A Survey of Large Language Model-Based Search Agents

Yunjia Xi[1], Jianghao Lin[*2], Yongzhao Xiao[1], Zheli Zhou[1], Rong Shan[1], Te Gao[3], Jiachen Zhu[1], Weiwen Liu[1], Yong Yu[1], Weinan Zhang[1]

[1] School of Computer Science, Shanghai Jiao Tong University, China
[2] Antai College of Economics and Management, Shanghai Jiao Tong University, China
[3] School of Computer Science and Engineering, Central South University, China
`{xiyunjia,linjianghao,wnzhang}@sjtu.edu.cn`

Abstract

The advent of Large Language Models (LLMs) has significantly revolutionized web search. The emergence of **LLM-based Search Agents** marks a pivotal shift towards deeper, dynamic, autonomous information seeking. These agents can comprehend user intentions and environmental context and execute multi-turn retrieval with dynamic planning, extending search capabilities far beyond the web. Leading examples like OpenAI's Deep Research highlight their potential for deep information mining and real-world applications. This survey provides the first systematic analysis of search agents. We comprehensively analyze and categorize existing works from the perspectives of architecture, optimization, application, and evaluation, ultimately identifying critical open challenges and outlining promising future research directions in this rapidly evolving field. Our repository is available on `https://github.com/YunjiaXi/Awesome-Search-Agent-Papers`.

1 Introduction

The advent of Large Language Models (LLMs) has ushered in a new era of natural language processing, fundamentally transforming numerous fields, including web search (Wang et al., 2024b; Zhao et al., 2023; Hadi et al., 2023; Xi et al., 2025c; Lin et al., 2025a, 2024; Xi et al., 2025b, 2024a). As shown in Figure 1, **Traditional Web Search** required users to manually select and consolidate relevant information from a list of results (Lin et al., 2021; Dai et al., 2021; Fu et al., 2023). With the rise of LLMs, **LLM-enhanced Search** emerged as a new paradigm, where LLMs rewrite user queries to improve search accuracy (Ma et al., 2023b; Liu and Mozafari, 2024; Xi et al., 2024b) or summarize search results for quicker comprehension, *i.e.*, traditional retrieval-augmented generation (RAG) (Gao et al., 2023; Fan et al., 2024). However, this integration tends to be static, as LLMs rely on single-turn or rule-based iterative search, which struggles to handle complex and dynamic context effectively.

*Corresponding author

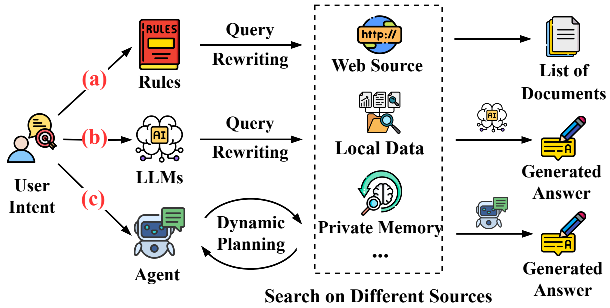

Figure 1: The evolution of search paradigm from (a) **Traditional Web Search** to (b) **LLM-enhanced Search**, and finally to (c) **Search Agents**.

The emergence of LLM agents marked a pivotal shift, leading to **Search Agents** (Zhang et al., 2024b). Endowed with autonomy, search agents can control the entire search process, leveraging context more effectively for adaptive reasoning and dynamic retrieval. In this paradigm, search becomes a proactive action and is no longer limited to the web, but extends to a broader range of information sources, *e.g.*, private databases and internal experiences within agents. Specifically, a search agent can be defined as *an LLM agent capable of comprehending user intentions and environment contexts, autonomously planning search strategies, executing multi-turn dynamic retrieval from diverse sources, and integrating information to provide comprehensive insights*. Leading industrial solutions, *e.g.*, Deep Research from OpenAI (OpenAI, 2025), Gemini (Gemini, 2025), and Perplexity (Perplexity, 2025), exemplify the potential of search agents in both deep information mining and commercialization.

Given these rapid advancements, we present the first systematic survey of search agents from multiple perspectives, analyzing them across the dimensions of *how to search*, *how to optimize*, *how to apply*, and *how to evaluate*. While recent related surveys have typically focused on a specific

Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 8244–8279

sub-domain or perspective, *e.g.*, Deep Research which emphasizes professional report generation from extensive information seeking (Xu and Peng, 2025; Huang et al., 2025b) or the integration of reasoning and RAG (Liang et al., 2025; Gao et al., 2025), our work comprehensively analyzes the holistic pipeline of search agents, including search structure, optimization, application, evaluation, and challenges. For each part, we provide a thorough analysis of representative works and developing tendencies.

Specifically, this paper is structured as follows: Sec. 2 introduces the task formulation for search agents. ***How to Search*** in Sec. 3 presents how agents scale up search turns and utilize complex search structures (*i.e.*, parallel, sequential, and hybrid) to determine query content. ***How to Optimize*** in Sec. 4 discusses various optimization methodologies for search agents, including tuning and non-tuning approaches. ***How to Apply*** in Sec. 5 delineates the extensive application areas of search agents, encompassing both internal agent enhancements (*e.g.*, reasoning, memory, and tool-use) and external applications (*e.g.*, math, medicine, and finance). ***How to Evaluate*** in Sec. 5 introduces evaluation of search agents, covering various datasets and metrics. Finally, Sec. 7 presents current challenges and promising future research directions.

2 Task Formulation

Given a user's intention q and context C, a search agent iteratively plans and acts to gather information and fulfill the user's intention.

Upon receiving intention q, the agent initiates a planning $\pi_0 = \text{Plan}(q, C)$ to conduct an information seeking trajectory. At each step t, the agent reflects on its current observation o_t and previous trajectory t and updates its plan $\pi_{t+1} = \text{Reflect}(o_t, h_t, \pi_t)$. It then performs an action $a_{t+1} = \text{Act}(\pi_{t+1})$ (*e.g.*, search for or browse certain content) yielding a new observation o_t, *e.g.*, the retrieved information. This process continues until sufficient information is acquired, forming a sequence of observations $O = \{o_1, o_2, \cdots, o_T\}$. From O, the agent extracts and ranks the most relevant data into an evidence set $E = \text{Select}(q, O)$ and generates a response $\hat{y}_q = \text{Generate}(q, E)$ to fulfill the user's intention.

3 How to Search

The core of a search agent lies in its ability to autonomously determine its actions based on user intent and environment context, deciding when and what to reason or search. This multi-turn process represents a significant shift towards "*scaling up test-time search*." Consequently, the traditional single-query has evolved into dynamic and context-dependent queries, where search queries are decided by sophisticated search structures (parallel, sequential, and hybrid) and search feedback.

3.1 Parallel Structure

Parallel search structures involve reformulating a single query into multiple distinct queries that can be processed simultaneously. This approach is often seen in earlier works, serving as a transitional phase from LLM-enhanced search.

Decomposition-based Parallel Search. When user intents are complex or vague, direct retrieval often fails. Decomposition-based work tackles this by a planning-execution-verification paradigm, breaking the original intent into smaller sub-queries, which are then executed in parallel and synthesized for a complete answer (Press et al., 2022). This approach focuses primarily on how to decompose the original query better. For instance, Khattab et al. (2022); Shi et al. (2024b); Zhang et al. (2025b) leverage strong LLMs through prompting to perform decomposition; Wang et al. (2024a) further learns decomposition strategies from structured knowledge graphs; while Li et al. (2023); Joshi et al. (2024) utilize fine-tuned smaller LLMs to generate retrieval plans and queries.

Diversification-based Parallel Search. Sometimes, a user's intent may correspond to multiple plausible queries. Thus, diversification-based approaches rewrite the original query into a diverse set of queries to be searched in parallel. This strategy ensures that the retrieved content captures a broader range of perspectives and interpretations. For instance, Kostric and Balog (2024) employs beam search to generate multiple candidate queries, while Dhole and Agichtein (2024) synthesizes diverse keywords by ensembling various prompts. Other approaches adopt a hybrid strategy, combining standard query rewriting, keyword extraction, and even the use of LLM-generated Pseudo-Answers as queries to enhance retrieval breadth and relevance (Li et al., 2024c; Abbasiantaeb et al.,

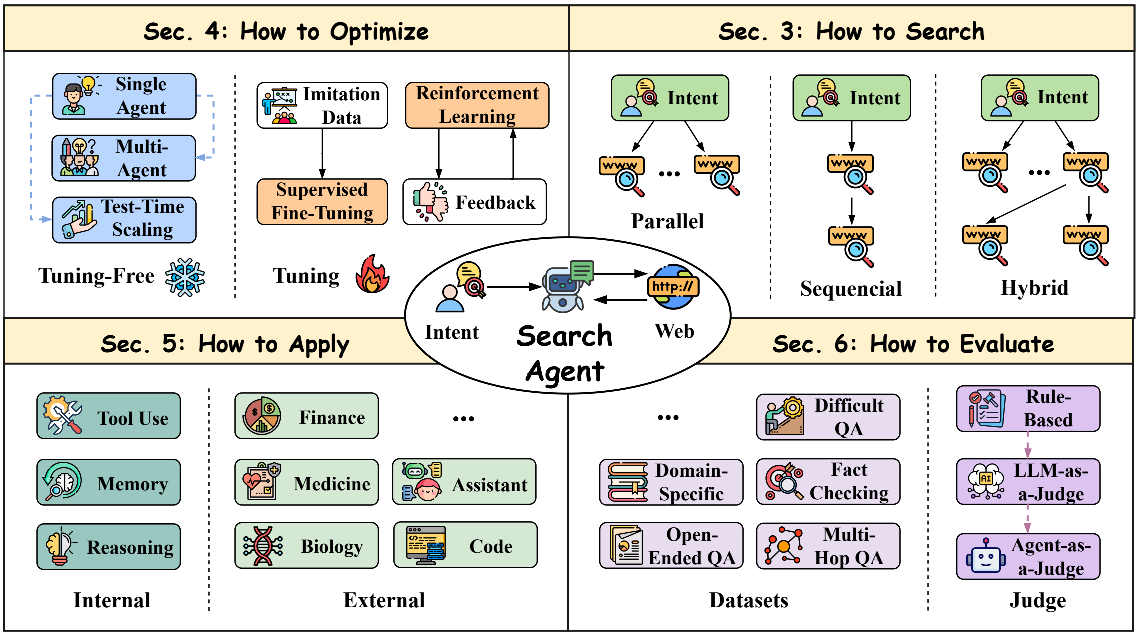

Figure 2: Structural overview of Search Agent – how to search, how to optimize, how to apply, and how to evaluate.

2024; Seo and Lee, 2025).

3.2 Sequential Structure

Parallel structure determines all search queries in advance, making it less adaptable to unexpected issues that may arise during search. In contrast, the sequential structure is more dynamic and flexible, allowing the agent to decide whether and what to search next based on results and reflections from prior steps. Note that early LLM-enhanced search also involves sequential structures; however, they are typically rule-based retrieval for each sentence or thought, rather than agent-driven mode (Yang et al., 2025b; Jiang et al., 2023; Shao et al., 2023; Trivedi et al., 2022a; Wang et al., 2024f).

Reflection-Driven Sequential Search. This approach commonly employs a loop-based mechanism where the agent performs a search, generates an answer based on search results, and then reflects on its quality and correctness (Hayashi et al., 2025; Xiao et al., 2025; Zhou et al., 2024; Lee et al., 2024; lee). Based on the reflection, the agent initiates further planning and search, iterating until a satisfactory response is achieved.

Proactivity-Driven Sequential Search. In this paradigm, the agent decides when to trigger a search and what to search based on the context, *e.g.*, user intent, search results, and previous reasoning. Such dynamic sequential decision-making can be guided by prompts that encode human-designed heuristics, allowing the agent to reason, act, and reflect step by step in response to intermediate outcomes (Li et al., 2025c; Jiang et al., 2025e; Huang et al., 2025a; Wu et al., 2025d). Alternatively, it can be learned through fine-tuning, *i.e.*, imitating expert trajectories (Asai et al., 2023; Islam et al., 2024; Yu et al., 2024; Aksitov et al., 2023) or exploring the environment autonomously to discover more effective strategies (Jin et al., 2025b; Song et al., 2025a; Chen et al., 2025b; Zheng et al., 2025; Wang et al., 2025c).

3.3 Hybrid Structure

The hybrid structure combines both parallel and sequential paradigms, enabling exploration along multiple paths simultaneously, increasing the likelihood of covering the correct answer. Based on the underlying structural properties, it can be categorized into tree-based and graph-based structures.

Tree-based Search Here, each node represents a retrieval step, and at each iteration, multiple successor nodes can be expanded in parallel from a given node. The final answer is then selected or synthesized by aggregating results from various paths to identify the most optimal outcome. Some approaches employ rule-based tree structures, where queries are first decomposed in parallel and then each path is explored independently before merging the results (Zhang et al., 2024c; Li et al., 2025g; Nguyen et al., 2025). Other methods adopt more dynamic strategies by using expansion functions to generate nodes and applying Monte Carlo Tree

Search (MCTS) algorithms, where the final answer is selected through mechanisms such as voting (Trinh et al., 2025a; Feng et al., 2025; Tran et al., 2024) or reward model (Xiong et al., 2025b; Li et al., 2024b; Ren et al., 2025).

Graph-based Search Graph-based structures allow arbitrary connections between nodes, enabling the search process to backtrack and revise earlier decisions. Some approaches decompose the problem into a directed acyclic graph (DAG), identifying dependencies between queries and traversing the graph dynamically (Li et al., 2024a). Others support dynamic node expansion (Chen et al., 2024; Hu et al., 2024) and shrinkage (Teng et al., 2025), allowing the agent to adaptively revise its reasoning and search direction.

Recent trends in search structure indicate a shift towards dynamism, from fixed sub-queries to contextually generated ones, and from parallel to sequential and hybrid structures. As shown in Table 1 and 2, hybrid structures are often preferred in tuning-free settings to cover more search paths and improve performance, while fine-tuned models tend to internalize this flexibility within sequential structures, thereby improving both efficiency and effectiveness.

4 How to Optimize

This section explores the key approaches to optimizing search agents, broadly categorized into tuning-free and tuning-based methods.

4.1 Tuning-Free Approaches

Tuning-free approaches primarily rely on human knowledge and predefined workflows to guide the agent's actions. While basic prompt-driven solutions often involve a single agent, there is a growing trend towards employing multi-agent architectures and test-time scaling to boost performance. See Table 1 for a more detailed comparison.

Single-Agent Architectures. In this architecture, a single agent handles the entire search process, including planning, query generation, and final answer synthesis. Since it's challenging for a single, prompt-driven agent to control every aspect of a complex process dynamically, these systems often adopt pre-defined, structured workflows. For example, some works involve an *iterative refinement loop*, *i.e.*, the agent first searches, then generates an answer, evaluates its quality, and finally decides whether to further search. This loop continues until a satisfactory answer is produced (Hayashi et al., 2025; Xiao et al., 2025; Lee et al., 2024; Zhou et al., 2024). Other works follow the *reflection chain*, where the agent reflects on its current progress or its intermediate results, dynamically adjusting its strategy (Wang et al., 2023; Zhang et al., 2024c; Xiong et al., 2025b; Li et al., 2025g).

Multi-Agent Architectures. While it is challenging for a single agent to handle multiple search tasks, multi-agent architectures decompose complex search tasks and distribute them among specialized agents. Common roles include *planner agent*, control the process, and terminate; *search agent*, gathering evidence from external sources; and *generation agent*, synthesizing the final answer with the collected evidence (Zhou et al., 2026; Chen et al., 2024; Huang et al., 2025a; Hu et al., 2024; Ma et al., 2025). Some work also involves *browser agent* (Huang et al., 2025a; Du et al., 2025), *evaluator agent* (Nguyen et al., 2025; Wang et al., 2025b; Trinh et al., 2025b), and *memory agent* (Wu et al., 2025d).

In terms of process control, most solutions employ a fixed execution order, where one agent passes its output to the next upon completion (Jiang et al., 2025e; Chen et al., 2024; Huang et al., 2025a; Hu et al., 2024; Wang et al., 2025b; Trinh et al., 2025b). Others adopt a centralized supervisor agent that dynamically allocates tasks and invokes other agents (Wu et al., 2025d; Ma et al., 2025).

Test-Time Scaling. Test-time scaling enhances agent performance by allocating more computation during inference to boost task performance, as previously demonstrated by models like OpenAI-o1 and DeepSeek-R1 (Zhang et al., 2025c). Recent work also confirms ***reasoning-centric scaling*** can significantly improve search agent (Zhang et al., 2025e; Lee et al., 2024; Wei et al., 2025).

As search agents interact with external environments, they offer an additional ***search-centric scaling***: increasing the number of interactions with external environments for better external knowledge exploration. Xi et al. (2025a) has observed that the performance of search agents scales smoothly with an increase in the maximum permissible number of search actions. Recent research is integrating both reasoning and search scaling, simultaneously deepening reasoning and allowing for more search actions (Feng et al., 2025; Tran et al., 2024; Jiang et al., 2024b; Ren et al., 2025). This combined approach leverages well-known inference scaling techniques such as Self-Consistency (SC), Best-of-N (BoN), and Monte Carlo Tree Search (MCTS) to achieve superior final results.

4.2 Tuning-based Approaches

Tuning-based approaches involve training the agent to automatically learn its next actions based on the current context through imitation and exploration. See Table 2 for a more detailed comparison.

Supervised Fine-Tuning (SFT). SFT directly trains LLMs on datasets comprising high-quality reasoning and search trajectories or actions, thereby internalizing these capabilities into the model. Recent works leverage SFT to enhance either components of search agents, *e.g.*, query rewriting and reflection, or the fully end-to-end reasoning-search pipeline. For ***component-level tuning***, annotation is often sourced from expert LLMs (Asai et al., 2023; Islam et al., 2024; Yu et al., 2024). For ***holistic training***, synthetic trajectories are generated via LLM-environment interactions and filtered by rejection sampling (Aksitov et al., 2023; Pan et al., 2023; Zhang et al., 2025g). Typical filtering criteria involve the correctness of the answer judged by an LLM against ground truth (Song et al., 2025b; Li et al., 2025d). More nuanced filtering criteria, including diversity, informativeness, and efficiency, are sometimes employed to further optimize training signal fidelity (Li et al., 2025d; Zhang et al., 2025a; Sun et al., 2025b). In some cases, general-purpose SFT data is blended to preserve agents' general capabilities (Lee et al., 2025).

SFT in this context primarily serves a few critical roles: (1) ***Distillation***, wherein data generated via elaborate prompting of powerful teacher models is transferred to smaller student models (Jiang et al., 2025a; Asai et al., 2023; Islam et al., 2024; Lee et al., 2025; Pan et al., 2023); (2) ***Self-Improvement***, involving iterative retraining on self-generated, high-quality trajectories filtered by rejection sampling (Wang et al., 2025a; Aksitov et al., 2023; Wu et al., 2025f); (3) ***Preparation for RL training***, where SFT serves as a crucial warm-up to initialize models with essential operational priors (Zhang et al., 2025g; Li et al., 2025d; Song et al., 2025b; Zhang et al., 2025a; Shi et al., 2025a), and also underpins reward model training for downstream reinforcement learning (Luo et al., 2025; Xiong et al., 2025a; Sun et al., 2025c).

Reinforcement Learning (RL). RL empowers search agents to learn flexible, optimal behaviors through exploration with environments. While some efforts target optimizing specific components of search agents, *e.g.*, retrieval (Jiang et al., 2025b; Hsu et al., 2024) and re-ranking (Xu et al., 2025a), a growing trend is to integrate RL into end-to-end training of the pipeline, including planning, searching, reflection, and generation (Jin et al., 2025b; Song et al., 2025a; Zheng et al., 2025). Commercial solutions, *e.g.*, OpenAI and Gemini's deep research, have also incorporated proprietary RL implementations. For ***RL algorithms***, search agent optimization largely relies on common RL algorithms like PPO (Schulman et al., 2017), GRPO (Shao et al., 2024b), and Reinforce++ (Hu, 2025). Though some studies analyze various algorithms for search agents (Song et al., 2025a; Jin et al., 2025a; Xiong et al., 2025a; Sun et al., 2025a; Jin et al., 2025b), there's no current consensus on which is best.

A crucial aspect of optimizing search agents with RL is constructing ***multi-objective reward functions***. While format adherence and answer correctness are nearly ubiquitous reward components, other objectives frequently integrated include efficiency (Wang et al., 2025c; Huang et al., 2025c), diversity (Mei et al., 2025; Dao and Le, 2025), evidence quality (Qian and Liu, 2025; Zhao et al., 2025), and retrieval gain (Wang et al., 2025d; Shi et al., 2025b). Additionally, penalties for redundancy and length are often incorporated to refine behavior (Wang et al., 2025d; Wu et al., 2025f; Song et al., 2025b). These rewards largely rely on rule-based verification, Outcome-based Reward Models (ORM), and Process-based Reward Models (PRM). Rule-based rewards are applied for verifiable results, *e.g.*, questions with standard answers. ORM typically uses LLMs to judge results without standard answers. PRM assesses utility from each individual step within the search trajectory (Zhang et al., 2025f; Xiong et al., 2025a; Sun et al., 2025c).

Mixed Approaches. These represent a robust strategy that combines multiple tuning methods. The prevailing methodology involves utilizing SFT as a warm-up phase for the RL stage (Zhang et al., 2025g; Li et al., 2025d; Song et al., 2025b; Zhang et al., 2025a; Shi et al., 2025a). This helps the model adapt to the task and correct format, providing a sound initialization for RL training.

Further optimizations implement an iterative training loop between SFT and RL (Zhang et al., 2025a), where SFT is optimized with high-quality rollouts from previous RL iterations, providing an enhanced initialization for the next RL training cycle. Shi et al. (2025c) employs a Generalized Expectation-Maximization framework for iterative trajectory exploration and optimization. Zhang et al. (2025g) integrates contrastive learning to learn when to trigger retrieval more effectively. Furthermore, Wu et al. (2025f) introduces a pre-training task, Retrieval-Augmented Mask Prediction, before the SFT and RL stages to boost the model's fundamental capabilities.

4.3 Discussion

The choice between tuning-free and tuning-based methods largely depends on available resources and deployment constraints. Tuning-free methods offer zero training cost, strong flexibility, and broad applicability, but may be limited in task-specific performance and inference efficiency. In contrast, tuning-based methods generally achieve superior results in domain-specific scenarios and support lightweight model deployment with higher inference efficiency, at the expense of greater training costs and potential risks of limited generalization. Below, we distill key design insights for each paradigm.

Improving Effectiveness of Tuning-Free Methods. Several complementary strategies can enhance tuning-free agents. First, *multi-agent design* assigns specialized roles such as planner, searcher, and browser to reduce the complexity of each agent's task (Wu et al., 2025b; Huang et al., 2025a). Second, *enhanced planning* decomposes complex problems into simpler sub-tasks for more efficient execution (Chen et al., 2024; Lee et al., 2024). Third, *advanced search structures* employ test-time scaling with more sophisticated structures such as trees or graphs, enabling deeper reasoning and broader exploration (Li et al., 2024b; Ren et al., 2025).

Improving Efficiency of Tuning-Free Methods. Efficiency can be improved through several dimensions. *Well-orchestrated multi-agent systems* introduce parallel execution (Chen et al., 2024) and employ summarizers to store useful information, thereby avoiding redundant planning (Jiang et al., 2024d). *Pruning and parallelization* techniques reduce unnecessary exploration while leveraging parallelism in complex structures (Li et al., 2024b). Furthermore, *balanced tool usage* avoids excessive reliance on external tools by encouraging stronger utilization of the model's own parametric knowledge. Finally, *adaptive complexity selection* dynamically selects methods of varying complexity (*e.g.*, direct search, parallel search, or planning-based search) depending on task difficulty (Wang and Xu, 2024).

Improving Effectiveness of Tuning-Based Methods. Several principles emerge from the tuning-based literature. The mainstream pipeline of *SFT for warm-up followed by RL fine-tuning* remains the most effective approach, with RL typically leveraging verifiable rewards (*e.g.*, format correctness and answer accuracy), mainly optimized using PPO or GRPO. *High-quality supervised trajectories* that emphasize diversity and quality control during data collection are essential (Wu et al., 2025a; Geng et al., 2025). *Multi-stage learning* gradually enhances the model's capabilities through staged training (Song et al., 2025a; Zhang et al., 2025g). Applying RL on *challenging tasks* pushes model robustness further (Li et al., 2025b; Geng et al., 2025). *Intermediate rewards* mitigate reward sparsity by incorporating process-level feedback (Wang et al., 2025d; Deng et al., 2025). Additionally, *multi-agent architectures* introduce collaborative frameworks to improve scalability and specialization (Hu et al., 2025b; Chen et al., 2025d).

5 How to Apply

Search Agents, with their flexibility and proactivity, have expanded information retrieval far beyond traditional web search. Externally, they conduct in-depth information seeking in diverse domains; internally, they enhance the agent's capabilities through targeted information search. More applications and comparisons are provided in Table 3.

5.1 External Applications

Search Agents are revolutionizing various industries and applications. A prominent example of this is their seamless integration into chatbots and **AI assistants**, *e.g.*, OpenAI, Gemini, Perplexity, and Gork. A particularly significant development within AI assistants is Deep Research (DR). DR systems are meticulously designed to conduct exhaustive searches across diverse sources, synthesize vast amounts of disparate information, and then present their findings in a well-organized, often professional report. Since 2025, there have been many successful commercial applications, *e.g.*, OpenAI DR (OpenAI, 2025), Perplexity DR (Perplexity, 2025), Gemini DR (Gemini, 2025). The field is also thriving with open-source projects like Jina AI node-DeepResearch (Jina, 2025) and Langchain's Open Deep Research (LangChain, 2025), alongside a growing body of academic research work (Yang et al., 2025c; Singh et al., 2025; Li et al., 2025d; Yuan et al., 2025).

Beyond AI assistants, Search Agents are also finding fertile ground in a variety of other specialized domains, including **e-commerce** (Bi et al., 2025; Lyu et al., 2025; Zhao et al., 2024; Wang et al., 2024e), **finance** (Li et al., 2024a; Lee et al., 2024; Vaghefi et al., 2025), **code** (Singh et al., 2025; Zhang et al., 2024a; Singh et al., 2025), **medicine** (Bhattacharyya, 2025; Chen et al., 2025f; Wang et al., 2025e), **biology** (Liu et al., 2024b; Al Dajani et al., 2025) and **chemistry** (Callahan et al., 2025; Li et al., 2025f). Besides, Search Agents serve as powerful teaching and **research assistants** by efficiently collecting materials across various fields (He et al., 2025; Brett and Myatt, 2025; Schneider et al., 2025).

5.2 Internal Applications

Beyond external sources, a crucial, often overlooked, retrieval source for an agent lies within its internal components: its memory, accumulated experiences, and available tools. Agentic search can be effectively introduced to significantly enhance its core capabilities:

Tool Use. As agents gain access to a growing arsenal of tools, identifying the most appropriate one for a given task becomes a pressing challenge. Agentic search offers a solution by multi-turn reasoning and search, enabling more dynamic and precise tool selection (Lumer et al., 2024; Du et al., 2024; Lumer et al., 2025; Xu et al., 2024a).

Memory. With user interactions accumulating, an agent's memory can become vast and unwieldy. Effectively navigating this information to pinpoint content is another area where agentic search excels. Agentic search can extract queries from complex and ambiguous user intentions and then conduct deep searches within the agent's memory to retrieve highly pertinent information (Xu et al., 2025b; Ocker et al., 2025; Tan et al., 2024b).

Reasoning. The experiences an agent learns through its self-evolution serve as an invaluable internal search source. Agentic search can dynamically retrieve relevant experiences from this internal reservoir, combining them with externally acquired knowledge to facilitate more robust and insightful reasoning (Wu et al., 2025d; Wang et al., 2024f; Wu et al., 2025g; Wang et al., 2023).

Currently, Search Agents are primarily used for Question Answering (QA) tasks, but are rapidly expanding to a broader spectrum of external domains and internal agent capabilities. The ultimate trajectory envisions a seamless and powerful integration of both external knowledge acquisition and internal self-optimization, creating more capable and adaptable agents for complex real-world scenarios.

6 How to Evaluate

Evaluating the performance of Search Agents is crucial for understanding their strengths, weaknesses, and overall effectiveness. This section outlines the evaluation process of search agents, including datasets, metrics, and judgment criteria. See Table 4 for more details.

6.1 Datasets for Evaluation

The datasets for evaluating search agents primarily include complex Question-Answering (QA), alongside challenging reasoning problems that require extensive information seeking.

Closed-ended QA. Closed-ended questions have specific, definite answers, which are easy to evaluate. However, simple closed-ended QA datasets like NQ (Kwiatkowski et al., 2019) and TriviaQA (Joshi et al., 2017) are often insufficient to effectively evaluate search agents in multi-turn search. Therefore, more challenging closed-ended QA datasets are developed, specifically focusing on multi-hop QA, challenging QA, and fact-checking. **Multi-hop QA** requires search agents to synthesize information from multiple sources, necessitating iterative information retrieval (Yang et al., 2018; Ho et al., 2020; Press et al., 2022; Trivedi et al., 2022b). **Challenging QA** is designed to present genuinely difficult problems, often featuring long-horizon questions (Wei et al., 2025; Zhou et al., 2025b) and involving long-tail knowledge and significant distracting information (Xi et al., 2025a; Pham et al., 2025). **Fact-checking task** evaluates a search agent's ability to verify factual claims (Wei et al., 2024; Wadden et al., 2020; Jiang et al., 2020; Wang et al., 2024d). It inherently involves iterative search, browsing, and comparative analysis of disparate information sources. See more details in Appendix A.

Open-ended QA Dataset. While closed-ended problems with definitive answers, many user inquiries are inherently open-ended, lacking a single, unambiguous correct response. Therefore, it is crucial to assess a search agent's performance on such open-ended questions. Our focus here is primarily on datasets designed for open-ended problems requiring deep information seeking, which is central to the objectives of Deep Research, a subcategory of search agents. These involve delving into broad topics to produce comprehensive, high-quality research reports (Coelho et al., 2025). Such datasets target multi-perspective or non-factual queries (Rosset et al., 2024) and expert-level research task (Du et al., 2025; Bosse et al., 2025; Tan et al., 2024a), with some incorporating multi-modal queries (Yang et al., 2025c).

Domain-Specific Dataset. The evaluation of search agents also involved solving domain-specific tasks with agentic information-seeking abilities. This encompasses two prominent subtypes: (1) Information Seeking in Specialized Fields: This involves iterative retrieval unique to domains like finance (Li et al., 2024a), business (Chen et al., 2025a; Choubey et al., 2025), Medicine (Pal et al., 2022; Chen et al., 2025c), and agriculture (Dongre et al., 2025), where search agents must integrate domain-specific knowledge to provide relevant, precise answers. (2) Hard Problems in Specific Domains: These involve highly complex challenges in fields like mathematics (He et al., 2024), physics (Rein et al., 2024), and code (Shi et al., 2024a), often resembling expert-level or Olympiad-style problems (Mialon et al., 2023; Phan et al., 2025). Solving such problems requires not only internal reasoning capabilities but also the ability to leverage external information and prior experiences, making agentic search essential.

6.2 Metrics & Judgment

This section introduces various metrics and judgment approaches for evaluating Search Agents.

Metrics. Most benchmarks assess the performance of search agents based on the *effectiveness* of the final output. For closed-ended QA, fact-checking, and most domain-specific tasks, the dominant metric is typically task success rate, *e.g.*, Exact Match (EM), F1 score, Accuracy, and Pass@k. Sometimes, the evaluation extends beyond the final output to the *intermediate processes*, such as evaluating the quality of the reasoning chain (Wu et al., 2024) and retrieval (Su et al., 2024; Yao et al., 2023; Wei et al., 2024; Eisenschlos et al., 2021). Since most benchmarks rely on static search environments, retrieval quality is often measured with standard metrics, *e.g.*, Precision, Recall, NDCG, and MAP, calculated against groundtruth documents. Recent work also explores retrieval quality in dynamic environments (Xi et al., 2025a). .

The evaluation of *open-ended tasks* for Deep Research presents significantly more complexity compared to closed-ended tasks. This complexity arises from the absence of a single "correct" answer, necessitating a nuanced evaluation of the agent's ability to synthesize information comprehensively. Their assessment typically requires multifaceted metrics, *e.g.*, key point coverage (Qi et al., 2024; Coelho et al., 2025), informativeness (Du et al., 2025), breadth and depth (Jiang et al., 2024c), coherence, organization (Yang et al., 2025c), readability (Du et al., 2025), and citation accuracy (Du et al., 2025). Furthermore, some studies adopt Arena-based evaluation (Win Rate) (Chandrahasan et al., 2025; Miroyan et al., 2025), where outputs from different search agents are presented side by side to human or LLM-based judges, who then determine which output is superior, tied, or inferior. This provides a comparative assessment of performance, particularly useful for nuanced, open-ended tasks where absolute metrics might fall short.

Judge. The judge of the above metrics has undergone significant evolution, progressing from simplistic rule-based metrics to more intricate LLM- and Agent-based judging paradigms. Initially, search agents were evaluated using **Rule-based Judges**, such as metrics Exact Match (EM) and F1 score, against predefined ground truth answers. However, these metrics fail to account for semantic variations, where factually correct answers may be phrased differently. Thus, **LLM-as-a-Judge** paradigm emerged, leveraging LLMs to assess the accuracy and quality of results. This approach is effective for both closed-ended tasks, where it correlates well with ground truth, and open-ended questions, where evaluation relies on predefined qualitative standards and expert reference reports. As agentic search involves multi-step reasoning and dynamic interaction, the **Agent-as-a-Judge** paradigm is gaining traction (Gou et al., 2025).

This approach utilizes specialized agents to evaluate the entire search process and its output, providing a deeper assessment than LLM-as-a-Judge. Besides, human evaluation remains the gold standard for nuanced judgment; however, its high cost limits its use, typically focusing on limited samples to validate LLM-based evaluation.

Although the scope of evaluation metrics and the precision of judging methodologies continue to improve, there is a compelling need for more complex and comprehensive evaluation dimensions (Zhu et al., 2025b). Current benchmarks primarily focus on efficiency and information accuracy, but the core competencies of search agents—particularly their ability to effectively retrieve, synthesize, and discriminate between information—should be central to any evaluation framework. Future evaluation paradigms must expand to include metrics that rigorously assess not only efficiency and accuracy but also source citation reliability and the agent's capacity to distinguish between credible and unreliable information. This will ensure a more thorough, holistic, and robust assessment of search agents.

7 Challenges and Future Directions

Despite recent advancements, search agents face significant challenges in handling the complexity and imperfections of real-world information. Due to page limitations, we briefly discuss these challenges and future directions here, and give a detailed discussion in Appendix B .

A primary hurdle of search agents is the need to ***broaden and fuse diverse information sources***, moving beyond external web data to integrate private databases and internal experiences and memory within agents. This integration is complicated by the problem of ***imperfect retrieval***, as external sources can be biased, factually incorrect, or misleading. Agents must be able to validate and synthesize information from conflicting sources. Furthermore, search agents need to transition ***from text to multi-modality***, a transformation that requires both new search infrastructure and more advanced agent comprehension, reasoning, and retrieval skills across various modalities.

The second set of challenges centers on the agent's core mechanics and underlying infrastructure. This requires building more autonomous and robust systems. A key area is developing ***customized reinforcement learning*** algorithms that are specifically optimized for the long-horizon, multi-step nature of search tasks, which are not always well-suited to general RL frameworks. We also need to build more ***resilient infrastructure***, from optimizing the efficiency of RL sampling to creating high-recall retrieval systems and intelligent scheduling mechanisms. The ultimate frontier for search agents is enabling true ***self-evolution***, where the agent can autonomously identify its own flaws, devise new strategies, and continuously improve over time without constant human guidance.

8 Conclusion

The evolution from traditional web search to LLM-enhanced search and Search Agents profoundly transforms information retrieval. These autonomous agents proactively leverage context and diverse sources, transforming search into a proactive, intelligent process. Our survey offers the first systematic analysis, dissecting their mechanisms, optimization, applications, and evaluation. This comprehensive view illuminates the vast potential of search agents for deep information mining and highlights challenges for fully realizing their transformative promise.

Limitations

This survey, while comprehensive and systematic, has some limitations. This work primarily focuses on academic research papers, which means it less extensively covers the intricacies of commercial applications. Since companies like OpenAI, Google (Gemini), and Perplexity often do not disclose the specific technical details of their deep research or search agent implementations, there's an inherent gap. This raises a crucial question: are the research directions and observed performance in academic settings truly aligned with the approaches and effectiveness seen in real-world commercial deployments?

Acknowledgment

The Shanghai Jiao Tong University team is partially supported by National Key RD Program of China (2022ZD0114804), Shanghai Municipal Science and Technology Major Project (2021SHZDZX0102) and National Natural Science Foundation of China (624B2096, 72542012, 72595872, 62322603, 62502310). The author Yunjia Xi is also supported by Wu Wen Jun Honorary Doctoral Scholarship.

References

Agent-g: An agentic framework for graph retrieval augmented generation.

Zahra Abbasiantaeb, Simon Lupart, and Mohammad Aliannejadi. 2024. Generating multi-aspect queries for conversational search. *arXiv preprint arXiv:2403.19302*.

Renat Aksitov, Sobhan Miryoosefi, Zonglin Li, Daliang Li, Sheila Babayan, Kavya Kopparapu, Zachary Fisher, Ruiqi Guo, Sushant Prakash, Pranesh Srinivasan, and 1 others. 2023. Rest meets react: Self-improvement for multi-step reasoning llm agent. *arXiv preprint arXiv:2312.10003*.

Saleem A Al Dajani, Abel Sanchez, and John R Williams. 2025. Deepseq: High–throughput single–cell rna sequencing data labeling via web search–augmented agentic generative ai foundation models. *bioRxiv*, pages 2025–06.

Rami Aly, Zhijiang Guo, Michael Schlichtkrull, James Thorne, Andreas Vlachos, Christos Christodoulopoulos, Oana Cocarascu, and Arpit Mittal. 2021. Feverous: Fact extraction and verification over unstructured and structured information. *arXiv preprint arXiv:2106.05707*.

Salaheddin Alzubi, Creston Brooks, Purva Chiniya, Edoardo Contente, Chiara von Gerlach, Lucas Irwin, Yihan Jiang, Arda Kaz, Windsor Nguyen, Sewoong Oh, and 1 others. 2025. Open deep search: Democratizing search with open-source reasoning agents. *arXiv preprint arXiv:2503.20201*.

Akari Asai, Zeqiu Wu, Yizhong Wang, Avirup Sil, and Hannaneh Hajishirzi. 2023. Self-rag: Learning to retrieve, generate, and critique through self-reflection. In *The Twelfth International Conference on Learning Representations*.

assafelovic. 2025. gpt-researcher. https://github.com/assafelovic/gpt-researcher, Accessed on 2025-7-1.

Roucher Aymeric. 2025. Open-source deepresearch – freeing our search agents. https://huggingface.co/blog/open-deep-research, Accessed on 2025-7-1.

Farima Fatahi Bayat, Lechen Zhang, Sheza Munir, and Lu Wang. 2024. Factbench: A dynamic benchmark for in-the-wild language model factuality evaluation. *arXiv preprint arXiv:2410.22257*.

Thomas Berkane, Marie-Laure Charpignon, and Maimuna Majumder. 2025. Llm-based web data collection for research dataset creation. *medRxiv*, pages 2025–05.

Kiran Bhattacharyya. 2025. Surgical information assistant: A technical report on an agentic information retrieval system for surgical information. *medRxiv*, pages 2025–05.

Manish Bhattarai, Miguel Cordova, Javier Santos, and Dan O'Malley. 2025. Arcs: Agentic retrieval-augmented code synthesis with iterative refinement. *arXiv preprint arXiv:2504.20434*.

Yuxi Bi, Yunfan Gao, and Haofen Wang. 2025. Steporec: Towards personalized outfit styling assistant via knowledge-guided multi-step reasoning. *arXiv preprint arXiv:2504.09915*.

Nikos I Bosse, Jon Evans, Robert G Gambee, Daniel Hnyk, Peter Mühlbacher, Lawrence Phillips, Dan Schwarz, Jack Wildman, and 1 others. 2025. Deep research bench: Evaluating ai web research agents. *arXiv preprint arXiv:2506.06287*.

David Brett and Anniek Myatt. 2025. Patience is all you need! an agentic system for performing scientific literature review. *arXiv preprint arXiv:2504.08752*.

EFFECT BUTTERFLY. 2025. Leave it to manus. https://manus.im/home, Accessed on 2025-6-30.

Bytedance. 2025. Deerflow. https://github.com/bytedance/deer-flow, Accessed on 2025-7-1.

Hongru Cai, Yongqi Li, Wenjie Wang, Fengbin Zhu, Xiaoyu Shen, Wenjie Li, and Tat-Seng Chua. 2025. Large language models empowered personalized web agents. In *Proceedings of the ACM on Web Conference 2025*, pages 198–215.

Tiffany J Callahan, Nathaniel H Park, and Sara Capponi. 2025. Agentic mixture-of-workflows for multi-modal chemical search. *arXiv preprint arXiv:2502.19629*.

Arie Cattan, Alon Jacovi, Ori Ram, Jonathan Herzig, Roee Aharoni, Sasha Goldshtein, Eran Ofek, Idan Szpektor, and Avi Caciularu. 2025. Dragged into conflicts: Detecting and addressing conflicting sources in search-augmented llms. *arXiv preprint arXiv:2506.08500*.

Sky CH-Wang, Darshan Deshpande, Smaranda Muresan, Anand Kannappan, and Rebecca Qian. 2025. Browsing lost unformed recollections: A benchmark for tip-of-the-tongue search and reasoning. *arXiv preprint arXiv:2503.19193*.

Prahaladh Chandrahasan, Jiahe Jin, Zhihan Zhang, Tevin Wang, Andy Tang, Lucy Mo, Morteza Ziyadi, Leonardo F. R. Ribeiro, Zimeng Qiu, Markus Dreyer, Akari Asai, and Chenyan Xiong. 2025. Deep research comparator: A platform for fine-grained human annotations of deep research agents. *Preprint*, arXiv:2507.05495.

Kaiyuan Chen, Yixin Ren, Yang Liu, Xiaobo Hu, Haotong Tian, Tianbao Xie, Fangfu Liu, Haoye Zhang, Hongzhang Liu, Yuan Gong, and 1 others. 2025a. xbench: Tracking agents productivity scaling with profession-aligned real-world evaluations. *arXiv preprint arXiv:2506.13651*.

Mingyang Chen, Tianpeng Li, Haoze Sun, Yijie Zhou, Chenzheng Zhu, Haofen Wang, Jeff Z Pan, Wen Zhang, Huajun Chen, Fan Yang, and 1 others. 2025b. Learning to reason with search for llms via reinforcement learning. *arXiv preprint arXiv:2503.19470*.

Shan Chen, Pedro Moreira, Yuxin Xiao, Sam Schmidgall, Jeremy Warner, Hugo Aerts, Thomas Hartvigsen, Jack Gallifant, and Danielle S Bitterman. 2025c. Medbrowsecomp: Benchmarking medical deep research and computer use. *arXiv preprint arXiv:2505.14963*.

Yiqun Chen, Lingyong Yan, Weiwei Sun, Xinyu Ma, Yi Zhang, Shuaiqiang Wang, Dawei Yin, Yiming Yang, and Jiaxin Mao. 2025d. Improving retrieval-augmented generation through multi-agent reinforcement learning. *arXiv preprint arXiv:2501.15228*.

Yiqun Chen, Lingyong Yan, Weiwei Sun, Xinyu Ma, Yi Zhang, Shuaiqiang Wang, Dawei Yin, Yiming Yang, and Jiaxin Mao. 2025e. Improving retrieval-augmented generation through multi-agent reinforcement learning. *arXiv preprint arXiv:2501.15228*.

Yixiang Chen, Penglei Sun, Xiang Li, and Xiaowen Chu. 2025f. Mrd-rag: Enhancing medical diagnosis with multi-round retrieval-augmented generation. *arXiv preprint arXiv:2504.07724*.

Zehui Chen, Kuikun Liu, Qiuchen Wang, Jiangning Liu, Wenwei Zhang, Kai Chen, and Feng Zhao. 2024. Mindsearch: Mimicking human minds elicits deep ai searcher. *arXiv preprint arXiv:2407.20183*.

Prafulla Kumar Choubey, Xiangyu Peng, Shilpa Bhagavath, Kung-Hsiang Huang, Caiming Xiong, and Chien-Sheng Wu. 2025. Benchmarking deep search over heterogeneous enterprise data. *Preprint*, arXiv:2506.23139.

João Coelho, Jingjie Ning, Jingyuan He, Kangrui Mao, Abhijay Paladugu, Pranav Setlur, Jiahe Jin, Jamie Callan, João Magalhães, Bruno Martins, and 1 others. 2025. Deepresearchgym: A free, transparent, and reproducible evaluation sandbox for deep research. *arXiv preprint arXiv:2505.19253*.

Consensus. 2025. Consensus: Ai search engine for research. https://consensus.app/home/blog/welcome-to-consensus/, Accessed on 2025-6-30.

Xinyi Dai, Jianghao Lin, Weinan Zhang, Shuai Li, Weiwen Liu, Ruiming Tang, Xiuqiang He, Jianye Hao, Jun Wang, and Yong Yu. 2021. An adversarial imitation click model for information retrieval. In *Proceedings of the Web Conference 2021*, pages 1809–1820.

Alan Dao and Thinh Le. 2025. Rezero: Enhancing llm search ability by trying one-more-time. *arXiv preprint arXiv:2504.11001*.

Yong Deng, Guoqing Wang, Zhenzhe Ying, Xiaofeng Wu, Jinzhen Lin, Wenwen Xiong, Yuqin Dai, Shuo Yang, Zhanwei Zhang, Qiwen Wang, and 1 others. 2025. Atom-searcher: Enhancing agentic deep research via fine-grained atomic thought reward. *arXiv preprint arXiv:2508.12800*.

Kaustubh D Dhole and Eugene Agichtein. 2024. Genqrensemble: Zero-shot llm ensemble prompting for generative query reformulation. In *European Conference on Information Retrieval*, pages 326–335. Springer.

Guanting Dong, Chenghao Zhang, Mengjie Deng, Yutao Zhu, Zhicheng Dou, and Ji-Rong Wen. 2024. Progressive multimodal reasoning via active retrieval. *arXiv preprint arXiv:2412.14835*.

Vardhan Dongre, Chi Gui, Shubham Garg, Hooshang Nayyeri, Gokhan Tur, Dilek Hakkani-Tür, and Vikram S Adve. 2025. Mirage: A benchmark for multimodal information-seeking and reasoning in agricultural expert-guided conversations. *arXiv preprint arXiv:2506.20100*.

Mingxuan Du, Benfeng Xu, Chiwei Zhu, Xiaorui Wang, and Zhendong Mao. 2025. Deepresearch bench: A comprehensive benchmark for deep research agents. *arXiv preprint arXiv:2506.11763*.

Yu Du, Fangyun Wei, and Hongyang Zhang. 2024. Anytool: Self-reflective, hierarchical agents for large-scale api calls. *arXiv preprint arXiv:2402.04253*.

dzhng. 2025. Open deep research. https://github.com/dzhng/deep-research, Accessed on 2025-7-1.

Julian Martin Eisenschlos, Bhuwan Dhingra, Jannis Bulian, Benjamin Börschinger, and Jordan Boyd-Graber. 2021. Fool me twice: Entailment from wikipedia gamification. *arXiv preprint arXiv:2104.04725*.

Siavash Hakim Elahi and Danielle Zyngier. One prompt to rule them all: Automated curve and image retrieval from pdfs and websites with agentic multimodal rag. In *World Environmental and Water Resources Congress 2025*, pages 158–170.

Abdelrahman Elewah and Khalid Elgazzar. 2025. Agentic search engine for real-time iot data. *arXiv preprint arXiv:2503.12255*.

Wenqi Fan, Yujuan Ding, Liangbo Ning, Shijie Wang, Hengyun Li, Dawei Yin, Tat-Seng Chua, and Qing Li. 2024. A survey on rag meeting llms: Towards retrieval-augmented large language models. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 6491–6501.

Wenfeng Feng, Chuzhan Hao, Yuewei Zhang, Jingyi Song, and Hao Wang. 2025. Airrag: Activating intrinsic reasoning for retrieval augmented generation via tree-based search. *arXiv preprint arXiv:2501.10053*.

James Ferguson, Matt Gardner, Hannaneh Hajishirzi, Tushar Khot, and Pradeep Dasigi. 2020. Iirc: A dataset of incomplete information reading comprehension questions. *arXiv preprint arXiv:2011.07127*.

Fosowl. 2025. Agenticseek: Private, local manus alternative. https://github.com/Fosowl/agenticSeek, Accessed on 2025-7-1.

FoundationAgents. 2025. Openmanus. https://github.com/FoundationAgents/OpenManus, Accessed on 2025-7-1.

Lingyue Fu, Jianghao Lin, Weiwen Liu, Ruiming Tang, Weinan Zhang, Rui Zhang, and Yong Yu. 2023. An f-shape click model for information retrieval on multi-block mobile pages. In *Proceedings of the sixteenth ACM international conference on web search and data mining*, pages 1057–1065.

Jingsheng Gao, Linxu Li, Weiyuan Li, Yuzhuo Fu, and Bin Dai. 2024. Smartrag: Jointly learn rag-related tasks from the environment feedback. *arXiv preprint arXiv:2410.18141*.

Yunfan Gao, Yun Xiong, Xinyu Gao, Kangxiang Jia, Jinliu Pan, Yuxi Bi, Yixin Dai, Jiawei Sun, Haofen Wang, and Haofen Wang. 2023. Retrieval-augmented generation for large language models: A survey. *arXiv preprint arXiv:2312.10997*, 2(1).

Yunfan Gao, Yun Xiong, Yijie Zhong, Yuxi Bi, Ming Xue, and Haofen Wang. 2025. Synergizing rag and reasoning: A systematic review. *arXiv preprint arXiv:2504.15909*.

Gemini. 2025. Gemini deep research. https://gemini.google/overview/deep-research/, Accessed on 2025-6-26.

Xinyu Geng, Peng Xia, Zhen Zhang, Xinyu Wang, Qiuchen Wang, Ruixue Ding, Chenxi Wang, Jialong Wu, Yida Zhao, Kuan Li, and 1 others. 2025. Webwatcher: Breaking new frontier of vision-language deep research agent. *arXiv preprint arXiv:2508.05748*.

Mor Geva, Daniel Khashabi, Elad Segal, Tushar Khot, Dan Roth, and Jonathan Berant. 2021a. Did aristotle use a laptop? a question answering benchmark with implicit reasoning strategies. *Transactions of the Association for Computational Linguistics*, 9:346–361.

Mor Geva, Daniel Khashabi, Elad Segal, Tushar Khot, Dan Roth, and Jonathan Berant. 2021b. Did aristotle use a laptop? a question answering benchmark with implicit reasoning strategies. *Transactions of the Association for Computational Linguistics*, 9:87–100.

Anna Goldie, Azalia Mirhoseini, Hao Zhou, Irene Cai, and Christopher D Manning. 2025. Synthetic data generation & multi-step rl for reasoning & tool use. *arXiv preprint arXiv:2504.04736*.

Gork. 2025. Grok 3 beta — the age of reasoning agents. https://x.ai/news/grok-3, Accessed on 2025-6-26.

Boyu Gou, Zanming Huang, Yuting Ning, Yu Gu, Michael Lin, Weijian Qi, Andrei Kopanev, Botao Yu, Bernal Jiménez Gutiérrez, Yiheng Shu, and 1 others. 2025. Mind2web 2: Evaluating agentic search with agent-as-a-judge. *arXiv preprint arXiv:2506.21506*.

Xinyan Guan, Jiali Zeng, Fandong Meng, Chunlei Xin, Yaojie Lu, Hongyu Lin, Xianpei Han, Le Sun, and Jie Zhou. 2025. Deeprag: Thinking to retrieval step by step for large language models. *arXiv preprint arXiv:2502.01142*.

Muhammad Usman Hadi, Rizwan Qureshi, Abbas Shah, Muhammad Irfan, Anas Zafar, Muhammad Bilal Shaikh, Naveed Akhtar, Jia Wu, Seyedali Mirjalili, and 1 others. 2023. A survey on large language models: Applications, challenges, limitations, and practical usage. *Authorea Preprints*, 3.

Kazuki Hayashi, Hidetaka Kamigaito, Shinya Kouda, and Taro Watanabe. 2025. Iterkey: Iterative keyword generation with llms for enhanced retrieval augmented generation. *arXiv preprint arXiv:2505.08450*.

Chaoqun He, Renjie Luo, Yuzhuo Bai, Shengding Hu, Zhen Leng Thai, Junhao Shen, Jinyi Hu, Xu Han, Yujie Huang, Yuxiang Zhang, and 1 others. 2024. Olympiadbench: A challenging benchmark for promoting agi with olympiad-level bilingual multimodal scientific problems. *arXiv preprint arXiv:2402.14008*.

Yichen He, Guanhua Huang, Peiyuan Feng, Yuan Lin, Yuchen Zhang, Hang Li, and 1 others. 2025. Pasa: An llm agent for comprehensive academic paper search. *arXiv preprint arXiv:2501.10120*.

Xanh Ho, Anh-Khoa Duong Nguyen, Saku Sugawara, and Akiko Aizawa. 2020. Constructing a multi-hop qa dataset for comprehensive evaluation of reasoning steps. *arXiv preprint arXiv:2011.01060*.

Sheryl Hsu, Omar Khattab, Chelsea Finn, and Archit Sharma. 2024. Grounding by trying: Llms with reinforcement learning-enhanced retrieval. *arXiv preprint arXiv:2410.23214*.

Chuanrui Hu, Shichong Xie, Baoxin Wang, Bin Chen, Xiaofeng Cong, and Jun Zhang. 2024. Level-navi agent: A framework and benchmark for chinese web search agents. *arXiv preprint arXiv:2502.15690*.

Chuanrui Hu, Shichong Xie, Baoxin Wang, Bin Chen, Xiaofeng Cong, and Jun Zhang. 2025a. Level-navi agent: A framework and benchmark for chinese web search agents. *arXiv preprint arXiv:2502.15690*.

Jian Hu. 2025. Reinforce++: A simple and efficient approach for aligning large language models. *arXiv preprint arXiv:2501.03262*.

Qisheng Hu, Quanyu Long, and Wenya Wang. 2025b. Coordinating search-informed reasoning and reasoning-guided search in claim verification. *arXiv preprint arXiv:2506.07528*.

Yunhai Hu, Yilun Zhao, Chen Zhao, and Arman Cohan. 2025c. Mcts-rag: Enhancing retrieval-augmented generation with monte carlo tree search. *arXiv preprint arXiv:2503.20757*.

Lisheng Huang, Yichen Liu, Jinhao Jiang, Rongxiang Zhang, Jiahao Yan, Junyi Li, and Wayne Xin Zhao. 2025a. Manusearch: Democratizing deep search in large language models with a transparent and open multi-agent framework. *arXiv preprint arXiv:2505.18105*.

Yuxuan Huang, Yihang Chen, Haozheng Zhang, Kang Li, Meng Fang, Linyi Yang, Xiaoguang Li, Lifeng Shang, Songcen Xu, Jianye Hao, Kun Shao, and Jun Wang. 2025b. Deep research agents: A systematic examination and roadmap. *Preprint*, arXiv:2506.18096.

Ziyang Huang, Xiaowei Yuan, Yiming Ju, Jun Zhao, and Kang Liu. 2025c. Reinforced internal-external knowledge synergistic reasoning for efficient adaptive search agent. *arXiv preprint arXiv:2505.07596*.

Shayekh Bin Islam, Md Asib Rahman, KSM Hossain, Enamul Hoque, Shafiq Joty, and Md Rizwan Parvez. 2024. Open-rag: Enhanced retrieval-augmented reasoning with open-source large language models. *arXiv preprint arXiv:2410.01782*.

Chelsi Jain, Yiran Wu, Yifan Zeng, Jiale Liu, Zhenwen Shao, Qingyun Wu, Huazheng Wang, and 1 others. 2025. Simpledoc: Multi-modal document understanding with dual-cue page retrieval and iterative refinement. *arXiv preprint arXiv:2506.14035*.

Dongzhi Jiang, Renrui Zhang, Ziyu Guo, Yanmin Wu, Jiayi Lei, Pengshuo Qiu, Pan Lu, Zehui Chen, Chaoyou Fu, Guanglu Song, and 1 others. 2024a. Mmsearch: Benchmarking the potential of large models as multi-modal search engines. *arXiv preprint arXiv:2409.12959*.

Jinhao Jiang, Jiayi Chen, Junyi Li, Ruiyang Ren, Shijie Wang, Wayne Xin Zhao, Yang Song, and Tao Zhang. 2024b. Rag-star: Enhancing deliberative reasoning with retrieval augmented verification and refinement. *arXiv preprint arXiv:2412.12881*.

Pengcheng Jiang, Lang Cao, Ruike Zhu, Minhao Jiang, Yunyi Zhang, Jimeng Sun, and Jiawei Han. 2025a. Ras: Retrieval-and-structuring for knowledge-intensive llm generation. *arXiv preprint arXiv:2502.10996*.

Pengcheng Jiang, Jiacheng Lin, Lang Cao, Runchu Tian, SeongKu Kang, Zifeng Wang, Jimeng Sun, and Jiawei Han. 2025b. Deepretrieval: Hacking real search engines and retrievers with large language models via reinforcement learning. *arXiv preprint arXiv:2503.00223*.

Pengcheng Jiang, Xueqiang Xu, Jiacheng Lin, Jinfeng Xiao, Zifeng Wang, Jimeng Sun, and Jiawei Han. 2025c. s3: You don't need that much data to train a search agent via rl. *arXiv preprint arXiv:2505.14146*.

Yi Jiang, Sendong Zhao, Jianbo Li, Haochun Wang, Lizhe Zhang, Yan Liu, and Bing Qin. 2025d. Cocoa: Collaborative chain-of-agents for parametric-retrieved knowledge synergy. *arXiv preprint arXiv:2508.01696*.

Yichen Jiang, Shikha Bordia, Zheng Zhong, Charles Dognin, Maneesh Singh, and Mohit Bansal. 2020. Hover: A dataset for many-hop fact extraction and claim verification. *arXiv preprint arXiv:2011.03088*.

Yucheng Jiang, Yijia Shao, Dekun Ma, Sina J Semnani, and Monica S Lam. 2024c. Into the unknown unknowns: Engaged human learning through participation in language model agent conversations. *arXiv preprint arXiv:2408.15232*.

Zhengbao Jiang, Frank F Xu, Luyu Gao, Zhiqing Sun, Qian Liu, Jane Dwivedi-Yu, Yiming Yang, Jamie Callan, and Graham Neubig. 2023. Active retrieval augmented generation. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 7969–7992.

Zhouyu Jiang, Mengshu Sun, Lei Liang, and Zhiqiang Zhang. 2024d. Retrieve, summarize, plan: Advancing multi-hop question answering with an iterative approach. *arXiv preprint arXiv:2407.13101*.

Zhouyu Jiang, Mengshu Sun, Lei Liang, and Zhiqiang Zhang. 2025e. Retrieve, summarize, plan: Advancing multi-hop question answering with an iterative approach. In *Companion Proceedings of the ACM on Web Conference 2025*, pages 1677–1686.

Bowen Jin, Jinsung Yoon, Priyanka Kargupta, Sercan O Arik, and Jiawei Han. 2025a. An empirical study on reinforcement learning for reasoning-search interleaved llm agents. *arXiv preprint arXiv:2505.15117*.

Bowen Jin, Hansi Zeng, Zhenrui Yue, Jinsung Yoon, Sercan Arik, Dong Wang, Hamed Zamani, and Jiawei Han. 2025b. Search-r1: Training llms to reason and leverage search engines with reinforcement learning. *arXiv preprint arXiv:2503.09516*.

Jiajie Jin, Xiaoxi Li, Guanting Dong, Yuyao Zhang, Yutao Zhu, Yang Zhao, Hongjin Qian, and Zhicheng Dou. 2025c. Decoupled planning and execution: A hierarchical reasoning framework for deep search. *arXiv preprint arXiv:2507.02652*.

AI Jina. 2025. node-deepresearch. `https://github.com/jina-ai/node-DeepResearch`, Accessed on 2025-7-1.

Ashutosh Joshi, Sheikh Muhammad Sarwar, Samarth Varshney, Sreyashi Nag, Shrivats Agrawal, and Juhi Naik. 2024. Reaper: Reasoning based retrieval planning for complex rag systems. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*, pages 4621–4628.

Mandar Joshi, Eunsol Choi, Daniel S Weld, and Luke Zettlemoyer. 2017. Triviaqa: A large scale distantly supervised challenge dataset for reading comprehension. *arXiv preprint arXiv:1705.03551*.

Omar Khattab, Keshav Santhanam, Xiang Lisa Li, David Hall, Percy Liang, Christopher Potts, and Matei Zaharia. 2022. Demonstrate-search-predict: Composing retrieval and language models for knowledge-intensive nlp. *arXiv preprint arXiv:2212.14024*.

Ivica Kostric and Krisztian Balog. 2024. A surprisingly simple yet effective multi-query rewriting method for conversational passage retrieval. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 2271–2275.

Satyapriya Krishna, Kalpesh Krishna, Anhad Mohananey, Steven Schwarcz, Adam Stambler, Shyam Upadhyay, and Manaal Faruqui. 2024. Fact, fetch, and reason: A unified evaluation of retrieval-augmented generation. *Proceedings of NAACL 2025*.

Tom Kwiatkowski, Jennimaria Palomaki, Olivia Redfield, Michael Collins, Ankur Parikh, Chris Alberti, Danielle Epstein, Illia Polosukhin, Jacob Devlin, Kenton Lee, and 1 others. 2019. Natural questions: a benchmark for question answering research. *Transactions of the Association for Computational Linguistics*, 7:453–466.

AI LangChain. 2025. open_deep_research. https://github.com/langchain-ai/open_deep_research, Accessed on 2025-7-1.

Myeonghwa Lee, Seonho An, and Min-Soo Kim. 2024. Planrag: A plan-then-retrieval augmented generation for generative large language models as decision makers. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 6537–6555.

Zhicheng Lee, Shulin Cao, Jinxin Liu, Jiajie Zhang, Weichuan Liu, Xiaoyin Che, Lei Hou, and Juanzi Li. 2025. Rearag: Knowledge-guided reasoning enhances factuality of large reasoning models with iterative retrieval augmented generation. *arXiv preprint arXiv:2503.21729*.

Cheng Li, Jiexiong Liu, Yixuan Chen, Qihang Zhou, and KunLun Meta. 2025a. Kunlunbaizerag: Reinforcement learning driven inference performance leap for large language models. *arXiv preprint arXiv:2506.19466*.

Jinzheng Li, Jingshu Zhang, Hongguang Li, and Yiqing Shen. 2024a. An agent framework for real-time financial information searching with large language models. *arXiv preprint arXiv:2502.15684*.

Kuan Li, Zhongwang Zhang, Huifeng Yin, Liwen Zhang, Litu Ou, Jialong Wu, Wenbiao Yin, Baixuan Li, Zhengwei Tao, Xinyu Wang, Weizhou Shen, Junkai Zhang, Dingchu Zhang, Xixi Wu, Yong Jiang, Ming Yan, Pengjun Xie, Fei Huang, and Jingren Zhou. 2025b. Websailor: Navigating super-human reasoning for web agent. *Preprint*, arXiv:2507.02592.

Xiaoxi Li, Guanting Dong, Jiajie Jin, Yuyao Zhang, Yujia Zhou, Yutao Zhu, Peitian Zhang, and Zhicheng Dou. 2025c. Search-o1: Agentic search-enhanced large reasoning models. *arXiv preprint arXiv:2501.05366*.

Xiaoxi Li, Jiajie Jin, Guanting Dong, Hongjin Qian, Yutao Zhu, Yongkang Wu, Ji-Rong Wen, and Zhicheng Dou. 2025d. Webthinker: Empowering large reasoning models with deep research capability. *arXiv preprint arXiv:2504.21776*.

Xingxuan Li, Weiwen Xu, Ruochen Zhao, Fangkai Jiao, Shafiq Joty, and Lidong Bing. 2024b. Can we further elicit reasoning in llms? critic-guided planning with retrieval-augmentation for solving challenging tasks. *arXiv preprint arXiv:2410.01428*.

Xingxuan Li, Ruochen Zhao, Yew Ken Chia, Bosheng Ding, Shafiq Joty, Soujanya Poria, and Lidong Bing. 2023. Chain-of-knowledge: Grounding large language models via dynamic knowledge adapting over heterogeneous sources. *arXiv preprint arXiv:2305.13269*.

Yuchen Li, Hengyi Cai, Rui Kong, Xinran Chen, Jiamin Chen, Jun Yang, Haojie Zhang, Jiayi Li, Jiayi Wu, Yiqun Chen, and 1 others. 2025e. Towards ai search paradigm. *arXiv preprint arXiv:2506.17188*.

Zhicong Li, Jiahao Wang, Zhishu Jiang, Hangyu Mao, Zhongxia Chen, Jiazhen Du, Yuanxing Zhang, Fuzheng Zhang, Di Zhang, and Yong Liu. 2024c. Dmqr-rag: Diverse multi-query rewriting for rag. *arXiv preprint arXiv:2411.13154*.

Zhuoqun Li, Haiyang Yu, Xuanang Chen, Hongyu Lin, Yaojie Lu, Fei Huang, Xianpei Han, Yongbin Li, and Le Sun. 2025f. Deepsolution: Boosting complex engineering solution design via tree-based exploration and bi-point thinking. *arXiv preprint arXiv:2502.20730*.

Zixuan Li, Wenxuan Liu, Long Bai, Chunmao Zhang, Wei Li, Fenghui Zhang, Quanxin Jin, Ruoyun He, Zhuo Chen, Zhilei Hu, and 1 others. 2025g. Knowcoder-v2: Deep knowledge analysis. *arXiv preprint arXiv:2506.06881*.

Jintao Liang, Gang Su, Huifeng Lin, You Wu, Rui Zhao, and Ziyue Li. 2025. Reasoning rag via system 1 or system 2: A survey on reasoning agentic retrieval-augmented generation for industry challenges. *arXiv preprint arXiv:2506.10408*.

Jianghao Lin, Xinyi Dai, Yunjia Xi, Weiwen Liu, Bo Chen, Hao Zhang, Yong Liu, Chuhan Wu, Xiangyang Li, Chenxu Zhu, and 1 others. 2025a. How can recommender systems benefit from large language models: A survey. *ACM Transactions on Information Systems*, 43(2):1–47.

Jianghao Lin, Weiwen Liu, Xinyi Dai, Weinan Zhang, Shuai Li, Ruiming Tang, Xiuqiang He, Jianye Hao, and Yong Yu. 2021. A graph-enhanced click model for web search. In *Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval*, pages 1259–1268.

Jianghao Lin, Rong Shan, Chenxu Zhu, Kounianhua Du, Bo Chen, Shigang Quan, Ruiming Tang, Yong Yu, and Weinan Zhang. 2024. Rella: Retrieval-enhanced large language models for lifelong sequential behavior comprehension in recommendation. In *Proceedings of the ACM Web Conference 2024*, pages 3497–3508.

Junyong Lin, Lu Dai, Ruiqian Han, Yijie Sui, Ruilin Wang, Xingliang Sun, Qinglin Wu, Min Feng, Hao Liu, and Hui Xiong. 2025b. Scirgen: Synthesize realistic and large-scale rag dataset for scientific research. *arXiv preprint arXiv:2506.11117*.

Jie Liu and Barzan Mozafari. 2024. Query rewriting via large language models. *arXiv preprint arXiv:2403.09060*.

Xinyu Liu, Wei Zhang, Zhiwei Li, and Lei Zhang. 2024a. Morehopqa: More than multi-hop reasoning. *arXiv preprint arXiv:2405.07437*.

Yungeng Liu, Zan Chen, Yu Guang Wang, and Yiqing Shen. 2024b. Toursynbio-search: A large language model driven agent framework for unified search method for protein engineering. In *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 5395–5400. IEEE.

Elias Lumer, Anmol Gulati, Vamse Kumar Subbiah, Pradeep Honaganahalli Basavaraju, and James A Burke. 2025. Scalemcp: Dynamic and auto-synchronizing model context protocol tools for llm agents. *arXiv preprint arXiv:2505.06416*.

Elias Lumer, Vamse Kumar Subbiah, James A Burke, Pradeep Honaganahalli Basavaraju, and Austin Huber. 2024. Toolshed: Scale tool-equipped agents with advanced rag-tool fusion and tool knowledge bases. *arXiv preprint arXiv:2410.14594*.

Haoran Luo, Yikai Guo, Qika Lin, Xiaobao Wu, Xinyu Mu, Wenhao Liu, Meina Song, Yifan Zhu, Luu Anh Tuan, and 1 others. 2025. Kbqa-o1: Agentic knowledge base question answering with monte carlo tree search. *arXiv preprint arXiv:2501.18922*.

Matteo Lupinacci, Francesco Blefari, Francesco Romeo, Francesco Aurelio Pironti, and Angelo Furfaro. 2025. Arcer: an agentic rag for the automated definition of cyber ranges. *arXiv preprint arXiv:2504.12143*.

Yougang Lyu, Xiaoyu Zhang, Lingyong Yan, Maarten de Rijke, Zhaochun Ren, and Xiuying Chen. 2025. Deepshop: A benchmark for deep research shopping agents. *arXiv preprint arXiv:2506.02839*.

Huanhuan Ma, Weizhi Xu, Yifan Wei, Liuji Chen, Liang Wang, Qiang Liu, and Shu Wu. 2023a. Ex-fever: A dataset for multi-hop explainable fact verification. *arXiv preprint arXiv:2310.09754*.

Tianyi Ma, Yiyue Qian, Zheyuan Zhang, Zehong Wang, Xiaoye Qian, Feifan Bai, Yifan Ding, Xuwei Luo, Shinan Zhang, Keerthiram Murugesan, and 1 others. 2025. Autodata: A multi-agent system for open web data collection. *arXiv preprint arXiv:2505.15859*.

Xinbei Ma, Yeyun Gong, Pengcheng He, Hai Zhao, and Nan Duan. 2023b. Query rewriting in retrieval-augmented large language models. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 5303–5315.

Alex Mallen, Akari Asai, Victor Zhong, Rajarshi Das, Daniel Khashabi, and Hannaneh Hajishirzi. 2022. When not to trust language models: Investigating effectiveness of parametric and non-parametric memories. *arXiv preprint arXiv:2212.10511*.

Reza Yousefi Maragheh, Pratheek Vadla, Priyank Gupta, Kai Zhao, Aysenur Inan, Kehui Yao, Jianpeng Xu, Praveen Kanumala, Jason Cho, and Sushant Kumar. 2025. Arag: Agentic retrieval augmented generation for personalized recommendation. *Preprint*, arXiv:2506.21931.

Jianbiao Mei, Tao Hu, Daocheng Fu, Licheng Wen, Xuemeng Yang, Rong Wu, Pinlong Cai, Xing Gao, Yu Yang, Chengjun Xie, and 1 others. 2025. O2-searcher: A searching-based agent model for open-domain open-ended question answering. *arXiv preprint arXiv:2505.16582*.

Grégoire Mialon, Clémentine Fourrier, Thomas Wolf, Yann LeCun, and Thomas Scialom. 2023. Gaia: a benchmark for general ai assistants. In *The Twelfth International Conference on Learning Representations*.

Microsoft. 2025. Microsoft 365 copilot blog. `https://techcommunity.microsoft.com/blog/microsoft365copilotblog/researcher-agent-in-microsoft-365-copilot/4397186`, Accessed on 2025-6-30.

Mihran Miroyan, Tsung-Han Wu, Logan King, Tianle Li, Jiayi Pan, Xinyan Hu, Wei-Lin Chiang, Anastasios N Angelopoulos, Trevor Darrell, Narges Norouzi, and 1 others. 2025. Search arena: Analyzing search-augmented llms. *arXiv preprint arXiv:2506.05334*.

AI Moonshot. 2025. Kimi-researcher. `https://moonshotai.github.io/Kimi-Researcher/`, Accessed on 2025-6-30.

Feiteng Mu, Yong Jiang, Liwen Zhang, Chu Liu, Wenjie Li, Pengjun Xie, and Fei Huang. 2024. Query routing for homogeneous tools: An instantiation in the rag scenario. In *Findings of the Association for Computational Linguistics: EMNLP 2024*.

Thang Nguyen, Peter Chin, and Yu-Wing Tai. 2025. Ma-rag: Multi-agent retrieval-augmented generation via collaborative chain-of-thought reasoning. *arXiv preprint arXiv:2505.20096*.

Tri Nguyen, Mir Rosenberg, Xia Song, Jianfeng Gao, Saurabh Tiwary, Rangan Majumder, and Li Deng. 2016. Ms marco: A human-generated machine reading comprehension dataset.

nickscamara. 2025. Open deep research. https://github.com/nickscamara/open-deep-research, Accessed on 2025-7-1.

Felix Ocker, Jörg Deigmöller, Pavel Smirnov, and Julian Eggert. 2025. A grounded memory system for smart personal assistants. *arXiv preprint arXiv:2505.06328*.

OpenAI. 2025. Introducing deep research. https://openai.com/index/introducing-deep-research/, Accessed on 2025-6-26.

Wojciech Ostrowski, Arnav Arora, Pepa Atanasova, and Isabelle Augenstein. 2020. Multi-hop fact checking of political claims. *arXiv preprint arXiv:2009.06401*.

Ought. 2025. Elicit: The ai research assistant. https://elicit.com/, Accessed on 2025-6-30.

Ankit Pal, Logesh Kumar Umapathi, and Malaikannan Sankarasubbu. 2022. Medmcqa: A large-scale multi-subject multi-choice dataset for medical domain question answering. In *Conference on health, inference, and learning*, pages 248–260. PMLR.

Haojie Pan, Zepeng Zhai, Hao Yuan, Yaojia Lv, Ruiji Fu, Ming Liu, Zhongyuan Wang, and Bing Qin. 2023. Kwaiagents: Generalized information-seeking agent system with large language models. *arXiv preprint arXiv:2312.04889*.

Perplexity. 2025. Introducing perplexity deep research. https://www.perplexity.ai/hub/blog/introducing-perplexity-deep-research, Accessed on 2025-6-26.

Fabio Petroni, Aleksandra Piktus, Angela Fan, Patrick Lewis, Majid Yazdani, Nicola De Cao, James Thorne, Yacine Jernite, Vladimir Karpukhin, Jean Maillard, and 1 others. 2020. Kilt: a benchmark for knowledge intensive language tasks. *arXiv preprint arXiv:2009.02252*.

Thinh Pham, Nguyen Nguyen, Pratibha Zunjare, Weiyuan Chen, Yu-Min Tseng, and Tu Vu. 2025. Sealqa: Raising the bar for reasoning in search-augmented language models. *arXiv preprint arXiv:2506.01062*.

Long Phan, Alice Gatti, Ziwen Han, Nathaniel Li, Josephina Hu, Hugh Zhang, Chen Bo Calvin Zhang, Mohamed Shaaban, John Ling, Sean Shi, and 1 others. 2025. Humanity's last exam. *arXiv preprint arXiv:2501.14249*.

Ofir Press, Muru Zhang, Sewon Min, Ludwig Schmidt, Noah Smith, and Mike Lewis. 2023. Measuring and narrowing the compositionality gap in language models. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 5687–5711, Singapore. Association for Computational Linguistics.

Ofir Press, Muru Zhang, Sewon Min, Ludwig Schmidt, Noah A Smith, and Mike Lewis. 2022. Measuring and narrowing the compositionality gap in language models. *arXiv preprint arXiv:2210.03350*.

Zehan Qi, Rongwu Xu, Zhijiang Guo, Cunxiang Wang, Hao Zhang, and Wei Xu. 2024. Long2rag: Evaluating long-context & long-form retrieval-augmented generation with key point recall. *arXiv preprint arXiv:2410.23000*.

Hongjin Qian and Zheng Liu. 2025. Scent of knowledge: Optimizing search-enhanced reasoning with information foraging. *arXiv preprint arXiv:2505.09316*.

Chidaksh Ravuru, Sagar Srinivas Sakhinana, and Venkataramana Runkana. 2024. Agentic retrieval-augmented generation for time series analysis. *arXiv preprint arXiv:2408.14484*.

David Rein, Betty Li Hou, Asa Cooper Stickland, Jackson Petty, Richard Yuanzhe Pang, Julien Dirani, Julian Michael, and Samuel R Bowman. 2024. Gpqa: A graduate-level google-proof q&a benchmark. In *First Conference on Language Modeling*.

Ruiyang Ren, Yuhao Wang, Junyi Li, Jinhao Jiang, Wayne Xin Zhao, Wenjie Wang, and Tat-Seng Chua. 2025. Holistically guided monte carlo tree search for intricate information seeking. *arXiv preprint arXiv:2502.04751*.

Corby Rosset, Ho-Lam Chung, Guanghui Qin, Ethan C Chau, Zhuo Feng, Ahmed Awadallah, Jennifer Neville, and Nikhil Rao. 2024. Researchy questions: A dataset of multi-perspective, decompositional questions for llm web agents. *arXiv preprint arXiv:2402.17896*.

Dongyu Ru, Lin Qiu, Xiangkun Hu, Tianhang Zhang, Peng Shi, Shuaichen Chang, Cheng Jiayang, Cunxiang Wang, Shichao Sun, Huanyu Li, and 1 others. 2024. Ragchecker: A fine-grained framework for diagnosing retrieval-augmented generation. *Advances in Neural Information Processing Systems*, 37:21999–22027.

Florian Schneider, Narges Baba Ahmadi, Niloufar Baba Ahmadi, Iris Vogel, Martin Semmann, and Chris Biemann. 2025. Collex–a multimodal agentic rag system enabling interactive exploration of scientific collections. *arXiv preprint arXiv:2504.07643*.

Julian Schnitzler, Xanh Ho, Jiahao Huang, Florian Boudin, Saku Sugawara, and Akiko Aizawa. 2024. Morehopqa: More than multi-hop reasoning. *arXiv preprint arXiv:2406.13397*.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.

Wonduk Seo and Seunghyun Lee. 2025. Qa-expand: Multi-question answer generation for enhanced query expansion in information retrieval. *arXiv preprint arXiv:2502.08557*.

Yijia Shao, Yucheng Jiang, Theodore A Kanell, Peter Xu, Omar Khattab, and Monica S Lam. 2024a. Assisting in writing wikipedia-like articles from scratch with large language models. *arXiv preprint arXiv:2402.14207*.

Zhihong Shao, Yeyun Gong, Yelong Shen, Minlie Huang, Nan Duan, and Weizhu Chen. 2023. Enhancing retrieval-augmented large language models with iterative retrieval-generation synergy. *arXiv preprint arXiv:2305.15294*.

Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, YK Li, Y Wu, and 1 others. 2024b. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. *arXiv preprint arXiv:2402.03300*.

Quan Shi, Michael Tang, Karthik Narasimhan, and Shunyu Yao. 2024a. Can language models solve olympiad programming? *arXiv preprint arXiv:2404.10952*.

Wenxuan Shi, Haochen Tan, Chuqiao Kuang, Xiaoguang Li, Xiaozhe Ren, Chen Zhang, Hanting Chen, Yasheng Wang, Lifeng Shang, Fisher Yu, and 1 others. 2025a. Pangu deepdiver: Adaptive search intensity scaling via open-web reinforcement learning. *arXiv preprint arXiv:2505.24332*.

Yaorui Shi, Shihan Li, Chang Wu, Zhiyuan Liu, Junfeng Fang, Hengxing Cai, An Zhang, and Xiang Wang. 2025b. Search and refine during think: Autonomous retrieval-augmented reasoning of llms. *arXiv preprint arXiv:2505.11277*.

Zhengliang Shi, Weiwei Sun, Shen Gao, Pengjie Ren, Zhumin Chen, and Zhaochun Ren. 2024b. Generate-then-ground in retrieval-augmented generation for multi-hop question answering. *arXiv preprint arXiv:2406.14891*.

Zhengliang Shi, Lingyong Yan, Dawei Yin, Suzan Verberne, Maarten de Rijke, and Zhaochun Ren. 2025c. Iterative self-incentivization empowers large language models as agentic searchers. *arXiv preprint arXiv:2505.20128*.

Ramneet Singh, Sathvik Joel, Abhav Mehrotra, Nalin Wadhwa, Ramakrishna B Bairi, Aditya Kanade, and Nagarajan Natarajan. 2025. Code researcher: Deep research agent for large systems code and commit history. *arXiv preprint arXiv:2506.11060*.

Huatong Song, Jinhao Jiang, Yingqian Min, Jie Chen, Zhipeng Chen, Wayne Xin Zhao, Lei Fang, and Ji-Rong Wen. 2025a. R1-searcher: Incentivizing the search capability in llms via reinforcement learning. *arXiv preprint arXiv:2503.05592*.

Huatong Song, Jinhao Jiang, Wenqing Tian, Zhipeng Chen, Yuhuan Wu, Jiahao Zhao, Yingqian Min, Wayne Xin Zhao, Lei Fang, and Ji-Rong Wen. 2025b. R1-searcher++: Incentivizing the dynamic knowledge acquisition of llms via reinforcement learning. *arXiv preprint arXiv:2505.17005*.

Gerion Spielberger, Florian M Artinger, Jochen Reb, and Rudolf Kerschreiter. 2025. Retrieval augmented generation for topic modeling in organizational research: An introduction with empirical demonstration. *arXiv preprint arXiv:2502.20963*.

Sakhinana Sagar Srinivas, Akash Das, Shivam Gupta, and Venkataramana Runkana. 2024. Accelerating manufacturing scale-up from material discovery using agentic web navigation and retrieval-augmented ai for process engineering schematics design. *arXiv preprint arXiv:2412.05937*.

Hongjin Su, Howard Yen, Mengzhou Xia, Weijia Shi, Niklas Muennighoff, Han-yu Wang, Haisu Liu, Quan Shi, Zachary S Siegel, Michael Tang, and 1 others. 2024. Bright: A realistic and challenging benchmark for reasoning-intensive retrieval. *arXiv preprint arXiv:2407.12883*.

Hao Sun, Zile Qiao, Jiayan Guo, Xuanbo Fan, Yingyan Hou, Yong Jiang, Pengjun Xie, Yan Zhang, Fei Huang, and Jingren Zhou. 2025a. Zerosearch: Incentivize the search capability of llms without searching. *arXiv preprint arXiv:2505.04588*.

Shuang Sun, Huatong Song, Yuhao Wang, Ruiyang Ren, Jinhao Jiang, Junjie Zhang, Fei Bai, Jia Deng, Wayne Xin Zhao, Zheng Liu, and 1 others. 2025b. Simpledeepsearcher: Deep information seeking via web-powered reasoning trajectory synthesis. *arXiv preprint arXiv:2505.16834*.

Zhongxiang Sun, Qipeng Wang, Weijie Yu, Xiaoxue Zang, Kai Zheng, Jun Xu, Xiao Zhang, Song Yang, and Han Li. 2025c. Rearter: Retrieval-augmented reasoning with trustworthy process rewarding. *arXiv preprint arXiv:2501.07861*.

Haochen Tan, Zhijiang Guo, Zhan Shi, Lu Xu, Zhili Liu, Yunlong Feng, Xiaoguang Li, Yasheng Wang, Lifeng Shang, Qun Liu, and 1 others. 2024a. Proxyqa: An alternative framework for evaluating long-form text generation with large language models. *arXiv preprint arXiv:2401.15042*.

John Chong Min Tan, Prince Saroj, Bharat Runwal, Hardik Maheshwari, Brian Lim Yi Sheng, Richard Cottrill, Alankrit Chona, Ambuj Kumar, and Mehul Motani. 2024b. Taskgen: A task-based, memory-infused agentic framework using strictjson. *arXiv preprint arXiv:2407.15734*.

Zhiwen Tan, Jiaming Huang, Qintong Wu, Hongxuan Zhang, Chenyi Zhuang, and Jinjie Gu. 2025. Rag-r1 : Incentivize the search and reasoning capabilities of llms through multi-query parallelism. *Preprint*, arXiv:2507.02962.

Xiangru Tang, Tianrui Qin, Tianhao Peng, Ziyang Zhou, Daniel Shao, Tingting Du, Xinming Wei, Peng Xia, Fang Wu, He Zhu, Ge Zhang, Jiaheng Liu, Xingyao Wang, Sirui Hong, Chenglin Wu, Hao Cheng, Chi Wang, and Wangchunshu Zhou. 2025. Agent kb: Leveraging cross-domain experience for agentic problem solving. *Preprint*, arXiv:2507.06229.

Yixuan Tang and Yi Yang. 2024a. Benchmarking retrieval-augmented generation for multi-hop queries. *CoRR*, abs/2401.15391.

Yixuan Tang and Yi Yang. 2024b. Multihop-rag: Benchmarking retrieval-augmented generation for multi-hop queries. *arXiv preprint arXiv:2401.15391*.

Fengwei Teng, Zhaoyang Yu, Quan Shi, Jiayi Zhang, Chenglin Wu, and Yuyu Luo. 2025. Atom of thoughts for markov llm test-time scaling. *arXiv preprint arXiv:2502.12018*.

Karishma Thakrar, Shreyas Basavatia, and Akshay Daftardar. 2025. Cultivating multimodal intelligence: Interpretive reasoning and agentic rag approaches to dermatological diagnosis. *Preprint*, arXiv:2507.05520.

Edan Toledo, Karen Hambardzumyan, Martin Josifoski, Rishi Hazra, Nicolas Baldwin, Alexis Audran-Reiss, Michael Kuchnik, Despoina Magka, Minqi Jiang, Alisia Maria Lupidi, Andrei Lupu, Roberta Raileanu, Kelvin Niu, Tatiana Shavrina, Jean-Christophe Gagnon-Audet, Michael Shvartsman, Shagun Sodhani, Alexander H. Miller, Abhishek Charnalia, and 6 others. 2025. Ai research agents for machine learning: Search, exploration, and generalization in mle-bench. *Preprint*, arXiv:2507.02554.

Hieu Tran, Zonghai Yao, Junda Wang, Yifan Zhang, Zhichao Yang, and Hong Yu. 2024. Rare: Retrieval-augmented reasoning enhancement for large language models. *arXiv preprint arXiv:2412.02830*.

Tam Trinh, Manh Nguyen, and Truong-Son Hy. 2025a. Towards robust fact-checking: A multi-agent system with advanced evidence retrieval. *arXiv preprint arXiv:2506.17878*.

Tam Trinh, Manh Nguyen, and Truong-Son Hy. 2025b. Towards robust fact-checking: A multi-agent system with advanced evidence retrieval. *Preprint*, arXiv:2506.17878.

Harsh Trivedi, Niranjan Balasubramanian, Tushar Khot, and Ashish Sabharwal. 2022a. Interleaving retrieval with chain-of-thought reasoning for knowledge-intensive multi-step questions. *arXiv preprint arXiv:2212.10509*.

Harsh Trivedi, Niranjan Balasubramanian, Tushar Khot, and Ashish Sabharwal. 2022b. Musique: Multi-hop questions via single-hop question composition. *Transactions of the Association for Computational Linguistics*, 10:539–554.

Hasan Md Tusfiqur Alam, Devansh Srivastav, Md Abdul Kadir, and Daniel Sonntag. 2024. Towards interpretable radiology report generation via concept bottlenecks using a multi-agentic rag. *arXiv e-prints*, pages arXiv–2412.

Saeid Ario Vaghefi, Aymane Hachcham, Veronica Grasso, Jiska Manicus, Nakiete Msemo, Chiara Colesanti Senni, and Markus Leippold. 2025. Ai for climate finance: Agentic retrieval and multi-step reasoning for early warning system investments. *arXiv preprint arXiv:2504.05104*.

Prakhar Verma, Sukruta Prakash Midigeshi, Gaurav Sinha, Arno Solin, Nagarajan Natarajan, and Amit Sharma. 2025. Plan*rag: Efficient test-time planning for retrieval augmented generation. In *Workshop on Reasoning and Planning for Large Language Models*.

David Wadden, Shanchuan Lin, Kyle Lo, Lucy Lu Wang, Madeleine van Zuylen, Arman Cohan, and Hannaneh Hajishirzi. 2020. Fact or fiction: Verifying scientific claims. *arXiv preprint arXiv:2004.14974*.

Junjie Wang, Mingyang Chen, Binbin Hu, Dan Yang, Ziqi Liu, Yue Shen, Peng Wei, Zhiqiang Zhang, Jinjie Gu, Jun Zhou, and 1 others. 2024a. Learning to plan for retrieval-augmented large language models from knowledge graphs. *arXiv preprint arXiv:2406.14282*.

Keheng Wang, Feiyu Duan, Sirui Wang, Peiguang Li, Yunsen Xian, Chuantao Yin, Wenge Rong, and Zhang Xiong. 2023. Knowledge-driven cot: Exploring faithful reasoning in llms for knowledge-intensive question answering. *arXiv preprint arXiv:2308.13259*.

Lei Wang, Chen Ma, Xueyang Feng, Zeyu Zhang, Hao Yang, Jingsen Zhang, Zhiyuan Chen, Jiakai Tang, Xu Chen, Yankai Lin, and 1 others. 2024b. A survey on large language model based autonomous agents. *Frontiers of Computer Science*, 18(6):186345.

Liang Wang, Haonan Chen, Nan Yang, Xiaolong Huang, Zhicheng Dou, and Furu Wei. 2025a. Chain-of-retrieval augmented generation. *arXiv preprint arXiv:2501.14342*.

Qiuchen Wang, Ruixue Ding, Zehui Chen, Weiqi Wu, Shihang Wang, Pengjun Xie, and Feng Zhao. 2025b. Vidorag: Visual document retrieval-augmented generation via dynamic iterative reasoning agents. *arXiv preprint arXiv:2502.18017*.

Qiuchen Wang, Ruixue Ding, Yu Zeng, Zehui Chen, Lin Chen, Shihang Wang, Pengjun Xie, Fei Huang, and Feng Zhao. 2025c. Vrag-rl: Empower vision-perception-based rag for visually rich information understanding via iterative reasoning with reinforcement learning. *arXiv preprint arXiv:2505.22019*.

Ruobing Wang, Daren Zha, Shi Yu, Qingfei Zhao, Yuxuan Chen, Yixuan Wang, Shuo Wang, Yukun Yan, Zhenghao Liu, Xu Han, and 1 others. 2024c. Retriever-and-memory: Towards adaptive note-enhanced retrieval-augmented generation. *arXiv preprint arXiv:2410.08821*.

Shengkang Wang, Hongzhan Lin, Ziyang Luo, Zhen Ye, Guang Chen, and Jing Ma. 2024d. Mfc-bench: Benchmarking multimodal fact-checking with large vision-language models. *arXiv preprint arXiv:2406.11288*.

Yaqi Wang and Haipei Xu. 2024. Srsa: A cost-efficient strategy-router search agent for real-world human-machine interactions. *arXiv preprint arXiv:2411.14574*.

Zhefan Wang, Yuanqing Yu, Wendi Zheng, Weizhi Ma, and Min Zhang. 2024e. Macrec: A multi-agent collaboration framework for recommendation. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 2760–2764.

Zihao Wang, Anji Liu, Haowei Lin, Jiaqi Li, Xiaojian Ma, and Yitao Liang. 2024f. Rat: Retrieval augmented thoughts elicit context-aware reasoning in long-horizon generation. *arXiv preprint arXiv:2403.05313*.

Ziliang Wang, Xuhui Zheng, Kang An, Cijun Ouyang, Jialu Cai, Yuhang Wang, and Yichao Wu. 2025d. Stepsearch: Igniting llms search ability via step-wise proximal policy optimization. *arXiv preprint arXiv:2505.15107*.

Ziqi Wang and Boqin Yuan. 2025. L-mars: Legal multi-agent workflow with orchestrated reasoning and agentic search. *arXiv preprint arXiv:2509.00761*.

Ziyue Wang, Junde Wu, Linghan Cai, Chang Han Low, Xihong Yang, Qiaxuan Li, and Yueming Jin. 2025e. Medagent-pro: Towards evidence-based multi-modal medical diagnosis via reasoning agentic workflow. *arXiv preprint arXiv:2503.18968*.

Jason Wei, Zhiqing Sun, Spencer Papay, Scott McKinney, Jeffrey Han, Isa Fulford, Hyung Won Chung, Alex Tachard Passos, William Fedus, and Amelia Glaese. 2025. Browsecomp: A simple yet challenging benchmark for browsing agents. *arXiv preprint arXiv:2504.12516*.

Jerry Wei, Chengrun Yang, Xinying Song, Yifeng Lu, Nathan Hu, Jie Huang, Dustin Tran, Daiyi Peng, Ruibo Liu, Da Huang, and 1 others. 2024. Long-form factuality in large language models. *arXiv preprint arXiv:2403.18802*.

Jialong Wu, Baixuan Li, Runnan Fang, Wenbiao Yin, Liwen Zhang, Zhengwei Tao, Dingchu Zhang, Zekun Xi, Yong Jiang, Pengjun Xie, and 1 others. 2025a. Webdancer: Towards autonomous information seeking agency. *arXiv preprint arXiv:2505.22648*.

Jialong Wu, Wenbiao Yin, Yong Jiang, Zhenglin Wang, Zekun Xi, Runnan Fang, Linhai Zhang, Yulan He, Deyu Zhou, Pengjun Xie, and 1 others. 2025b. Webwalker: Benchmarking llms in web traversal. *arXiv preprint arXiv:2501.07572*.

Jian Wu, Linyi Yang, Zhen Wang, Manabu Okumura, and Yue Zhang. 2024. Cofca: A step-wise counterfactual multi-hop qa benchmark. *arXiv preprint arXiv:2402.11924*.

Jian Wu, Linyi Yang, Zhen Wang, Manabu Okumura, and Yue Zhang. 2025c. Cofca: A step-wise counterfactual multi-hop qa benchmark. *arXiv preprint arXiv:2402.11924*.

Junde Wu, Jiayuan Zhu, and Yuyuan Liu. 2025d. Agentic reasoning: Reasoning llms with tools for the deep research. *arXiv preprint arXiv:2502.04644*.

Peilin Wu, Mian Zhang, Xinlu Zhang, Xinya Du, and Zhiyu Zoey Chen. 2025e. Search wisely: Mitigating sub-optimal agentic searches by reducing uncertainty. *arXiv preprint arXiv:2505.17281*.

Weiqi Wu, Xin Guan, Shen Huang, Yong Jiang, Pengjun Xie, Fei Huang, Jiuxin Cao, Hai Zhao, and Jingren Zhou. 2025f. Masksearch: A universal pre-training framework to enhance agentic search capability. *arXiv preprint arXiv:2505.20285*.

Wenjie Wu, Yongcheng Jing, Yingjie Wang, Wenbin Hu, and Dacheng Tao. 2025g. Graph-augmented reasoning: Evolving step-by-step knowledge graph retrieval for llm reasoning. *arXiv preprint arXiv:2503.01642*.

Yunjia Xi, Jianghao Lin, Menghui Zhu, Yongzhao Xiao, Zhuoying Ou, Jiaqi Liu, Tong Wan, Bo Chen, Weiwen Liu, Yasheng Wang, and 1 others. 2025a. Infodeepseek: Benchmarking agentic information seeking for retrieval-augmented generation. *arXiv preprint arXiv:2505.15872*.

Yunjia Xi, Weiwen Liu, Jianghao Lin, Xiaoling Cai, Hong Zhu, Jieming Zhu, Bo Chen, Ruiming Tang, Weinan Zhang, and Yong Yu. 2024a. Towards open-world recommendation with knowledge augmentation from large language models. In *Proceedings of the 18th ACM Conference on Recommender Systems*, pages 12–22.

Yunjia Xi, Weiwen Liu, Jianghao Lin, Bo Chen, Ruiming Tang, Weinan Zhang, and Yong Yu. 2024b. Memocrs: Memory-enhanced sequential conversational recommender systems with large language models. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*, pages 2585–2595.

Yunjia Xi, Hangyu Wang, Bo Chen, Jianghao Lin, Menghui Zhu, Weiwen Liu, Ruiming Tang, Zhewei Wei, Weinan Zhang, and Yong Yu. 2025b. Efficiency unleashed: Inference acceleration for llm-based recommender systems with speculative decoding. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 1891–1901.

Yunjia Xi, Muyan Weng, Wen Chen, Chao Yi, Dian Chen, Gaoyang Guo, Mao Zhang, Jian Wu, Yuning Jiang, Qingwen Liu, and 1 others. 2025c. Bursting filter bubble: Enhancing serendipity recommendations with aligned large language models. *arXiv preprint arXiv:2502.13539*.

Liang Xiao, Wen Dai, Shuai Chen, Bin Qin, Chongyang Shi, Haopeng Jing, and Tianyu Guo. 2025. Retrieval-augmented generation by evidence retroactivity in llms. *arXiv preprint arXiv:2501.05475*.

Guangzhi Xiong, Qiao Jin, Xiao Wang, Yin Fang, Haolin Liu, Yifan Yang, Fangyuan Chen, Zhixing Song, Dengyu Wang, Minjia Zhang, and 1 others. 2025a. Rag-gym: Optimizing reasoning and search agents with process supervision. *arXiv preprint arXiv:2502.13957*.

Guanming Xiong, Haochen Li, and Wen Zhao. 2025b. Mcts-kbqa: Monte carlo tree search for knowledge base question answering. *arXiv preprint arXiv:2502.13428*.

Mingjun Xu, Jinhan Dong, Jue Hou, Zehui Wang, Sihang Li, Zhifeng Gao, Renxin Zhong, and Hengxing Cai. 2025a. Mm-r5: Multimodal reasoning-enhanced reranker via reinforcement learning for document retrieval. *arXiv preprint arXiv:2506.12364*.

Qiancheng Xu, Yongqi Li, Heming Xia, and Wenjie Li. 2024a. Enhancing tool retrieval with iterative feedback from large language models. *arXiv preprint arXiv:2406.17465*.

Renjun Xu and Jingwen Peng. 2025. A comprehensive survey of deep research: Systems, methodologies, and applications. *arXiv preprint arXiv:2506.12594*.

Shicheng Xu, Liang Pang, Huawei Shen, Xueqi Cheng, and Tat-Seng Chua. 2024b. Search-in-the-chain: Interactively enhancing large language models with search for knowledge-intensive tasks. In *Proceedings of the ACM Web Conference 2024*, pages 1362–1373.

Wujiang Xu, Kai Mei, Hang Gao, Juntao Tan, Zujie Liang, and Yongfeng Zhang. 2025b. A-mem: Agentic memory for llm agents. *arXiv preprint arXiv:2502.12110*.

Xiwei Xu, Dawen Zhang, Qing Liu, Qinghua Lu, and Liming Zhu. 2025c. Agentic rag with human-in-the-retrieval. In *2025 IEEE 22nd International Conference on Software Architecture Companion (ICSA-C)*, pages 498–502. IEEE.

Shuo Yang, Yuqin Dai, Guoqing Wang, Xinran Zheng, Jinfeng Xu, Jinze Li, Zhenzhe Ying, Weiqiang Wang, and Edith CH Ngai. 2025a. Realfactbench: A benchmark for evaluating large language models in real-world fact-checking. *arXiv preprint arXiv:2506.12538*.

Yingxuan Yang, Huacan Chai, Yuanyi Song, Siyuan Qi, Muning Wen, Ning Li, Junwei Liao, Haoyi Hu, Jianghao Lin, Gaowei Chang, and 1 others. 2025b. A survey of ai agent protocols. *arXiv preprint arXiv:2504.16736*.

Zhaorui Yang, Bo Pan, Han Wang, Yiyao Wang, Xingyu Liu, Minfeng Zhu, Bo Zhang, and Wei Chen. 2025c. Multimodal deepresearcher: Generating text-chart interleaved reports from scratch with agentic framework. *arXiv preprint arXiv:2506.02454*.

Zhilin Yang, Peng Qi, Saizheng Zhang, Yoshua Bengio, William W Cohen, Ruslan Salakhutdinov, and Christopher D Manning. 2018. Hotpotqa: A dataset for diverse, explainable multi-hop question answering. *arXiv preprint arXiv:1809.09600*.

Barry Menglong Yao, Aditya Shah, Lichao Sun, Jin-Hee Cho, and Lifu Huang. 2023. End-to-end multimodal fact-checking and explanation generation: A challenging dataset and models. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 2733–2743.

Tian Yu, Shaolei Zhang, and Yang Feng. 2024. Auto-rag: Autonomous retrieval-augmented generation for large language models. *arXiv preprint arXiv:2411.19443*.

Huaying Yuan, Zheng Liu, Junjie Zhou, Ji-Rong Wen, and Zhicheng Dou. 2025. Videodeepresearch: Long video understanding with agentic tool using. *arXiv preprint arXiv:2506.10821*.

Zhenrui Yue, Honglei Zhuang, Aijun Bai, Kai Hui, Rolf Jagerman, Hansi Zeng, Zhen Qin, Dong Wang, Xuanhui Wang, and Michael Bendersky. 2024. Inference scaling for long-context retrieval augmented generation. *arXiv preprint arXiv:2410.04343*.

Dingchu Zhang, Yida Zhao, Jialong Wu, Baixuan Li, Wenbiao Yin, Liwen Zhang, Yong Jiang, Yufeng Li, Kewei Tu, Pengjun Xie, and 1 others. 2025a. Evolvesearch: An iterative self-evolving search agent. *arXiv preprint arXiv:2505.22501*.

Kechi Zhang, Jia Li, Ge Li, Xianjie Shi, and Zhi Jin. 2024a. Codeagent: Enhancing code generation with tool-integrated agent systems for real-world repo-level coding challenges. *arXiv preprint arXiv:2401.07339*.

Ningning Zhang, Chi Zhang, Zhizhong Tan, Xingxing Yang, Weiping Deng, and Wenyong Wang. 2025b. Credible plan-driven rag method for multi-hop question answering. *arXiv preprint arXiv:2504.16787*.

Qiyuan Zhang, Fuyuan Lyu, Zexu Sun, Lei Wang, Weixu Zhang, Wenyue Hua, Haolun Wu, Zhihan Guo, Yufei Wang, Niklas Muennighoff, and 1 others. 2025c. A survey on test-time scaling in large language models: What, how, where, and how well? *arXiv preprint arXiv:2503.24235*.

Wei Zhang, Xinyu Liu, Jiajun Li, Zhiwei Wang, Zhiwei Li, and Lei Zhang. 2025d. Vidorag: Visual document retrieval-augmented generation via dynamic iterative reasoning agents. *arXiv preprint arXiv:2502.18017*.

Weinan Zhang, Junwei Liao, Ning Li, Kounianhua Du, and Jianghao Lin. 2024b. Agentic information retrieval. *arXiv preprint arXiv:2410.09713*.

Weizhi Zhang, Yangning Li, Yuanchen Bei, Junyu Luo, Guancheng Wan, Liangwei Yang, Chenxuan Xie, Yuyao Yang, Wei-Chieh Huang, Chunyu Miao, and 1 others. 2025e. From web search towards agentic deep research: Incentivizing search with reasoning agents. *arXiv preprint arXiv:2506.18959*.

Wenlin Zhang, Xiangyang Li, Kuicai Dong, Yichao Wang, Pengyue Jia, Xiaopeng Li, Yingyi Zhang, Derong Xu, Zhaocheng Du, Huifeng Guo, and 1 others. 2025f. Process vs. outcome reward: Which is better for agentic rag reinforcement learning. *arXiv preprint arXiv:2505.14069*.

Xiaoming Zhang, Ming Wang, Xiaocui Yang, Daling Wang, Shi Feng, and Yifei Zhang. 2024c. Hierarchical retrieval-augmented generation model with rethink for multi-hop question answering. *arXiv preprint arXiv:2408.11875*.

Yuxiang Zhang, Yuqi Yang, Jiangming Shu, Xinyan Wen, and Jitao Sang. 2025g. Agent models: Internalizing chain-of-action generation into reasoning models. *arXiv preprint arXiv:2503.06580*.

Qingfei Zhao, Ruobing Wang, Dingling Xu, Daren Zha, and Limin Liu. 2025. R-search: Empowering llm reasoning with search via multi-reward reinforcement learning. *arXiv preprint arXiv:2506.04185*.

Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, Zican Dong, and 1 others. 2023. A survey of large language models. *arXiv preprint arXiv:2303.18223*, 1(2).

Yuyue Zhao, Jiancan Wu, Xiang Wang, Wei Tang, Dingxian Wang, and Maarten De Rijke. 2024. Let me do it for you: Towards llm empowered recommendation via tool learning. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 1796–1806.

Yuxiang Zheng, Dayuan Fu, Xiangkun Hu, Xiaojie Cai, Lyumanshan Ye, Pengrui Lu, and Pengfei Liu. 2025. Deepresearcher: Scaling deep research via reinforcement learning in real-world environments. *arXiv preprint arXiv:2504.03160*.

Chenyu Zhou, Huacan Chai, Wenteng Chen, Zihan Guo, Rong Shan, Yuanyi Song, Tianyi Xu, Yingxuan Yang, Aofan Yu, Weiming Zhang, and 1 others. 2026. Externalization in llm agents: A unified review of memory, skills, protocols and harness engineering. *arXiv preprint arXiv:2604.08224*.

Junting Zhou, Wang Li, Yiyan Liao, Nengyuan Zhang, Tingjia Miao, Zhihui Qi, Yuhan Wu, and Tong Yang. 2025a. Scholarsearch: Benchmarking scholar searching ability of llms. *Preprint*, arXiv:2506.13784.

Peilin Zhou, Bruce Leon, Xiang Ying, Can Zhang, Yifan Shao, Qichen Ye, Dading Chong, Zhiling Jin, Chenxuan Xie, Meng Cao, and 1 others. 2025b. Browsecomp-zh: Benchmarking web browsing ability of large language models in chinese. *arXiv preprint arXiv:2504.19314*.

Yujia Zhou, Zheng Liu, Jiajie Jin, Jian-Yun Nie, and Zhicheng Dou. 2024. Metacognitive retrieval-augmented large language models. In *Proceedings of the ACM Web Conference 2024*, pages 1453–1463.

Andrew Zhu, Alyssa Hwang, Liam Dugan, and Chris Callison-Burch. 2024. Fanoutqa: A multi-hop, multi-document question answering benchmark for large language models. In *Proceedings of ACL 2024 (Short Papers)*.

Changtai Zhu, Siyin Wang, Ruijun Feng, Kai Song, and Xipeng Qiu. 2025a. Convsearch-r1: Enhancing query reformulation for conversational search with reasoning via reinforcement learning. *arXiv preprint arXiv:2505.15776*.

Jiachen Zhu, Menghui Zhu, Renting Rui, Rong Shan, Congmin Zheng, Bo Chen, Yunjia Xi, Jianghao Lin, Weiwen Liu, Ruiming Tang, and 1 others. 2025b. Evolutionary perspectives on the evaluation of llm-based ai agents: A comprehensive survey. *arXiv preprint arXiv:2506.11102*.

Yuchen Zhuang, Yue Yu, Kuan Wang, Haotian Sun, and Chao Zhang. 2023. Toolqa: A dataset for llm question answering with external tools. *Advances in Neural Information Processing Systems*, 36:50117–50143.

A Classification of Closed-ended QA

Multi-hop QA Dataset. Unlike straightforward QA questions, *e.g.*, NQ (Kwiatkowski et al., 2019), TriviaQA (Joshi et al., 2017), and MS MARCO (Nguyen et al., 2016), Multi-hop QA demands that agents piece together information from multiple sources, necessitating multi-step reasoning and iterative information retrieval. Frequently employed datasets for this purpose include HotpotQA (Yang et al., 2018), 2WikiMultiHopQA (Ho et al., 2020), Bamboogle (Press et al., 2022), and MuSiQue (Trivedi et al., 2022b). Recent research has further explored implicit multi-hop questions (Geva et al., 2021a), reasoning over information from diverse sources (Tang and Yang, 2024b), and evaluating the reliability of the reasoning chain (Wu et al., 2024). Additionally, some efforts extend multi-hop QA to integrate various reasoning types (Schnitzler et al., 2024) and multi-modal information (Wang et al., 2025b), pushing the boundaries of complexity.

Challenging QA Dataset. With the expansion of LLMs' internal knowledge, many multi-hop questions are now solvable by these models only with their parametric memory, thereby failing to sufficiently engage iterative planning and retrieval capabilities of search agents. To address this, some new challenging benchmarks are emerging, crafted to present genuinely challenging QA problems set within real-world web environments. Examples like BrowseComp (Wei et al., 2025), BrowseComp-ZH (Zhou et al., 2025b), and Mind2Web2 (Gou et al., 2025) feature long-horizon problems that require extended search durations. Datasets such as InfoDeepSeek (Xi et al., 2025a), ORION (Huang et al., 2025a), and SealQA (Pham et al., 2025) target long-tail knowledge and questions with substantial distracting information. Some method further ensures question difficulty by filtering out the questions that LLMs can easily overcome with a single-turn search (Xi et al., 2025a)

Fact-Checking Dataset. Fact-checking constitutes another critical task for evaluating search agents. The process of verifying factual claims inherently demands iterative search, browsing, and the comparative analysis of disparate information sources. This task is evolving from simple text-based verification to the analysis of long-form factuality (Wei et al., 2024), unstructured data (Aly et al., 2021), multi-modal content (Wang et al., 2024d; Yao et al., 2023; Yang et al., 2025a), and complex multi-hop fact-checking scenarios (Wadden et al., 2020; Jiang et al., 2020; Eisenschlos et al., 2021; Ostrowski et al., 2020), pushing the boundaries of search agent reliability and trustworthiness in sourcing accurate information.

B Detailed Challenges and Future Directions

Despite the significant advancements in Search Agent capabilities, several formidable challenges remain, particularly as we push towards more sophisticated and autonomous systems.

Broaden and Fuse Information Sources. While current search agents primarily leverage browsers and public web data, the next frontier involves integrating more private and proprietary datasets, extending from the agent's external environment to its internal knowledge bases. This integration demands sophisticated methods for combining heterogeneous data formats (text, images, structured data, etc.) and, crucially, resolving conflicts or inconsistencies that arise when information from multiple sources contradicts. Developing robust mechanisms to reconcile conflicting facts or synthesize disparate perspectives is paramount for producing coherent and reliable outputs.

Imperfect Retrieval. Search agents frequently operate in environments where retrieved information is imperfect, containing noise, biases, or even outright falsehoods. The internet, for instance, is replete with unreliable information. A significant challenge is to enhance the agent's ability to discern and validate external information, moving beyond mere retrieval to critical evaluation. This necessitates developing advanced information verification techniques and robust fact-checking mechanisms within the agent's pipeline. Improving an agent's "skepticism" and its capacity for critical assessment is vital for elevating the quality and trustworthiness of its generated outputs.

From Text to Multi-Modality. The majority of current search agents are text-centric. However, the real world is inherently multimodal. A major challenge is to transition search agents from purely text-based understanding to incorporating and processing diverse modalities. This involves two key aspects: first, enhancing the underlying search infrastructure to support better multimodal search queries (e.g., searching for images based on textual descriptions, or videos based on actions). Second, it requires improving the search agent's multimodal comprehension and reasoning abilities – its capacity to understand, synthesize, and reason across text, images, audio, and video to provide more holistic and contextually rich answers.

Customized Reinforcement Learning. While general RL algorithms have shown promise, a significant challenge lies in developing customized reinforcement learning algorithms specifically optimized for search agents. The unique characteristics of search tasks, such as long-horizon planning, imperfect feedback, and the knowledge boundary of search agent, often do not align perfectly with standard RL frameworks. This calls for novel algorithms that can effectively manage sparse rewards in iterative search, learn optimal query generation strategies, and make efficient decisions about when and how to interact with external tools or information sources. Tailored RL approaches can lead to more flexible and robust agent behaviors that adapt dynamically to varied search scenarios.

Crucially, reward modeling for search agents presents its own set of complexities. Current RL-based Search Agents often rely on reward signals derived from easily verifiable, closed-ended problems, such as factual QA, where short, definitive answers allow for straightforward validation. However, many real-world user queries are open-ended information-seeking problems, lacking a single, clear-cut answer. Designing effective reward functions and comprehensive reinforcement learning schemes for these nuanced, open-ended scenarios remains a substantial hurdle. This involves defining what constitutes a "good" answer when there's no single ground truth and how to reward the process of sophisticated information discovery and synthesis. Tailored RL approaches and innovative reward structures are essential for robust agent behaviors that adapt dynamically to varied and ambiguous search needs.

Robust Infrastructure. The ambitious goals for search agents necessitate substantial advancements in underlying infrastructure. A critical challenge is enhancing the efficiency of RL sampling, which can be computationally intensive, to accelerate training. More broadly, optimizing the entire support infrastructure for search agents is crucial. This includes developing high-recall approximate retrieval systems for faster and more relevant information access, implementing priority-aware scheduling to manage complex, concurrent search tasks efficiently, and designing systems that can dynamically schedule requests based on real-time task status, ensuring responsiveness and optimal resource allocation.

Knowledge Boundary Dilemma. Search agents focus on information seeking, yet LLMs also store a vast amount of parametric knowledge and do not need to invoke external search tools for every query. Striking the balance between internal knowledge retrieval and external tool invocation is critical for both efficiency and accuracy. Identifying this *knowledge boundary*, *i.e.*, when to rely on parametric knowledge and when to call external search, remains a crucial challenge. Recent work has explored methods for better integrating these internal and external knowledge sources (Jiang et al., 2025d) and designing reward functions that make agents more aware of this boundary (Huang et al., 2025c).

Information Veracity and Distrust. The open web is rife with misinformation, outdated facts, and adversarial content. LLMs can be misled by incorrect externally retrieved information even when their own parametric knowledge is correct (Xi et al., 2025a). Therefore, building mechanisms for information vetting and critical reasoning is not an optional feature but a core requirement for robust search agents. Potential approaches include using multiple retrieval sources with verification loops (Wang and Yuan, 2025), as well as enhancing the agent's intrinsic critical thinking abilities.

Tool Unreliability and Heterogeneity. A key distinction for search agents is that their primary tool, *i.e.*, the search engine, is inherently unreliable and non-deterministic. Unlike tools that return correct outputs if invoked properly (*e.g.*, calculators or code interpreters), a well-formulated search query can still yield irrelevant or no results. Moreover, different search engines may exhibit heterogeneous capabilities, excelling in different domains (Xi et al., 2025a; Mu et al., 2024). The agent must therefore not only learn *how* to query but also learn *which* search engine to invoke through techniques like query routing (Mu et al., 2024).

Search Agent Self-Evolution. The ultimate frontier for search agents involves achieving true self-evolution. This presents a profound challenge: enabling agents to continuously learn, adapt, and improve their search strategies and capabilities autonomously over time, without constant human intervention. This involves developing mechanisms for agents to identify their own limitations, generate novel hypotheses about how to improve, and then test those hypotheses through interaction with the environment. Such self-evolving agents would possess an unprecedented capacity for discovery and problem-solving, marking a new era in artificial intelligence.

C Detailed Tables

Table 1 presents the detailed classification and comparison of non-tuning methods, showing their search structures (parallel, sequential, or hybrid), their detailed sub-structures, whether they adopt multi-agent architectures and test-time scaling (TTS) strategies, as well as their evaluation methods and metrics.

Table 2 summarizes tuning-based methods, comparing their search structures and substructures, evaluation approaches, and tuning strategies. It specifically analyzes reinforcement learning (RL) methods in terms of training algorithms (RL algo.), supervision signals(RL supv.), and reward functions.

Table 3 illustrates the diverse application domains of search agents, including internal capabilities such as memory, reasoning, and tool use, as well as external domains such as mathematics, coding, finance, and healthcare. It covers commercial deployments, open-source projects, and academic research.

Table 4 compares various datasets, analyzing their categories, scales, modalities, construction methods, evaluation environments, and strategies.

Table 1: Overview of tuning-free methods. ***TTS*** is short for ***Test Time Scaling***.

Model Name	Search Structure	Sub-Structure	Multi-Agent	TTS	Evaluation	Metrics
MAS (Trinh et al., 2025b)	Hybrid	Tree	Yes	No	Rules, LLM	Macro F1
AI Search Paradigm (Li et al., 2025e)	Hybrid	Graph	Yes	No	Human, Online A/B Test	Offline metric: Normalized Win Rate Online metric: Change query rate, Number of page views, Number of daily active users, Dwell time
KnowCoder-V2 (Li et al., 2025g)	Hybrid	Tree	Yes	No	Rules, LLM	QA: Hits@1 Report: Completeness, Thoroughness, Factuality, Coherence, Insight
Multimodal Deep-Researcher (Yang et al., 2025c)	Sequential	Reflection	No	No	Human, LLM	Informativeness and Depth, Coherence and Organization, Verifiability, Visualization Quality, Visualization Consistency
Agentic Deep Research (Zhang et al., 2025e)	Hybrid	Tree	No	Yes	Rules	Accuracy
AutoData (Ma et al., 2025)	Hybrid	Graph	Yes	No	Rules	F1 score, Precision, Recall, Task finishing time
ManuSearch (Huang et al., 2025a)	Hybrid	Tree	Yes	No	LLM	Pass@1
Code Re-searcher (Singh et al., 2025)	Sequential	Proactivity	No	No	Rules	Pass, Crash Resolution Rate
MA-RAG (Nguyen et al., 2025)	Hybrid	Tree	Yes	No	Rules	QA: Exact Match Fact checking: Accuracy
IterKey (Hayashi et al., 2025)	Sequential	Reflection	No	No	Rules	Exact Match, Recall
ODS (Alzubi et al., 2025)	Sequential	Proactivity	No	No	LLM	Accuracy
MCTS-RAG (Hu et al., 2025c)	Hybrid	Tree	No	Yes	Rules	Accuracy
N/A (Xu et al., 2025c)	Hybrid	Tree	Yes	No	N/A	N/A
HG-MCTS (Ren et al., 2025)	Hybrid	Tree	No	Yes	Rules	String accuracy, ROUGE, Exact Match, F1 score, Cover Exact Match
ViDoRAG (Wang et al., 2025b)	Sequential	Proactivity	Yes	Yes	Rules, LLM	Accuracy, Recall
Agentic Reason-ing (Wu et al., 2025d)	Sequential	Proactivity	Yes	Yes	Rules	Accuracy
FinSearch (Li et al., 2024a)	Hybrid	Graph	No	No	Rules	Accuracy, Processing time
SolutionRAG (Li et al., 2025f)	Hybrid	Tree	No	Yes	LLM	Analytical Score, Technical Score

Model Name	Search Structure	Sub-Structure	Multi-Agent	TTS	Evaluation	Metrics
MCTS-KBQA (Xiong et al., 2025b)	Hybrid	Tree	No	Yes	Rules	F1 score, Accuracy, Random Hits@1
Search-o1 (Li et al., 2025c)	Sequential	Proactivity	No	No	–	Exact Match, F1, Pass@1
AirRAG (Feng et al., 2025)	Hybrid	Tree	No	Yes	–	Exact Match, F1 score, Accuracy
ReARTeR (Sun et al., 2025c)	Hybrid	Tree	No	Yes	Rules, LLM	Accuracy
RetroRAG (Xiao et al., 2025)	Sequential	Loop	No	No	Rules	Exact Match, Token-level F1, Precision, Recall
Level-Navi Agent (Hu et al., 2024)	Hybrid	Graph	Yes	No	Rules, LLM	Correctness Scores, Semantic Similarity Scores, Relevance Scores, Searcher Count
RAG-Star (Jiang et al., 2024b)	Hybrid	Tree	No	Yes	Rules	Exact Match, F1 score, Cover Exact Match
AR-MCTS (Dong et al., 2024)	Hybrid	Tree	No	Yes	Rules	Accuracy
SRSA (Wang and Xu, 2024)	Hybrid	Tree	No	No	LLM	Informativeness, Completeness, Novelty, Actionability
PlanRAG (Verma et al., 2025)	Hybrid	Graph	No	No	Rules	Accuracy
CR-Planner (Li et al., 2024b)	Hybrid	Tree	No	Yes	Rules	Accuracy, NDCG@10
DRAG, Iter-RAG (Yue et al., 2024)	Sequential	Loop	No	Yes	Rules	EM, F1, accuracy
Agent-G (lee)	Sequential	Proactivity	No	No	Rules	Hit, Recall, MRR, Accuracy, Hallucination, Missing
Co-STORM (Jiang et al., 2024c)	Sequential	Reflection	Yes	No	Human, LLM	Report Quality: Relevance, Broad Coverage, Depth, Novelty; Discourse Quality: Novelty, Intent Alignment, No Repetition
HiRAG (Zhang et al., 2024c)	Hybrid	Tree	Yes	No	Rules	EM; Token-level F1 score, Precision, Recall
MindSearch (Chen et al., 2024)	Hybrid	Graph	Yes	No	Human, LLM	Depth, Breadth, Factuality, Accuracy
ReSP (Jiang et al., 2025e)	Sequential	Proactivity	Yes	Yes	Rules	F1 score
PlanRAG (Lee et al., 2024)	Sequential	Loop	No	No	Rules	Accuracy
STORM (Shao et al., 2024a)	Sequential	Reflection	No	No	Rules, Human	Heading Soft Recall, Heading Entity Recall
MetaRAG (Zhou et al., 2024)	Sequential	Proactivity	Yes	No	Rules	EM, Precision, F1, Recall
KD-CoT (Wang et al., 2023)	Sequential	Proactivity	Yes	No	Rules	Hits@1, F1

Model Name	Search Structure	Sub-Structure	Multi-Agent	TTS	Evaluation	Metrics
SearChain (Xu et al., 2024b)	Hybrid	Tree	No	No	Rules	ROUGE, Cover EM
HiRA (Jin et al., 2025c)	Hybrid	Tree	Yes	No	LLM	Accuracy

Table 2: Overview of tuning-based methods.

Model Name	Search Structure	Sub-structure	Tuning	Training Env.	RL Algo.	RL Supv.	RL Reward	Metrics
HARIS (Hu et al., 2025b)	Hybrid	Graph	RL	Simulated	GRPO	Rule+ORM	Format Answer Accuracy Decision Accuracy	F1-score Accuracy
CoRAG (Wang et al., 2025a)	Sequential	Proactive	SFT	Simulated	/	/	/	F1-score EM
Self-RAG (Asai et al., 2023)	Hybrid	Tree	SFT	Simulated	/	/	/	EM, F1, FactScore, MAUVE Citation Precision and Recall
Open-RAG (Islam et al., 2024)	Hybrid	Tree	SFT	Simulated	/	/	/	EM, F1, Recall
Auto-RAG (Yu et al., 2024)	Sequential	Proactive	SFT	Simulated	/	/	/	EM, F1, Acc
RAS (Jiang et al., 2025a)	Hybrid	Graph	SFT	Simulated	/	/	/	EM, F1, ROUGE
ReST (Aksitov et al., 2023)	Sequential	Proactive	mixed	Real-world	ReST	ORM	/	Acc
Kwai (Pan et al., 2023)	Sequential	Proactive	SFT	Real-world	/	/	/	ROUGE-L, EM
SimpleDeepSearcher (Sun et al., 2025b)	Hybrid	Tree	mixed	Real-world	DPO, Re-inforce++	Rule+ORM	Format Answer F1	F1, LLM-as-Judge
ReaRAG (Lee et al., 2025)	Sequential	Proactive	SFT	Real-world	/	/	/	F1, EM
EXSEARCH (Shi et al., 2025c)	Sequential	Proactive	RL	Simulated	GEM	PRM	Trajectory Quality Utility	F1, EM
KBQA-o1 (Luo et al., 2025)	Hybrid	Tree	SFT	Simulated	/	/	/	F1, EM
Search-R1 (Jin et al., 2025b)	Sequential	Proactive	RL	Simulated	PPO, GRPO	Rule+ORM	EM	EM
DeepNote (Wang et al., 2024c)	Sequential	Proactive	RL	Simulated	DPO	ORM	/	F1, EM, Acc
R1-Searcher (Song et al., 2025a)	Sequential	Proactive	RL	Simulated	Reinforce++, GRPO	ORM	Format Answer Acc	EM, LLM-as-Judge
ReSearch (Chen et al., 2025b)	Sequential	Proactive	RL	Simulated	GRPO	ORM	Format Answer Acc	EM, LLM-as-Judge
DeepResearcher (Zheng et al., 2025)	Sequential	Proactive	RL	Real-world	GRPO	Rule+ORM	Format Answer F1	F1, LLM-as-Judge
AutoCOA (Zhang et al., 2025g)	Sequential	Proactive	Mixed	Simulated	GRPO	Rule+ORM	Format Answer F1	EM, LLM-as-Judge
SWiRL (Goldie et al., 2025)	Sequential	Proactive	RL	Simulated	Policy Gradient	PRM	LLM(Gemini)	PM,Acc,F1,EM LLM-as-judge
O^2-Searcher (Mei et al., 2025)	Sequential	Proactive	RL	Simulated	GRPO	ORM	Format, Diversity, Factual	EM, F1, LLM Finding Similarity
ZeroSearch (Sun et al., 2025a)	/	/	Mixed	Simulated	PPO,GRPO Reinforce++	Rule	Answer Accuracy	EM
StepSearch (Wang et al., 2025d)	Sequential	Proactive	RL	Simulated	PPO	PRM	Format,Accuracy Search Key Information Gains Redundancy Penalty	EM, F1
VRAG-RL (Wang et al., 2025c)	Sequential	Proactive	RL	Simulated	GRPO	Rule+ORM	Retrieval Efficiency Pattern Consistency Model-Based outcome	Acc
WebThinker (Li et al., 2025d)	Sequential	Proactive	RL	Real-world	DPO	PRM	likelihood of Trajectory	LLM-as-Judge
WebDancer (Wu et al., 2025a)	Sequential	Proactive	Mixed	Real-world	DAPO	Rule+ORM	Format LLM(Answer)	LLM-as-Judge
MaskSearch (Wu et al., 2025f)	Sequential	Proactive	Mixed	Simulated	DAPO	ORM	LLM(Answer)	Token-level Recall

Model Name	Search Structure	Sub-structure	Tuning	Training Env.	RL Algo.	RL Supv.	RL Reward	Metrics
DeepDiver (Shi et al., 2025a)	Sequential	Proactive	Mixed	Real-world	GRPO	Rule+ORM	Format, Accuracy, Extra Search Call Rewards, Loose&Strict Scheduler	LLM-as-Judge
R1-Searcher++ (Song et al., 2025b)	Sequential	Proactive	Mixed	Simulated	Reinforce++	ORM	Format Answer Group (Less Retrieval)	F1, LLM-as-Judge
ReasonRAG (Zhang et al., 2025f)	Sequential	Proactive	RL	Simulated	DPO	PRM	Shortest Path Reward Estimation	EM, F1
ConvSearch-R1 (Zhu et al., 2025a)	Sequential	Proactive	Mixed	Simulated	GRPO	ORM	Rank-Incentive Reward	MRR,NDCG,Recall
EvolveSearch (Zhang et al., 2025a)	Sequential	Proactive	Mixed	Real-world	GRPO	ORM	Format LLM(Answer)	LLM-as-Judge
Inforage (Qian and Liu, 2025)	Sequential	Proactive	Mixed	Simulated	PPO	PRM	Outcome Information Gain Efficiency	EM
AutoRefine (Shi et al., 2025b)	Sequential	Proactive	RL	Simulated	GRPO	ORM+PRM	Answer F1 Retrieval EM	EM
MMOA-RAG (Chen et al., 2025e)	Hybrid	Tree	Mixed	Simulated	MAPPO	ORM	Answer F1 various penalty	EM,F1,Acc
R-search (Zhao et al., 2025)	Sequential	Proactive	RL	Simulated	GRPO	ORM	Answer F1, Model-based Evidence, EM,Format	EM,F1
ReZero (Dao and Le, 2025)	Sequential	Proactive	RL	Simulated	GRPO	Rule+ORM PRM	Acc,Format,Search (Retry/Acc/Diversity)	Acc
s3 (Jiang et al., 2025c)	Sequential	Proactive	RL	Simulated	PPO	ORM	Gain Beyond RAG	Span Match, LLM-as-Judge
/ (Jin et al., 2025a)	Sequential	Proactive	RL	Simulated	PPO/GRPO	ORM	Format Answer F1	F1
Re²Search++ (Xiong et al., 2025a)	Sequential	Proactive	RL	Simulated	DPO/PPO	PRM	Simulated	EM,F1,Acc, LLM-as-Judge
ReARTeR (Sun et al., 2025c)	Sequential	Proactive	RL	Simulated	/	PRM	Monte Carlo Score	Acc, LLM-as-Judge
SmartRAG (Gao et al., 2024)	Sequential	Proactive	Mixed	Real-world	PPO	ORM	Answer Reward Retrieval Penalty	EM, F1, Hit
β-GRPO (Wu et al., 2025e)	Sequential	Proactive	RL	Simulated	β-GRPO	PRM	Min Prob of Search tags	EM
DeepRAG (Guan et al., 2025)	Hybrid	Binary Tree	Mixed	Simulated	GRPO	ORM	Answer acc Retrieval Cost	EM, F1
IEKA (Huang et al., 2025c)	Sequential	Proactive	RL	Simulated	GRPO	ORM	Answer EM Retrieval Counts Penalty	EM, Search Valid Rate
KunLunBaizeRAG (Li et al., 2025a)	Sequential	Proactive	Mixed	Simulated	DAPO	ORM	Answer EM Format,Length Search Effiency	EM, LLM-as-Judge
RAG-R1 (Tan et al., 2025)	Parallel	Chain	Mixed	Simulated	PPO	ORM	Answer EM	EM
Web-Sailor (Li et al., 2025b)	Hybrid	Graph	Mixed	Real-world	DUPO	ORM	Format Answer F1	Pass@1 LLM-Acc

Table 3: Different applications of Search Agents.

Model Name	Scope	Domain	Use Cases	Tuning	Multi-Agent
OpenAI Deep Research (OpenAI, 2025)	external	AI assistant	Commercial	Yes	Yes
Perplexity Deep Research (Perplexity, 2025)	external	AI assistant	Commercial	Yes	Unknown
Gemini Deep Research (Gemini, 2025)	external	AI assistant	Commercial	Yes	Unknown
Grok DeepSearch (Gork, 2025)	external	AI assistant	Commercial	Yes	Unknown
Researcher agent in Copilot (Microsoft, 2025)	external	AI assistant	Commercial	Yes	Unknown
Kimi Researcher (Moonshot, 2025)	external	AI assistant	Commercial	Yes	No
Elicit (Ought, 2025)	external	AI assistant	Commercial	Yes	No
Consensus (Consensus, 2025)	external	AI assistant	Commercial	Yes	No
Manus (BUTTERFLY, 2025)	external	AI assistant	Commercial	Yes	Yes
node-DeepResearch (Jina, 2025)	external	AI assistant	Open Source Project	Yes	Unknown
open deep research (LangChain, 2025)	external	AI assistant	Open Source Project	Yes	Yes
GPT Researcherh (assafelovic, 2025)	external	AI assistant	Open Source Project	Yes	Yes
Open-source DeepResearch (Aymeric, 2025)	external	AI assistant	Open Source Project	Yes	Yes
Open Deep Research (dzhng, 2025)	external	AI assistant	Open Source Project	Yes	Unknown
open-deep-research (nickscamara, 2025)	external	AI assistant	Open Source Project	Yes	No
AgenticSeek (Fosowl, 2025)	external	AI assistant	Open Source Project	Yes	Yes
DeerFlow (Bytedance, 2025)	external	AI assistant	Open Source Project	Yes	Yes
OpenManus (FoundationA-gents, 2025)	external	AI assistant	Open Source Project	Yes	Yes
SimpleDoc (Jain et al., 2025)	external	Document Visual	Research	Yes	Yes
StePO-Rec (Bi et al., 2025)	external	E-commerce	Research	No	No
DeepShop (Lyu et al., 2025)	external	E-commerce	Research	Yes	No
ARAG (Maragheh et al., 2025)	external	Recommendation	Research	No	Yes
MACRec (Wang et al., 2024e)	external	Recommendation	Research	No	Yes
ToolRec (Zhao et al., 2024)	external	Recommendation	Research	No	Unknown
PUMA (Cai et al., 2025)	external	E-commerce	Research	No	Unknown
FinSearch (Li et al., 2024a)	external	Finance	Research	Yes	Yes
PlanRAGh (Lee et al., 2024)	external	Business	Research	Yes	Unknown
Glass-Box Agent (Vaghefi et al., 2025)	external	Finance	Research	Yes	Yes
Code Researcher (Singh et al., 2025)	external	Code	Research	Yes	No
ARCeR (Lupinacci et al., 2025)	external	Code	Research	Yes	No
CodeAgent (Zhang et al., 2024a)	external	Code	Research	Yes	No

Model Name	Scope	Domain	Use Cases	Tuning	Multi-Agent
ARCS (Bhattarai et al., 2025)	external	Code	Research	Yes	No
Agentic Retrieval-Augmented (Ravuru et al., 2024)	external	Time Series Analysis	Research	No	Yes
Agentic RAG (Spielberger et al., 2025)	external	Topic Modeling	Research	No	No
DeRetSyn (Bhattacharyya, 2025)	external	Medicine	Research	No	No
MRD-RAG (Chen et al., 2025f)	external	Medical Diagnosis	Research	No	No
MedAgent-Pro (Wang et al., 2025e)	external	Medical Diagnosis	Research	No	No
CBMs and Multi-Agent RAG (Tusfiqur Alam et al., 2024)	external	Medicine	Research	No	Yes
DeepSeq (Al Dajani et al., 2025)	external	Biology	Research	Yes	No
TourSynbio-Search (Liu et al., 2024b)	external	Biology	Research	Yes	Yes
LLM-based system (Brett and Myatt, 2025)	external	Scientific Reserach	Research	Yes	Yes
PaSa (He et al., 2025)	external	Scientific Reserach	Research	Yes	Yes
CollEX (Schneider et al., 2025)	external	Scientific Reserach	Research	Yes	Yes
Claude 3.7 Sonnet (Berkane et al., 2025)	external	Scientific Reserach	Research	Yes	No
IoT-ASE (Elewah and Elgazzar, 2025)	external	Internet of Things	Research	Yes	Yes
CRAG-MoW (Callahan et al., 2025)	external	Chemistry	Research	No	Yes
AAWN and Graph RAG Framework (Srinivas et al., 2024)	external	Chemistry	Research	No	Yes
SolutionRAG (Li et al., 2025f)	external	Engineering	Research	No	No
Agentic Multimodal RAG (Elahi and Zyngier)	external	Water Resources	Research	No	Yes
CMI (Thakrar et al., 2025)	external	Medical Diagnosis	Research	No	No
AIDE (Toledo et al., 2025)	external	Code	Research	No	No
Toolshed (Lumer et al., 2024)	internal	Tool Use	Research	No	No
AnyTool (Du et al., 2024)	internal	Tool Use	Research	No	Yes
ScaleMCP (Lumer et al., 2025)	internal	Tool Use	Research	No	No
TR-Feedback (Xu et al., 2024a)	internal	Tool Learning	Research	No	No
TaskGen (Tan et al., 2024b)	internal	Memory	Research	No	Yes
A-MEM (Xu et al., 2025b)	internal	Memory	Research	No	No
Grounded Memory System (Ocker et al., 2025)	internal	Memory	Research	Yes	No
Agentic Reasoning (Wu et al., 2025d)	internal	Memory	Research	Yes	Yes
AGENT KB (Tang et al., 2025)	internal	Memory	Research	No	Yes
RAT (Wang et al., 2024f)	internal	Reasoning	Research	Yes	No
KG-RAR (Wu et al., 2025g)	internal	Reasoning	Research	Yes	No
KD-CoT (Wang et al., 2023)	internal	Reasoning	Research	No	No
Search-o1 (Li et al., 2025c)	internal	Reasoning	Research	Yes	No

Model Name	Scope	Domain	Use Cases	Tuning	Multi-Agent
CR-Planner (Li et al., 2024b)	internal	Reasoning	Research	Yes	No
AR-MCTS (Dong et al., 2024)	internal	Reasoning	Research	Yes	No
SWiRL (Goldie et al., 2025)	internal	Reasoning	Research	Yes	No
WebThinker (Li et al., 2025d)	internal	Reasoning	Research	Yes	No

Table 4: Overview of datasets.

Dataset Name	Category	Scale	Modality	Constr-uction	Environ-ments	Evaluation	Metrics
HotpotQA (Yang et al., 2018)	Multi-hop QA	113k	Text	Rules, Manual	Static	Rules	EM, F1
2WikiMultiHopQA (Ho et al., 2020)	Multi-hop QA	192606	Text	Rules, Manual	Static	Rules, Human, LLM	EM,F1
Bamboogle (Press et al., 2023)	Multi-hop QA	8600	Text	Rules, Manual	Static	Human, LLM	Acc, Gap Ratio, Compositionality
MuSiQue-Ans (Trivedi et al., 2022b)	Multi-hop QA	25000	Text	Manual	N/A	Human	EM,F1
StrategyQA (Geva et al., 2021b)	Multi-hop QA	2780	Text	Manual	N/A	Human	Acc
FRAMES (Krishna et al., 2024)	Multi-hop QA	824	Text	Manual	Dynamic	Human	Acc
MultiHop-RAG (Tang and Yang, 2024a)	Multi-hop QA	/	Text	Manual	Dynamic	Human	N/A
HoVer (Jiang et al., 2020)	Fast-Checking	/	Text	Manual	Dynamic	Human	Acc
FanOutQA (Zhu et al., 2024)	Multi-hop QA	8339	Text	Manual	Dynamic	Human	Acc
Web24 (Hu et al., 2025a)	Others	/	Text	Manual	Dynamic	Human	N/A
ViDoSeek (Zhang et al., 2025d)	Others	/	Text, Image	Manual	Dynamic	Human	N/A
MoreHopQA (Liu et al., 2024a)	Multi-hop QA	1118	Text	Manual	N/A	Human	Acc, Reasoning Step Acc
CofCA (Wu et al., 2025c)	Multi-hop QA	1800	Text	Manual	Static	Human	EM, F1, PM, Reasoning Chain Acc
BrowseComp (Wei et al., 2025)	Challeging QA	1266	Text	Manual	Dynamic	LLM	Acc, CE
InfoDeepSeek (Xi et al., 2025a)	Challeging QA	245	Text	Manual	Dynamic	LLM, Human	AnsAcc, InfoAcc, EffEvidUtil, Info-Compactness
ORION (Huang et al., 2025a)	Challeging QA	310	Text	LLM, Manual	Dynamic	LLM	Pass@1 Acc
BrowseComp-ZH (Zhou et al., 2025b)	Challeging QA	289	Text	Manual	Dynamic	LLM, Human	Acc, CE
PopQA (Mallen et al., 2022)	Challeging QA	14k	Text	Rules	Static	Rules	Acc
WebPuzzle (Shi et al., 2025a)	Challeging QA	24275	Text	LLM, Manual	Dynamic	LLM	Acc, CE
BLUR (CH-Wang et al., 2025)	Challeging QA	573	Text, Image, Video, Audio	Manual	Dynamic	Rules, LLM	Acc
BRIGHT (Su et al., 2024)	Challeging QA	1384	Text	LLM, Manual	Dynamic	Rules	nDCG@10, Precision@10, Recall@10

Dataset Name	Category	Scale	Modality	Constr-uction	Environ-ments	Evaluation	Metrics
SealQA (Pham et al., 2025)	Challeging QA	619	Text	LLM, Manual	Dynamic	LLM	Acc
MMSearch (Jiang et al., 2024a)	Challeging QA	300	Text, Image	Manual	Dynamic	Rules	F1, ROUGE-L, BLEU-1, Acc
ScholarSearch (Zhou et al., 2025a)	Challeging QA	223	Text	LLM, Manual	Dynamic	LLM	Acc
Mind2Web 2(Gou et al., 2025)	Challeging QA	130	Text, Image	LLM, Manual	Dynamic	LLM, Human	PartComp, SuccRate, Pass@3, AvgTime, AnsLen
O2-QA (Mei et al., 2025)	Open-ended QA	300	Text	LLM, Manual	Static	Rules, LLM	F1, EM, LFS
Researchy Questions (Rosset et al., 2024)	Open-ended QA	96k	Text	Rules, LLM, Manual	Static	LLM	Acc, Score
MultimodalReportBench (Yang et al., 2025c)	Open-ended QA	100	Text, Image	Manual	Dynamic	LLM, Human	InfoDepth, OrgCoh, Verif, VisQual, VisCons
DeepResearchGym (Coelho et al., 2025)	Open-ended QA	unknown	Text	Rules	Static	LLM	KP-Rec, KP-Contra, Cit-Prec, Cit-Rec, Clarity, Insight
Deep Research Bench (Bosse et al., 2025)	Open-ended QA	89	Text	Manual	Static	Rules, LLM, Human	BinScore, Recall(Rec), F1, AbsDiff
DeepResearch Bench (Du et al., 2025)	Open-ended QA	100	Text	Manual	Static	Rules, LLM	Comprehensiveness, Insight/Depth, InstFollowing ...
WildSeek (Jiang et al., 2024c)	Open-ended QA	100	Text	Rules, LLM, Manual	Dynamic	Rules, LLM, Human	Relevance, Breadth, Depth, Novelty, Consistency...
ProxyQA (Tan et al., 2024a)	Open-ended QA	100	Text	Manual	N/A	LLM	Acc, SelfAR, HumAR
Long2RAG (Qi et al., 2024)	Open-ended QA	280	Text	LLM, Manual	Dynamic	LLM	KP-Rec, KP-F1, KP-Precision(KP-Pre)
Mocheg (Yao et al., 2023)	Fact-Checking	61593	Text, Image	Rules, Manual	Static	Rules	Prec, Rec, NDCG, MAP, S-Rec
MFC-Bench (Wang et al., 2024d)	Fact-Checking	35k	Text, Image	Rules, LLM, Manual	Static	Rules	Acc, F1
RealFactBench (Yang et al., 2025a)	Fact-Checking	6k	Text, Image	LLM, Manual	Dynamic	Rules	F1, MCC, UnkRate, ExpQuality
LongFact (Wei et al., 2024)	Fact-Checking	2280	Text	LLM, Manual	Dynamic	LLM	F1@K, Prec, Rec
PolitiHop (Ostrowski et al., 2020)	Fact-Checking	500	Text	Manual	Static	Rules	macro-F1, Acc, F1, Prec, FEVER score

Dataset Name	Category	Scale	Modality	Construction	Environments	Evaluation	Metrics
FM2 (Eisenschlos et al., 2021)	Fact-Checking	12968	Text	Manual	Static	Rules	R-Prec, Rec@5, Rec@10, Acc, F1, FEVER score
HoVer (Jiang et al., 2020)	Fact-Checking	26171	Text	Manual	Static	Rules	Acc, F1, EM, R-Prec, Rec@k, HOVER Score
SCIFACT	Fact-Checking	1409	Text	Rules, Manual	Static	Rules	F1, Acc, Rec, FEVER Score
EX-FEVER (Ma et al., 2023a)	Fact-Checking	60k+	Text	Manual	Static	Rules	EM, Hit@k, Acc, Rouge score
FEVEROUS (Aly et al., 2021)	Fact-Checking	87026	Text, Table	Manual	Static	Rules	FEVER Score, EM, Hit@k, Acc, F1, Rec
FactBench (Bayat et al., 2024)	Fact-Checking	1000	Text	Rules, LLM, Manual	Dynamic	LLM	Hallucination Score, Fact-Prec, Acc, CHumJ
FinSearchBench-24 (Li et al., 2024a)	Domain-Specific	1500	Text	Rules, LLM, Manual	Dynamic	Rules	Acc, Time
MIRAGE (Dongre et al., 2025)	Domain-Specific	35306	Text, Image	LLM, Manual	Static	LLM	Reasoning Score, Acc, IdentAcc, Relevance, Completeness...
xbench (Chen et al., 2025a)	Domain-Specific	100	Text	Manual	Dynamic	LLM	Matching Rate, AccScore
SolutionBench (Li et al., 2025f)	Domain-Specific	950	Text	LLM, Manual	Static	LLM	Analytical Score, Technical Score
DQA (Lee et al., 2024)	Domain-Specific	301	Text	Rules	Static	Rules	Acc
MedMCQA (Pal et al., 2022)	Domain-Specific	193k+	Text	Rules, Manual	Static	LLM	Acc
MedBrowseCom (Chen et al., 2025c)	Domain-Specific	1k+	Text	Manual	Dynamic	LLM, Human	Acc, ECErr
GPQA (Rein et al., 2024)	Domain-Specific	546	Text	Manual	N/A	Human	Acc, ECErr, Post-hoc agreement
ScIRGen-Geo (Lin et al., 2025b)	Domain-Specific	61k	Text	LLM, Manual	Static	Rules	Entailment Model Acc, Recall@k, MRR@100, ROUGE-L
OlympiadBench (He et al., 2024)	Domain-Specific	8476	Text, Image	Rules, LLM, Manual	Static	Rules, Human	Micro-average Acc, Acc
DeepShop (Lyu et al., 2025)	Domain-Specific	600	Text, Image	LLM, Manual	Dynamic	Rules, LLM, Human	Fine-grained SuccRate, Holistic Task SuccRate
USACO (Shi et al., 2024a)	Domain-Specific	307	Text	Rules, LLM, Manual	Static	Rules	Execution SuccRate, Pass@1 Acc
GAIA (Mialon et al., 2023)	Domain-Specific	466	Text, Image	Manual	Dynamic	Rules	Acc

Dataset Name	Category	Scale	Modality	Construction	Environments	Evaluation	Metrics
HLE (Phan et al., 2025)	Domain-Specific	2500	Text, Image	LLM, Manual	Static	LLM	Acc, RMS
HERB (Choubey et al., 2025)	Domain-Specific	40704	Text	LLM, Manual	Static	Rules, LLM	F1, Prec, Rec, AvgPerform-Score, Likert Scale Rating
RAGChecker (Ru et al., 2024)	Others	4162	Text	LLM, Manual	Static	LLM	Prec, Rec, F1
ToolQA (Zhuang et al., 2023)	Others	1530	Text, Table	Rules, LLM, Manual	Static	Rules	EM
KILT (Petroni et al., 2020)	Others	3.2M	Text	Rules, Manual	Static	Rules	Acc, EM, F1, ROUGE-L, R-Prec, Recall@k, KILT scores
CONFLICTS (Cattan et al., 2025)	Others	458	Text	Manual	Dynamic	LLM	Fact-Ground, AnsRec, ExpBeh
Search Arena (Miroyan et al., 2025)	Others	36721	Text	Manual	Dynamic	Human	WinRate, Bradley-Terry model coefficients
WebWalkerQA (Wu et al., 2025b)	Others	680	Text	LLM, Manual	Dynamic	LLM, Rules	Acc, Action count
IIRC (Ferguson et al., 2020)	Others	13441	Text	Manual	Dynamic	Rules	EM, F1
Instruct2DS (Ma et al., 2025)	Others	234	Text	Manual	Dynamic	Rules	F1, Prec, Rec, Exe-Prec, Exe-Rec
DRComparator (Chandrahasan et al., 2025)	Others	/	Text	Manual	Dynamic	Human	BT Score