LONGSEEKER: ELASTIC CONTEXT ORCHESTRATION FOR LONG-HORIZON SEARCH AGENTS

Anonymous authors
Paper under double-blind review

ABSTRACT

Long-horizon search agents must manage a rapidly growing working context as they reason, call tools, and observe information. Naively accumulating all intermediate content can overwhelm the agent, increasing costs and the risk of errors. We propose that effective context management should be adaptive: parts of the agent's trajectory are maintained at different levels of detail depending on their current relevance to the task. To operationalize this principle, we introduce **Context-ReAct**, a general agentic paradigm for *elastic context orchestration* that integrates reasoning, context management, and tool use in a unified loop. Context-ReAct provides five atomic operations: *Skip*, *Compress*, *Rollback*, *Snippet* and *Delete*, which allow the agent to dynamically reshape its working context, preserving important evidence, summarizing resolved information, discarding unhelpful branches, and controlling context size. Together, these operations form a fine-grained action space spanning abstractive compression, exact extraction, deletion, structural rollback, and no modification. Building on this paradigm, we develop **LongSeeker**, a long-horizon search agent fine-tuned from Qwen3-30B-A3B on 10k synthesized trajectories. Across four representative search benchmarks, LongSeeker achieves **61.5%** on BrowseComp and **62.5%** on BrowseComp-ZH, substantially outperforming Tongyi DeepResearch (43.4% and 46.7%) and AgentFold (36.2% and 47.3%). These results highlight the potential of adaptive context management, showing that agents can achieve more reliable and efficient long-horizon reasoning by actively shaping their working memory.

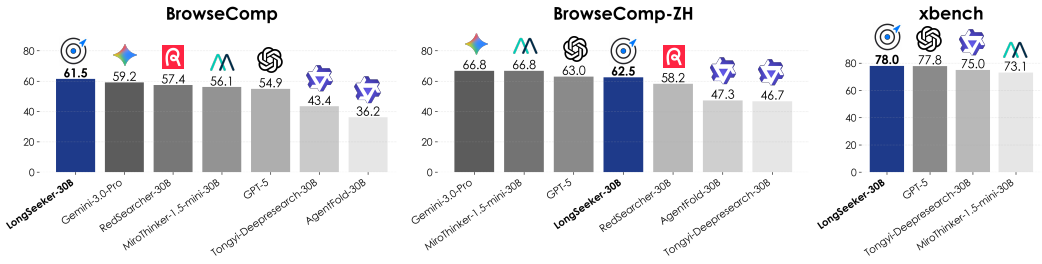


Figure 1: LongSeeker-30B delivers strong results on challenging long-horizon benchmarks, matching or surpassing several foundation models and search agents.

1 INTRODUCTION

The emergence of search agents has transformed how humans retrieve and synthesize information from the web. Tasks once requiring manual query iteration can now be delegated end-to-end to an AI agent via a single instruction. Agentic search has thus become a cornerstone capability pursued by AI labs, exemplified by the trajectory from OpenAI's Deep Research systems (OpenAI, 2025a) to today's top-tier large language models supporting multi-step, tool-augmented reasoning built upon the ReAct paradigm (Yao et al., 2023).

However, long-horizon search agents built on the ReAct paradigm face an inherent context bottleneck: as observations, reasoning traces, and tool calls accumulate, the working context becomes increasingly

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

noisy, redundant, and eventually too long to retain in full. Existing remedies remain partial: sliding-window truncation is importance-agnostic (Team et al., 2026); threshold-triggered re-starting disrupts reasoning continuity (DeepSeek-AI, 2025); periodic summarization suffers from fixed-granularity compression and accumulating abstraction errors (Zhou et al., 2025b; Yu et al., 2025; Lu et al., 2025); and proactive curation is still limited in where or how it can intervene (Ye et al., 2025; Yao et al., 2026). As a result, existing methods cannot provide precise, on-demand control over the evolving shape of the agent's context.

Addressing these, our key insight is that effective context management requires an *elastic* working context, one that can compress, preserve, discard, and restructure different parts of the agent's context according to its current state. That is, during long-horizon search, information should exist in different forms as the task evolves: fresh evidence may need to remain intact for verification, resolved evidence can be distilled into conclusions, precision-critical details may survive as snippets, and failed branches should be removed or rolled back. This *state-dependent fidelity* ensures that the agent maintains the right level of detail for each part of its agentic reasoning trajectory.

Following this insight, we propose **Context-ReAct**, a general agentic paradigm for *elastic context orchestration* in long-horizon search agents. At each reasoning turn, the agent jointly produces its reasoning trace, a set of context meta-operations, and the next tool call in a single autoregressive pass. The meta-operations are applied before the next observation is appended, allowing the agent to actively determine *when* to update its context, *where* in the trajectory to intervene, and *how* each part of the past should be represented.

Based on this paradigm, we design a meta-operation vocabulary consisting of five atomic actions. (1) *Skip* leaves the context unchanged when it is already compact and informative. (2) *Compress* replaces any contiguous range of historical steps with an abstractive summary. (3) *Snippet* preserves an exact substring from a step, retaining precision-critical evidence such as numbers, entity names, quotations, or code without abstractive distortion. (4) *Delete* removes a step which no longer carries residual value. (5) *Rollback* abandons an unproductive branch by reverting the context to an earlier state while recording the reason for backtracking and any transferable insight. Together, these fine-grained operations maintain a multi-resolution working context in which different parts of the context can remain verbatim, be compressed, be partially quoted, be removed, or be structurally rolled back.

To instantiate and evaluate Context-ReAct, we build **LongSeeker**, a long-horizon search agent fine-tuned from Qwen3-30B-A3B on 10k synthesized search trajectories. We evaluate LongSeeker on four representative search benchmarks: BrowseComp (Wei et al., 2025), BrowseComp-ZH (Zhou et al., 2025a), xbench (Chen et al., 2025), and GAIA (Mialon et al., 2023). LongSeeker achieves scores of **61.5%** and **62.5%** on BrowseComp and BrowseComp-ZH respectively, significantly outperforming competitive baselines such as Tongyi DeepResearch (43.4% and 46.7%) (Team et al., 2025) and AgentFold (36.2% and 47.3%) (Ye et al., 2025). These results suggest that elastic context orchestration is a scalable path toward more capable long-horizon agents, shifting context management from a peripheral engineering heuristic to a core component of agentic reasoning.

Our main contributions are:

- Paradigm. We propose Context-ReAct, a general agentic paradigm for elastic context orchestration that lets agents decide *when*, *where*, and *how* to reshape their working context during ReAct-style search.
- Operations. We introduce five meta-operations, *Skip*, *Compress*, *Rollback*, *Snippet*, and *Delete*, forming a fine-grained operation set for multi-resolution context control.
- Experiments. We train **LongSeeker** on 10k synthesized trajectories and achieve **61.5%** on BrowseComp and **62.5%** on BrowseComp-ZH, outperforming strong long-horizon search baselines.

2 Related Work

Search Agents. LLM-based search agents have transformed information retrieval from static query-response matching into dynamic, multi-step reasoning processes. Central to this transformation is the ReAct paradigm (Yao et al., 2023), which structures agent behavior as an iterative cycle of reasoning, action execution, and observation integration. OpenAI's Deep Research (OpenAI, 2025a) pioneers the fully closed-source path, followed by a series of proprietary agents; meanwhile,

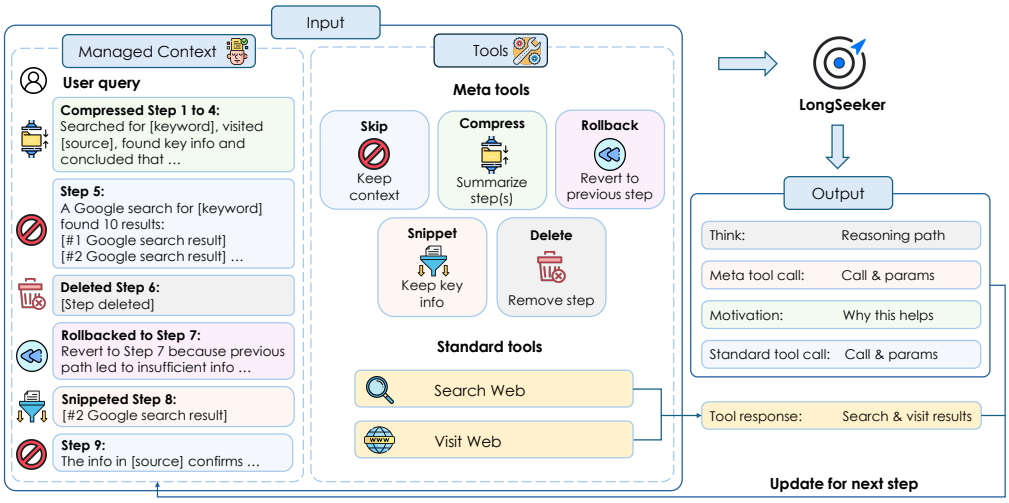


Figure 2: Overview of the **Context-ReAct** paradigm. Unlike standard ReAct, which passively accumulates history, and unlike prior proactive curation methods (Ye et al., 2025; Yao et al., 2026) that operate at a coarse granularity, Context-ReAct introduces a *fine-grained* meta-action layer. At each step, the agent co-generates meta-operations (SKIP, COMPRESS, ROLLBACK, SNIPPET, DELETE) alongside standard tool calls, enabling elastic context orchestration that spans the full spectrum from lossless extraction to structural backtracking.

open-source efforts such as WebSailor (Li et al., 2025) and Tongyi DeepResearch (Team et al., 2025) push capabilities forward through large-scale trajectory synthesis and post-training optimization. Yet these advances retain a fundamental limitation: they follow the conventional ReAct pattern of unconditionally accumulating observations, causing progressive degradation in context quality and heightened risk of exceeding context windows during extended tasks.

Context Management for Agents. Managing growing context has attracted considerable recent attention, with existing methods falling into four categories. *Sliding-window* heuristics such as keep-last-k—adopted by the MiroThinker series (Team et al., 2026)—retain only recent steps and discard older content regardless of importance, while *discard-all* variants flush the entire context once thresholds are reached, as in DeepSeek-V3.2 (DeepSeek-AI, 2025) and GLM-4.7 (Zhipu AI, 2025). *Periodic summarization* methods such as MEM1 (Zhou et al., 2025b) train agents to maintain compact internal states, while MemAgent (Yu et al., 2025) processes documents in segments with fixed-size memory buffers. *Proactive curation* approaches such as AgentFold (Ye et al., 2025) and ARC (Yao et al., 2026) enable agents to actively decide what and when to compress, yet they still lack surgical operations and cannot revisit earlier history to purge outdated content.

Our Approach. Context-ReAct advances beyond prior work by defining a fine-grained set of five atomic meta-operations—*Skip*, *Compress*, *Rollback*, *Snippet* and *Delete*—that are co-generated with standard tool calls at every step. Unlike fixed-rule truncation, each decision is content-aware; unlike periodic summarization, operations are invoked only when necessary; unlike proactive curation, our operation set spans the spectrum from exact extraction to structural backtracking. All decisions are learned end-to-end from synthesized long-horizon search trajectories.

3 METHOD

We present **Context-ReAct**, a general paradigm that augments the standard ReAct loop with an explicit *meta-action* layer for on-demand context management. As illustrated in Figure 2, our approach enables the agent to actively curate its working memory by co-generating meta-operations alongside standard reasoning and tool calls. Section 3.1 establishes the formal definition; Section 3.2 defines the five atomic meta-operations; and Section 3.3 details the data synthesis and training pipeline.

3.1 AGENTIC PARADIGM

In this section, we introduce our proposed Context-ReAct with a formal definition.

Standard ReAct. In the standard ReAct paradigm (Yao et al., 2023), each step S_i^{std} is defined as

$$S_i^{\text{std}} = (r_i,\ c_i,\ o_i), \tag{1}$$

where r_i is the chain-of-thought reasoning trace, c_i is the tool call, and o_i is the environment observation returned by the tool. The context history at time t is the concatenation $H_t = [S_1^{\text{std}}, \ldots, S_t^{\text{std}}]$, which grows monotonically and without bound under this *append-only* design. As irrelevant observations accumulate, the signal-to-noise ratio of H_t degrades, and $|H_t|$ eventually risks exceeding the model's context limit.

Context-ReAct. To preserve the generality of the standard ReAct paradigm while equipping the agent with dynamic context management, we augment each step by inserting several *meta operations* between the reasoning trace and the standard tool call. The resulting step structure is

$$S_i^{\text{meta}} = (r_i,\ M_i,\ c_i,\ o_i), \tag{2}$$

where $M_i = [op_1^{(i)}, op_2^{(i)}, \ldots, op_k^{(i)}]$ is a list of meta-operations generated by the agent to transform the context *before* the next step begins. Formally, the effective context H'_t used at step $t{+}1$ is

$$H'_t \ = \ T(H_t,\ M_t), \tag{3}$$

where T is the composition of the individual operations in M_t, each drawn from the primitive set (defined in Section 3.2). This mechanism enables the agent to maintain a compact and relevant working memory ($|H'_t| \ll |H_t|$ in typical long-horizon tasks) without any external trigger or architectural modification.

The design principle is that meta-operations are *co-generated* with the reasoning trace and tool call in a single, end-to-end generation step, rather than being triggered by an external heuristic such as a length threshold. The agent therefore learns *when*, *where*, and *how* to intervene in its own context as an integral part of its policy; see Figure 2 (right) for illustration.

3.2 ATOMIC META-OPERATIONS

To enable flexible and task-aware management of the working context, Context-ReAct equips the agent with a set of atomic meta-operations. These operations define the full set of primitive actions that can be applied to the context at each reasoning step. Throughout, let $H = [S_1, \ldots, S_n]$ denote the current history.

(1) SKIP is the identity operator. That is, the agent takes no action on the context:

$$\textsc{Skip}(H) \ = \ H, \tag{4}$$

which is issued when the current context is already compact, without modifying the current context.

(2) COMPRESS performs abstractive summarization over any contiguous range of steps $[a, b]$ ($a \leq b$), replacing the context with a summarized step $S_{a:b}$, where Σ is the summarized string:

$$\textsc{Compress}(H,\ a,\ b,\ \Sigma) \ = \ [S_1,\ \ldots,\ S_{a-1},\ S_{a:b} = \Sigma,\ S_{b+1},\ \ldots,\ S_n]. \tag{5}$$

Crucially, $[a, b]$ need not be a recent window (Ye et al., 2025): the agent can *retroactively* recognize that steps from early in the trajectory have become compressible. For example, earlier searches may no longer need to be preserved in full once their useful information has been captured by later evidence, allowing the agent to compress them on the fly. This look-back flexibility is unavailable in sliding-window approaches, which can only truncate from one end of the history.

(3) ROLLBACK reverts the context to step k by discarding all subsequent steps $S_k, \ldots, S_n$ and appending a summarized step that records the reason for backtracking and any transferable insight:

$$\textsc{Rollback}(H,\ k,\ \Sigma) \ = \ [S_1,\ \ldots,\ S_k = \Sigma]. \tag{6}$$

ROLLBACK models the structural intuition of branch abandonment in tree-based search (DFS/MCTS) (Shi et al., 2025): when the agent recognizes that a reasoning path has reached a

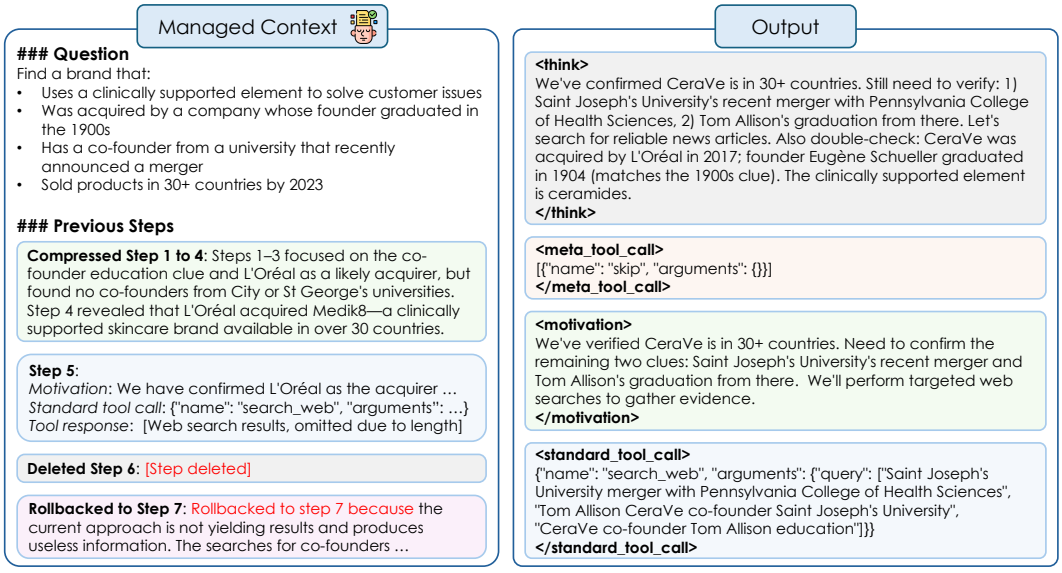


Figure 3: **Managed context and structured output at a single Context-ReAct step.** *Left:* The curated context after applying meta-operations to the raw trajectory. Steps 1–4 are consolidated via COMPRESS into a summary preserving essential findings. Step 5 remains unchanged via SKIP. Step 6 is DELETEd to eliminate noise from a redundant query. The trajectory is ROLLBACKed to Step 7, discarding unproductive exploration. The resulting context is minimal yet information-dense. *Right:* The four-field structured output containing reasoning, meta-operations, motivation, and the standard tool call.

dead end, it discards the failed sub-trajectory while preserving its causal explanation, preventing the same mistake from being repeated.

(4) SNIPPET replaces the observation o_k of the k-th step with the verbatim substring delimited by the anchor strings *pre* and *suf*:

$$\textsc{Snippet}(H,\, k,\, \textit{pre},\, \textit{suf}) \;=\; \big[S_1,\; \ldots,\; (r_k,\, c_k,\, o_k[\textit{pre}{:}\textit{suf}]),\; \ldots,\; S_n\big]. \tag{7}$$

Unlike generative summarization, SNIPPET is *lossless* with respect to the retained segment: it performs pointer-based substring extraction rather than token regeneration, preserving precise numerical values, entity names, URLs, or code without abstractive rewriting.

(5) DELETE removes the k-th step, discarding its reasoning trace, tool call, and observation:

$$\textsc{Delete}(H,\, k) \;=\; \big[S_1,\; \ldots,\; S_{k-1},\; S_{k+1},\; \ldots,\; S_n\big]. \tag{8}$$

This operation is appropriate when an entire step is uninformative and leaves no useful trace—e.g., a failed or redundant query whose result, reasoning, and call all warrant complete removal to reduce noise.

Composite application. Multiple meta-operations may be composed within a single step by listing them sequentially in M_i. This compositionality allows **LongSeeker** to, for instance, DELETE a noisy interaction step and simultaneously COMPRESS a longer historical segment in one step.

Figure 3 illustrates the combined effect of these meta-operations on a live trajectory (see Appendix C for a complete example). Rather than receiving the full raw trajectory, the model sees a compact, curated view of its history. As shown in the left half of the figure, Steps 1–4 have been consolidated via COMPRESS into a single summary sentence that preserves the essential findings while discarding verbose intermediate observations. Step 5 remains unchanged via SKIP. Step 6 has been DELETEd to eliminate noise from a redundant query. Most significantly, the trajectory has been ROLLBACKed to Step 7, discarding an unproductive sub-chain of exploration. The resulting managed context is minimal yet information-dense, ensuring the model can focus on the remaining open sub-questions.

Together, these fine-grained operations allow the agent to select the appropriate level of fidelity and structural change for each part of its trajectory.

3.3 Data Synthesis and Training

Trajectory synthesis. Training **LongSeeker** via **Context-ReAct** requires trajectories that contain not only correct final answers but also high-quality *context management decisions* at intermediate steps—a supervision signal absent from all existing datasets. We construct a corpus of 10,000 annotated trajectories through a two-stage pipeline.

Stage 1: Seed question collection. We sample 10,000 complex, multi-hop questions from OpenSeeker (Du et al., 2026b;a), comprising 9,000 English and 1,000 Chinese questions, filtering for questions that require substantive multi-step reasoning to answer.

Stage 2: Context-ReAct trajectory rollout. Each question is solved by DeepSeek-V3.2 (DeepSeek-AI, 2025) acting as the teacher model and operating under the full Context-ReAct paradigm. At every step, the teacher directly generates the complete four-field structured output—`<think>`, `<meta_tool_call>`, `<motivation>`, and `<standard_tool_call>`—in a single pass, producing context management decisions and the next tool call jointly. Trajectories with correct format constitute the final training set.

Supervised fine-tuning. We fine-tune Qwen 3 30B-A3B (Yang et al., 2025) on the annotated corpus via standard next-token prediction:

$$\mathcal{L}_{\text{SFT}} = -\sum_{t=1}^{T}\sum_{j}\log p_\theta\left(x_j^{(t)} \;\middle|\; x_{<j}^{(t)},\ H'_{t-1}\right), \tag{9}$$

where $x^{(t)}$ is the full structured output at step t—including the chain-of-thought, meta-tool call, motivation, and standard tool call—and H'_{t-1} is the context after applying the meta-operations from the previous step. Computing the loss over the *entire* structured output forces the model to jointly learn *which* meta-operation to invoke, *when* to invoke it, and *how* to use the standard tool given the current context.

4 Experiments

4.1 Experimental Setup

Evaluations. We evaluate our **LongSeeker** on four key benchmarks spanning targeted information-seeking and general agent capabilities. **BrowseComp** (Wei et al., 2025) and **BrowseComp-ZH** (Zhou et al., 2025a) evaluate multi-step navigation and hard information retrieval in English and Chinese, respectively. Due to resource constraints, we randomly sample 400 questions from BrowseComp; the remaining benchmarks are evaluated on their original evaluation sets. **xbench** (Chen et al., 2025) assesses complex deep research capabilities including planning, reasoning, and synthesis across profession-aligned real-world tasks. Finally, **GAIA** (Mialon et al., 2023) (text-only subset) evaluates general agent capabilities requiring combined web browsing, tool use, and multi-step reasoning. We set the max tool call as 300 for all benchmarks. For BrowseComp and BrowseComp-ZH, we also apply the discard-all technique and allow for 5 rounds at maximum following MiroThinker (Team et al., 2026). All LongSeeker results are averaged over three runs.

Baselines. To validate the effectiveness of **LongSeeker**, we compare it against several state-of-the-art systems categorized into two groups: (1) *foundation models with tools*, comprising frontier proprietary systems such as GPT-5 (OpenAI, 2025b), Gemini-3.0-Pro (Google DeepMind, 2025), Claude-Opus-4.5 (Anthropic, 2025), and Seed-2.0-Pro (ByteDance Seed Team, 2026), alongside open-weight models DeepSeek-V3.2 (DeepSeek-AI, 2025) and GLM-4.7 (Zhipu AI, 2025); and (2) *search agents*, which serve as direct, comparable-scale benchmarks at 30B parameters, including MiroThinker series (Team et al., 2026), REDSearcher (Zheng et al., 2026), IterResearch (Chen et al., 2026), AgentFold (Ye et al., 2025), Tongyi-DeepResearch (Team et al., 2025), and OpenSeeker (Du et al., 2026b). This diverse baseline set covers representative paradigms in contemporary agentic search. All baseline results are sourced from official publications or publicly available evaluation platforms.

Table 1: Main results. **LongSeeker**, trained under the **Context-ReAct** paradigm, outperforms GPT-5 and Gemini-3.0-Pro on BrowseComp despite having only 30B parameters, highlighting the effectiveness and potential of Context-ReAct. Scores marked with * denote ReAct-based agents without context management on BrowseComp and BrowseComp-ZH, while "–" indicates unavailable or unknown results.

Model	Param	Training	BrowseComp	BrowseComp-ZH	xbench-2505	GAIA-text
			Foundation Model with Tools			
GPT-5	-	-	54.9	63.0	77.8	76.4
Gemini-3.0-Pro	-	-	59.2	66.8	-	74.8
Claude-Opus-4.5	-	-	67.8	62.4*	-	71.5
Seed-2.0-Pro	-	-	77.3	82.4	-	78.6
DeepSeek-V3.2	671B	-	67.6	65.0*	78.0	75.1
GLM-4.7	358B	-	67.5	66.6*	72.0	-
			Search Agent			
MiroThinker-1.7-mini	30B	CPT + SFT + RL	67.9	72.3	-	80.3
MiroThinker-1.5-mini	30B	CPT + SFT + RL	56.1	66.8	73.1	72.0
RedSearcher	30B	CPT + SFT + RL	57.4	58.2	-	80.1
IterResearch	30B	SFT + RL	37.3	45.2	71.0	72.8
AgentFold	30B	SFT	36.2	47.3	-	67.0
Tongyi-DeepResearch	30B	CPT + SFT + RL	43.4*	46.7*	75.0	70.9
OpenSeeker	30B	SFT	29.5*	48.4*	74.0	-
LongSeeker	30B	SFT	**61.5**	**62.5**	**78.0**	**77.7**

Table 2: Controlled comparison of Context-ReAct supervision. Both models use the same 10K training questions, teacher model, base model, and SFT configuration, and differ only in the agent scaffold and the resulting supervision trajectories. BrowseComp and BrowseComp-ZH are evaluated without discard-all.

Method	BrowseComp	BrowseComp-ZH	xbench-2505	GAIA-text
ReAct-SFT	43.5	56.0	78.0	74.5
Context-ReAct-SFT	**56.0**	**57.3**	78.0	**77.7**

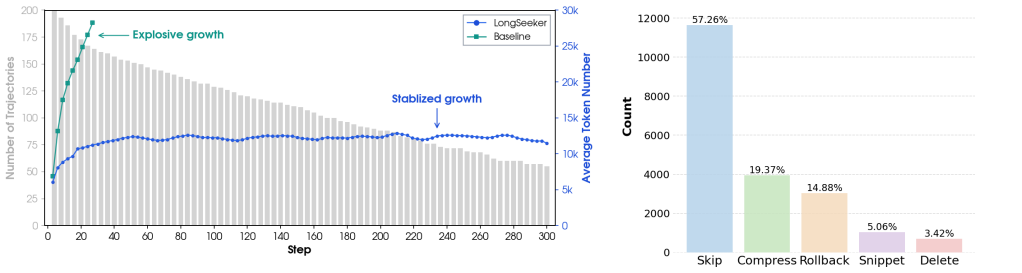
4.2 Results and Analysis

Main results. Table 1 presents the primary evaluation results. **LongSeeker** achieves **61.5** on BrowseComp and **62.5** on BrowseComp-ZH (56.0 and 57.3 without discard-all, respectively), establishing strong performance among 30B-scale open-source search agents. Notably, LongSeeker exceeds MiroThinker-1.5-mini (56.1), Tongyi-DeepResearch (43.4), IterResearch (37.3), AgentFold (36.2), and OpenSeeker-v1 (29.5) on BrowseComp. Extending evaluation to xbench and GAIA, LongSeeker achieves 78.0 on xbench-2505 and 77.7 on GAIA-text. These scores confirm that the benefits of Elastic Context Orchestration generalize beyond purely information-seeking tasks to broader agent capabilities.

Controlled Comparison with ReAct-SFT. To evaluate the effect of Context-ReAct supervision, we compare LongSeeker with a ReAct-SFT baseline. Both models use the same 10K training questions, DeepSeek-V3.2 teacher, Qwen3-30B-A3B base model, and SFT configuration; they differ only in the agent scaffold and the resulting supervision trajectories.

As shown in Table 2, Context-ReAct improves BrowseComp from 43.5 to 56.0 ($+12.5$), BrowseComp-ZH from 56.0 to 57.3 ($+1.3$), and GAIA-text from 74.5 to 77.7 ($+3.2$). These results demonstrate that Context-ReAct supervision provides clear performance benefits beyond those obtained from the same training questions, teacher model, base model, and SFT configuration.

Token Efficiency and Context Growth. To empirically evaluate token efficiency, we compare Context-ReAct-SFT with the matched ReAct-SFT baseline on 200 questions sampled from BrowseComp. ReAct-SFT averages 63.65 tool calls and 3,180.6K total tokens per trajectory, whereas Context-ReAct-SFT averages 132.50 tool calls and 1,799.8K total tokens. Although Context-ReAct-SFT performs more tool calls, it reduces total token consumption by 43.4% while achieving higher



(a) Context Growth Dynamics of LongSeeker (b) Meta-operation distribution of LongSeeker

Figure 4: Analysis of LongSeeker's context management on 200 trajectories sampled from BrowseComp. (a) Context-ReAct-SFT maintains a stable and well-bounded active context length over long horizons, while the matched ReAct-SFT baseline grows continuously as observations accumulate. (b) LongSeeker learns through training to utilize all five meta-operations and effectively invoke and compose them to solve long-horizon tasks.

accuracy, showing that its context management enables longer searches with lower overall token consumption.

As depicted in Figure 4a, ReAct-SFT passively appends observations and therefore processes an increasingly long context at each turn, whereas **LongSeeker** maintains a stable and concise working memory, with its active context length plateauing below 15k tokens even at 300 steps. This stabilized growth results from our *fine-grained* meta-operations: the model can purge failed branches (ROLLBACK), discard irrelevant retrievals (DELETE), retain essential evidence (SNIPPET), and abstract verbose history (COMPRESS). Consequently, the working context remains compact and information-dense, leaving substantial headroom for longer and more complex tasks.

Meta-operation usage. Figure 4b shows the usage distribution of the five meta-operations of LongSeeker. We observe that LongSeeker effectively leverages the full set of atomic operations—*Skip*, *Compress*, *Rollback*, *Snippet*, and *Delete*—across trajectories, suggesting that it has acquired a robust strategy for invoking and composing meta-operation to handle complex long-horizon tasks. This ability to coherently and purposefully navigate the context-manipulation space contributes to its strong overall performance.

We also observe a mild imbalance, where *Snippet* and *Delete* are used less frequently. This likely stems from the nature of long-horizon search: early in the process, it is difficult to confidently identify irrelevant information, so LongSeeker tends to preserve more context and adopts a conservative pruning strategy.

4.3 ABLATION STUDIES

Comparison of Context Management Strategies. To evaluate the effectiveness of the Context-ReAct paradigm, we conduct controlled experiments on BrowseComp under a unified setup, where all methods are built upon the same base model (DeepSeek-V3.2) and share identical configurations. We compare Context-ReAct with three context management strategies: (1) *Summary*, which compresses the overflowed trajectory into a summary and resumes the rollout from the condensed context; (2) *Discard-all*, which resets the context by removing all previous tool-call history, similar to reinitializing with a fresh context; and (3) *Compress*, an AgentFold-style strategy that manages context using *Compress* alone. Due to computational constraints, we do not synthesize a separate Compress-only training corpus and fine-tune an additional model. As shown in Figure 5, Context-ReAct achieves the best overall performance, demonstrating its effectiveness in long-horizon tasks through elastic context control that enables adaptive management.

5 CONCLUSION

We propose **Context-ReAct**, a general agentic paradigm for elastic context orchestration that enables agents to jointly generate reasoning, context meta-operations, and tool calls at each step. Context-

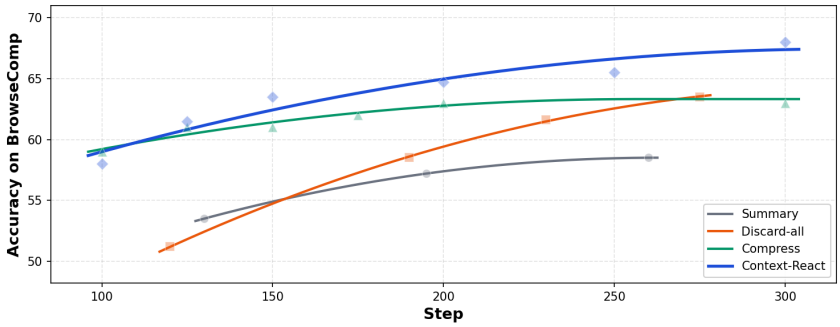


Figure 5: Effectiveness of the Context-ReAct paradigm on BrowseComp compared to other context management strategies. Context-ReAct achieves the best performance under the same step budget.

ReAct defines five atomic operations—*Skip*, *Compress*, *Rollback*, *Snippet* and *Delete*. Together, these operations provide a *fine-grained* mechanism for multi-resolution control over the evolving working context. We train **LongSeeker-30B** based on this paradigm and demonstrate competitive performance on long-horizon search benchmarks, notably surpassing Tongyi DeepResearch and AgentFold on BrowseComp. Our experiments empirically validate that the application of these atomic tools yields significantly more compact and efficient context management compared to append-only or coarse-grained truncation approaches, enabling sustained high performance at long horizons with strong potential for harder tasks.

Future work. Our current implementation leverages SFT on synthesized trajectories without rejection sampling or advanced exploration strategies. One direction involves applying RL to optimize meta-operation usage, enabling agents to explore the action spaces. Furthermore, we envision Context-ReAct as a foundational architectural paradigm rather than a search-specific solution. Its core philosophy of state-dependent fidelity is inherently domain-agnostic, offering a scalable blueprint for other long-horizon challenges such as autonomous software engineering, large-scale legal discovery, and multi-modal scientific reasoning, where the ability to fluidly restructure massive working contexts is critical.

AI USE STATEMENT

In this work, we used generative AI tools for language polishing, assistance with LaTeX and code debugging. We have not used generative AI tools for generating experimental results, data, or unverified citations, and other required disclosure tasks are not applicable to this work. Additionally, we used generative AI tools for improving the clarity and organization of the manuscript. We have reviewed all AI-assisted work. We manually verified relevant literature and citations, and checked AI-assisted code and technical suggestions for correctness. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This work does not involve human subjects, sensitive data, or other significant ethical concerns. We are not aware of any conflicts of interest, privacy, security, or fairness issues arising from this work.

REPRODUCIBILITY STATEMENT

We provide sufficient details to reproduce our results, including the model and training configurations, implementation details, evaluation settings, and data processing procedures in the main paper and appendix. Our code will be publicly released to further support reproducibility.

REFERENCES

Anthropic. Claude opus 4.5 system card, 2025. URL https://www.anthropic.com/claude-opus-4-5-system-card.

ByteDance Seed Team. Seed2.0 model card: Towards intelligence frontier for real-world complexity, 2026. URL https://lf3-static.bytednsdoc.com/obj/eden-cn/lapzild-tss/ljhwZthlaukjlkulzlp/seed2/0214/Seed2.0%20Model%20Card.pdf.

Guoxin Chen, Zile Qiao, Xuanzhong Chen, Donglei Yu, Haotian Xu, Wayne Xin Zhao, Ruihua Song, Wenbiao Yin, Huifeng Yin, Liwen Zhang, Kuan Li, Minpeng Liao, Yong Jiang, Pengjun Xie, Fei Huang, and Jingren Zhou. Iterresearch: Rethinking long-horizon agents with interaction scaling. *arXiv preprint arXiv:2511.07327*, 2026.

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

DeepSeek-AI. Deepseek-v3.2: Pushing the frontier of open large language models, 2025. URL https://arxiv.org/abs/2512.02556.

Yuwen Du, Rui Ye, Shuo Tang, Keduan Huang, Xinyu Zhu, Yuzhu Cai, and Siheng Chen. Openseeker-v2: Pushing the limits of search agents with informative and high-difficulty trajectories. *arXiv preprint arXiv:2605.04036*, 2026a.

Yuwen Du, Rui Ye, Shuo Tang, Xinyu Zhu, Yijun Lu, Yuzhu Cai, and Siheng Chen. Openseeker: Democratizing frontier search agents by fully open-sourcing training data. *arXiv preprint arXiv:2603.15594*, 2026b.

Google DeepMind. Model evaluation - approach, methodology & results, gemini 3 pro, 2025. URL https://storage.googleapis.com/deepmind-media/gemini/gemini_3_pro_model_evaluation.pdf.

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. Websailor: Navigating super-human reasoning for web agent, 2025. URL https://arxiv.org/abs/2507.02592.

Miao Lu, Weiwei Sun, Weihua Du, Zhan Ling, Xuesong Yao, Kang Liu, and Jiecao Chen. Scaling llm multi-turn rl with end-to-end summarization-based context management, 2025. URL https://arxiv.org/abs/2510.06727.

Grégoire Mialon, Clémentine Fourrier, Craig Swift, Thomas Wolf, Yann LeCun, and Thomas Scialom. Gaia: a benchmark for general ai assistants, 2023.

OpenAI. Deep research system card, February 2025a. URL https://cdn.openai.com/deep-research-system-card.pdf.

OpenAI. Gpt-5 system card, 2025b. URL https://arxiv.org/abs/2601.03267.

Zijing Shi, Meng Fang, and Ling Chen. Monte carlo planning with large language model for text-based game agents, 2025. URL https://arxiv.org/abs/2504.16855.

MiroMind Team, S. Bai, L. Bing, L. Lei, R. Li, X. Li, X. Lin, E. Min, L. Su, B. Wang, L. Wang, L. Wang, S. Wang, X. Wang, Y. Zhang, Z. Zhang, G. Chen, L. Chen, Z. Cheng, Y. Deng, Z. Huang, D. Ng, J. Ni, Q. Ren, X. Tang, B.L. Wang, H. Wang, N. Wang, C. Wei, Q. Wu, J. Xia, Y. Xiao, H. Xu, X. Xu, C. Xue, Z. Yang, Z. Yang, F. Ye, H. Ye, J. Yu, C. Zhang, W. Zhang, H. Zhao, and P. Zhu. Mirothinker-1.7 & h1: Towards heavy-duty research agents via verification, 2026. URL https://arxiv.org/abs/2603.15726.

Tongyi DeepResearch Team, Baixuan Li, Bo Zhang, Dingchu Zhang, Fei Huang, Guangyu Li, Guoxin Chen, Huifeng Yin, Jialong Wu, Jingren Zhou, Kuan Li, Liangcai Su, Litu Ou, Liwen Zhang, Pengjun Xie, Rui Ye, Wenbiao Yin, Xinmiao Yu, Xinyu Wang, Xixi Wu, Xuanzhong Chen, Yida Zhao, Zhen Zhang, Zhengwei Tao, Zhongwang Zhang, Zile Qiao, Chenxi Wang, Donglei Yu, Gang Fu, Haiyang Shen, Jiayin Yang, Jun Lin, Junkai Zhang, Kui Zeng, Li Yang, Hailong Yin, Maojia Song, Ming Yan, Minpeng Liao, Peng Xia, Qian Xiao, Rui Min, Ruixue Ding, Runnan Fang, Shaowei Chen, Shen Huang, Shihang Wang, Shihao Cai, Weizhou Shen, Xiaobin Wang, Xin Guan, Xinyu Geng, Yingcheng Shi, Yuning Wu, Zhuo Chen, Zijian Li, and Yong Jiang. Tongyi deepresearch technical report, 2025. URL https://arxiv.org/abs/2510.24701.

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

An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, Chujie Zheng, Dayiheng Liu, Fan Zhou, Fei Huang, Feng Hu, Hao Ge, Haoran Wei, Huan Lin, Jialong Tang, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Yang, Jiaxi Yang, Jing Zhou, Jingren Zhou, Junyang Lin, Kai Dang, Keqin Bao, Kexin Yang, Le Yu, Lianghao Deng, Mei Li, Mingfeng Xue, Mingze Li, Pei Zhang, Peng Wang, Qin Zhu, Rui Men, Ruize Gao, Shixuan Liu, Shuang Luo, Tianhao Li, Tianyi Tang, Wenbiao Yin, Xingzhang Ren, Xinyu Wang, Xinyu Zhang, Xuancheng Ren, Yang Fan, Yang Su, Yichang Zhang, Yinger Zhang, Yu Wan, Yuqiong Liu, Zekun Wang, Zeyu Cui, Zhenru Zhang, Zhipeng Zhou, and Zihan Qiu. Qwen3 technical report, 2025. URL https://arxiv.org/abs/2505.09388.

Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik Narasimhan, and Yuan Cao. React: Synergizing reasoning and acting in language models, 2023. URL https://arxiv.org/abs/2210.03629.

Yilun Yao, Shan Huang, Elsie Dai, Zhewen Tan, Zhenyu Duan, Shousheng Jia, Yanbing Jiang, and Tong Yang. Arc: Active and reflection-driven context management for long-horizon information seeking agents, 2026. URL https://arxiv.org/abs/2601.12030.

Rui Ye, Zhongwang Zhang, Kuan Li, Huifeng Yin, Zhengwei Tao, Yida Zhao, Liangcai Su, Liwen Zhang, Zile Qiao, Xinyu Wang, Pengjun Xie, Fei Huang, Siheng Chen, Jingren Zhou, and Yong Jiang. Agentfold: Long-horizon web agents with proactive context management, 2025. URL https://arxiv.org/abs/2510.24699.

Hongli Yu, Tinghong Chen, Jiangtao Feng, Jiangjie Chen, Weinan Dai, Qiying Yu, Ya-Qin Zhang, Wei-Ying Ma, Jingjing Liu, Mingxuan Wang, and Hao Zhou. Memagent: Reshaping long-context llm with multi-conv rl-based memory agent, 2025. URL https://arxiv.org/abs/2507.02259.

Chu Zheng, Xiao Wang, Jack Hong, Huiming Fan, Yuqi Huang, Yue Yang, Guohai Xu, Chenxiao Zhao, Cheng Xiang, Shengchao Hu, Dongdong Kuang, Ming Liu, Bing Qin, and Xing Yu. Redsearcher: A scalable and cost-efficient framework for long-horizon search agents. *arXiv preprint arXiv:2602.14234*, 2026.

Zhipu AI. Glm-4.7: Advancing the coding capability, December 2025. URL https://z.ai/blog/glm-4.7.

Peilin Zhou, Bruce Leon, Xiang Ying, Can Zhang, Yifan Shao, Qichen Ye, Dading Chong, Zhiling Jin, Chenxuan Xie, Meng Cao, Yuxin Gu, Sixin Hong, Jing Ren, Jian Chen, Chao Liu, and Yining Hua. Browsecomp-zh: Benchmarking web browsing ability of large language models in chinese, 2025a. URL https://arxiv.org/abs/2504.19314.

Zijian Zhou, Ao Qu, Zhaoxuan Wu, Sunghwan Kim, Alok Prakash, Daniela Rus, Jinhua Zhao, Bryan Kian Hsiang Low, and Paul Pu Liang. Mem1: Learning to synergize memory and reasoning for efficient long-horizon agents, 2025b. URL https://arxiv.org/abs/2506.15841.

A Training Details

We implement SFT using Slime and initialize the model from Qwen3-30B-A3B. We train on the 10,000 synthesized Context-ReAct trajectories for 3 epochs with a global batch size of 64. The optimizer is Adam with a learning rate of 5×10^{-5}, cosine decay, a warmup ratio of 0.1, a minimum learning rate of 1×10^{-6}, and weight decay of 0.1. We use a tensor model parallel size of 2 and a context parallel size of 8 for long-context training. Tool responses are masked from the training loss. Training is conducted on 16 NVIDIA H200 GPUs.

B Evaluation Details

Due to computational constraints, we randomly sample 400 questions from the full BrowseComp benchmark using random seed 42. All other evaluation settings follow Section 4.1. Final-answer correctness is judged by DeepSeek-V3.2 using the following prompt.

Prompt for Final Answer Evaluation

System:
You are an evaluator. Your ONLY task is to decide: does the Predicted Answer refer to the SAME entity or value as the Ground Truth?

User:
Question: {query}
Ground Truth: {gt}
Predicted Answer: {pred}

Output only a JSON object with this format:
`{"correct": true/false, "reason": "brief explanation"}`

C Case Study

Figure 6 shows the managed context after applying meta-operations, and Figure 7 shows the corresponding structured output from **LongSeeker**.

Managed Context

Question
Give me the first and last name of the first Prime Minister of this country by identifying the country based on the following clues, as of 2023:
- it is a landlocked country and two of its neighbors are also landlocked countries
- its literacy rate in 2022 increased by 1.16% as compared to 2014
- it is home to a species from the Family Chamaeleonidae named after the individual who collected its holotype.

Previous Steps
[Compressed Step 1~5]
We performed initial searches to understand the clues. We identified that there are 44 landlocked countries, and many have at least two landlocked neighbors. We searched for chameleon species named after their holotype collector and found examples like Trioceros jacksonii (Jackson's chameleon) named after F.J. Jackson, but this species is from Kenya and Tanzania (not landlocked). We also searched for literacy rate increase of 1.16% between 2014 and 2022, but did not find a specific country yet. We need to narrow down by finding a landlocked country that hosts a chameleon species named after its holotype collector and has a literacy rate increase of exactly 1.16% from 2014 to 2022.

[Rollback to Step 6]
The previous steps identified Uganda as a candidate based on the chameleon clue, but the literacy rate increase does not match (12.08% vs 1.16%). We must consider other landlocked countries with at least two landlocked neighbors and verify the chameleon species and literacy rate clues more carefully. We also need to ensure we are using the correct literacy rate metric (possibly youth literacy or adult literacy from a specific source). We will start by listing all landlocked countries, then filter by those with at least two landlocked neighbors, then search for chameleon species named after their holotype collector in those countries, and finally check literacy rate changes.

[Compressed Step 7~13]
We attempted to find the country by searching for the precise literacy rate increase of 1.16% from 2014 to 2022, but the search results did not directly yield a country. We also searched for chameleon species named after their holotype collector, finding Jackson's chameleon (Trioceros jacksonii) named after F.J. Jackson, but it is found in Kenya and Tanzania (not landlocked). We also considered Johnston's chameleon (Trioceros johnstoni) but need to confirm if it is named after the holotype collector. We then listed landlocked African countries with at least two landlocked neighbors: Botswana, Burkina Faso, Central African Republic, Chad, Mali, Niger, Rwanda, South Sudan, Uganda, Zambia, Zimbabwe. We need to verify the literacy rate increase and chameleon clue for each candidate.

[Deleted Step 14]
[Step deleted]

[Compressed Step 15]
We searched for literacy rate data (2014 and 2022) for each candidate landlocked African country with at least two landlocked neighbors. The results show that for most countries, the increase from 2014 to 2022 is not 1.16%. However, Zimbabwe's literacy rate increased from 88.7% in 2014 to 89.85% in 2022, a rise of 1.15% (very close to 1.16%, possibly rounding). Other candidates like Uganda, Rwanda, Zambia, etc., show increases far above 1.16%. Thus, Zimbabwe emerges as a strong candidate based on the literacy rate clue.

[Snippeted Step 16]
Motivation
We have identified Zimbabwe as a candidate based on its ~1.16% literacy rate increase. We have already confirmed it's landlocked with at least two landlocked neighbors (Zambia, Botswana). The remaining clue is a Chamaeleonidae species found in Zimbabwe and named after its holotype collector. We will search for that species, then identify Zimbabwe's first Prime Minister (likely post-independence).
Standard tool call
{"name": "search_web", "arguments": {"query": ["chameleon species Zimbabwe named after collector", "Zimbabwe chameleon holotype collector", "Chamaeleonidae Zimbabwe named after", "chameleon named after collector found in Zimbabwe"]}}
Tool response
3. [Marshall's pygmy chameleon – Wikipedia] The specific name, marshalli, is in honor of British entomologist Guy Anstruther Knox Marshall, who collected the holotype. ... Zimbabwe and the adjacent upland ...
4. [Chamaeleo dilepis - The Reptile Database - Reptarium] Taxonomic database that provides basic information about all living reptile species, such as turtles, snakes, lizards, and crocodiles, as well as tuataras ...

Figure 6: Complete case study showing managed context at a reasoning step. The trajectory demonstrates the combined effect of COMPRESS, ROLLBACK, DELETE, and SNIPPET operations.

Output

<think>
We have identified Zimbabwe as a candidate based on literacy rate increase (approx 1.15-1.16%). Now we need to verify the chameleon clue: "it is home to a species from the Family Chamaeleonidae named after the individual who collected its holotype." We need to find a chameleon species found in Zimbabwe that is named after the person who collected its holotype. The search results mention "Marshall's pygmy chameleon" (Rhampholeon marshalli) named after Guy Anstruther Knox Marshall, who collected the holotype. However, we need to confirm that this species is found in Zimbabwe. The Wikipedia page says it is found in Zimbabwe and Mozambique. That seems promising. But we should verify that the species is indeed from the Family Chamaeleonidae (it is a chameleon). Also, we need to ensure that the naming is indeed after the collector of the holotype. The Wikipedia snippet says "The specific name, marshalli, is in honor of British entomologist Guy Anstruther Knox Marshall, who collected the holotype." That matches the clue.

We also need to confirm the literacy rate increase exactly. In step 14, we found Zimbabwe's literacy rate increased from 88.7% in 2014 to 89.85% in 2022, which is a 1.15% increase. The clue says 1.16% increase. Possibly rounding difference. We should verify the exact numbers from a reliable source.

Additionally, we need to identify the first Prime Minister of Zimbabwe. The question says \"the first Prime Minister of this country by identifying the country based on the following clues, as of 2023.\" That likely means the first Prime Minister after independence (or the first in the country's history). Zimbabwe's first Prime Minister was Robert Mugabe (took office 1980). However, there is also the position of Prime Minister of Southern Rhodesia (before independence). But the question says \"as of 2023\", meaning we are considering the country as it is now. The first Prime Minister of independent Zimbabwe is Robert Mugabe. But we need to confirm.

We should also verify that Zimbabwe is indeed landlocked and has at least two landlocked neighbors (Zambia and Botswana). That is correct.

Thus, we need to gather confirming evidence for each clue. Let's visit relevant pages.

First, visit the Wikipedia page for Marshall's pygmy chameleon to confirm the naming and distribution.

Second, visit a reliable source for Zimbabwe literacy rates (e.g., World Bank, CIA World Factbook, or other statistical databases) to confirm the 1.16% increase.

Third, confirm the first Prime Minister of Zimbabwe.
</think>

<meta_tool_call>
[{"name": "skip", "arguments": {}}]
</meta_tool_call>

<motivation>
We have identified Zimbabwe as a strong candidate based on the literacy rate increase (approximately 1.15-1.16%) and the fact that it is landlocked with at least two landlocked neighbors. Now we need to verify the remaining clue about the chameleon species named after its holotype collector. The search results suggest Marshall's pygmy chameleon (Rhampholeon marshalli) is named after collector Guy Anstruther Knox Marshall and is found in Zimbabwe. We should visit the Wikipedia page to confirm this. Additionally, we need to verify the literacy rate increase with precise data and confirm the first Prime Minister of Zimbabwe. We'll use visit_web to gather this information.
</motivation>

<standard_tool_call>
{"name": "visit_web", "arguments": {"url": [omitted due to length], "goal": "1. Confirm that Marshall's pygmy chameleon is named after the collector of its holotype and is found in Zimbabwe. 2. Find literacy rate data for Zimbabwe in 2014 and 2022 to verify 1.16% increase. 3. Identify the first Prime Minister of Zimbabwe."}}
</standard_tool_call>

Figure 7: Complete structured output from **LongSeeker**, including reasoning, meta-tool calls, motivation, and standard tool call.