Proposal: The 2nd Workshop on Foundation Models in the Wild

Tagline: Making foundation models more adaptive, reliable and efficient to be deployed in the wild.
Modality: In-person
Anticipated audience size: We expected more than 1000 attendees.
Contacts: Xinyu Yang: xinyuya2@andrew.cmu.edu; Huaxiu Yao: huaxiu@cs.unc.edu

Workshop Summary

In the era of AI-driven transformations, foundation models (FMs) have become pivotal in various applications, from natural language processing to computer vision. These models, with their immense capabilities, reshape the future of scientific research and the broader human society, but also introduce challenges in their in-the-wild/real-world deployments. The Workshop on FMs in the wild delves into the urgent need for these models to be useful when deployed in our societies. The significance of this topic cannot be overstated, as the real-world implications of these models impact everything from daily information access to critical decision-making in fields like medicine and finance. Stakeholders, from developers to end-users, care deeply about this because the successful integration of FMs into in-the-wild frameworks necessitates a careful consideration of many properties, including adaptivity, reliability, efficiency, and reasoning ability. Some of the fundamental questions that this workshop aims to address are:

- **In-the-wild Adaptation:** How can we leverage techniques such as Retrieval-Augmented Generation (RAG), In-context Learning (ICL), or Fine-tuning (FT) to adapt FMs for specific domains, such as drug discovery, education, or clinical health?
- **Reasoning and Planning:** How can FMs be enhanced to tackle more complex in-the-wild tasks that require multi-step reasoning or decision-making, such as multi-hop question answering, mathematical problem-solving, theorem proving, code generation, or robot planning scenarios?
- **Reliability and Responsibility:** How can FMs work reliably outside their training distribution? And how can we address issues like hallucination, fairness, ethics, safety and privacy within the society?
- **Practical Limitations in Deployment:** How can FMs tackle challenges in practical applications, such as system constraints, memory requirements, response time demands, data acquisition barriers, and computational costs for inference-time scaling and long-context input?

As prospective participants, we primarily target machine learning researchers and industry partitioners interested in the questions and foci outlined above. Our target audience includes professionals deeply involved with FMs and their in-the-wild applications, especially for those who focus on the real-world deployments of these models. We also welcome submissions from researchers in the natural sciences (e.g., physics, chemistry, biology) and social sciences (e.g., pedagogy, sociology) to offer attendees a more comprehensive perspective. In summary, our topics of interest include, but are not limited to:

- Innovations in techniques for customizing models to individual user preferences, tasks, or domains
- Advancements in the reasoning and planning abilities of FMs in complex real-world challenges
- Theoretical and empirical investigations into the reliability and responsibility of various FMs
- Strategies for overcoming practical limitations (e.g., memory, time, data) of FMs in broad applications
- Methods for integrating multiple modalities (e.g., text, images, action) into a unified in-the-wild framework
- Discussions on FM agents that perform intricate tasks through interaction with the environment
- In-depth discussions exploring the in-the-wild deployments and applications of FMs
- Benchmark methodologies for assessing FMs in real-world settings

Invited Speakers

We are pleased that a group of researchers with diverse backgrounds, affiliations, and areas of expertise have agreed to give invited talks at our workshop. Each speaker will bring a unique perspective to current developments in foundation models and their applications in the wild. We aim to provide titles for all talks prior to the event. **Confirmed speakers include:**

1. Anima Anandkumar (California Institute of Technology, female), FMs for science and theorem proving
2. Xinyun Chen (Google DeepMind, female), Limitations of LLMs for in-the-wild applications
3. Chelsea Finn (Stanford University, female), embodied FMs for robots and physically-actuated devices
4. Tatsunori Hashimoto (Stanford University, male), Evaluations and scaling in the context of FM agents
5. Reza Shokri (National University of Singapore & Microsoft, male), FM safety, security, and privacy
6. Jie Tang (Tsinghua University, male), Reliable and scalable LLM training
7. Yuandong Tian (Meta AI Research, male), LLM reasoning, planning, and decision-making
8. René Vidal (University of Pennsylvania, male), Robustness, interpretability, and faithfulness of FMs

Diversity Commitment

In the **selection of organizers and speakers**, we actively promoted diversity in all its forms. The final roster of organizers and speakers comprises individuals from varied genders, races, countries, affiliations, and scientific backgrounds. Among the organizers and speakers, we have ensured representation across the full spectrum of scientific seniority, including Ph.D. students, assistant professors, associate professors, full professors, and industry researchers.

To further enhance the accessibility of our workshop to a broader audience, we also plan to offer a **registration fee grant** for those who might otherwise be unable to register. To support these initiatives, we will seek sponsorships from leading companies such as Google Deepmind, Meta AI, OpenAI, and Amazon.

Our **review process** for submitted materials will be double-blind (conducted via OpenReview) to mitigate institutional and author biases. The program will be curated to ensure a wide representation of research areas while upholding the standards of quality set by the double-blind review process. Consequently, our workshop will benefit from a diverse cohort of participants, and we will invite several contributors to speak alongside our primary invitees.

We are launching a **reviewing mentorship program** designed to foster the growth and development of junior reviewers. Within this initiative, junior reviewers will be paired with senior reviewers. These mentor-mentee relationships aim to ensure that junior reviewers receive real-time feedback, guidance, and mentorship as they navigate the complexities of crafting insightful and constructive reviews for workshop submissions. By facilitating this collaborative and educational process, we hope not only to elevate the quality of reviews but also to cultivate the next generation of expert reviewers in the field.

Recognizing the challenges posed by **varying time zones** in a hybrid meeting format, we will incorporate a blend of synchronous and asynchronous activities to ensure wide participation. Specifically, we will ask both invited speakers and authors of accepted papers to provide pre-recorded videos in advance, enabling registered attendees to access the content flexibly. Live sessions, such as panel discussions and Q&A segments for invited talks or spotlights, will be facilitated through platforms like sli.do

Tentative Schedule

This workshop will adopt an in-person format. Specifically, our workshop will feature **eight 30-minute invited talks (comprising a 25-minute presentation followed by a 5-minute Q&A session), three 15-minute contributed talks selected from submissions, two 1-hour poster sessions, and a 1-hour panel discussion** to delve into the future of FMs in the wild. The poster session should be attended in person. To enhance the experience for attendees, we will utilize a dedicated channel on a chat platform like Rocket.Chat to facilitate interactions among workshop participants.

Tentative Schedule of Paper Submission. We will follow the suggested dates by ICLR.

- Workshop paper submission deadline: February 3, 2025
- Workshop paper notification date: March 5, 2025
- Final workshop program, camera-ready, videos uploaded: March 20, 2025

Tentative Workshop Schedule.

Morning:

- 08:50 – 09:00 Introduction and opening remarks
- 09:00 - 09:30 Invited Talk 1
- 09:30 - 10:00 Invited Talk 2
- 10:00 - 10:15 Contributed Talk 1
- 10:15 - 11:15 Poster Session 1
- 11:15 - 11:45 Invited Talk 3
- 11:45 - 12:15 Invited Talk 4
- 12:15 - 13:30 *Break*

Afternoon:

- 13:30 - 14:00 Invited Talk 5
- 14:00 - 14:30 Invited Talk 6
- 14:30 - 14:45 Contributed Talk 2
- 14:45 - 15:45 Poster Session 2
- 15:45 - 16:15 Invited Talk 7
- 16:15 - 16:30 Contributed Talk 3
- 16:30 - 17:00 Invited Talk 8
- 17:00 - 18:00 Panel discussion

Previous Related Workshops

Our workshop builds upon the “Foundation Models in the Wild” workshop held at ICML 2024. While continuing to address adaptability, reliability, and efficiency of FM applications in the wild, we will place additional emphasis on the reasoning and planning capabilities of these models in real-world scenarios. We’re particularly interested in exploring how these advanced features can be applied to solve complex machine learning challenges like theorem proving and code generation, and how they can contribute to the practical development of FM agents and embodied FMs. Unlike other workshops that focus on a singular aspect of FMs’ capabilities or applications, we seek to stimulate discussions on the in-the-wild applications of foundation models from multiple perspectives, thereby fostering a comprehensive dialogue about the necessary properties for their utilization in scientific research and broader human society. Some notable workshops focus on FMs’ capabilities include:“Workshop on Scalable Continual Learning for Lifelong Foundation Models” (NeurIPS 2024), “Adaptive Foundation Models: Evolving AI for Personalized and Efficient Learning” (NeurIPS 2024), “Safe Generative AI” (NeurIPS 2024), “Long-Context Foundation Models” (ICML 2024), “Next Generation of AI Safety” (ICML 2024), “Secure and Trustworthy Large Language Models” (ICLR 2024), “Workshop on Reliable and Responsible Foundation Models” (ICLR 2024). In terms of the applications of FMs in specific domains, workshops such as “Foundation Models for Science: Progress, Opportunities, and Challenges” (NeurIPS 2024), “The First Workshop on Large Foundation Models for Educational Assessment” (NeurIPS 2024), “AI for Science: Scaling in AI for Scientific Discovery” (ICML 2024), “Workshop on Large Language Models for Agents” (ICLR 2024), “Generative Models for Decision Making” (ICLR 2024).

Organizers and Biographies

Xinyu Yang (Carnegie Mellon University)

- Email: xinyuya2@andrew.cmu.edu
- Webpage: https://xinyuyang.me/
- Google Scholar: https://scholar.google.com/citations?user=Fvq2R14AAAAJ

- Bio: Xinyu Yang is a second-year PhD student at CMU. His research is centered on useful, efficient, and adaptive foundation model systems for scientific research and human society, with a specific focus on long-context reasoning problems.

Huaxiu Yao (UNC-Chapel Hill)

- Email: huaxiu@cs.unc.edu
- Webpage: https://www.huaxiuyao.io/
- Google Scholar: https://scholar.google.com/citations?hl=en&user=A20BZnQAAAAJ
- Bio: Huaxiu Yao is a tenure-track Assistant Professor at the Department of Computer Science with a joint appointment in the School of Data Science and Society, UNC-Chapel Hill. He was a Postdoctoral Scholar in Computer Science at Stanford University. Huaxiu earned his Ph.D. degree from Pennsylvania State University. Currently, focuses on both the theoretical and applied aspects of building reliable and responsible foundation models. He is also dedicated to applying foundation models to solve real-world scientific and social applications, such as healthcare, transportation, and education. He has organized and co-organized workshops at ICML and NeurIPS and has served as a tutorial speaker at conferences such as KDD, AAAI, and IJCAI. Additionally, Huaxiu has extensive industry experience, having interned at companies such as Amazon Science, and Salesforce Research.

Mohit Bansal (UNC-Chapel Hill)

- Email: mbansal@cs.unc.edu
- Webpage: https://www.cs.unc.edu/ mbansal/
- Google Scholar: https://scholar.google.com/citations?user=DN8QtscAAAAJ&hl=en
- Bio: Mohit Bansal is the John R. & Louise S. Parker Professor and the Director of the MURGe-Lab (UNC-NLP Group) in the Computer Science department at UNC-Chapel Hill. Prior to this, he was a research assistant professor (3-year endowed position) at TTI-Chicago. He received his Ph.D. in 2013 from the University of California at Berkeley (where he was advised by Dan Klein) and his B.Tech. from the Indian Institute of Technology at Kanpur in 2008. His research expertise is in natural language processing and multimodal machine learning, with a particular focus on grounded and embodied semantics, language generation and Q&A/dialogue, and interpretable and generalizable deep learning. He is a recipient of IIT Kanpur Young Alumnus Award, DARPA Director's Fellowship, NSF CAREER Award, Google Focused Research Award, Microsoft Investigator Fellowship, Army Young Investigator Award (YIP), DARPA Young Faculty Award (YFA), and outstanding paper awards at ACL, CVPR, EACL, COLING, and CoNLL. He has been a keynote speaker for the AACL 2023 and INLG 2022 conferences. His service includes ACL Executive Committee, ACM Doctoral Dissertation Award Committee, CoNLL Program Co-Chair, ACL Americas Sponsorship Co-Chair, and Associate/Action Editor for TACL, CL, IEEE/ACM TASLP, and CSL journals.

Beidi Chen (Carnegie Mellon University & Meta)

- Email: beidic@andrew.cmu.edu
- Webpage: https://www.andrew.cmu.edu/user/beidic/
- Google Scholar: https://scholar.google.com/citations?user=jCNJhFcAAAAJ
- Bio: Beidi Chen is an Assistant Professor at the Department of Electrical and Computer Engineering at Carnegie Mellon University. Previously, she was a postdoc researcher at Stanford working with Dr. Chris Ré. Beidi received my Ph.D. in Computer Science from Rice University under the supervision of Dr. Anshumali Shrivastava in 2020. Her research focuses on large-scale machine learning. Specifically, Beidi design and optimize randomized algorithms (algorithm-hardware co-design) to accelerate large machine learning systems for real-world problems.

Junlin Han (Meta & University of Oxford)

- Email: junlinhcv@gmail.com
- Webpage: https://junlinhan.github.io/
- Google Scholar: https://scholar.google.com/citations?user=5L0Uj_IAAAAJ
- Bio: Junlin Han is full-time researcher at GenAI Llama team, Meta, and a first-year PhD student at University of Oxford, working with Prof. Philip Torr. His research focuses on computer vision, deep learning, and artificial intelligence, with a particular concentration on studying data.

Pavel Izmailov (Anthropic & New York University)

- Email: pi390@nyu.edu
- Webpage: https://izmailovpavel.github.io/
- Google Scholar: https://scholar.google.com/citations?user=AXxTpGUAAAAJ
- Bio: Pavel Izmailov is a Researcher at Anthropic. He is primarily interested in reasoning, AI for science and AI alignment. Starting in Fall 2025, he will be joining NYU as an Assistant Professor in the Tandon CSE department, and Courant CS department by courtesy. He is also a member of the NYU CILVR Group. Previously, he worked on reasoning and superintelligent AI alignment at OpenAI.

Jinqi Luo (University of Pennsylvania)

- Email: jinqiluo@upenn.edu
- Webpage: https://peterljq.github.io/
- Google Scholar: https://scholar.google.com/citations?user=a55zN_4AAAAJ
- Bio: Jinqi Luo is a PhD student of CIS in the UPenn, affiliated with GRASP Lab and IDEAS Center, advised by Prof. René Vidal and collaborating with Prof. Chris Callison-Burch. His research focuses on deep generative modeling, multimodal ML, and foundation model trustworthiness.

Pang Wei Koh (University of Washington)

- Email: pangwei@cs.washington.edu
- Webpage: https://koh.pw/
- Google Scholar: https://scholar.google.com/citations?user=Nn990CkAAAAJ
- Bio: Pang Wei Koh is an assistant professor in the Allen School of Computer Science and Engineering at the University of Washington. His research interests are in the theory and practice of building reliable and interactive machine learning systems. His research has been published in Nature and Cell, featured in media outlets such as The New York Times and The Washington Post, and recognized by the MIT Technology Review Innovators Under 35 Asia Pacific award and best paper awards at ICML and KDD. He received his PhD and BS in Computer Science from Stanford University. Prior to his PhD, he was the 3rd employee and Director of Partnerships at Coursera. He has organized several sessions of NeurIPS Workshops on Distribution Shifts.

Weijia Shi (University of Washington)

- Email: swj0419@uw.edu
- Webpage: https://weijiashi.notion.site/
- Google Scholar: https://scholar.google.com/citations?user=zt73PHcAAAAJ
- Bio: Weijia Shi is a PhD student in Computer Science at the University of Washington advised by Luke Zettlemoyer and Noah A. Smith. Her research focuses on natural language processing and machine learning, with a particular interest in retrieval-augmented LMs and trustworthy generative AI.

Philip Torr (University of Oxford)

- Email: philip.torr@eng.ox.ac.uk
- Webpage: https://www.robots.ox.ac.uk/phst/
- Google Scholar: https://scholar.google.com/citations?user=kPxa2w0AAAAJ
- Bio: Philip Torr is a Full Professor at the University of Oxford and the founder of the Torr Vision group. His research focuses on Computer Vision and Machine Learning. He received his Ph.D. from the University of Oxford and worked as a research scientist for Microsoft Research. He has received numerous awards, including the Marr prize and the Royal Society Wolfson Research Merit Award. He was elected Fellow of the Royal Academy of Engineering (FREng) in 2019, and Fellow of the Royal Society (FRS) in 2021. He also serves as a director of OxSight and the Chief Scientific Advisor for Five AI.

Songlin Yang (Massachusetts Institute of Technology)

- Email: yangsl66@mit.edu
- Webpage: https://sustcsonglin.github.io/
- Google Scholar: https://scholar.google.com/citations?user=1chlis0AAAAJ
- Bio: Songlin Yang is a second-year PhD student at MIT CSAIL, advised by Prof. Yoon Kim. Her research is centered on the intersection of machine learning system and large language model, with a specific focus on the hardware-aware algorithm design for efficient sequence modeling.

Luke Zettlemoyer (University of Washington & Meta)

- Email: lsz@cs.washington.edu
- Webpage: https://www.cs.washington.edu/people/faculty/lsz
- Google Scholar: https://scholar.google.com/citations?user=UjpbO6IAAAAJ
- Bio: Luke Zettlemoyer is a Professor in the Allen School of Computer Science & Engineering at the University of Washington, and also a Research Scientist at Meta. His honors include multiple paper awards, and being named an PECASE Awardee and an Allen Distinguished Investigator. Previously, he did postdoctoral research at the University of Edinburgh and earned a Ph.D. at MIT. His research is in the intersections of natural language processing, machine learning, and decision making under uncertainty. He is particularly interested in designing learning algorithms for recovering representations of the meaning of natural language text

Jiaheng Zhang (National University of Singapore)

- Email: jhzhang@nus.edu.sg
- Webpage: https://zjhzjh123.github.io/
- Google Scholar: https://scholar.google.com/citations?hl=en&user=vh90–IAAAAJ
- Bio: Jiaheng Zhang is an assistant professor in School of Computing at National University of Singapore. He obtained Ph.D. degree in Computer Science at UC Berkeley, advised by Prof. Dawn Song. His research interests lie in LLM/AI security and privacy, applied cryptography, especially zero-knowledge proofs and applications on blockchain and machine learning models.