Adaptation of Agentic AI: A Unified Framework for Enhancing Autonomous Systems

A CHAPEAUX NOTE

By Dr. Syed Muntasir Mamun

In the dynamic field of artificial intelligence as of late 2025, agentic AI represents a significant advancement, utilizing foundation models such as large language models to develop autonomous systems that can perceive environments, plan actions, employ external tools, and carry out multi-step tasks. These systems, grounded in models from series like GPT or multimodal equivalents, exhibit impressive capabilities across areas from scientific exploration to routine automation. Nevertheless, agentic AI encounters ongoing obstacles, including inconsistent tool utilization resulting in mistakes, restricted ability for extended planning in varying conditions, gaps in domain-specific reasoning that impair effectiveness in niche sectors, issues with robustness in unpredictable or real-world settings, and inadequate generalization to new or unfamiliar tasks.

The paper "Adaptation of Agentic AI" by Pengcheng Jiang and collaborators, available on arXiv as 2512.16301v2 and released on November 4, 2025, tackles these challenges directly, advocating adaptation as a key approach for improving agentic systems. The main argument is that while foundation models possess substantial power, they necessitate specific adjustments to attain superior performance, dependability, and adaptability. To systematize the growing body of work in this domain, the authors introduce a cohesive framework that classifies adaptations into two main categories—agent adaptation and tool adaptation—subdivided into four paradigms: A1, involving agent optimization through verifiable tool results; A2, guided by assessments of the agent's ultimate outputs; T1, featuring tools trained separately; and T2, where tools are adjusted using cues from a static agent. This classification, depicted in the paper's Figure 1, which overviews adaptations illustrating interactions between agents and tools, establishes a theoretical structure that integrates varied techniques and underscores compromises.

Accompanying the paper is a supporting GitHub repository at https://github.com/pat-jj/Awesome-Adaptation-of-Agentic-AI, compiling a broad array of adaptation techniques, datasets, and evaluation metrics. Updated as recently as December 22, 2025, it incorporates recent developments such as Orion, a co-adaptation system for multi-agent environments, and QAgent, a subagent optimized for queries in search-intensive tasks, demonstrating the area's vitality and positioning the repository as an evolving asset for users.

The paper's organization, shown in Figure 2 as a hierarchical outline from foundational elements to prospects, proceeds logically from basics to sophisticated uses and future outlooks. It starts with background material, defining agentic AI systems as self-governing entities centered on large language models or multimodal models acting as reasoning hubs. Essential elements encompass planning units for breaking down tasks, like static approaches such as chain-of-thought or dynamic ones like ReAct, tool employment for outside engagements such as APIs or code runners, and memory for preserving context. Adaptation types are presented as prompt engineering, modifying actions through contextual directives without altering parameters, and fine-tuning, refining model weights with targeted data.

The overview formalizes the framework with mathematical symbols, portraying agents as functions converting inputs to actions or results, tools as invocable units, and adaptations as enhancements over parameters or strategies. The paradigms are elaborated: A1 employs tool execution cues like code outputs or retrieval metrics for agent refinement; A2 uses comprehensive reviews of agent outputs, including answer precision; T1 trains tools autonomously, such as pre-trained retrievers; and T2 refines tools with agent-provided signals, like feedback for rerankers. Concrete instances anchor these, for example, A1 in code creation where execution environments indicate errors, or T2 in flexible memory where agent reviews enhance retrieval.

Agent adaptation explores A1 and A2. For A1, initial efforts involve supervised fine-tuning on tool-interaction records and off-policy techniques gathering data from varied explorations. Methods based on reinforcement learning with verifiable rewards enhance agents using outcome-oriented incentives, as seen in Toolformer or retrieval-enhanced configurations. A2 concentrates on agent outputs: without tools, it applies preference ratings or self-assessments for reasoning improvement; with tools, it includes final response evaluations after tool integration, as in outcome-guided fine-tuning for planning.

Tool adaptation addresses T1 and T2. T1, independent of agents, covers base systems like HuggingGPT, coordinating pre-trained models as tools, and ViperGPT for visual duties. Groups and training approaches include retrievers, specialized domain models, and subagents trained through supervised or self-supervised goals on extensive datasets. T2, overseen by agents, includes early techniques like reward-guided adjustment, subagent-as-tool where compact models are refined on agent input such as search subagents, and agentic memory with modules optimizing recall based on agent evaluations.

Comparisons examine the paradigms: A1 versus A2 contrasts precise, detailed signals in A1 with infrequent, comprehensive ones in A2, highlighting A1's overfitting potential; T1 versus T2 emphasizes T1's broad applicability but absence of agent tailoring against T2's customized efficiency. Strategic advice directs choices according to assets, favoring T1/T2 for budget-limited cases. Applications demonstrate the paradigms: deep research employs mixed A1/T2 for repeated hypothesis evaluation; software creation uses A2 for code enhancement; computer operation utilizes T1 for browser instruments; and drug development merges A1/T1 for molecular modeling. Opportunities investigate co-adaptation for combined agent-tool refinement, continual adaptation for shifting environments, safe adaptation to reduce hazards, and efficient adaptation for limited-resource contexts.

The paper's strengths reside in its classification, which elucidates a disjointed domain by merging adaptation tactics under a consistent structure. It explicitly outlines compromises that were formerly tacit, such as A1's dependence on verifiable tool indicators for accurate optimization compared to A2's holistic assessments for wider synchronization, which may encounter feedback scarcity. This lucidity is novel, linking conventional fine-tuning with nascent reinforcement-style techniques, offering a perspective to consider efforts like ReAct as A1 models or Reflexion as A2. Practical direction is a further advantage: the authors stress combined methods, observing how leading systems integrate paradigms, such as T1 pre-trained retrievers with T2 flexible subagents and A1 tuned reasoners in layered designs. This confronts practical issues, as evidenced in applications where erratic tool use is alleviated through A1's execution input in software creation, or domain discrepancies are spanned via T1's dedicated models in drug development. The prospective opportunities segment pinpoints vital unresolved matters, like attributing progress in co-adaptation and catastrophic forgetting in continual adaptation, providing a guide extending past existing studies.

In general, the framework's modularity—viewing tools as interchangeable or agent-customized—advances versatility, while the focus on empirical standards across fields highlights the paper's thoroughness. By encompassing mathematical formalisms, it attracts theorists, yet stays reachable via examples, rendering it an essential consolidation.

Despite its thoroughness, the paper openly recognizes constraints in agentic AI, such as erratic tool use arising from mismatched calls or interpretation faults, and confined long-term planning where agents stumble in chained dependencies. Robustness concerns surface in turbulent settings, where disturbances interrupt reasoning, and weak generalization obstructs shift to unseen duties. In agent adaptation, A1 and A2 hazard overfitting: A1 to particular tool runs without moderation, A2 to scarce signals that might disregard intermediate flaws. Tool adaptations confront their own barriers: T1's standalone training may produce general tools unsuited to certain agents, while T2's dependence on fixed agents restricts range if the agent's cues are prejudiced or partial. Computational expenses are recurrent, with agent fine-tuning requiring considerable assets for vast-parameter models, opposing tool adaptations' thrift but possibly at depth's cost.

The challenges section emphasizes profound issues: co-adaptation's stability-plasticity conflict, where shared optimization threats instability or loss of earlier knowledge; continual adaptation's management of changing data flows, worsening forgetting; safe adaptation's susceptibilities like specification gaming or hazardous exploration in critical areas; and efficient adaptation's needs for on-device customization amid resource limits, such as peripheral computing in portable agents.

Extending the paper's examination with further viewpoints enriches the discourse. Ethically, adaptation endangers bias escalation if training data incorporates social biases, potentially aggravating disparities in uses like drug development. Privacy worries arise in T2's on-device tools, where agent-overseen adaptations could disclose user information in feedback cycles. Accountability in arising behaviors—unforeseen interplays between adapted agents and tools—demands examination, as mixed systems might produce erratic results. Theoretically, shortcomings encompass convergence assurances for co-adaptation algorithms, where shared gradients could diverge, or uniform benchmarks for fluid, multi-agent situations beyond single-agent emphases. Broadening to multi-agent setups, adaptations might include coordination indicators, such as a human-signaled variant merging supervision to tackle A2's infrequent feedback, inspired by human-in-the-loop models.

Hardware amalgamation provides sustainability perspectives: efficient adaptations could exploit neuromorphic processors emulating biologically inspired plasticity for continual learning, lessening energy demands in extended tasks. Interdisciplinary connections are plentiful—neural plasticity from biology could guide anti-forgetting systems in A1/A2, while socio-economic effects, like employment shifts in software creation, highlight the necessity for fair implementation. These concepts augment the framework by suggesting expansions, such as blended human-tool signals in T2 for sturdiness in low-signal settings.

Drawing on the opportunities segment, trends beyond 2026 may intensify, bolstered by the GitHub repository's revisions, like Orion's multi-agent co-adaptation for joint duties and QAgent's query enhancement in T2, directing toward expandable, dispersed adaptations. Unified co-adaptation structures could incorporate meta-learning for sparse-data scenarios, facilitating swift customization with fairness-conscious goals to lessen biases. Theoretical progress might formalize adaptation mechanics, supplying limits on generalization similar to PAC learning for agent-tool duos. Societally, adapted agentic AI vows change: in drug development, hybrid paradigms could hasten treatments, while in research, they nurture innovations. However, excessive dependence threats reducing human proficiency, prompting issues about AI oversight. The paper's function as a fundamental guide—assembling tactics and compromises—places it to shape these paths, encouraging empirical confirmations in varied, practical benchmarks.

"Adaptation of Agentic AI" emerges as a pertinent amalgamation, condensing the domain's intricacies into a practical framework amid 2025's swift progress. Its classification and counsel enable scholars and practitioners to maneuver adaptations, cultivating more proficient systems. As the associated repository advances, the paper's impact will expand, but fulfilling its promise requires stringent empirical examination of paradigms in developing contexts. In the end, it delineates a route toward sturdy, principled agentic AI, equilibrating advancement with prudence.

References

Achiam, J. et al. (2017) Constrained policy optimization. In International conference on machine learning. PMLR, pp. 22–31.

Achiam, J. et al. (2023) GPT-4 technical report. arXiv preprint arXiv:2303.08774.

Agarwal, R. et al. (2025) Training language models to self-correct via reinforcement learning. In The Thirteenth International Conference on Learning Representations.

Aljundi, R. et al. (2019) Gradient based sample selection for online continual learning. Advances in neural information processing systems, 32.

Amodei, D. et al. (2016) Concrete problems in ai safety. arXiv preprint arXiv:1606.06565.

Anandkumar, A. et al. (2024) Lean copilot: Large language models as copilots for theorem proving in lean. arXiv preprint arXiv:2404.12534.

Anandkumar, A. et al. (2025) Leandojo: Theorem proving with retrieval-augmented language models. Advances in Neural Information Processing Systems, 36, pp. 21573–21612.

Anandkumar, A. et al. (2025) Leanprogress: Guiding search for neural theorem proving via proof progress prediction. arXiv preprint arXiv:2502.17925.

Anokhin, P. et al. (2024) Arigraph: Learning knowledge graph world models with episodic memory for llm agents. arXiv preprint arXiv:2407.04363.

Anthropic (2025a) Claude code: Deep coding at terminal velocity. Available at: https://www.anthropic.com/claude-code (Accessed: 29 October 2025).

Anthropic (2025b) Claude takes research to new places. Available at: https://www.anthropic.com/news/research (Accessed: 28 October 2025).

Anthropic (2025c) Code execution with mcp: Building more efficient ai agents. Available at: https://www.anthropic.com/engineering/code-execution-with-mcp (Accessed: 4 November 2025).

Arani, E. et al. (2022) Learning fast, learning slow: A general continual learning method based on complementary learning system. In International Conference on Learning Representations.

Arani, E. et al. (2024) A comprehensive survey of direct preference optimization: Datasets, theories, variants, and applications. arXiv preprint arXiv:2410.15595.

Arora, S. et al. (2024) Simple linear attention language models balance the recall-throughput tradeoff. arXiv preprint arXiv:2402.18668.

Asai, A. et al. (2024) Self-rag: Learning to retrieve, generate, and critique through self-reflection. In The Twelfth International Conference on Learning Representations.

Asawa, P. et al. (2025) Qedcartographer: Automating formal verification using reward-free reinforcement learning. In Proceedings of the IEEE/ACM 47th International Conference on Software Engineering, pp. 307–320.

Asawa, P. et al. (2025) Toolexpander: Extending the frontiers of tool-using reinforcement learning to weak llms. arXiv preprint arXiv:2510.07737.

Bai, G. et al. (2025) Fedspallm: Federated pruning of large language models. In Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers), pp. 8361–8373.

Bang, J. et al. (2021) Rainbow memory: Continual learning with a memory of diverse samples. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 8218–8227.

Bang, J. et al. (2025) Why do multi-agent llm systems fail? arXiv preprint arXiv:2503.13657.

Belcak, P. et al. (2025) Small language models are the future of agentic ai. arXiv preprint arXiv:2506.02153.

Bhatia, A. et al. (2024) Hypothetical minds: Scaffolding theory of mind for multi-agent tasks with large language models. arXiv preprint arXiv:2407.07086.

Bondarenko, A. et al. (2025) Demonstrating specification gaming in reasoning models. arXiv preprint arXiv:2502.13295.

Caron, M. et al. (2021) Unsupervised dense information retrieval with contrastive learning. arXiv preprint arXiv:2112.09118.

Chen, D. et al. (2020) Dense passage retrieval for open-domain question answering. In EMNLP (1), pp. 6769–6781.

Chen, Q. et al. (2024) Chatdev: Communicative agents for software development. In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 15174–15186.

Chen, S. et al. (2024) Advancing tool-augmented large language models: Integrating insights from errors in inference trees. Advances in Neural Information Processing Systems, 37, pp. 106555–106581.

Chen, Z. et al. (2024) Dualnet: Continual learning, fast and slow. Advances in Neural Information Processing Systems, 34, pp. 16131–16144.

Chen, Z. et al. (2025) A2fm: An adaptive agent foundation model for tool-aware hybrid reasoning. arXiv e-prints, arXiv–2510.

Chen, Z. et al. (2025) Think before you retrieve: Learning test-time adaptive search with small language models. arXiv preprint arXiv:2511.07581.

Cheng, D. et al. (2024) Colbert: Efficient and effective passage search via contextualized late interaction over bert. In Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval, pp. 39–48.

Citron, D. (2025) Deep research is now available on gemini 2.5 pro experimental. Available at: https://blog.google/products/gemini/deep-research-gemini-2-5-pro-experimental/ (Accessed: 2025).

Cogntiion (2025) Introducing swe-grep and swe-grep-mini: Rl for multi-turn, fast context retrieval. Available at: https://cognition.ai/blog/swe-grep (Accessed: 29 October 2025).

Cursor (2025) Cursor - the ai code editor. Available at: https://www.cursor.com/ (Accessed: 29 October 2025).

De Lange, M. et al. (2021) A continual learning survey: Defying forgetting in classification tasks. IEEE transactions on pattern analysis and machine intelligence, 44(7), pp. 3366–3385.

Dettmers, T. et al. (2023) Qlora: Efficient finetuning of quantized llms. Advances in neural information processing systems, 36, pp. 10088–10115.

Ding, K. et al. (2025) Scitoolagent: A knowledge-graph-driven scientific agent for multitool integration. Nature Computational Science, pp. 1–11.

Dou, Z-Y. et al. (2024) Re-rest: Reflection-reinforced self-training for language agents. In Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, pp. 15394–15411.

Du, Y. et al. (2025) Retrieve-plan-generation: An iterative planning and answering framework for knowledge-intensive llm generation. In Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, pp. 4683–4702.

Dulac-Arnold, G. et al. (2019) Challenges of real-world reinforcement learning. arXiv preprint arXiv:1904.12901.

Ellis, E. et al. (2025) Training llm agents to empower humans. arXiv preprint arXiv:2510.13709.

Fang, J. et al. (2025) A comprehensive survey of self-evolving ai agents: A new paradigm bridging foundation models and lifelong agentic systems. arXiv preprint arXiv:2508.07407.

Fang, Y. et al. (2023) Knowledge graph-enhanced molecular contrastive learning with functional prompt. Nature Machine Intelligence, pp. 1–12.

Fang, Y. et al. (2025) Retool: Reinforcement learning for strategic tool use in llms. arXiv preprint arXiv:2504.11536.

Feng, J. et al. (2025) Retool: Reinforcement learning for strategic tool use in llms. arXiv preprint arXiv:2504.11536.

Feng, X. et al. (2024) A survey on large language model based autonomous agents. Frontiers of Computer Science, 18(6), p. 186345.

Gao, S. et al. (2024) Empowering biomedical discovery with ai agents. Cell, 187(22), pp. 6125–6151.

Gao, S. et al. (2025) A survey of self-evolving agents: On path to artificial super intelligence. arXiv preprint arXiv:2507.21046.

Gao, S. et al. (2025) Democratizing ai scientists using tooluniverse. arXiv preprint arXiv:2509.23426.

García, J. and Fernández, F. (2015) A comprehensive survey on safe reinforcement learning. Journal of Machine Learning Research, 16(1), pp. 1437–1480.

Gehring, J. et al. (2024) Rlef: Grounding code llms in execution feedback with reinforcement learning. arXiv preprint arXiv:2410.02089.

Gehring, J. et al. (2025) Rlef: Grounding code llms in execution feedback with reinforcement learning. arXiv preprint arXiv:2410.02089.

Gesi, J. et al. (2025) Deepseek-prover-v2: Advancing formal mathematical reasoning via reinforcement learning for subgoal decomposition. arXiv preprint arXiv:2504.21801.

Greshake, K. et al. (2023) Not what you've signed up for: Compromising real-world llm-integrated applications with indirect prompt injection. In Proceedings of the 16th ACM workshop on artificial intelligence and security, pp. 79–90.

Gu, A. et al. (2025) Challenges and paths towards ai for software engineering. arXiv preprint arXiv:2503.22625.

Gu, S. et al. (2024) A review of safe reinforcement learning: Methods, theories and applications. IEEE Transactions on Pattern Analysis and Machine Intelligence.

Gu, X. et al. (2025) metatextgrad: Automatically optimizing language model optimizers. arXiv preprint arXiv:2505.18524.

Guo, D. et al. (2025) Deepseek-r1 incentivizes reasoning in llms through reinforcement learning. Nature, 645(8081), pp. 633–638.

Hillis, W.D. (1990) Co-evolving parasites improve simulated evolution as an optimization procedure. Physica D: Nonlinear Phenomena, 42(1-3), pp. 228–234.

Hong, S. et al. (2023) Metagpt: Meta programming for a multi-agent collaborative framework. In The Twelfth International Conference on Learning Representations.

Hong, S. et al. (2025) Advances and challenges in foundation agents: From brain-inspired intelligence to evolutionary, collaborative, and safe systems. arXiv preprint arXiv:2504.01990.

Hu, E.J. et al. (2022) Lora: Low-rank adaptation of large language models. ICLR, 1(2), p. 3.

Hu, J. et al. (2024) Camel: Communicative agents for” mind” exploration of large language model society. Advances in Neural Information Processing Systems, 36, pp. 51991–52008.

Huang, C. et al. (2025) Sft doesn’t always hurt general capabilities: Revisiting domain-specific fine-tuning in llms. arXiv preprint arXiv:2509.20758.

Huang, C. et al. (2025) Self-improving llm agents at test-time. arXiv preprint arXiv:2510.07841.

Huang, L. et al. (2024) Dora: Weight-decomposed low-rank adaptation. In Forty-first International Conference on Machine Learning.

Huang, L. et al. (2025) Hierarchical decomposition of prompt-based continual learning: Rethinking obscured sub-optimality. Advances in Neural Information Processing Systems, 36, pp. 69054–69076.

Huang, L. et al. (2025) Mitigating catastrophic forgetting in online continual learning by modeling previous task interrelations via pareto optimization. In Forty-first international conference on machine learning.

Hubert, T. et al. (2025) Olympiad-level formal mathematical reasoning with reinforcement learning. Nature, pp. 1–3.

Izacard, G. et al. (2023) Uprise: Universal prompt retrieval for improving zero-shot evaluation. In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, pp. 12318–12337.

Jackson, J. et al. (2025) Improving cursor tab with online reinforcement learning. Available at: https://cursor.com/blog/tab-rl (Accessed: 29 September 2025).

Jia, X. et al. (2025) A survey on multi-agent reinforcement learning and its application. Journal of Automation and Intelligence, 3(2), pp. 73–91.

Jiang, N. et al. (2022) Text embeddings by weakly-supervised contrastive pre-training. arXiv preprint arXiv:2212.03533.

Jiang, P. et al. (2025a) Bi-level contrastive learning for knowledge-enhanced molecule representations. In Proceedings of the Thirty-Ninth AAAI Conference on Artificial Intelligence.

Jiang, P. et al. (2025b) Deepretrieval: Hacking real search engines and retrievers with large language models via reinforcement learning. In The Second Conference on Language Modeling.

Jiang, P. et al. (2025c) Ras: Retrieval-and-structuring for knowledge-intensive llm generation. arXiv preprint arXiv:2502.10996.

Jiang, P. et al. (2025d) s3: You don't need that much data to train a search agent via rl. In Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing.

Jiang, P. et al. (2025e) Training llms for ehr-based reasoning tasks via reinforcement learning. arXiv preprint arXiv:2505.24105.

Jin, B. et al. (2025a) Search-r1: Training llms to reason and leverage search engines with reinforcement learning. In The Second Conference on Language Modeling.

Jin, B. et al. (2025b) Swe-bench: Can language models resolve real-world github issues? arXiv preprint arXiv:2310.06770.

Jin, B. et al. (2025c) Swe-bench: Can language models resolve real-world github issues? In The Twelfth International Conference on Learning Representations.

Jin, N. et al. (2024) Matching patients to clinical trials with large language models. Nature communications, 15(1), p. 9074.

Jin, W. et al. (2017) Predicting organic reaction outcomes with weisfeiler-lehman network. arXiv preprint arXiv:1709.04555.

Jin, W. et al. (2018) Learning multimodal graph-to-graph translation for molecular optimization. arXiv preprint arXiv:1812.01070.

Jin, W. et al. (2025) A generative deep learning approach to de novo antibiotic design. Cell.

Kale, S. and Dhami, D.S. (2025) Knowrl: Teaching language models to know what they know. arXiv preprint arXiv:2510.11407.

Kamanuru, V. et al. (2025) Agentic context engineering: Evolving contexts for self-improving language models. arXiv preprint arXiv:2510.04618.

Karbasi, A. and Kassianik, P. (2025) Evaluating security risk in deepseek and other frontier reasoning models. Cisco Blogs, Cisco Systems, 31.

Kassianik, P. and Karbasi, A. (2025) Evaluating security risk in deepseek and other frontier reasoning models. Cisco Blogs.

Ke, Z. et al. (2024) Bridging the preference gap between retrievers and llms. In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 10438–10451.

Kirkpatrick, J. et al. (2017) Overcoming catastrophic forgetting in neural networks. Proceedings of the national academy of sciences, 114(13), pp. 3521–3526.

Kirillov, A. et al. (2023) Segment anything. In Proceedings of the IEEE/CVF international conference on computer vision, pp. 4015–4026.

Krakovna, V. et al. (2020) Specification gaming: the flip side of ai ingenuity. DeepMind Blog.

Krishnan, A. et al. (2025) A generative deep learning approach to de novo antibiotic design. Cell.

Kumar, A. et al. (2025) Training language models to self-correct via reinforcement learning. In The Thirteenth International Conference on Learning Representations.

Kumarappan, A. et al. (2024) Leanagent: Lifelong learning for formal theorem proving. arXiv preprint arXiv:2410.06209.

Lee, J. et al. (2024) Putnambench: A multilingual competition-mathematics benchmark for formal theorem-proving. In AI for Math Workshop@ ICML 2024.

Li, G. et al. (2023) Camel: Communicative agents for” mind” exploration of large language model society. Advances in Neural Information Processing Systems, 36, pp. 51991–52008.

Li, H. et al. (2024) Stella: Self-evolving llm agent for biomedical research. arXiv preprint arXiv:2507.02004.

Li, H. et al. (2025) Blade: Enhancing black-box large language models with small domain-specific models. In Proceedings of the AAAI Conference on Artificial Intelligence, 39, pp. 24422–24430.

Li, H. et al. (2025) Trialgenie: Empowering clinical trial design with agentic intelligence and real world data. medRxiv, pp. 2025–04.

Li, L. et al. (2024) Dora: Weight-decomposed low-rank adaptation. In Forty-first International Conference on Machine Learning.

Li, L. et al. (2025a) Flashrl: 8bit rollouts, full power rl. Available at: https://fengyao.notion.site/flash-rl.

Li, L. et al. (2025b) Mobillm: Enabling llm fine-tuning on the mobile device via server assisted side tuning. arXiv preprint arXiv:2502.20421.

Li, L. et al. (2025c) Pocketllm: Enabling on-device fine-tuning for personalized llms. arXiv preprint arXiv:2407.01031.

Li, Z. and Hoiem, D. (2017) Learning without forgetting. IEEE transactions on pattern analysis and machine intelligence, 40(12), pp. 2935–2947.

Li, Z. et al. (2025) In-the-flow agentic system optimization for effective planning and tool use. arXiv preprint arXiv:2510.05592.

Liang, X. et al. (2023) Unleashing infinite-length input capacity for large-scale language models with self-controlled memory system. arXiv preprint arXiv:2304.13343, 10.

Lin, J. et al. (2024) Rec-r1: Bridging generative large language models and user-centric recommendation systems via reinforcement learning. Transactions on Machine Learning Research.

Lin, J. et al. (2025) Training llms for ehr-based reasoning tasks via reinforcement learning. arXiv preprint arXiv:2505.24105.

Liu, A. et al. (2024) Tuning language models by proxy. In First Conference on Language Modeling.

Liu, B. et al. (2025) Advances and challenges in foundation agents: From brain-inspired intelligence to evolutionary, collaborative, and safe systems. arXiv preprint arXiv:2504.01990.

Liu, J. and Zhang, L. (2025) Code-r1: Reproducing r1 for code with reliable rewards. Available at: https://github.com/ganler/code-r1.

Liu, L. et al. (2023) Chatdb: Augmenting llms with databases as their symbolic memory. arXiv preprint arXiv:2306.03901.

Liu, L. et al. (2025) Flashrl: 8bit rollouts, full power rl. Available at: https://fengyao.notion.site/flash-rl.

Luo, J. et al. (2025) Large language model agent: A survey on methodology, applications and challenges. arXiv preprint arXiv:2503.21460.

Luo, R. et al. (2025) Gui-r1: A generalist r1-style vision-language action model for gui agents. arXiv preprint arXiv:2504.10458.

Lu, J. et al. (2023) Memochat: Tuning llms to use memos for consistent long-range open-domain conversation. arXiv preprint arXiv:2308.08239.

Lu, Z. et al. (2025) Agent lightning: Train any ai agents with reinforcement learning. arXiv preprint arXiv:2508.03680.

Lu, Z. et al. (2025) Webgen-agent: Enhancing interactive website generation with multi-level feedback and step-level reinforcement learning. arXiv preprint arXiv:2509.22644.

Luo, X. et al. (2025) Agent lightning: Train any ai agents with reinforcement learning. arXiv preprint arXiv:2508.03680.

Lyu, Y. et al. (2024) Retrieve-plan-generation: An iterative planning and answering framework for knowledge-intensive llm generation. In Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, pp. 4683–4702.

Lyu, Y. et al. (2025) Zerosearch: Incentivize the search capability of llms without searching. arXiv preprint arXiv:2505.04588.

Ma, P. et al. (2025) Sql-r1: Training natural language to sql reasoning model by reinforcement learning. arXiv preprint arXiv:2504.08600.

Manheim, D. and Garrabrant, S. (2018) Categorizing variants of goodhart's law. arXiv preprint arXiv:1803.04585.

Marczak, D. et al. (2024) Magmax: Leveraging model merging for seamless continual learning. In European Conference on Computer Vision. Springer, pp. 379–395.

Mathlib Community (2020) The lean mathematical library. In Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs, POPL '20. ACM, p. 367–381.

Mei, L. et al. (2025) Ai-searchplanner: Modular agentic search via pareto-optimal multi-objective reinforcement learning. arXiv preprint arXiv:2508.20368.

Modarressi, A. et al. (2023) Ret-llm: Towards a general read-write memory for large language models. arXiv preprint arXiv:2305.14322.

Murahari, V. et al. (2024) Personagym: Evaluating persona agents and llms. arXiv preprint arXiv:2407.18416.

Ni, A. et al. (2024) Next: teaching large language models to reason about code execution. In Proceedings of the 41st International Conference on Machine Learning, pp. 37929–37956.

Ning, Z. and Xie, L. (2024) A survey on multi-agent reinforcement learning and its application. Journal of Automation and Intelligence, 3(2), pp. 73–91.

OpenAI (2025a) Codex. Available at: https://openai.com/codex/ (Accessed: 29 October 2025).

OpenAI (2025b) Computer-using agent. Available at: https://openai.com/index/computer-using-agent/ (Accessed: 28 October 2025).

OpenAI (2025c) Introducing deep research. Available at: https://openai.com/index/introducing-deep-research/ (Accessed: 2025).

Packer, C. et al. (2023) Memgpt: Towards llms as operating systems. arXiv preprint arXiv:2310.08560.

Pan, B. et al. (2025) Introducing swe-grep and swe-grep-mini: Rl for multi-turn, fast context retrieval. Available at: https://cognition.ai/blog/swe-grep (Accessed: 29 October 2025).

Panait, L. and Luke, S. (2005) Cooperative multi-agent learning: The state of the art. Autonomous agents and multi-agent systems, 11(3), pp. 387–434.

Panait, L. and Rasmussen, C.E. (2000) Cooperative coevolution: An architecture for evolving co-adapted subcomponents. Evolutionary Computation, 8(1), pp. 1–29.

Park, J.S. et al. (2023) Generative agents: Interactive simulacra of human behavior. In Proceedings of the 36th annual acm symposium on user interface software and technology, pp. 1–22.

Patil, S.G. et al. (2024) Gorilla: Large language model connected with massive apis. Advances in Neural Information Processing Systems, 37, pp. 126544–126565.

Pham, Q. et al. (2021) Dualnet: Continual learning, fast and slow. Advances in Neural Information Processing Systems, 34, pp. 16131–16144.

Piao, H. et al. (2024) Federated continual learning via prompt-based dual knowledge transfer. In Forty-first International Conference on Machine Learning.

Platonov, A. et al. (2025) A systematic survey of prompt engineering in large language models: Techniques and applications. arXiv preprint arXiv:2402.07927.

Plaat, A. et al. (2025) Agentic large language models, a survey. arXiv preprint arXiv:2503.23037.

Polu, S. and Sutskever, I. (2020) Generative language modeling for automated theorem proving. arXiv preprint arXiv:2009.03393.

Polu, S. et al. (2023) Generative language modeling for automated theorem proving. arXiv preprint arXiv:2009.03393.

Potter, M.A. and De Jong, K.A. (2000) Cooperative coevolution: An architecture for evolving co-adapted subcomponents. Evolutionary Computation, 8(1), pp. 1–29.

Poznanski, J. et al. (2025a) olmocr 2: Unit test rewards for document ocr. arXiv preprint arXiv:2510.19817.

Poznanski, J. et al. (2025b) olmocr: Unlocking trillions of tokens in pdfs with vision language models. arXiv preprint arXiv:2502.18443.

Qin, Y. et al. (2023) Toolllm: Facilitating large language models to master 16000+ real-world apis. In ICLR.

Qin, Y. et al. (2024) Tool learning with foundation models. In ACM Computing Surveys.

Qin, Y. et al. (2025) A survey on large language model reasoning failures. In 2nd AI for Math Workshop@ ICML 2025.

Qin, Y. et al. (2025) A survey on self-evolution of large language models. arXiv preprint arXiv:2404.14387.

Rafailov, R. et al. (2023) Direct preference optimization: Your language model is secretly a reward model. Advances in neural information processing systems, 36, pp. 53728–53741.

Rafailov, R. et al. (2024) Direct preference optimization: Your language model is secretly a reward model. Advances in neural information processing systems, 36, pp. 53728–53741.

Rosin, C.D. and Miikkulainen, R. (2004) Competitive coevolution through evolutionary complexification. Journal of Artificial Intelligence Research, 21, pp. 63–100.

Schick, T. et al. (2023) Toolformer: Language models can teach themselves to use tools. Advances in Neural Information Processing Systems, 36, pp. 68539–68551.

Schulman, J. et al. (2017) Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347.

Schulman, J. et al. (2025) Lora without regret. Thinking Machines Lab: Connectionism. doi: 10.64434/tml.20250929.

Shang, J. et al. (2025) Llm-based agents for tool learning: A survey. Data Science and Engineering, pp. 1–31.

Shi, W. et al. (2024) Replug: Retrieval-augmented black-box language models. In Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers), pp. 8364–8377.

Shi, Z. et al. (2024) Dualedit: Dual editing for knowledge updating in vision-language models. arXiv preprint arXiv:2506.13638.

Shi, Z. et al. (2025a) A comprehensive survey of continual learning: Theory, method and application. ACM Computing Surveys.

Shi, Z. et al. (2025b) Generalized tensor-based parameter-efficient fine-tuning via lie group transformations. arXiv preprint arXiv:2504.00851.

Shi, Z. et al. (2025c) Tool learning in the wild: Empowering language models as automatic tool agents. In Proceedings of the ACM on Web Conference 2025, pp. 2222–2237.

Shinn, N. et al. (2023) Reflexion: Language agents with verbal reinforcement learning. Advances in Neural Information Processing Systems, 36, pp. 8634–8652.

Si, C. et al. (2024) Flora: Low-rank core space for n-dimension. arXiv preprint arXiv:2405.14739, 10.

Si, C. et al. (2025) Unleashing the power of task-specific directions in parameter efficient fine-tuning. In The Thirteenth International Conference on Learning Representations.

Song, P. et al. (2025) A survey on large language model reasoning failures. In 2nd AI for Math Workshop@ ICML 2025.

Song, P. et al. (2025) Lean copilot: Large language models as copilots for theorem proving in lean. arXiv preprint arXiv:2404.12534.

Su, H. et al. (2024) Source-free domain adaptation with frozen multimodal foundation model. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 23711–23720.

Sun, J. et al. (2025a) Dynamicrag: Leveraging outputs of large language model as feedback for dynamic reranking in retrieval-augmented generation. arXiv preprint arXiv:2505.07233.

Sun, J. et al. (2025b) Grace: Generative representation learning via contrastive policy optimization. arXiv preprint arXiv:2510.04506.

Surís, D. et al. (2023) Vipergpt: Visual inference via python execution for reasoning. In Proceedings of the IEEE/CVF international conference on computer vision, pp. 11888–11898.

Suzgun, M. et al. (2025) Dynamic cheatsheet: Test-time learning with adaptive memory. arXiv preprint arXiv:2504.07952.

Swanson, K. et al. (2024) Generative ai for designing and validating easily synthesizable and structurally novel antibiotics. Nature machine intelligence, 6(3), pp. 338–353.

Swanson, K. et al. (2025) The virtual lab of ai agents designs new sars-cov-2 nanobodies. Nature, pp. 1–3.

Tang, Q. et al. (2023) Toolalpaca: Generalized tool learning for language models with 3000 simulated cases. arXiv preprint arXiv:2306.05301.

Tang, S. et al. (2025) Search and refine during think: Autonomous retrieval-augmented reasoning of llms. arXiv e-prints, arXiv–2505.

Tao, Z. et al. (2024) A survey on self-evolution of large language models. arXiv preprint arXiv:2404.14387.

Tong, X. et al. (2024) Deep representation learning of chemical-induced transcriptional profile for phenotype-based drug discovery. Nature Communications, 15(1), p. 5378.

Trivedi, H. et al. (2024) Appworld: A controllable world of apps and people for benchmarking interactive coding agents. In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 16022–16076.

Tsoukalas, G. et al. (2024) Putnambench: A multilingual competition-mathematics benchmark for formal theorem-proving. In AI for Math Workshop@ ICML 2024.

Tsoukalas, G. et al. (2025) Co-evolving parasites improve simulated evolution as an optimization procedure. In Artificial Life II. Addison-Wesley, volume X, pp. 313–324.

Tsang, H.T. et al. (2025) Autograph-r1: End-to-end reinforcement learning for knowledge graph construction. arXiv preprint arXiv:2510.15339.

Wang, L. et al. (2023) Learning to retrieve in-context examples for large language models. In Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 1752–1767.

Wang, L. et al. (2024a) A comprehensive survey of continual learning: Theory, method and application. IEEE transactions on pattern analysis and machine intelligence, 46(8), pp. 5362–5383.

Wang, L. et al. (2024b) Hierarchical decomposition of prompt-based continual learning: Rethinking obscured sub-optimality. Advances in Neural Information Processing Systems, 36, pp. 69054–69076.

Wang, L. et al. (2024c) Learning to prompt for continual learning. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 139–149.

Wang, L. et al. (2024d) Memory replay with data compression for continual learning. In International Conference on Learning Representations.

Wang, L. et al. (2024e) Text embeddings by weakly-supervised contrastive pre-training. arXiv preprint arXiv:2212.03533.

Wang, L. et al. (2025a) A perspective for adapting generalist ai to specialized medical ai applications and their challenges. NPJ Digital Medicine, 8(1), p. 429.

Wang, L. et al. (2025b) Accelerating clinical evidence synthesis with large language models. npj Digital Medicine, 8(1), p. 509.

Wang, L. et al. (2025c) A foundation model for human-ai collaboration in medical literature mining. Nature Communications, 16(1), p. 8361.

Wang, L. et al. (2025d) Making large language models reliable data science programming copilot for biomedical research. Nature Biomedical Engineering.

Wang, X. et al. (2024a) Executable code actions elicit better llm agents. In International Conference on Machine Learning. PMLR, pp. 50208–50232.

Wang, X. et al. (2024b) Openhands: An open platform for AI software developers as generalist agents. In The Thirteenth International Conference on Learning Representations.

Wang, X. et al. (2024c) Swe-bench: Can language models resolve real-world github issues? Advances in Neural Information Processing Systems, 37, pp. 50528–50652.

Wang, X. et al. (2024d) Unified embeddings for multimodal retrieval via frozen llms. In Findings of the Association for Computational Linguistics: EACL 2024, pp. 1537–1547.

Wang, Z. et al. (2024) Sam-clip: Merging vision foundation models towards semantic and spatial understanding. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3635–3647.

Wang, Z. et al. (2025a) Cba: Improving online continual learning via continual bias adaptor. In Proceedings of the IEEE/CVF international conference on computer vision, pp. 19082–19092.

Wang, Z. et al. (2025b) Stepsearch: Igniting llms search ability via step-wise proximal policy optimization. arXiv preprint arXiv:2505.15107.

Wei, J. et al. (2022) Chain-of-thought prompting elicits reasoning in large language models. Advances in neural information processing systems, 35, pp. 24824–24837.

Wei, J. et al. (2022) Finetuned language models are zero-shot learners. In International Conference on Learning Representations.

Wei, J. et al. (2025) Llm-based agents for tool learning: A survey. Data Science and Engineering, pp. 1–31.

Wei, X. et al. (2025) Not what you've signed up for: Compromising real-world llm-integrated applications with indirect prompt injection. In Proceedings of the 16th ACM workshop on artificial intelligence and security, pp. 79–90.

Wu, H. et al. (2023) Autogen: Enabling next-gen llm applications via multi-agent conversations. In First Conference on Language Modeling.

Wu, H. et al. (2023) Visual chatgpt: Talking, drawing and editing with visual foundation models. arXiv preprint arXiv:2303.04671.

Wu, Q. et al. (2024) Autogen: Enabling next-gen llm applications via multi-agent conversations. In First Conference on Language Modeling.

Wu, Y. et al. (2024a) Federated continual learning via prompt-based dual knowledge transfer. In Forty-first International Conference on Machine Learning.

Wu, Y. et al. (2024b) Meta continual learning revisited: Implicitly enhancing online hessian approximation via variance reduction. In The Twelfth international conference on learning representations, volume 2.

Wu, Y. et al. (2024c) Sd-lora: Scalable decoupled low-rank adaptation for class incremental learning. In ICLR.

Wu, Y. et al. (2025) A survey on the memory mechanism of large language model-based agents. ACM Transactions on Information Systems, 43(6), pp. 1–47.

Xu, B. et al. (2024) Smiles-mamba: Chemical mamba foundation models for drug admet prediction. arXiv preprint arXiv:2408.05696.

Xu, R. and Peng, J. (2025) A comprehensive survey of deep research: Systems, methodologies, and applications. arXiv preprint arXiv:2506.12594.

Xu, W. et al. (2025) Llm-based agents for tool learning: A survey. Data Science and Engineering, pp. 1–31.

Xu, X. et al. (2025) Scaling up multi-turn off-policy rl and multi-agent tree search for llm step-provers. arXiv preprint arXiv:2509.06493.

Yang, K. et al. (2023) Tree of thoughts: Deliberate problem solving with large language models. Advances in neural information processing systems, 36, pp. 11809–11822.

Yang, K. et al. (2024) Formal mathematical reasoning: A new frontier in ai. arXiv preprint arXiv:2412.16075.

Yang, K. et al. (2025) Goedel-prover-v2: Scaling formal theorem proving with scaffolded data synthesis and self-correction. arXiv preprint arXiv:2508.03613.

Yao, S. et al. (2022) React: Synergizing reasoning and acting in language models. In The eleventh international conference on learning representations.

Yao, S. et al. (2023) Tree of thoughts: Deliberate problem solving with large language models. Advances in neural information processing systems, 36, pp. 11809–11822.

Yao, S. et al. (2024) Tau-bench: A benchmark for tool-agent-user interaction in real-world domains. In The Thirteenth International Conference on Learning Representations.

Yao, S. et al. (2025) Swe-bench: Can language models resolve real-world github issues? arXiv preprint arXiv:2310.06770.

Yih, W-t. et al. (2024) Replug: Retrieval-augmented black-box language models. In Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers), pp. 8364–8377.

Yu, Z. et al. (2023) Augmentation-adapted retriever improves generalization of language models as generic plug-in. In The 61st Annual Meeting Of The Association For Computational Linguistics.

Yu, Z. et al. (2025) Augmentation-adapted retriever improves generalization of language models as generic plug-in. In The 61st Annual Meeting Of The Association For Computational Linguistics.

Zhang, Q. et al. (2023) Hugginggpt: Solving ai tasks with chatgpt and its friends in hugging face. Advances in neural information processing systems, 36, pp. 38154–38180.

Zhang, S. et al. (2025a) Nemotron-research-tool-n1: Exploring tool-using language models with reinforced reasoning. arXiv preprint arXiv:2505.00024.

Zhang, S. et al. (2025b) Parameter-efficient fine-tuning for large models: A comprehensive survey. Transactions on Machine Learning Research. ISSN 2835-8856.

Zhang, Z. et al. (2025) A survey on the memory mechanism of large language model-based agents. ACM Transactions on Information Systems, 43(6), pp. 1–47.

Zhao, A. et al. (2024) Expel: Llm agents are experiential learners. In Proceedings of the AAAI Conference on Artificial Intelligence, 38, pp. 19632–19642.

Zhao, H. et al. (2025) Goedel-prover-v2: Scaling formal theorem proving with scaffolded data synthesis and self-correction. arXiv preprint arXiv:2508.03613.

Zhou, L. et al. (2025) Multi-agent tool-integrated policy optimization. arXiv preprint arXiv:2510.04678.

Zhou, S. et al. (2023) Webarena: A realistic web environment for building autonomous agents. arXiv preprint arXiv:2307.13854.

Zhou, S. et al. (2024a) Visualwebarena: Evaluating multimodal agents on realistic visual web tasks. In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 881–905.

Zhou, S. et al. (2024b) Webarena: A realistic web environment for building autonomous agents. In The Twelfth International Conference on Learning Representations.

Zhou, S. et al. (2025) Llm-based agents for tool learning: A survey. Data Science and Engineering, pp. 1–31.

Zhu, Z. and Li, D. (2017) Learning without forgetting. IEEE transactions on pattern analysis and machine intelligence, 40(12), pp. 2935–2947.

Zhu, Z. et al. (2023) Evolutionary-scale prediction of atomic-level protein structure with a language model. Science, 379(6637), pp. 1123–1130.

Zhuang, Y. et al. (2024) Bbox-adapter: Lightweight adapting for black-box large language models. arXiv preprint arXiv:2402.08219.

Zhuang, Y. et al. (2025) Toolkengpt: Augmenting frozen language models with massive tools via tool embeddings. Advances in neural information processing systems, 36, pp. 45870–45894.

Zitnik, M. et al. (2024) Empowering biomedical discovery with ai agents. Cell, 187(22), pp. 6125–6151.

Zou, J. et al. (2025) Optimizing generative ai by backpropagating language model feedback. Nature, 639(8055), pp. 609–616.