zishun Yu, Tengyu Xu, Di Jin, Karthik Abinav Sankararaman, Yun He, Wenxuan Zhou, Zhouhao Zeng, Eryk Helenowski, Chen Zhu, Sinong Wang, Hao Ma, Han Fang* · [ICML](#) (2025)
- Step-KTO: Optimizing Mathematical Reasoning through Stepwise Binary Feedback *Yen-Ting Lin, Di Jin, Tengyu Xu, Tianhao Wu, Sainbayar Sukhbaatar, Chen Zhu, Yun He, Yun-Nung Chen, Jason Weston, Yuandong Tian, Arash Rahnama, Sinong Wang, Hao Ma, Han Fang* · [arXiv](#) (2025)
- Improving model factuality with fine-grained critique-based evaluator. *Yiqing Xie, Wenxuan Zhou, Pradyot Prakash, Di Jin, Yuning Mao, Quintin Fettes, Arya Talebzadeh, Sinong Wang, Han Fang, Carolyn Rose, Daniel Fried, Hejia Zhang* · [ACL](#) (2024)
- Multi-if: Benchmarking llms on multi-turn and multilingual instructions following. *Yun He, Di Jin, Chaoqi Wang, Chloe Bi, Karishma Mandyam, Hejia Zhang, Chen Zhu, Ning Li, Tengyu Xu, Hongjiang Lv, Shruti Bhosale, Chenguang Zhu, Karthik Abinav Sankararaman, Eryk Helenowski, Melanie Kambadur, Aditya Tayade, Hao Ma, Han Fang, Sinong Wang* · [arXiv](#) (2024)
- The Perfect Blend: Redefining RLHF with mixture of judges. *Tengyu Xu, Eryk Helenowski, Karthik Abinav Sankararaman, Di Jin, Kaiyan Peng, Eric Han, Shaoliang Nie, Chen Zhu, Hejia Zhang, Wenxuan Zhou, Zhouhao Zeng, Yun He, Karishma Mandyam, Arya Talabzadeh, Madian Khabsa, Gabriel Cohen, Yuandong Tian, Hao Ma, Sinong Wang, Han Fang* · [arXiv](#) (2024)
- Effective long-context scaling of foundation models. *Wenhan Xiong, Jingyu Liu, Igor Molybog, Hejia Zhang, Prajwala Bhargava, Rui Hou, Louis Martin, Rashi Rungta, Karthik Abinav Sankararaman, Barlas Oguz, Madian Khabsa, Han Fang, Yashar Mehdad, Sharan Narang, Kshitiz Malik, Angela Fan, Shruti Bhosale, Sergey Edunov, Mike Lewis, Sinong Wang, Hao Ma* · [NAACL](#) (2024)
- Representation deficiency in masked language modeling. *Yu Meng, Jitin Krishnan, Sinong Wang, Qifan Wang, Yuning Mao, Han Fang, Marjan Ghazvininejad, Jiawei Han, Luke Zettlemoyer* · [ICLR](#) (2023)
- Improved adaptive algorithm for scalable active learning with weak labeler. *Yifang Chen, Karthik Sankararaman, Alessandro Lazaric, Matteo Pirotta, Dmytro Karamshuk, Qifan Wang, Karishma Mandyam, Sinong Wang, Han Fang* · [arXiv](#) (2022)
- Bayesformer: Transformer with uncertainty estimation. *Karthik Abinav Sankararaman, Sinong Wang, Han Fang* · [arXiv](#) (2022)
- Microestimates of wealth for all low-and middle-income countries. *Guanghua Chi, Han Fang, Sourav Chatterjee, Joshua E Blumenstock* · [PNAS](#) (2021)
- Reducing target group bias in hate speech detectors. *Darsh J Shah, Sinong Wang, Han Fang, Hao Ma, Luke Zettlemoyer* · [arXiv](#) (2021)
- Entailment as few-shot learner. *Sinong Wang, Han Fang, Madian Khabsa, Hanzi Mao, Hao Ma* · [arXiv](#) (2021)
- Linformer: Self-Attention with Linear Complexity. *Sinong Wang, Belinda Z Li, Madian Khabsa, Han Fang, Hao Ma* · [arXiv](#) (2020)

In PhD:

- Graphical and machine learning algorithms for large-scale genomics data. *Han Fang* · Dissertation (2017)
- Complex rearrangements and oncogene amplifications revealed by long-read DNA and RNA sequencing of a breast cancer cell line. *Maria Nattestad, Sara Goodwin, Karen Ng, Timour Baslan, Fritz J Sedlazeck, Philipp Rescheneder, Tyler Garvin, Han Fang, et al, W Richard McCombie, Michael C Schatz* · [Genome Research](#) (2018)
- Scikit-ribo enables accurate estimation and robust modeling of translation dynamics at codon resolution. *Han Fang, Yi-Fei Huang, Aditya Radhakrishnan, Adam Siepel, Gholson J Lyon, Michael C Schatz* · [Cell Systems](#) (2018)
- Whole genome sequencing of one complex pedigree illustrates challenges with genomic medicine. *Han Fang, Yiyang Wu, Hui Yang, Margaret Yoon, Laura T Jiménez-Barrón, David Mittelman, Reid Robison, Kai Wang, Gholson J Lyon* · [BMC Medical Genomics](#) (2017)
- GenomeScope: fast reference-free genome profiling from short reads. *Gregory W Vurture, Fritz J Sedlazeck, Maria Nattestad, Charles J Underwood, Han Fang, James Gurtowski, Michael C Schatz* · [Bioinformatics](#) (2017)
- Accurate detection of complex structural variations using single molecule sequencing. *Fritz J Sedlazeck, Philipp Rescheneder, Moritz Smolka, Han Fang, Maria Nattestad, Arndt von Haeseler, Michael C Schatz* · [Nature Method](#) (2017)

- Indel variant analysis of short-read sequencing data with Scalpel. *Han Fang, Ewa A Grabowska, et al, Michael C Schatz, Giuseppe Narzisi* · [Nature Protocols](#) (2016)
- Reducing INDEL calling errors in whole genome and exome sequencing data. *Han Fang, Yiyang Wu, Giuseppe Narzisi, et al, Michael C Schatz, Gholsen J Lyon* · [Genome Medicine](#) (2014)
- Accurate de novo and transmitted indel detection in exome-capture data using microassembly. *Giuseppe Narzisi, Jason A O'rave, Ivan Iossifov, Han Fang, et al, Michael Wigler, Michael C Schatz* · [Nature Methods](#) (2014)

CONFERENCE

Platform Talk Presentations:

- *Meta AI: Building Meta's next gen AI experiences with Llama*
Meta Connect Conference, Menlo Park, CA 2023
- *Tetris Planner: Optimizing Facebook Data Warehouse Data Placement*
INFORMS Annual Meeting, Seattle, WA 2019
- *Machine Learning and graph partitioning for Facebook Data Warehouse*
The Data Science Conference, Chicago, IL 2018
- *Scikit-ribo reveals precise codon-level translational control by dissecting ribosome pausing and codon elongation.*
Biological Data Science Meeting, Cold Spring Harbor, NY 2016
- *Scikit-ribo reveals precise codon-level translational control by dissecting ribosome pausing and codon elongation.*
Advances in Genome Biology and Technology(AGBT) Meeting, Orlando, FL 2016
- *Scikit-ribo: Accurate A-site prediction and robust modeling of translation control from Riboseq & RNAseq data.*
Genome Informatics Meeting, Cold Spring Harbor, NY 2015
- *Reducing INDEL calling errors in whole genome and exome sequencing data.*
Biological Data Science Meeting, Cold Spring Harbor, NY 2014

First-author Poster Presentations:

- *Scikit-ribo: Accurate estimation and robust modelling of translation dynamics at codon resolution*
Biology of Genome Meeting, Cold Spring Harbor, NY 2017
- *Scikit-ribo reveals precise codon-level translational control by dissecting ribosome pausing and codon elongation.*
Advances in Genome Biology and Technology(AGBT) Meeting, Hollywood, FL 2017
- *Scikit-ribo reveals precise codon-level translational control by dissecting ribosome pausing and codon elongation.*
Genome Informatics Meeting, Cambridge, UK 2016
- *Scikit-ribo reveals precise codon-level translational control by dissecting ribosome pausing and codon elongation.*
Translational control Meeting, Cold Spring Harbor, NY 2016
- *Scikit-ribo: Accurate A-site prediction and robust modeling of translation control from Riboseq & RNAseq data.*
Probabilistic Modeling in Genomics Meeting, Cold Spring Harbor, NY 2015
- *Indel variant analysis of short-read sequencing data with Scalpel. (Reviewers' Choice)*
American Society of Human Genetics Annual Meeting, Baltimore, MD 2015
- *Reducing INDEL calling errors in whole genome and exome sequencing data.*
Personal Genomes Meeting, Cold Spring Harbor, NY 2014
- *Whole genome analysis of a pedigree with Prader-Willi syndrome, hereditary hemochromatosis, and dysautonomia.*
Personal Genomes Meeting, Cold Spring Harbor, NY 2014
- *Reducing INDEL calling errors in whole genome and exome sequencing data.*
American Society of Human Genetics Annual Meeting, San Diego, CA 2014

PROFESSIONAL ACTIVITIES

- 2021, ECML-PKDD, Program Committee
- 2019-2025, University of Washington Data Science Program, Industrial Advisory Board
- 2018-Present, The Data Science Conference (TDSC), Advisory Board
- 2018 International Conference on Computing and Information Systems (ICCIS), Program Committee
- 2018 International Conference on Intelligent Information Technologies (ICIIT), Reviewer
- 2018 International Conference on Research in Computational Molecular Biology (RECOMB), Reviewer
- 2017 Nucleic Acids Research (NAR), Reviewer
- 2017 Workshop on Algorithms in Bioinformatics (WABI), Reviewer