ICLR 2026 Workshop on AI with Recursive Self-Improvement

1 Workshop Summary

Recursive self-improvement (RSI) is moving from thought experiments to deployed AI systems: LLM agents now rewrite their own codebases or prompts (Yang et al., 2023; Soni et al., 2024; Zhuge et al., 2024a), scientific discovery pipelines schedule continual fine-tuning (Starace et al., 2025), and robotics stacks patch controllers from streaming telemetry (Ghasemipour et al., 2025), and even improve the product-level codes (Novikov et al., 2025).

The **ICLR 2026 Workshop on AI with Recursive Self-Improvement** brings together researchers to discuss a simple question with big consequences: how do we build the algorithmic foundations for powerful and reliable self-improving AI systems? As loops that update weights, rewrite prompts, or adapt controllers move from labs into production, we will surface the methods that work—how to design, evaluate, and govern these loops without hand-waving. This workshop examines algorithms for self-improvement across experience learning (Silver and Sutton, 2025), synthetic data pipelines, multimodal agentic systems, weak-to-strong generalization, and inference-time scaling, and will discuss and refine methods for recursive self-improvement. In short, we care about loops that actually get better—and can show it.

To give the workshop a clear spine, we organize contributions around five lenses: change targets inside the system, temporal regime of adaptation, mechanisms and drivers, operating contexts, and evidence of improvement. This framing synthesizes recent perspectives on self-evolving agents while grounding them in practical, auditable deployment settings. We are paradigm-agnostic: we welcome work on foundation models, agent frameworks, robots, and also learning algorithms and optimizers, control and program synthesis, as well as data/infra systems and evaluation tooling that enable recursive self-improvement.

2 Motivation and Objectives

Recent advances in LLMs have turned RSI from a thought experiment into an engineering reality. Training loops that modify their own prompts, objectives, or weights are now integral to foundation-model development and a range of downstream applications.

Current achievements and bottlenecks of RSI. Yet what constitutes self-improvement remains conceptually unsettled. For example, many current systems emphasize in-benchmark optimization rather than genuine, open-ended improvement through interaction with dynamic environments. The field still lacks shared criteria for distinguishing transient performance gains from true advances in general capability and autonomy. Meanwhile, practical know-how is scattered across papers and internal reports. By unifying emerging criteria and encouraging collective exploration, this area now faces a major opportunity for progress.

We see three recurring blockers to principled RSI research:

- the absence of clear definitions for improvement operators, leaving open what constitutes genuine self-improvement versus task-specific optimization;
- limited theoretical understanding of how improvement loops interact with resource constraints—compute, data, and feedback—which often leads to brittle or unsustainable progress; and
- fragmented evaluation practices, with little consensus on how to measure improvement beyond benchmark performance or to document failures and regressions.

Foundations and Current Landscape of Recursive Self-Improvement. Addressing these blockers calls for connecting the foundations of online and continual learning—streaming updates, hierarchical adaptation between fast and slow weights, and instrumented evaluation loops—building on early ideas such as OOPS (Schmidhuber, 2004) and the Gödel Machine (Schmidhuber, 2009; Schaul and Schmidhuber, 2010), as well as recent automated-evaluation pipelines (Zhuge et al., 2024b; Starace et al., 2025). Together, these ingredients point toward a unified view of RSI as measurable adaptation under uncertainty, where improvement is not assumed but demonstrated. In today's **LLM-heavy landscape**, RSI manifests as self-editing language agents, critique-guided test-time improvement, and open-ended exploration in virtual or embodied environments (Guo et al., 2025; Lee et al., 2025; Shinn et al., 2024; Wang et al., 2023). Yet these systems sustain genuine self-improvement only when their feedback loops are carefully instrumented—rewards must be logged in real time, adaptations must stay within guardrails, and memories must remain auditable (Best Ideas Collective, 2025; Zhong et al., 2024).

Concretely, we scope RSI along five complementary axes: change targets (parameters, memory/context, tools/workflows, or agent architecture), adaptation timing (within a task episode versus between tasks or deployments), adaptation mechanisms (e.g., reward- or critique-based, imitation/demonstration, or population/evolutionary search), operating contexts (web/UI, embodied robotics, scientific discovery, or enterprise stacks), and evidence and assurance beyond benchmarks (long-horizon stability, regression risk, safety and alignment). We will highlight exemplar methods across these axes—from critique-guided loops and iterative refinement (Shinn et al., 2024; Madaan et al., 2023) to online curricula and web agents (Qi et al., 2024; Xie et al., 2024)—as testbeds for principled, auditable self-improvement.

Focus and Approach. This workshop aims to crystallize the emerging science of recursive self-improvement—how learning systems rewrite, test, and govern their own updates. We expand RSI beyond narrow feedback loops to include algorithmic, representational, and collective forms of adaptation: models that refine internal objectives, agents that coordinate across modalities and time horizons, and research ecosystems that continuously incorporate new data, code, and evaluation signals. Discussion will revolve around three interlocking layers of an RSI loop: (1) improvement operators, mechanisms that transform feedback into stable updates under compute and safety constraints; (2) intrinsic diagnostics, metrics and probes that quantify whether learning continues or stalls; and (3) governed adaptation, frameworks that ensure human-aligned oversight, rollback triggers, and memory management. Rather than treating these as engineering details, the workshop treats them as the algorithmic substrate of self-developing intelligence, linking online learning, meta-optimization, and embodied reasoning into one continuum.

Expected Outcomes. RSI research now spans language models, reinforcement learning, robotics, and scientific discovery. Loops that rewrite prompts, weights, or hypotheses already operate inside foundation-model pipelines and embodied agents, yet their behavior remains poorly characterized. Evaluation, safety, and governance tools lag behind algorithmic progress, leaving unclear whether systems improve, drift, or merely overfit benchmarks. This workshop responds by establishing the conceptual and empirical foundations for reliable self-improvement—drawing from continual and online learning, adaptive control, meta-optimization, and evaluation science. Sessions will define reproducible baselines for improvement cycles, propose intrinsic metrics that distinguish genuine capability gain from optimization artifacts, and share deployment patterns that align autonomous adaptation with accountability. Building on these results, the community will release open assets—evidence cards, intrinsic-metric kits, safety checklists, and instrumentation baselines—while connecting theory, empirical studies, and real-world deployments. By bridging algorithmic insight, measurement practice, and governance, this workshop positions RSI as a unifying frontier across machine learning disciplines: how learning systems learn to improve themselves, safely and verifiably.

3 Scope and Topics of Interest

We welcome contributions across algorithms, models, and systems—not only agents. This includes optimization methods, core algorithms, reinforcement learning, LLM training and post-training, systems/infra, and evaluation. The list below is illustrative rather than exhaustive; cross-cutting work is encouraged.

Core Dimensions These lenses help position a contribution without constraining methodology. Authors can briefly map their work onto one or more dimensions—change targets, adaptation timing, mechanisms and drivers, operating contexts, and evidence/safeguards—to make scope, evaluation, and comparisons across submissions clearer.

- **Change targets:** updates to model parameters and adapters; memories and retrieval; tools/workflows and codebases; agent architectures and coordination; plus surrounding optimization/control systems (optimizers, schedulers, data engines, evaluation tooling).
- **Adaptation timing:** within-episode, between-episode, deployment-time, and pre-/post-training cycles; in-context shaping, supervised updates, and RL in offline/online regimes (Guo et al., 2025).
- **Mechanisms and drivers:** reward- and critique-driven signals, imitation/demonstration, bandits/active learning, and population/evolutionary search; with diagnostics for stability and regressions (Shinn et al., 2024; Madaan et al., 2023).
- **Operating contexts:** web/UI automation (Qi et al., 2024), embodied/robotics (Wang et al., 2023), scientific discovery, and enterprise.
- **Evidence and safeguards:** long-horizon and regression-aware metrics, reliability under shift, failure-aware evidence, and governance/guardrails; OS-scale environments and community suites (Xie et al., 2024; MLCommons Association, 2025).

Representative Topics The items below illustrate typical fits rather than strict boundaries. Submissions may focus on a single area or combine several (e.g., optimization + RL + systems). We welcome both theoretical contributions and practice-driven case studies when claims are backed by clear, reproducible evidence.

- **Optimization and algorithms:** online/streaming optimization, meta-optimization, curricula/schedulers, long-horizon credit assignment, gradient-free and population-based methods.
- **Learning paradigms:** supervised/instruction tuning, unsupervised post-training, RL (single/multi-agent; offline/online), bandits/active learning; critique/reward-driven training.
- **LLM training and post-training:** preference/alignment methods, adapters/LoRA, inference-time scaling, and specification-driven decoding.
- **Test-time adaptation:** in-context shaping, memory edits, self-editing prompts/tools, iterative refinement and deliberate reasoning (Madaan et al., 2023; Yao et al., 2023; Shinn et al., 2024).
- **Memory and knowledge:** retrieval augmentation, persistent memory stores, long-term context management, knowledge editing and unlearning.
- **Systems and infrastructure:** compilers/schedulers, distributed training/evaluation, data pipelines and monitoring, rollback and safety gating.
- **Evaluation and benchmarks:** long-horizon/regression-aware metrics, reliability and human factors; OS/web/embodied suites (Xie et al., 2024; MLCommons Association, 2025; Qi et al., 2024).

Reporting Guidelines These prompts are lightweight and non-prescriptive. Please include what is feasible in your setting; brief notes on constraints are perfectly fine. The goal is to help reviewers and readers interpret self-improvement claims and reproduce the core results with minimal friction.

- **Measurement and diagnostics:** trajectory-level logs (decisions, state summaries, rewards/critique); stability probes across seeds/horizons; critique-to-change attribution; and clear rollback readiness for self-edits (Xie et al., 2024; Qi et al., 2024).
- **Efficiency and data practices:** improvement-per-token/FLOP and wall-clock/energy (when available); data freshness/licensing; provenance for generated/scraped content; de-duplication and unlearning policies where applicable (MLCommons Association, 2025).

4 Prior Related Workshops and Points of Difference

Recent workshops such as World Models @ ICLR 2025, LLM Agents and Tool Use @ ICLR 2025, and NeurIPS 2024's "Autonomous Discovery" series have explored adjacent themes, from world-model-driven planning to multimodal agent orchestration and automated evaluation. The ICLR 2025 workshop on "Scaling Self-Improving Foundation Models" further underscored synthetic data pipelines for large models, spotlighting data scarcity and pre-training recipes. These events typically emphasized single modalities, frontier-model scale, or algorithmic knobs without detailing deployment instrumentation and governance.

Our workshop differs in five concrete ways:

- **Evidence across the entire loop.** Every contribution must provide improvement-operator cards, replayable artifacts, and governance checklists so reported gains are auditable.
- **Cross-domain comparisons.** We convene language, vision, robotics, scientific discovery, and policy communities to analyse shared self-improvement mechanisms instead of siloed application reports.
- **Instrumentation-first programming.** Lightning failure clinics, artifact labs, and compliance-aligned review templates ensure deployment instrumentation and safety practice remain central, filling gaps left by workshops focused mainly on synthetic data scaling.
- **Taxonomy-driven scope.** The program is explicitly organized around change targets, adaptation timing, mechanisms, operating contexts, and evidence dimensions (Guo et al., 2025), enabling principled comparisons across methods and domains rather than ad-hoc collections of case studies.
- **Paradigm-agnostic remit.** We welcome algorithms, models, and systems work—not only agent frameworks—including classical ML, control, program synthesis, distributed infrastructure, and evaluation tooling that contribute to recursive self-improvement.

5 Invited Speakers and Panelists (alphabet rank)

We will first consider speakers who can attend on-site for a better audience experience, and note that both on-site and online participation are workable for panelist sessions.

Our speakers bridge theory, large-scale deployments, and governance. They span academic research on self-improving methods and industry-scale deployments; for example, the lineup includes Google DeepMind researchers behind algorithm-discovery agents such as AlphaEvolve (Novikov et al., 2025):

- **Arman Cohan (Yale University) (he/him), confirmed** is an Assistant Professor in the Department of Computer Science at Yale University, where he leads the Yale NLP Lab. Before joining Yale, he was a Research Scientist at the Allen Institute for AI (AI2). His research lies at the intersection of natural language processing and machine learning, focusing on language modeling, representation learning, information retrieval, and applications of NLP in specialized domains. His group's work on representation, retrieval, and model adaptation informs methods for continual adaptation and self-improving systems.
- **Bang Liu (Université de Montréal) (he/him), confirmed** is an Assistant Professor in the Department of Computer Science and Operations Research (DIRO) at the Université de Montréal, a core member of the RALI Lab for Applied Research in Computational Linguistics, and an associate academic member of Mila – Quebec AI Institute, where he also holds a Canada CIFAR AI Chair. He received his BEng from the University of Science and Technology of China in 2013, and his MSc and PhD from the University of Alberta in 2015 and 2020, respectively. His research spans natural language processing, multimodal and embodied learning, the theory and techniques of artificial general intelligence—particularly understanding and improving large language models—and the use of AI for scientific domains such as health, materials science, and extended reality (XR). His recent work on foundation agents and LLM improvement connects directly to recursive self-improvement.
- **Chelsea Finn (Stanford) (she/her), confirmed** is an Assistant Professor of Computer Science and Electrical Engineering at Stanford University and co-founder of Pi. She leads the IRIS Lab, which studies intelligence through large-scale robotic interaction and is affiliated with the Stanford AI Lab (SAIL) and the Machine Learning Group. Her research focuses on enabling robots and other agents to develop broadly intelligent behavior through learning and interaction. Finn received her PhD in Computer Science from the University of California, Berkeley, and her BS in Electrical Engineering and Computer Science from MIT. Before joining Stanford, she was a Research Scientist at Google Brain (now DeepMind). She is the recipient of numerous honors, including the 2025 Presidential Early Career Award for Scientists and Engineers (PECASE), the Sloan Research Fellowship, and the NSF CAREER Award, among others. Her group's large-scale interactive learning in robotics provides core testbeds for continual adaptation and self-improving agents.
- **Graham Neubig (Carnegie Mellon University) (he/him), confirmed** is an Associate Professor at the Language Technologies Institute in the School of Computer Science at Carnegie Mellon University and Chief Scientist at All Hands AI, where he develops AI agents for software development. He leads NeuLab, a research group focused on machine learning and natural language processing. His research centers on large language models, with particular emphasis on question answering, code generation, multilingual processing, and evaluation and interpretability. His recent work on code-centric agents and evaluation provides practical loops for self-improving developer systems.
- **Matej Balog (Google DeepMind) (he/him), confirmed** is a Staff Research Scientist at Google DeepMind in the AI for Science unit, where he and his team build AI agents that discover new algorithms for solving complex computational problems. He has contributed to major projects including DeepCoder, AlphaTensor, FunSearch, and AlphaEvolve (Novikov et al., 2025)—a next-generation LLM-based coding agent that has produced novel scientific discoveries in mathematics and computer science, as well as algorithms now deployed at Google in Gemini, Transformer models, TPU hardware design, and data-center optimization. Balog received his PhD from the University of Cambridge and his MSc in Mathematics and Computer Science from the University of Oxford. His work explores AI systems capable of advancing both theoretical science and large-scale engineering, aiming to create agents that continually expand the frontier of algorithmic discovery.
- **Yu Su (Ohio State University) (he/him), confirmed** is an Associate Professor and Innovation Scholar in the Department of Computer Science and Engineering at The Ohio State University, where he co-directs the OSU NLP Group, co-leads the Foundational AI Team in the ICICLE AI Institute, and leads the Machine Learning Foundations Team in the Imageomics Institute. He received his PhD in Computer Science from the University of California, Santa Barbara, and a BSc from Tsinghua University, and previously worked as a researcher at Microsoft Semantic Machines. A 2025 Sloan Research Fellow, he has received multiple Best Paper and Outstanding Paper Awards from CVPR and ACL. His research

Table 1 Tentative schedule (subject to change).

Morning

Time	Planned Event
08:45–09:00	Opening Remarks
09:00–09:30	Invited Talk 1
09:30–10:00	Invited Talk 2
10:00–11:00	Poster Session 1
11:00–11:30	Invited Talk 3
11:30–12:00	Invited Talk 4
12:00–13:30	Break

Afternoon

Time	Planned Event
13:30–14:00	Invited Talk 5
14:00–14:45	Contributed Talks (×3)
14:45–15:45	Poster Session 2
15:45–16:15	Invited Talk 6
16:15–16:45	Contributed Talks (×2)
16:45–17:45	Panel Discussion
17:45–18:00	Closing Remarks

broadly explores artificial intelligence, particularly the role of language as a medium for reasoning, communication, and planning in AI agents. His current interests include language-based agents, multimodal reasoning and grounding, planning and world models, continual learning and memory, evaluation benchmarks, and AI for scientific discovery. These topics tie directly to memory, planning, and continual self-improvement.

- **Yuandong Tian (Meta Superintelligence) (he/him), confirmed** is a Research Scientist Director at Meta FAIR, where he leads research on decision-making, reinforcement learning, planning, and the theoretical understanding of large language models. He has led or mentored projects including OpenGo, an efficient single-GPU reproduction of AlphaZero; StreamingLLM and GaLore, which improve LLM training and inference; and Coconut, exploring continuous latent reasoning. He co-led reasoning efforts for Llama 4 during his time at Meta GenAI. Tian received his PhD from the Robotics Institute at Carnegie Mellon University and previously worked on the Google Self-Driving Car team. His work has earned honors such as the ICML 2021 Outstanding Paper Honorable Mention, the ICCV 2013 Marr Prize Honorable Mention, and the CGO 2022 Distinguished Paper Award, and he has served as Area Chair for NeurIPS, ICML, ICLR, AAAI, CVPR, and AISTATS. His research advances optimization and reasoning mechanisms central to self-improving model behavior.

Speakers will remain on site during poster blocks and the governance panel to critique artifacts, and we will announce additional industry panelists (policy, safety, evaluation) in the CFP supplement to ground discussion in deployed systems. Come ready to share both wins and rough edges.

Interaction is built into three anchors:

- **Lightning Failure Clinic (14:00–14:45).** Spotlight talks on negative results followed by moderated cross-review and live polling to surface follow-up work.
- **Poster and Artifact Labs (10:00–11:00, 14:45–15:45).** Poster sessions double as instrumentation pods where authors demo loops, release operator cards, and take hybrid Q&A.
- **Governance and Deployment Roundtable (16:45–17:45).** Structured prompts and breakout questions drive commitments on safety triggers, rollback policies, and evaluation standards for recursive systems.

Table 1 captures the cadence, pairing technical depth with dialogue-heavy formats that connect theory, deployment, and governance threads.

6 Submission Tracks and Timeline

We align our CFP with ICLR 2026 deadlines while giving contributors enough time for artifact preparation. Submissions open on 15 November 2025 with templates for the improvement-operator card, artifact statement, and governance checklist. We accept three complementary tracks. If your work spans more than one, pick the closest fit and note the crossovers in a sentence:

- **Research Papers** (up to 8 pages + references) introducing new RSI algorithms, intrinsic diagnostics, or governance frameworks backed by replayable evidence.
- **Systems and Deployment Papers** (up to 6 pages) documenting production loops, instrumentation stacks, or evaluation infrastructure with compute- and cost-aware reporting. This includes LLM online/post-training pipelines and data engines that integrate synthetic data generation, filtering, and continual updates to retrieval/memory or adapters, as well as compilers, schedulers, and distributed training/evaluation systems that enable safe, efficient self-improvement.

Table 2 Key dates aligned with ICLR 2026 workshop requirements. We mirror the official ICLR 2026 milestones and document them in advance so organizers, reviewers, and authors can plan. All times are AoE.

Date	Milestone
12 Jan 2026	First mentoring clinic and checklist walkthrough (recording shared for asynchronous access).
30 Jan 2026	Recommended submission date so reviewers receive full papers before the official discussion window.
7 Feb 2026	Final submission deadline (AoE) with artifacts and governance checklist uploaded.
14 Feb 2026	Reviews due; ethics and safety flagging window opens for organizers.
21 Feb 2026	Author rebuttal and discussion close; meta-reviewers consolidate decisions.
24 Feb 2026	Final decisions signed off internally to allow notification preparation.
1 Mar 2026	Authors notified and accepted papers made public on OpenReview, matching the official ICLR requirement.
3 Mar 2026	Camera-ready packages (paper, operator card, artifact links) due for proceedings layout.
11 Mar 2026	Metadata and PDFs imported into iclr.cc and the central ICLR schedule, satisfying the compliance checklist.

- **Tiny Papers / Practice Notes** (2–4 pages) for focused failure analyses, new datasets, evaluation checklists, or intermediate lessons that deserve early feedback.

Each submission packet includes three standardised components to make results auditable and reusable:

- **Improvement-operator card:** objective/feedback definition, operator type, scope of change, stability constraints and rollback triggers, data/compute budgets, and expected failure modes.
- **Artifact statement:** code and scripts with seeds/configs, environment specifications, a minimal reproducer, licensing of assets, and compute/energy reporting sufficient for re-execution.
- **Governance and safety checklist:** misuse risks and mitigations, guardrails/monitoring, data governance, human-in-the-loop/approval gates, and audit logging/retention plans.

Each submission should include the improvement-operator card, artifact statement, and governance checklist; substantial omissions may be grounds for desk rejection. We aim to be helpful rather than bureaucratic—if something doesn't fit these templates, tell us. We will host December and mid-January office-hour clinics on artifact packaging and safety documentation to support newer groups. Our internal timeline matches the official ICLR 2026 milestones; Table 2 calls out the shared **1 March 2026** author notification/public release and the **11 March 2026** iclr.cc import dates on the compliance checklist.

Systems-Track Reporting for LLM/Data Engines. Where submissions focus on foundation-model adaptation pipelines, we ask authors to consider reporting, as applicable: (i) update cadence and gating rules; (ii) sources of data (human, synthetic, programmatic), with provenance/retention policies; (iii) efficiency accounting (tokens/FLOPs per percentage point gain and wall-clock/energy usage); and (iv) safeguards for generated content and rollback-on-regression triggers. If some details are not available, a short note explaining constraints is perfectly fine.

Representative Domains and Task Formats. We particularly welcome reproducible studies in: (i) web/UI automation with curriculum or feedback-driven agents (Qi et al., 2024); (ii) embodied/robotic settings where adaptation impacts control and safety (Wang et al., 2023); and (iii) open-ended scientific or engineering workflows where hypotheses/code evolve across iterations. Submissions should make clear how adaptation timing and mechanisms map onto the task structure. Algorithmic and theoretical contributions that are not tied to a single domain are also in scope when they advance the foundations of recursive self-improvement.

7 Policy on LLM Usage and Tiny/Short Paper Tracks

We will follow the ICLR 2026 LLM policy across every track. Authors must include an LLM usage statement, and the Tiny Papers / Practice Notes track forbids machine-generated prose beyond light copy-editing. Missing disclosures or undisclosed assistive use may lead to desk rejection. Reviewers receive explicit guidance to flag potential violations, and organizers may conduct occasional spot checks using logs, prompt histories, or diffs when authors choose to provide them.

We're not here to catch you out; transparent usage simply helps readers interpret results and reviewers calibrate credit.

For submissions that study LLM-augmented improvement loops, authors must delineate human-authored versus AI-produced artifacts and share prompts or scripts sufficient for audit. Poster and artifact labs allocate time for teams to demonstrate their guardrails, ensuring compliance is observable rather than stated only in text.

Generated-Data Governance and Unlearning. Where adaptation involves generated or scraped data, authors should, where applicable, disclose data sources, licensing, filtering (toxicity/PII), and retention schedules; describe de-duplication and provenance tracking; and outline procedures for data removal/unlearning upon request. For safety-critical domains, we encourage authors to specify criteria for halting updates and rolling back model states when regressions are detected.

Common Failure Modes and Guardrails. We ask authors to highlight known risks in self-improving agents, including: reward/specification hacking, memory drift and stale context, brittle self-edits, and unbounded exploration leading to regressions. Mitigations can include layered approvals for high-impact edits, confidence- or uncertainty-aware update triggers, fallbacks with safe baselines, and structured self-critique pipelines (Shinn et al., 2024; Madaan et al., 2023; Wang et al., 2025).

8 Audience, Access, and Inclusion

Audience and Outreach. We expect roughly 260 in-person attendees and 220 remote participants, extrapolating from ICLR 2025 workshops on world models and agentic LLMs plus current demand for RSI deployments. The audience spans foundation-model researchers, RL and control theorists, robotics teams, evaluation scientists, and policy groups dealing with self-improving systems, with special effort to engage early-career researchers and production engineers. Outreach includes a dedicated website and OpenReview hub amplified through KAUST, MILA and aligned WeChat/Slack lists; teaser seminars and artifact-prep office hours with partner labs; coordinated social-media features (X, LinkedIn, etc) that highlight speakers; and collaboration with ICLR communications to appear in conference newsletters and daily programming.

Inclusion and Support. The organizing team, speakers, and reviewers already span academia, industry, policy, and regions across North America, Europe, the Middle East, and Asia. We will extend coverage by recruiting reviewers from Latin America and Africa and reserving lightning-talk slots for emerging RSI communities. Submission forms may include optional demographic fields to help us understand representation; any use will be limited to high-level aggregates if feasible. We aim to include a mix of academic and industry session chairs and diverse perspectives. We will point attendees to ICLR travel and childcare grants; if sponsorship permits, we will explore a small number of microgrants. For accessibility, we will encourage best practices for artifacts (contrast, alt text, captions where feasible) and request platform captioning when available. The workshop adopts the ICLR code of conduct, shares an anti-harassment statement on the website, and provides a contact for confidential reports.

Hybrid Access and Artifacts. All plenary sessions stream on the ICLR platform with live captioning, and poster sessions run in mirrored physical and GatherTown rooms so remote attendees can join Q&A. Within 72 hours we will upload recordings, slides, improvement-operator cards, and artifact repositories to the workshop site, with DOI-tagged proceedings on OpenReview. An open-licensed "RSI Field Guide" will package insights, failure cases, governance takeaways, session notes, and poll results for reuse by those who miss the live event.

9 Review Process and Evaluation

We will run double-blind review on OpenReview. Each paper receives at least three expert reviews and a meta-review from an organizer or senior PC member. Rubrics cover evidence quality, replayability, governance readiness, and ethical considerations, mirroring the improvement-operator card and artifact checklist. During rebuttal we pair reviewers with authors for targeted spot checks of artifacts, compute/energy claims, and human oversight assertions; flagged ethics or safety issues route to a specialist reviewer.

We encourage accepted papers to include a replication-readiness checklist (code executability, dataset licensing, risk disclosures). After the workshop we will release anonymized review statistics, decision rationales, and the checklists to document evaluation practice for future RSI events.

Replication-Readiness Details. We recommend including: (i) exact commit hashes, configs, and seeds; (ii) minimal scripts for re-running core loops and ablations; and (iii) a red-team note on failure cases/regressions and how they are surfaced by the metrics. When this level of packaging is not feasible, please provide the smallest viable reproducer and a brief note on constraints. Concise, runnable evidence beats heavyweight packaging.

OpenReview's COI system (recent co-authorship, institutional overlap in the last two years, advisor/advisee relationships, and family ties) will be enforced for all organizers, reviewers, and area chairs. The primary matching step will exclude conflicts automatically, and organizers will manually audit edge cases before assignments go out. We cap reviewer load at four papers (three for first-time reviewers) and ensure organizers do not review submissions where they have any perceived conflict. All organizers confirm compliance with ICLR's participation rule (no individual involved in more than two workshops); we will publish a statement on the website and re-verify once the full program is announced. Reviewers must confirm they are not serving on more than three ICLR workshops, and we will coordinate with the workshops chairs if reassignment is needed. Authors may only be lead presenter on one accepted paper to avoid oversubscription of session slots, while co-authorship across tracks remains welcome.

Efficiency and Evolution Reproducibility. As a strength (not a requirement), reviewers will look favorably on evidence of efficiency (e.g., improvement-per-token/FLOP, energy use, wall-clock) and, when applicable, clear population/evolutionary details (mutation/crossover operators, selection criteria, and population logs) that aid reproduction. Authors are encouraged—but not required—to report stability under repeated runs and any rollback events observed during evaluation. These items will be treated as bonus signals rather than gating criteria.

10 Organising Team and Program Committee

The organising committee blends researchers running state-of-the-art RSI systems with policy and governance expertise:

- **Mingchen Zhuge (KAUST) (he/him)** is a PhD candidate advised by Prof. Jürgen Schmidhuber, specializing in multimodal agents and world models, with a focus on training and post-training methods and their applications. He has authored over twenty top-tier publications, which have collectively received more than 3,300 citations, with the majority from first-author work. His research have been selected as oral presentations such as ICLR and ICML and have earned a Best Paper Award at the NeurIPS Ro-FoMo Workshop. Zhuge is also an active open-source contributor to frameworks including MetaGPT, GPTSwarm, OpenDevin (OpenHands), and Agent-as-a-Judge, and was recognized with a WAIC Future Star Nomination (2025) and as an Outstanding Reviewer at CVPR 2023. He contributed the workshop organization of World Models@ICLR 2025, MAS@ICML 2025.
- **Ailing Zeng (Anuttacon) (she/her)** is a technical staff researcher at Anuttacon, dedicated to advancing humanistic general intelligence and the emerging paradigm of AI soul casting. Her research centers on developing multimodal, human-like intelligent agents that integrate perception, reasoning, and creativity. She has served as an Area Chair for CVPR 2024 and 2025 and has published over 50 papers in top-tier conferences such as ICLR, ICML, NeurIPS, CVPR, ICCV, and AAAI with over 8,000 citations. One of her papers on time-series forecasting was selected as the Most Influential Paper of AAAI 2023.
- **Sherry Yang (NYU&DeepMind) (she/her)** is an Assistant Professor of Computer Science at the Courant Institute of Mathematical Sciences, New York University, and a Staff Research Scientist at Google DeepMind. Her research lies at the intersection of machine learning, reinforcement learning, and generative modeling, with a focus on foundation models for decision-making, world modeling, and robotics. She was a postdoctoral researcher at Stanford University working with Prof. Percy Liang, received her Ph.D. from the University of California, Berkeley advised by Prof. Pieter Abbeel, and earned her B.S. and M.Eng. degrees from MIT. Her work UniSim: Learning Interactive Real-World Simulators received the Outstanding Paper Award at ICLR 2024, and she is co-organizing the World Models in Robotics workshop at CoRL 2025.
- **Deyao Zhu (ByteDance) (he/him)** is a Research Scientist at ByteDance. His research explores multimodal large language models, learning from experience, and reinforcement learning with the goal of advancing artificial general intelligence. Zhu focuses on building multimodal language models capable of reasoning and decision-making, with interests spanning video–language understanding, planning via large language models, motion forecasting, and related deep learning approaches.
- **Vikas Chandra (Meta Reality Labs) (he/him)** leads an applied AI research and engineering organization at Meta Reality Labs, where his team develops multimodal and generative AI systems, including on-device large language models and vision-language models. His work spans the full machine learning stack—from algorithms and architectures to model optimization and deployment—enabling advanced AI use cases across Meta's products. Chandra received his PhD in Computer Engineering from Carnegie Mellon University and has served as Visiting Scholar (2011–2014) and Visiting Faculty (2016–2017) at Stanford University. He has authored over 125 research publications, holds more than 50 US and international patents, and is a recipient of the ACM-SIGDA Technical Leadership Award. He is a Senior Member of IEEE and was invited to the US National Academy of Engineering's Frontiers of Engineering Symposium in 2017.
- **Jürgen Schmidhuber (KAUST / IDSIA) (he/him)** is a pioneering computer scientist widely recognized for foundational contributions to modern deep learning and artificial intelligence. Since his teenage years, he has pursued the goal of creating a self-improving AI more intelligent than himself. His laboratory's work on neural networks in the early 1990s—including the development of Long Short-Term Memory (LSTM) and other architectures—revolutionized machine learning. These methods became the backbone of applications such as speech recognition, machine translation, and computer vision, eventually running on billions of devices worldwide. Schmidhuber's research also introduced key ideas

Table 3 Pool of Candidate Reviewers.

Name	Affiliation / Position	Institution
Wenyi Wang	PhD Candidate	KAUST
Haozhe Liu	PhD Candidate	KAUST
Dylan R. Ashley	PhD Candidate	IDSIA
Shuming Liu	PhD Candidate	KAUST
Tenglong Ao	PhD Candidate	PKU
Xuan Ju	PhD Candidate	CUHK
Zhiyang Dou	PhD Candidate	MIT
Jing Lin	PhD Candidate	NTU
Yue Ma	PhD Candidate	HKUST
Yuxuan Bian	PhD Candidate	CUHK
Tianhe Ren	PhD Candidate	HKU
Yuhang Yang	PhD Candidate	USTC
Junming Che	PhD Candidate	HKUST
Shunlin Lu	PhD Candidate	CUHK(SZ)
Zhiyu Zhao	PhD Candidate	Bristol
David Fox	PhD Candidate	Bristol
Xiangning Yu	PhD Candidate	TJU
Sam Bowyer	PhD Candidate	Bristol
Jiefeng Li	Research Scientist	Nvidia Research
Guochen Gorden Qian	Research Scientist	Snap Research
Jun Chen	Research Scientist	Meta
Jie Yang	Research Scientist	Tencent
Junshu Tang	Research Scientist	Tencent
Xin Chen	Research Scientist	ByteDance
Siyuan Huang	Research Scientist	BIGAI
Yu Li	Principal Researcher	IDEA
Xihui Liu	Assistant Professor	HKU
Xiaoguang Han	Assistant Professor	CUHK(SZ)
Yong-Lu Li	Associate Professor	SJTU
Ziwei Liu	Associate Professor	NTU
Ruimao Zhang	Associate Professor	Sun Yat-sen University

behind generative and self-supervised learning, including artificial curiosity, generative adversarial networks (1990), self-supervised pre-training (1991), and early versions of transformer-like models. He has authored more than 400 peer-reviewed papers and co-founded multiple AI companies. His formal theory of creativity, curiosity, and fun connects computational learning with art and scientific discovery. Schmidhuber is currently Co-Chair of the KAUST Center for Generative AI and Scientific Director of the Swiss AI Lab IDSIA, among other positions. His numerous honors include the IEEE Neural Networks Pioneer Award, the Helmholtz Award, and the Steiger Award, recognizing his sustained influence on the theory and practice of artificial intelligence.

The pool of candidate reviewers is listed in Table 3.

References

Best Ideas Collective. Online learning as the next llm paradigm. https://example.com/best-ideas-online-learning, July 2025. Accessed: 2025-07-20.

Seyed Kamyar Seyed Ghasemipour, Ayzaan Wahid, Jonathan Tompson, Pannag Sanketi, and Igor Mordatch. Self-improving embodied foundation models. arXiv preprint arXiv:2509.15155, 2025.

Tianyu Guo, Kexin Chen, Yixuan Li, et al. A survey of self-evolving agents: On path to artificial super intelligence. arXiv preprint arXiv:2507.21046, 2025.

Soobin Lee, Minji Park, Hao Zhang, Yanqi Zhou, and Chen Sun. Enabling self-improving agents to learn at test time with critique-guided signals. arXiv preprint arXiv:2507.17131, 2025.

Aman Madaan, Niket Tandon, Prakhar Gupta, Skyler Hallinan, Luyu Gao, Sarah Wiegreffe, Uri Alon, Nouha Dziri, Shrimai Prabhumoye, Yiming Yang, Shashank Gupta, Bodhisattwa Prasad Majumder, Katherine Hermann, Sean Welleck, Amir Yazdanbakhsh, and Peter Clark. Self-refine: Iterative refinement with self-feedback. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, editors, Advances in Neural Information Processing Systems, volume 36, pages 46534–46594. Curran Associates, Inc., 2023.

MLCommons Association. Mlcommons announces agent reliability evaluation suite (ares). https://mlcommons.org/2025/06/ares-announce/, 2025. Accessed: 2025-08-10.

Alexander Novikov, Ngân Vũ, Marvin Eisenberger, Emilien Dupont, Po-Sen Huang, Adam Zsolt Wagner, Sergey Shirobokov, Borislav Kozlovskii, Francisco JR Ruiz, Abbas Mehrabian, et al. Alphaevolve: A coding agent for scientific and algorithmic discovery. arXiv preprint arXiv:2506.13131, 2025.

Zehan Qi, Xiao Liu, Iat Long Iong, Hanyu Lai, Xueqiao Sun, Jiadai Sun, Xinyue Yang, Yu Yang, Shuntian Yao, Wei Xu, Jie Tang, and Yuxiao Dong. Webrl: Training llm web agents via self-evolving online curriculum reinforcement learning. arXiv preprint arXiv:2411.02337, 2024.

Tom Schaul and J"urgen Schmidhuber. Metalearning. Scholarpedia, 5(6):4650, 2010.

J"urgen Schmidhuber. Optimal ordered problem solver. Machine Learning, 54:211–254, 2004.

J"urgen Schmidhuber. Ultimate cognition à la g"odel. Cognitive Computation, 1(2):177–193, 2009.

Noah Shinn, Federico Cassano, Ashwin Gopinath, Karthik Narasimhan, and Shunyu Yao. Reflexion: Language agents with verbal reinforcement learning. Advances in Neural Information Processing Systems, 36, 2024.

David Silver and Richard S. Sutton. Welcome to the era of experience. Preprint for Designing an Intelligence, MIT Press, 2025. Accessed 2025-02-01.

Achint Soni, Sreyas Venkataraman, Abhranil Chandra, Sebastian Fischmeister, Percy Liang, Bo Dai, and Sherry Yang. Videoagent: Self-improving video generation. arXiv preprint arXiv:2410.10076, 2024.

Giulio Starace, Oliver Jaffe, Dane Sherburn, James Aung, Jun Shern Chan, Leon Maksin, Rachel Dias, Evan Mays, Benjamin Kinsella, Wyatt Thompson, et al. Paperbench: Evaluating ai's ability to replicate ai research. arXiv preprint arXiv:2504.01848, 2025.

Guanzhi Wang, Yuqi Xie, Yunfan Jiang, Ajay Mandlekar, Chaowei Xiao, Yuke Zhu, Linxi Fan, and Anima Anandkumar. Voyager: An open-ended embodied agent with large language models. arXiv preprint arXiv:2305.16291, 2023.

Zihan Wang, Kangrui Wang, Qineng Wang, Pingyue Zhang, Linjie Li, Zhengyuan Yang, Xing Jin, Kefan Yu, Minh Nhat Nguyen, Licheng Liu, Eli Gottlieb, Yiping Lu, Kyunghyun Cho, Jiajun Wu, Li Fei-Fei, Lijuan Wang, Yejin Choi, and Manling Li. Ragen: Understanding self-evolution in llm agents via multi-turn reinforcement learning. arXiv preprint arXiv:2504.20073, 2025.

Tianbao Xie, Danyang Zhao, Kai Yu, Uri Alon, Mark Riedl, Karthik Narasimhan, Tatsunori Hashimoto, et al. Osworld: Benchmarking multimodal agents for open-ended tasks in real computer environments. arXiv preprint arXiv:2404.07972, 2024.

Chengrun Yang, Xuezhi Wang, Yifeng Lu, Hanxiao Liu, Quoc V Le, Denny Zhou, and Xinyun Chen. Large language models as optimizers. In The Twelfth International Conference on Learning Representations, 2023.

Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Tom Griffiths, Yuan Cao, and Karthik Narasimhan. Tree of thoughts: Deliberate problem solving with large language models. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, editors, Advances in Neural Information Processing Systems, volume 36, pages 11809–11822. Curran Associates, Inc., 2023.

Wanjun Zhong, Lianghong Guo, Qiqi Gao, He He, and Youyang Qiao. Memorybank: Enhancing large language models with long-term memory. arXiv preprint arXiv:2305.10250, 2024.

Mingchen Zhuge, Wenyi Wang, Louis Kirsch, Francesco Faccio, Dmitrii Khizbullin, and Jürgen Schmidhuber. Gptswarm: Language agents as optimizable graphs. In Forty-first International Conference on Machine Learning, 2024a.

Mingchen Zhuge, Changsheng Zhao, Dylan Ashley, Wenyi Wang, Dmitrii Khizbullin, Yunyang Xiong, Zechun Liu, Ernie Chang, Raghuraman Krishnamoorthi, Yuandong Tian, et al. Agent-as-a-judge: Evaluate agents with agents. arXiv preprint arXiv:2410.10934, 2024b.