Lifelong Agents: Learning, Aligning, Evolving

**Cheng Qian[1], Emre Can Acikgoz[1], Hongru Wang[2], Zhenfei Yin[3], Manling Li[4],
Yun-Nung (Vivian) Chen[5], Mengdi Wang[6], Caiming Xiong[7]**

[1]University of Illinois Urbana-Champaign, [2]University of Edinburgh, [3]University of Oxford,
[4]Northwestern University, [5]National Taiwan University, [6]Princeton University,
[7]Salesforce AI Research

Website: `https://lifelongagent.github.io`
Contact: `lifelongagent@googlegroups.com`

1 Workshop Summary

Artificial intelligence has reached a pivotal stage. Recent advances in reinforcement learning (Shao et al., 2024), large language models (LLMs) (Mialon et al., 2023), and embodied systems (Li et al., 2022) have pushed the frontier of what autonomous agents can achieve (Luo et al., 2025). Yet, the prevailing paradigm of *train once, deploy once* exposes a critical limitation: today's agents excel in static benchmarks, but show deficiency when deployed in dynamic, evolving environments (Silver & Sutton). They often fail to adapt to shifting user goals, changing contexts, or long-term operation, restricting their robustness, trustworthiness, and societal utility (Acikgoz et al., 2025; Liu et al., 2025; Qian et al., 2025).

The concept of the **lifelong agent** offers a paradigm shift. Such agents are not static artifacts but dynamic processes: they continuously accumulate knowledge, refine skills, and evolve capabilities across time (Gao & et al., 2025). A lifelong agent must not only learn and adapt, but also align with human preferences and operate responsibly in changing environments and under resource constraints. Building such agents is a pressing challenge if we want AI systems that are sustainable, trustworthy, and impactful in the real world (Bengio et al., 2025).

Research communities have explored many of the necessary ingredients, such as continual learning, reinforcement learning, alignment frameworks, post-training adaptations, and efficiency methods for resource-constrained agents (Kumar et al., 2025; Acikgoz et al., 2025; Belcak et al., 2025). Yet these discussions remain fragmented. This workshop aims to bring these strands together in the first unified forum, reframing intelligence as a process that **learns, aligns, and evolves** across an agent's lifespan. We envision a comprehensive discussion that establishes foundations, highlights shared principles, and surfaces open challenges.

We propose to structure the workshop around three core dimensions of lifelong agents, with cross-cutting requirements of efficiency and sustainability. Each dimension includes guiding questions and subtopics for discussion:

1. **Learning Across Time**: How can agents continuously acquire and consolidate knowledge and skills without catastrophic forgetting?
   - Methods for continual learning under non-stationary distributions.
   - Techniques for balancing plasticity (adaptation) and stability (retention).
   - Incorporating multimodal and embodied data streams in dynamic environments.
   - Memory systems for integrating short-term experience with long-term knowledge.
   - Benchmarks for evaluating adaptation across diverse and evolving tasks.

2. **Alignment Over Lifespans**: How can agents remain trustworthy and value-consistent while adapting over time?
   - Preference learning and alignment under evolving user goals.
   - Oversight mechanisms and safeguards for long-term autonomous adaptation.
   - Personalization while preserving safety, fairness, and accountability.
   - Ethical considerations of self-modifying agents and persistent deployment.
   - Techniques for measuring and auditing alignment drift over time.
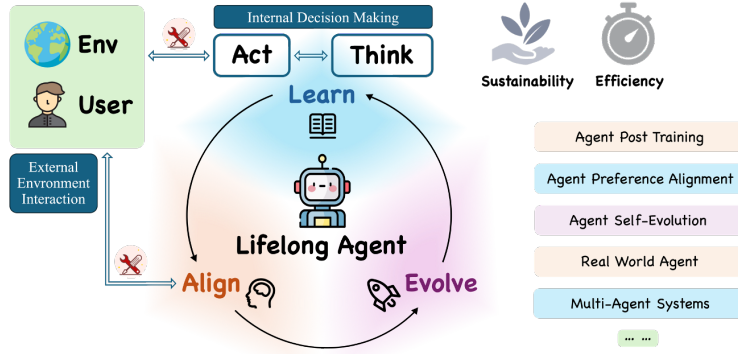
Figure 1: Conceptual Illustration of Lifelong Agents.

3. **Evolving Capabilities**: How can agents not only adapt but also self-improve and expand their functionality?
   - Mechanisms for autonomous refinement of reasoning, representations, or tools.
   - Discovering new strategies, skills, or sub-modules in open-ended settings.
   - Leveraging collective intelligence in multi-agent and human-agent ecosystems.
   - Synergizing LLMs with smaller, specialized models for efficiency and agility.
   - Long-term evaluations: when and how does an agent demonstrate true growth?

4. **Cross-cutting Requirements**: Efficiency, sustainability, and evaluation frameworks.
   - Efficient continual learning under limited computational and energy budgets.
   - Resource-aware deployment of lifelong agents in real-world settings.
   - Datasets and testbeds for assessing lifelong learning, alignment, and self-improvement.
   - Standardized evaluation metrics for persistence, adaptability, and long-term safety.

Topics of Interest. The workshop encourages paper submissions on topics including, but not limited to, the following research directions.

- **Agentic Reinforcement Learning:** memory-augmented RL, continual exploration, and resource-aware adaptation, aiming to support agents that remain effective over extended lifespans.

- **Post-training of Agentic LLMs:** continual fine-tuning, instruction alignment, domain shift adaptation, and tool use, enabling agents to remain competent beyond their initial training.

- **Dynamic Alignment:** long-term value learning, oversight mechanisms, and methods for detecting and correcting drift, ensuring agents adapt without losing trustworthiness.

- **Self-Evolving Agents:** autonomous refinement of reasoning strategies, self-modifying representations, and emergent behaviors in multi-agent systems, pointing toward agents that self-improve.

- **Embodied and Real-World Agents:** lifelong robotics, adaptive perception-action loops, and operation under uncertainty, advancing agents that learn directly from complex environments.

- **Efficient and Sustainable Agents:** token- and compute-efficient inference, adaptive resource allocation, and energy-aware design, critical for scalable and long-term deployment.

- **Multi-Agent Lifelong Systems:** cooperation, competition, negotiation, and persistent collectives of adaptive agents, highlighting challenges of coordination across time.

- **AI Agents for Science:** autonomous hypothesis generation, experimental design, and knowledge discovery, reimagining agents as active participants in scientific progress.

- **Evaluation and Benchmarking:** metrics that capture adaptability, persistence, and alignment across long horizons, providing a foundation for rigorous comparison of approaches.

- **Applications and Case Studies:** lessons from deployment in robotics, healthcare, science, education, and decision-making domains, showing how lifelong agents perform in practice.

In addition to invited talks, we will organize panel discussions to stimulate cross-disciplinary dialogue, and we will feature a poster session to highlight emerging directions. By convening diverse

communities under the unifying theme of lifelong agents, this workshop seeks to not only showcase cutting-edge research but also articulate a roadmap toward AI agents that are not only intelligent, but intelligently persistent and adaptive.

2 TENTATIVE SCHEDULE

Workshop Schedule. We propose a full-day hybrid workshop that blends invited talks, contributed presentations, poster sessions, and panel discussions in order to maximize both intellectual depth and community engagement. Our schedule is guided by three principles: (1) balance between keynote presentations from senior researchers and contributions from emerging voices; (2) generous allocation of time for poster sessions, coffee breaks, and interactive discussions to promote exchange of ideas; and (3) accessibility and inclusivity, ensuring that both in-person and virtual participants can participate fully.

The program will include six invited talks (30 minutes each, with time for Q&A), three oral presentation slots highlighting accepted submissions, and two 1-hour poster sessions that will run in parallel with coffee breaks to stimulate informal conversations. Two panel discussions will allow for cross-cutting dialogue on the challenges and opportunities of lifelong agents. Finally, an awards ceremony will conclude the day, recognizing outstanding contributions. In total, 3.5 hours (about 50% of the program, excluding lunch) are dedicated to structured interactions, ensuring that participants not only learn but also connect meaningfully with each other.

Activity Type	Time
Opening Remarks	8:45-9:00
Invited Talk 1: Graham Neubig	9:00-9:30
Invited Talk 2: Azalia Mirhoseini	9:30-10:00
Coffee Break & Poster Session 1	10:00-11:00
Oral Presentations (10 min each)	11:00-11:30
Invited Talk 3: Siva Reddy	11:30-12:00
Panel Discussion I	12:00-12:30
Lunch	12:30-13:30
Invited Talk 4: Sergey Levine	13:30-14:00
Invited Talk 5: Asli Celikyilmaz	14:30-15:00
Coffee Break & Poster Session II	15:00-16:00
Invited Talk 6: Yu Su	16:00-16:30
Panel Discussion II	16:30-17:00
Awards and Closing Remarks	17:00-17:15

Table 1: Single-Day Workshop Schedule (Tentative).

Interaction and Access. In line with our commitment to inclusivity, the workshop will be conducted in a hybrid format. All invited talks, panels, and oral presentations will be livestreamed, with interactive Q&A to allow virtual participants to contribute in real time. Poster sessions will be supported through both in-person presentation and dedicated online platforms, enabling asynchronous participation across time zones. A dedicated chat space will facilitate continuous discussion throughout the day, bridging the gap between in-person and online audiences. After the workshop, slides and recordings (with permission) will be made available on the workshop website, democratizing access for those unable to attend synchronously.

3 LOGISTICS, REVIEW, AND SUPPORT

Paper Submission and Review Process. We will accept full papers (up to 9 pages) and short papers (2–4 pages) in accordance with ICLR 2026 guidelines. The short paper track is designed to encourage participation from early-career researchers, underrepresented groups, and those with limited resources, as well as to accommodate late-breaking ideas. Suitable submissions for this track

include implementations of novel ideas, modest but self-contained theoretical results, follow-up experiments, re-analyses of existing work, and fresh perspectives on prior publications. AI-generated submissions will not be permitted, but AI-assisted writing is allowed when clearly identified, following the ICLR 2026 proper use of AI policy.

To comply with ICLR's notification deadline (March 1, 2026), we propose the following timeline:

- Workshop paper submission deadline: **February 2, 2026**
- Notification of acceptance: **February 26, 2026**
- Camera-ready and final workshop program: **March 15, 2026**

This timeline ensures fairness and transparency while leaving sufficient time for authors to prepare final versions and for organizers to integrate accepted papers into the program. All accepted papers will be featured prominently through oral or poster sessions, and an awards session at the conclusion of the workshop will highlight the best contributions.

Award Setting. We will present up to three Best Paper Awards, with winners featured in the oral sessions. This recognition is intended to highlight exceptional contributions and incentivize high-quality submissions.

Funding Support. We have secured **$2,000** in sponsorship from Oumi and **$5,000** in sponsorship from Salesforce. We are also actively pursuing additional support from Amazon to fund student travel awards. Our organizing team has a strong track record of managing such funds through academic affiliations, ensuring smooth distribution. We are also in discussions to provide compute and API credit vouchers from platforms like Cohere and OpenAI as additional rewards for outstanding contributions. These efforts reflect our commitment to supporting emerging researchers and fostering an inclusive academic community.

Inclusion Support. We will introduce workshop attendance awards tailored for K–12 students from underrepresented communities in Rio de Janeiro, as well as undergraduate students from local institutions. This initiative aims to lower financial barriers, spark interest in AI research, and provide valuable exposure to professional networks at ICLR. By actively engaging local schools and communities, we hope to empower the next generation of researchers and innovators in Brazil.

4   Invited Speakers and Panelists

All of the following invitees have agreed to participate as speakers or panelists in our workshop, with either confirmed or tentatively confirmed commitments.

Sergey Levine (University of California, Berkeley)

Sergey Levine is an Associate Professor in EECS at UC Berkeley. His research centers on algorithms that enable autonomous agents to acquire complex behaviors through learning, with emphasis on deep reinforcement learning for decision making and control in robotics and other domains. He received his B.S., M.S., and Ph.D. in Computer Science from Stanford University, and joined Berkeley's faculty in 2016. He leads the Robotic Artificial Intelligence and Learning (RAIL) Lab at Berkeley, where his group advances offline RL, imitation learning, and real-world robotic learning.

Azalia Mirhoseini (Stanford University; Google DeepMind)

Azalia Mirhoseini is an Assistant Professor of Computer Science at Stanford University, where she leads the Scaling Intelligence Lab, and a Senior Staff Research Scientist at Google DeepMind. Her research focuses on learning-based methods for systems and chip design (e.g. RL for chip floorplanning) and self-improving scalable AI. Before academia, she held roles at Anthropic and Google Brain, where she co-founded their "ML for Systems" effort. She earned her Ph.D. in Electrical & Computer Engineering from Rice University (Best PhD Thesis Award) and her B.Sc. in Electrical Engineering from Sharif

University of Technology. Her honors include MIT Technology Review's 35 Under 35 and wide recognition for her work in flagship venues like Nature.

SIVA REDDY (MCGILL UNIVERSITY; MILA; CANADA CIFAR AI CHAIR)

Siva Reddy is an Assistant Professor jointly in Computer Science and Linguistics at McGill University, a Canada CIFAR AI Chair, and a core academic member at Mila. He co-leads the McGill NLP group, with research focusing on language understanding for QA and conversational systems that integrate symbolic/linguistic structure with deep learning; prior roles include a postdoctoral position at Stanford and recognition as a Google PhD Fellow during his doctoral studies at the University of Edinburgh. He also served as a research scientist at ServiceNow Research, complementing his academic leadership.

GRAHAM NEUBIG (CARNEGIE MELLON UNIVERSITY)

Graham Neubig is an Associate Professor at the Language Technologies Institute, Carnegie Mellon University, affiliated with CMU's Machine Learning Department, and Chief Scientist at All Hands AI. His research covers NLP and LLMs, from core modeling to applications (including software agents), with a longstanding focus on multilingual technologies and practical, open tools that translate advances into robust systems. His long-term goal is to break down barriers in human–human and human–computer communication through NLP.

ASLI CELIKYILMAZ (META FAIR)

Asli Celikyilmaz is a Senior Research Manager at Meta FAIR and previously a Senior Principal Researcher at Microsoft Research. Her work has significantly influenced natural language generation and evaluation, as well as memory-augmented and reasoning-augmented agents that ground language in social and physical contexts; she also serves the community through many editorial and leadership roles. She completed earlier research appointments at UC Berkeley and Microsoft, and maintains active publication and service portfolios across major NLP/AI venues.

YU SU (OHIO STATE UNIVERSITY)

Yu Su is an Associate Professor of Computer Science and Engineering at The Ohio State University, where he co-directs the OSU NLP group and helps lead AI institute efforts (Foundational AI team at ICICLE and ML Foundations at Imageomics). Previously, he was a Senior Researcher at Microsoft Semantic Machines. He received his Ph.D. in Computer Science from the University of California, Santa Barbara, and his B.S. from Tsinghua University; his recognitions include the UCSB Outstanding Dissertation Award and a Schmidt Sciences AI2050/Innovation-style distinction. His research spans language agents, knowledge grounding, and the interface between language and structured knowledge.

EMMANOUIL KOUKOUMIDIS (OUMI PBC)

Emmanouil Koukoumidis is the CEO and co-founder of Oumi PBC, an open-source AI platform for foundation model development and deployment. Previously, he led major generative AI efforts at Google Cloud, including the PaLM/Gemini initiative, overseeing engineering, model quality, and Responsible AI. He received his Ph.D. in Electrical Engineering from Princeton University and his undergraduate engineering degree from the National Technical University of Athens, with earlier research experience at MIT, Microsoft Research, CERN, and SMART (Singapore). His work has spanned mobile/cloud systems, NLP, and deep learning, with thousands of citations, and he has been recognized through awards such as the CODiE finalist distinction for Oumi's tools. He also helped secure $10M in seed funding for Oumi, advancing the vision of open, community-driven frontier AI.

Why these speakers. These speakers collectively embody the multidisciplinary expertise essential for realizing the vision of *lifelong agents*: AI systems that learn, align, and evolve across time. **Sergey**

Levine brings foundational advances in reinforcement learning and autonomous skill acquisition, illuminating how agents can continuously improve through experience. **Azalia Mirhoseini** connects this to the systems frontier, designing scalable, self-improving architectures that sustain long-term adaptation. **Siva Reddy** and **Yu Su** advance the integration of language, knowledge, and reasoning, enabling agents to ground their learning in structured, factual worlds. **Graham Neubig** contributes expertise in practical LLM systems, bridging research and deployment in dynamic environments, while **Asli Celikyilmaz** drives progress in generation, memory, and evaluation—key components for agents that retain and refine capabilities. Complementing these perspectives, **Emmanouil Koukoumidis** provides insight into open, community-driven infrastructures that make continual learning and responsible evolution feasible at scale. Together, their perspectives span the algorithmic, linguistic, and systemic dimensions of building AI that truly learns and lives across its lifespan.

5 ORGANIZERS AND BIOGRAPHIES

5.1 ORGANIZERS

Cheng Qian ([Email] [Website])   is a Ph.D. student at the University of Illinois Urbana-Champaign, where his research focuses on large language model agents with an emphasis on tool use, reasoning, and creativity. His representative work includes CREATOR and ToolRL. He has also been honored with the Capital One ASKS Center Fellowship and served as Area Chair in the ARR ACL 2025 and EMNLP 2025 cycles. He obtained his bachelor degree from Tsinghua University.

Emre Can Acikgoz ([Email] [Website])   is a PhD candidate in Computer Science at UIUC advised by Dilek Hakkani-Tur and Gokhan Tur. His research focuses on developing interactive Conversational (Language) Agents. Some of his representative works include CoALM and Desiderata for Conversational Agents. He is DSTC12 Workshop Session Chair at SIGDial 2025 and currently an Applied Scientist Intern at Amazon Alexa AI.

Hongru Wang ([Email] [Website])   is an incoming postdoc at University of Edinburgh. He received his PhD degree from The Chinese University of Hong Kong. His work on dialogue system and language agents won 2 best paper awards at 2023 International Doctoral Forum and SIGHAN@ACL 2024. Some of his representative works include Self-DC, OTC-PO and Theory of Agent. He organized the first Tool Learning tutorial (ToolsMeetLLM) at SIGIR 2024 and co-founded the NLP Academic Exchange Platform (NICE), a forum for sharing and discussing cutting-edge research across career stages.

Zhenfei Yin ([Email] [Website])   is a postdoctoral researcher at the University of Oxford. His primary research interests include multimodal learning, AI agents, multi-agent systems, and embodied agents. He has organized Trustworthy Multimodal Foundation Models and AI Agents (TiFA) workshop at ICML 2024; Multi-modal Foundation Model meets Embodied AI workshop at ICML 2024; Multi-Agent Systems in the Era of Foundation Models workshop at ICML 2025; Multi-Modal Reasoning for Agentic Intelligence workshop at ICCV 2025.

5.2 SENIOR ORGANIZERS

Manling Li ([Email] [Website])   is an Assistant Professor at Northwestern University and an Amazon Scholar, specializing in multimodal and embodied AI agent. Her research has been recognized including the ACL 2025 Inaugural Best Desseratation Award Honorable Mention, ACL 2024 Outstanding Paper Award, ACL 2020 and NAACL 2021 Best Demo Paper Awards, EECS Rising Star, MIT Tech Review Innovators Under 35, etc. She was lead organizers of Towards Knowledgeable Foundation Models (ACL 2024/2025, AAAI 2025) and Foundation Models meet Embodied Agents (CVPR 2025), and has served in committee such as Virtual Infrastructure Chairs for ACL 2025, Demonstrations Chairs for EMNLP 2024. She also delivered related tutorials at ACL, NAACL, etc.

Yun-Nung (Vivian) Chen ([Email] [Website])   is a Professor at National Taiwan University. Her research on dialogue systems and conversational agents has earned major honors including Google Faculty Research Awards, Amazon AWS Machine Learning Research Awards, and the MOST Young Scholar Fellowship. She has co-organized workshops such as NLP for Conversational

AI (NLP4ConvAI) series (NeurIPS 2020-2023, ACL 2021/2022) and Dialog System Technology Challenge (DSTC) workshops (AAAI 2020-2025). She has also delivered tutorials on conversational AI and multimodal dialogue at premier conferences including ACL, NeurIPS, ICASSP, and IJCNLP.

Mengdi Wang ([Email] [Website])   is a Professor at Princeton University. Her research focused on stochastic optimization, reinforcement learning, and AI for science, and has earned major honors including the NSF CAREER Award, Google Faculty Award, MIT Tech Review Innovators Under 35 (China), WAIC YunFan Award, and the AACC Donald Eckman Award. She has co-organized the Workshop on Multi-Agent Reinforcement Learning and Bandit Learning at the Simons Institute (2022), and served as Program Chair for ICLR 2023 and Senior AC for NeurIPS, ICML, and COLT.

Caiming Xiong ([Email] [Website])   is Senior Vice President leading AI Research & Applied Research at Salesforce, where he directs teams working on LLMs, agents, NLP, and vision; prior to industry, he was a postdoctoral researcher at UCLA after earning his Ph.D. in Computer Science and Engineering from the State University of New York at Buffalo. His group has released widely used research artifacts such as CodeGen/XGen and enterprise-focused model/tooling, with publications across ACL, NeurIPS, CVPR, ICLR, etc. He has served on the steering committee for the Interactive and Executable Semantic Parsing (IntEx-SemPar) workshop at EMNLP, and frequently contributes to conference programs and industry research forums.

5.3   ADVISORY BOARD

Due to the workshop's policy limiting the number of official organizers to eight, we have established an Advisory Board to include several senior professors who have been instrumental in shaping the theme and scope of the event. While advisory board members will not be directly involved in the organization or logistical coordination, their experience and insight have provided valuable guidance on the workshop's vision, structure, and academic direction.

Heng Ji ([Email] [Website])   is a Professor at University of Illinois Urbana-Champaign and an Amazon Scholar. Her research interests focus on NLP, Data Mining and AI for Science. The awards she received include "AI's 10 to Watch" Award, NSF CAREER award, Google Research Award, IBM Watson Faculty Award, Bosch Research Award, and Amazon AWS Award, ACL2020 Best Demo Paper Award, and NAACL2021 Best Demo Paper Award. She has coordinated the NIST TAC Knowledge Base Population task since 2010. She is elected as the North American Chapter of the Association for Computational Linguistics (NAACL) secretary 2020-2021.

Kam-Fai Wong ([Email] [Website])   is a Professor at The Chinese University of Hong Kong. He is a member of the ACM, Senior Member of IEEE as well as Fellow of ACL, BCS (UK), IET (UK) and HKIE. He is the founding Editor-In-Chief of ACM Transactions on Asian Language Processing, and serves as associate editor of International Journal on Computational Linguistics and Chinese Language Processing.

Philip Torr ([Email] [Website])   is a Professor at the University of Oxford and holds the Five AI / Royal Academy of Engineering Research Chair in Computer Vision and Machine Learning. His research focuses on AI, machine learning, and computer vision, with seminal contributions to object recognition, segmentation, 3D reconstruction, and tracking. His accolades include the Marr Prize (1998), Royal Society Wolfson Merit Award, and fellowships with both the Royal Academy of Engineering and the Royal Society. He is also a Turing AI World-Leading Researcher Fellow, recognized for academic excellence and real-world impact.

Jun Wang ([Email] [Website])   is Chair Professor of Computer Science at University College London and Director of the UCL Centre for Artificial Intelligence. His research focuses on reinforcement learning, multi-agent systems, computational advertising, and generative models, with applications ranging from science to the arts. He has received multiple best paper awards and previously earned his PhD from Delft University of Technology.

Dilek Hakkani-Tür ([Email] [Website])   is a Professor at the University of Illinois Urbana-Champaign, and formerly a leading researcher at Amazon Alexa AI, Google Research, Microsoft

Research, ICSI, and AT&T Labs (2001–2023) on dialogue systems and natural language understanding. She has earned IEEE, ISCA, and ACL Fellowships (in 2024) and served in major editorial and leadership roles, including Editor-in-Chief of IEEE/ACM Transactions on Audio, Speech and Language Processing (2019–2021), co-Editor-in-Chief of Transactions of the Association for Computational Linguistics (since 2024), and President of ACL's SIGDIAL (2023–2025).

Gokhan Tur ([Email] [Website]) is a Research Professor at the University of Illinois Urbana-Champaign and an Amazon Scholar. He was a founding member of the Microsoft Cortana team, and later worked on conversational AI at Apple Siri, Google Research, Uber AI, and Amazon Alexa AI. His research has been recognized with the IEEE SPS Best Paper Award (2020) and multiple international best paper awards. Prof. Tur is a Fellow of IEEE and has served as senior area chair for EMNLP 2024 and NAACL 2024, as well as associate editor for IEEE Transactions on Audio, Speech, and Language Processing. He received his Ph.D. in Computer Science from Bilkent University, Turkey in 2000.

6 EXPECTED AUDIENCE SIZE AND OUTREACH PLAN

Expected Audience Size. We anticipate sustained in-person attendance of **200–300** participants throughout the day, with peak sessions reaching **300–400** attendees. With hybrid participation and day-long livestreaming, we expect a **total unique reach of 500–600** across in-person and virtual audiences. These estimates are grounded in the demonstrated draw of our invited speakers, the cross-disciplinary scope (e.g. agentic AI for scientific discovery), organizers' collective social following (∼20k), and attendance patterns observed at recent top-tier conferences (ICLR/ICML/NeurIPS) for closely related workshops.

How we will attract the audience. Our outreach strategy blends targeted community channels with broad social amplification to ensure both depth and diversity of participation. We will (1) launch an official website and social media presence (X, LinkedIn, Facebook) to centralize updates, post early speaker bios/abstracts, and announce the call for papers; (2) distribute the Call of Paper (CFP) through established academic venues (e.g., WikiCFP, OpenReview mailing lists), major community forums, and relevant Google Groups; (3) engage affinity organizations and mailing lists (e.g., WiML, LatinX in AI, Black in AI, Queer in AI, MusIML) to broaden inclusivity; (4) coordinate announcement waves via organizers' and speakers' networks to compound reach; and (5) collaborate with local universities and institutes to invite students, faculty, and industry labs. To maximize accessibility and engagement, the workshop will be hybrid with live Q&A for virtual participants, captioned streams, and post-event availability of slides and recordings; accepted papers will appear on the workshop site with an option to post to arXiv acknowledging the event.

Access, Broadcast, and Promotion. We will maintain a dedicated website with rolling updates (speakers, program, CFP timeline), run staged social campaigns (save-the-date → CFP → speaker reveals → program highlights), and encourage co-promotion by speakers/organizers. The event will be livestreamed for remote attendees with interactive Q&A and captioning; talk slides and recordings will be posted promptly after the workshop to extend impact beyond the conference.

7 DIVERSITY COMMITMENT

Diversity of Topics. Our workshop embraces the full breadth of research directions required to advance lifelong agents, ensuring that diverse methodological perspectives are represented. The program spans natural language processing, reinforcement learning, embodied AI, multimodal reasoning, alignment and safety, optimization, and large-scale system building. This breadth reflects our commitment to fostering innovation by encouraging dialogue between traditionally isolated communities. By uniting perspectives from symbolic reasoning, deep learning, human-computer interaction, and scientific discovery domains, we aim to cultivate a holistic agenda that inspires participants to think beyond subfields and identify shared principles that can shape future agentic AI.

Diversity of Organizers. Our organizing committee represents a wide range of institutions and backgrounds, including the University of Illinois Urbana-Champaign, University of Edinburgh,

University of Oxford, Northwestern University, National Taiwan University, Princeton University, and industry like Oumi and Salesforce. The team includes researchers at multiple career stages: from senior PhD students, postdoctoral researchers, to assistant, associate, and full professors. Importantly, the team is gender-balanced and geographically diverse, with members spanning North America, Europe, and Asia. This diversity ensures that the workshop benefits from a broad set of perspectives in its design and execution, and models inclusivity for the next generation of researchers.

Diversity of Speakers. We are proud to host an exceptionally diverse lineup of invited speakers and panelists from both academia and industry, including Ohio State University, Meta AI, Mila/McGill University, UC Berkeley, Stanford, Carnegie Mellon University, and Oumi PBC. The speaker and panelist team includes senior faculty, early-career professors, and industry leaders, with representation across genders, races, and ethnicities. By combining thought leaders in open-source AI, foundational models, reinforcement learning, commonsense reasoning, and governance, the workshop provides not only technical depth but also visible role models for a wide spectrum of researchers. Their varied experiences will enrich the event and highlight the importance of inclusive excellence in AI research.

Diversity of Participants. We are committed to cultivating a vibrant and inclusive participant base. Our call for papers will explicitly encourage submissions from researchers in underrepresented groups and from diverse subfields, including reinforcement learning, data science, vision, and scientific applications (e.g., biochemistry, physics, climate science). We will also conduct outreach to affinity groups such as WiML, LatinX in AI, Black in AI, Queer in AI, and MusIML, alongside local institutions to engage students and early-career researchers. To lower barriers to entry, we ensure accessibility for virtual participants through livestreams and online archiving of materials. By welcoming participants from varied cultural, disciplinary, and demographic backgrounds, we aim to create an environment where diverse voices are heard, connections are fostered, and future leaders in agentic AI are empowered.

8 LLM USAGE POLICY

In line with ICLR 2026 guidelines, our workshop permits the use of large language models (LLMs) only as supportive tools (e.g., for language polishing, formatting, or coding assistance) but not as primary authors or reviewers. All submitted work must be fundamentally human-authored, with AI-generated papers strictly prohibited, particularly in the short and tiny paper tracks. Organizers may use LLMs for clerical support (e.g., drafting communication or summaries), but all substantive reviewing and decision-making will be performed by human experts with full oversight. This ensures integrity, originality, and fairness in both submissions and the workshop process.

9 PREVIOUS AND RELATED WORKSHOPS

Related Events. Our workshop builds on a rich sequence of recent efforts while carving out its own focus. The ICLR 2024 Workshop on Large Language Models for Agents concentrated on how LLMs can act as agents through tool use, planning, and interaction with external systems. The ICLR 2025 Workshop on Reasoning and Planning for LLMs targeted reasoning and decision-making skills within LLMs, exploring how reinforcement learning and post-training can improve structured planning. The NeurIPS 2024 Workshop on Scalable Continual Learning for Lifelong Foundation Models emphasized continual adaptation and memory efficiency for foundation models across different modalities. The ICLR 2025 World Models workshop focused on model-based representations of the world and their applications to robotics, video, and science. Long-running efforts such as the Embodied AI Workshop at CVPR and the NeurIPS 2021 Robot Learning Workshop have pushed advances in embodiment, self-supervised learning, and sim-to-real transfer. For continuity, our team also includes organizers with leadership in these adjacent spaces, including CVPR 2025 Foundation Models Meet Embodied Agents (Manling Li) and ICML workshops on Trustworthy Multimodal Foundation Models and AI Agents (Zhenfei Yin).

How this workshop stands out. While these prior workshops each focused on a specific dimension (reasoning and planning, continual learning, world models, or embodiment) our event centers on the **integrated challenge of lifelong agents**. We bring together the full cycle of *learning across*

time, maintaining alignment while adapting, and evolving capabilities under real-world resource constraints, with evaluation and governance built in from the start. Concretely, we (1) make alignment over long lifespans a first-class theme, emphasizing evaluation of preference drift and oversight beyond static benchmarks; (2) connect NLP, RL, multi-agent, and embodied communities through lifelong learning use cases; (3) highlight the role of open-source platforms and scalable infrastructure in making persistent agents reproducible and deployable; and (4) dedicate nearly half of the program to structured interaction (poster sessions, orals, panel) along with a short/tiny paper track to elevate early-career and under-resourced researchers. This integrated framing of lifelong learning $\rightarrow$ alignment $\rightarrow$ self-evolution, measured over extended horizons, moves beyond past debates and toward shared principles, guidelines, and reproducible practices for building truly persistent agentic systems.

REFERENCES

Emre Can Acikgoz, Cheng Qian, Hongru Wang, Vardhan Dongre, Xiusi Chen, Heng Ji, Dilek Hakkani-Tür, and Gokhan Tur. A desideratum for conversational agents: Capabilities, challenges, and future directions. *arXiv preprint arXiv:2504.16939*, 2025.

Peter Belcak, Greg Heinrich, Shizhe Diao, Yonggan Fu, Xin Dong, Saurav Muralidharan, Yingyan Celine Lin, and Pavlo Molchanov. Small language models are the future of agentic ai. *arXiv preprint arXiv:2506.02153*, 2025.

Yoshua Bengio, Sören Mindermann, Daniel Privitera, Tamay Besiroglu, Rishi Bommasani, Stephen Casper, Yejin Choi, Philip Fox, Ben Garfinkel, Danielle Goldfarb, et al. International ai safety report. *arXiv preprint arXiv:2501.17805*, 2025.

Huanang Gao and Jiayi Geng et al. A survey of self-evolving agents: On path to artificial super intelligence, 2025. URL `https://arxiv.org/abs/2507.21046`.

Komal Kumar, Tajamul Ashraf, Omkar Thawakar, Rao Muhammad Anwer, Hisham Cholakkal, Mubarak Shah, Ming-Hsuan Yang, Phillip HS Torr, Fahad Shahbaz Khan, and Salman Khan. Llm post-training: A deep dive into reasoning large language models. *arXiv preprint arXiv:2502.21321*, 2025.

Shuang Li, Xavier Puig, Chris Paxton, Yilun Du, Clinton Wang, Linxi Fan, Tao Chen, De-An Huang, Ekin Akyürek, Anima Anandkumar, et al. Pre-trained language models for interactive decision-making. *Advances in Neural Information Processing Systems*, 35:31199–31212, 2022.

Bang Liu, Xinfeng Li, Jiayi Zhang, Jinlin Wang, Tanjin He, Sirui Hong, Hongzhang Liu, Shaokun Zhang, Kaitao Song, Kunlun Zhu, et al. Advances and challenges in foundation agents: From brain-inspired intelligence to evolutionary, collaborative, and safe systems. *arXiv preprint arXiv:2504.01990*, 2025.

Junyu Luo, Weizhi Zhang, Ye Yuan, Yusheng Zhao, Junwei Yang, Yiyang Gu, Bohan Wu, Binqi Chen, Ziyue Qiao, Qingqing Long, et al. Large language model agent: A survey on methodology, applications and challenges. *arXiv preprint arXiv:2503.21460*, 2025.

Grégoire Mialon, Roberto Dessi, Maria Lomeli, Christoforos Nalmpantis, Ramakanth Pasunuru, Roberta Raileanu, Baptiste Roziere, Timo Schick, Jane Dwivedi-Yu, Asli Celikyilmaz, Edouard Grave, Yann LeCun, and Thomas Scialom. Augmented language models: a survey. *Transactions on Machine Learning Research*, 2023. ISSN 2835-8856. URL `https://openreview.net/forum?id=jh7wH2AzKK`. Survey Certification.

Cheng Qian, Zuxin Liu, Akshara Prabhakar, Jielin Qiu, Zhiwei Liu, Haolin Chen, Shirley Kokane, Heng Ji, Weiran Yao, Shelby Heinecke, et al. Userrl: Training interactive user-centric agent via reinforcement learning. *arXiv preprint arXiv:2509.19736*, 2025.

Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, YK Li, et al. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. *arXiv preprint arXiv:2402.03300*, 2024.

David Silver and Richard S Sutton. Welcome to the era of experience.