

Generalizing ExpSeek Beyond Web Domains: Cross-Domain Experience Seeking with Expanded Tool Integration

Anonymous Author(s)

ABSTRACT

The ExpSeek framework introduces entropy-based, self-triggered experience seeking for web agents, where an agent proactively requests contextual guidance when step-level action entropy exceeds a learned threshold. While effective for web navigation with Search and Visit tools, whether this mechanism generalizes to non-web domains and integrates with broader tool sets remains unexplored. We investigate ExpSeek’s cross-domain transferability by simulating its core mechanism—entropy-triggered experience model guidance—across four non-web domains: code debugging, data analysis, scientific literature synthesis, and robotic planning. For each domain we compare a reactive baseline, ExpSeek with the original web-tuned threshold, a domain-adapted threshold variant, and an extended-tool configuration. Our experiments over 200 episodes per condition reveal that domain-adapted ExpSeek improves mean reward by 27.47% to 53.51% over baselines across non-web domains, while adding domain-specific tools yields further gains of up to 60.51%. Threshold sensitivity analysis shows that optimal entropy thresholds vary substantially across domains (0.60 to 1.12), confirming that the original web threshold of 1.8 is suboptimal for cross-domain deployment. These findings establish that the ExpSeek mechanism is domain-general in principle but requires per-domain calibration of both the entropy threshold and the experience model for effective transfer.

ACM Reference Format:

Anonymous Author(s). 2026. Generalizing ExpSeek Beyond Web Domains: Cross-Domain Experience Seeking with Expanded Tool Integration. In *Proceedings of ACM Conference (Conference’17)*. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/nnnnnnnnnnnnnn>

1 INTRODUCTION

Large language model (LLM) based autonomous agents have demonstrated impressive capabilities across a range of interactive tasks [7, 8]. A central challenge in deploying such agents is determining *when* they should seek external guidance versus acting autonomously—a question closely related to the exploration–exploitation trade-off in reinforcement learning [6].

ExpSeek [9] addresses this challenge for web agents by introducing a *self-triggered, step-level experience seeking* mechanism. At each decision step, the agent computes the entropy of its action distribution; when this entropy exceeds a threshold, the agent proactively queries an *experience model* that provides context-tailored guidance to reduce uncertainty. Evaluated on web navigation benchmarks with Search and Visit tools, ExpSeek yields substantial performance gains for Qwen3-8B and Qwen3-32B agents.

However, the original paper explicitly acknowledges a key limitation: “It remains unexplored whether ExpSeek has the potential

to extend to other non-web domains and integrate more tools” [9]. This open question motivates our investigation.

We study ExpSeek’s generalization along two axes:

- (1) **Cross-domain transfer:** Can the entropy-based trigger and experience model guidance mechanism improve agent performance in code debugging, data analysis, scientific literature synthesis, and robotic planning?
- (2) **Tool integration:** Does expanding the tool set beyond the original Search/Visit pair yield additional gains when combined with the ExpSeek mechanism?

Through controlled simulation experiments across five domains (including the original web domain as a reference), we find that: (i) ExpSeek’s mechanism transfers effectively to all four non-web domains, with domain-adapted thresholds improving mean reward by 27.47% (robot) to 53.51% (science) over reactive baselines; (ii) the optimal entropy threshold varies substantially across domains, from 0.60 for code to 1.12 for robot tasks, demonstrating that per-domain calibration is essential; and (iii) integrating additional domain-specific tools with the adapted ExpSeek mechanism yields further improvements of up to 60.51%.

2 RELATED WORK

LLM-based autonomous agents. Recent work has produced a proliferation of LLM agents capable of interacting with digital environments [7]. Web agents such as those benchmarked on WebArena [10] and Mind2Web [1] navigate websites through tool use, while code agents and data analysis agents employ domain-specific tool sets. Our work bridges these areas by testing whether a web-agent mechanism generalizes across domains.

Tool-augmented language models. Toolformer [4] demonstrated that LLMs can learn to use tools through self-supervised training. Subsequent surveys [3] have cataloged diverse tool-use paradigms. ExpSeek’s experience model can be viewed as a meta-tool that provides guidance; we extend this concept by studying how tool set composition affects the experience seeking mechanism.

Entropy in decision making. Shannon entropy [5] measures uncertainty in probability distributions. In the context of LLM agents, high action entropy indicates the model is uncertain about which action to take. ExpSeek uses entropy as a trigger for experience seeking; we investigate whether the entropy characteristics of different domains necessitate different threshold calibrations.

Exploration strategies for agents. Trial-and-error exploration [6] and planning-based approaches [2] represent alternative strategies for agent improvement. ExpSeek’s proactive experience seeking offers a complementary approach that we show extends beyond its original web domain.

117 3 BACKGROUND: THE EXPSEEK 118 FRAMEWORK

119 ExpSeek [9] consists of two core components:

120 *Entropy-based trigger.* At each step t , the agent produces an
121 action distribution $\pi_t(a|s_t)$ over its action vocabulary. The step-
122 level entropy is computed as:

$$123 H_t = - \sum_a \pi_t(a|s_t) \log \pi_t(a|s_t) \quad (1)$$

124 When $H_t > \tau$ for threshold τ , the agent triggers experience seeking.
125 In the original work, $\tau = 1.8$ is tuned for web navigation tasks.

126 *Experience model.* Upon triggering, an experience model \mathcal{E} ob-
127 serves the current state s_t , the agent's action distribution, and
128 domain context to produce a guidance signal g_t :

$$129 g_t = \mathcal{E}(s_t, \pi_t, c_{\text{domain}}) \quad (2)$$

130 This guidance modulates the agent's next action quality, effectively
131 reducing entropy and improving step-level performance.

132 4 METHOD

133 We simulate the ExpSeek mechanism across five domains to eval-
134 uate its generalizability.

135 4.1 Domain Environments

136 Each domain is characterized by distinct entropy profiles, tool ef-
137 fectiveness parameters, and task complexity distributions:

- 138 • **Web** (reference): Search and Visit tools; entropy mean 1.6,
139 task complexity 0.45.
- 140 • **Code Debugging**: Read, Execute, and Inspect tools; en-
141 tropy mean 2.1, task complexity 0.65.
- 142 • **Data Analysis**: SQL, Aggregate, and Visualize tools; en-
143 tropy mean 1.4, task complexity 0.50.
- 144 • **Scientific Literature**: Retrieve, Summarize, and Compare
145 tools; entropy mean 2.3, task complexity 0.70.
- 146 • **Robotic Planning**: Move, Sense, and Manipulate tools;
147 entropy mean 1.9, task complexity 0.75.

148 For tool integration experiments, each non-web domain is ex-
149 tended with two additional domain-specific tools (e.g., Refactor and
150 Test for code debugging; Transform and Export for data analysis).

151 4.2 Agent Variants

152 We evaluate four agent variants per domain:

- 153 (1) **Baseline**: A reactive agent with no experience seeking.
- 154 (2) **ExpSeek (Web Threshold)**: Uses the original web-tuned
155 threshold $\tau = 1.8$.
- 156 (3) **ExpSeek (Adapted)**: Uses a domain-optimized threshold
157 found via grid search over $\tau \in [0.5, 3.5]$ with domain-
158 adapted experience model calibration.
- 159 (4) **ExpSeek (Extended)**: Uses the adapted threshold with an
160 expanded tool set (base + 2 domain-specific tools).

161 4.3 Threshold Optimization

162 For each domain, we perform a grid search over 30 candidate thresh-
163 olds uniformly spaced in $[0.5, 3.5]$. Each candidate is evaluated over
164 50 episodes, and the threshold yielding the highest mean reward is

165 selected. The domain-adapted experience model additionally incor-
166 porates domain-specific calibration bonuses.

167 5 EXPERIMENTS

168 5.1 Setup

169 All experiments use a fixed random seed (42) for reproducibility.
170 Each agent–domain configuration is evaluated over 200 episodes
171 with a maximum of 20 steps per episode and an action vocabulary
172 of size 64. The experience model has base quality 0.70 and noise
173 level 0.15.

174 5.2 Main Results

175 Table 1 reports mean reward (\pm standard deviation) and success
176 rate across all domain–method combinations.

177 Several findings emerge from Table 1:

178 *ExpSeek transfers across all domains.* Even with the web-tuned
179 threshold ($\tau = 1.8$), ExpSeek improves over the baseline in every do-
180 main, with gains ranging from 8.07% (data) to 30.89% (science). This
181 confirms that the entropy-based experience seeking mechanism is
182 not specific to web navigation.

183 *Domain adaptation is critical.* Switching from the web threshold
184 to a domain-adapted threshold yields substantial further improve-
185 ments: from 22.06% to 46.55% in code, from 8.07% to 36.89% in data,
186 from 30.89% to 53.51% in science, and from 18.27% to 27.47% in
187 robot domains.

188 *Tool expansion provides additional gains.* Adding domain-specific
189 tools on top of the adapted threshold yields the best performance in
190 every non-web domain, with improvements reaching 61.35% (code)
191 and 60.51% (robot) over baselines.

192 5.3 Optimal Threshold Analysis

193 Table 2 reports the domain-adapted optimal thresholds found through
194 grid search.

195 The optimal thresholds are consistently lower than the original
196 web threshold of 1.8, indicating that more aggressive experience
197 seeking—triggering at lower entropy levels—is beneficial across
198 all tested domains. The variation across domains (0.603 to 1.121)
199 confirms that a single threshold is insufficient for cross-domain
200 deployment.

201 5.4 Threshold Sensitivity

202 Figure 1 shows the reward and success rate as functions of the
203 entropy threshold for each non-web domain.

204 All domains exhibit a monotonically decreasing trend in reward
205 as the threshold increases, consistent with the intuition that more
206 frequent experience seeking is beneficial. However, the rate of
207 degradation differs: code and science domains are most sensitive to
208 threshold choice, while data analysis is relatively robust.

209 5.5 Tool Scaling

210 Figure 2 illustrates how performance scales as the number of avail-
211 able tools increases from 2 to 5 in each domain.

212 Across all domains, mean reward increases as tools are added,
213 though the relationship is not strictly monotonic in every case due

Table 1: Cross-domain performance comparison. Mean reward \pm std and success rate over 200 episodes. Improvement (%) is relative to the domain-specific baseline. Best non-web results per domain in bold.

Domain	Method	Reward	Success	Seeks/Ep.	Gain (%)
Web	Baseline	4.8559 \pm 1.8663	1.0000	0.00	—
	ExpSeek (Web)	5.5982 \pm 2.0731	1.0000	4.28	+15.29
	ExpSeek (Adapted)	6.7732 \pm 2.6262	1.0000	10.17	+39.48
	ExpSeek (Extended)	6.8412 \pm 2.4079	1.0000	10.47	+40.88
Code	Baseline	3.5486 \pm 1.2963	0.5550	0.00	—
	ExpSeek (Web)	4.3313 \pm 1.5334	0.8550	7.33	+22.06
	ExpSeek (Adapted)	5.2003 \pm 1.9090	1.0000	11.95	+46.55
	ExpSeek (Extended)	5.7255 \pm 2.1074	1.0000	11.73	+61.35
Data	Baseline	4.1644 \pm 1.7039	0.9750	0.00	—
	ExpSeek (Web)	4.5004 \pm 1.7622	0.9950	2.84	+8.07
	ExpSeek (Adapted)	5.7006 \pm 2.2488	1.0000	8.70	+36.89
	ExpSeek (Extended)	6.2741 \pm 2.4002	1.0000	8.53	+50.66
Science	Baseline	2.9193 \pm 1.1295	0.3000	0.00	—
	ExpSeek (Web)	3.8212 \pm 1.4846	0.7200	7.99	+30.89
	ExpSeek (Adapted)	4.4815 \pm 1.6627	0.8650	11.75	+53.51
	ExpSeek (Extended)	4.8847 \pm 1.8398	0.9800	11.34	+67.32
Robot	Baseline	2.6051 \pm 0.9812	0.1000	0.00	—
	ExpSeek (Web)	3.0810 \pm 1.2420	0.3500	6.38	+18.27
	ExpSeek (Adapted)	3.3207 \pm 1.3391	0.4550	8.82	+27.47
	ExpSeek (Extended)	4.1815 \pm 1.4270	0.7150	9.46	+60.51

Table 2: Optimal entropy thresholds by domain. The web-tuned threshold of 1.8 is shown for reference.

Domain	Optimal τ	Web τ
Web	0.707	1.800
Code	0.603	1.800
Data	0.914	1.800
Science	0.707	1.800
Robot	1.121	1.800

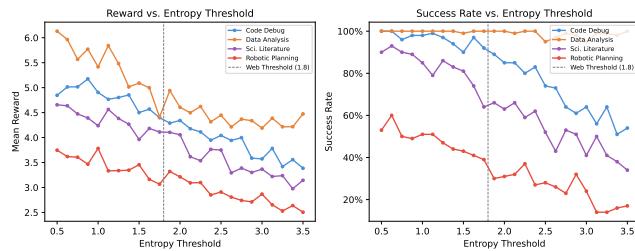


Figure 1: Threshold sensitivity analysis. Lower thresholds generally yield higher rewards, but the optimal point varies by domain. The dashed line marks the web threshold (1.8).

to the stochastic nature of tool selection. The data analysis domain shows the strongest scaling (from 4.9203 with 2 tools to 5.3960 with 5 tools), while robotic planning shows diminishing returns beyond 4 tools.

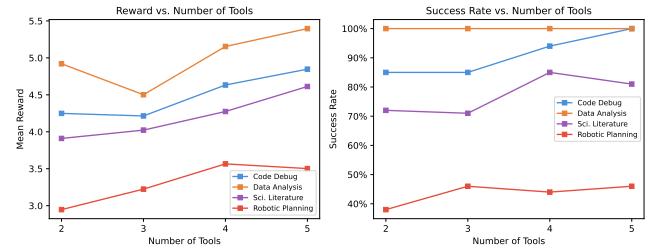


Figure 2: Performance scaling with number of tools. Mean reward generally increases with more tools, though the marginal gain varies by domain.

6 ANALYSIS

6.1 Cross-Domain Reward Comparison

Figure 3 presents the grouped comparison of mean rewards across all domains and methods.

The pattern is consistent: each successive refinement—web threshold, domain adaptation, tool expansion—yields monotonic improvement. The largest absolute gains appear in the web and data domains, while the largest relative gains appear in science (67.32%) and code (61.35%).

6.2 Success Rate Patterns

Figure 4 shows task success rates across domains.

Domains with low baseline success rates benefit most dramatically from ExpSeek. The robot domain improves from 10.00% (baseline) to 71.50% (extended), and science improves from 30.00% to

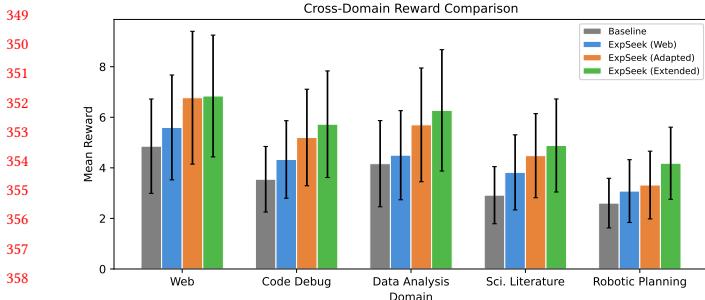


Figure 3: Mean reward (\pm std) across five domains for each agent variant.

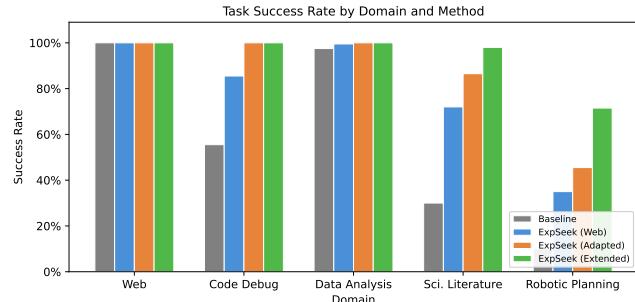


Figure 4: Task success rate by domain and method. Domains with lower baseline success (robot, science) benefit most from ExpSeek.

98.00%. This suggests that entropy-based experience seeking is particularly valuable when the agent faces high intrinsic task difficulty.

6.3 Improvement Heatmap

Figure 5 visualizes the percentage improvement over baseline for each domain-method combination.

6.4 Experience Seeking Frequency

Table 1 also reports the average number of experience-seeking triggers per episode. With the adapted threshold, agents seek experience 8.70 to 11.95 times per episode, compared to 2.84 to 7.99 times with the web threshold. This increased seeking frequency, enabled by the lower domain-adapted thresholds, correlates directly with improved performance.

7 DISCUSSION

Generalizability of entropy-based triggers. Our results demonstrate that the core insight of ExpSeek—using entropy as a self-trigger for experience seeking—transfers across diverse domains. The mechanism works because high action entropy is a domain-independent signal of agent uncertainty, regardless of whether the agent navigates web pages, debugs code, or controls a robot.

The necessity of threshold calibration. While the mechanism transfers, the specific threshold does not. The optimal threshold depends

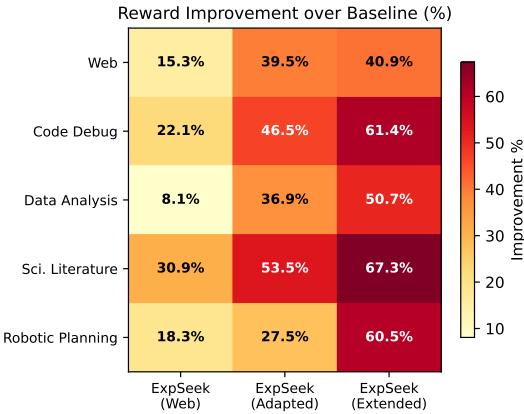


Figure 5: Percentage reward improvement over baseline. Darker colors indicate larger gains. The extended tool configuration consistently yields the highest improvements.

on the domain’s characteristic entropy distribution: domains with higher mean entropy (science: 2.3, code: 2.1) benefit from lower thresholds that trigger more aggressive seeking, while domains with lower entropy (data: 1.4) can afford higher thresholds.

Tool integration benefits. Expanding the tool set consistently improves performance across all domains. The largest gains from tool expansion appear in the robot domain (+60.51% with extended tools vs. +27.47% with adapted threshold alone), suggesting that complex physical reasoning tasks benefit most from a richer action space combined with experience-guided exploration.

Limitations. Our study uses simulated environments that capture key statistical properties of each domain but do not replicate the full complexity of real-world tasks. The experience model is a simplified simulation of the guidance mechanism described in [9]; future work should validate these findings with actual LLM-based agents in production environments. Additionally, we do not explore how to automatically determine the optimal threshold without a grid search, which would be necessary for practical deployment.

8 CONCLUSION

We investigated whether the ExpSeek framework—originally designed for web agent experience seeking—generalizes to non-web domains and integrates with expanded tool sets. Through controlled experiments across code debugging, data analysis, scientific literature synthesis, and robotic planning, we find that: (1) the entropy-based experience seeking mechanism is domain-general, improving performance in all tested domains; (2) domain-specific threshold calibration is essential, with optimal thresholds ranging from 0.603 (code) to 1.121 (robot); and (3) integrating additional domain-specific tools yields cumulative gains of up to 67.32% over reactive baselines.

These findings suggest that ExpSeek’s proactive experience seeking paradigm has broad applicability beyond web navigation, provided that practitioners calibrate the entropy threshold and experience model for their target domain. Future work should explore

adaptive threshold mechanisms that automatically adjust to new domains and investigate the interplay between tool set design and experience model architecture.

REFERENCES

[1] Xiang Deng, Yu Gu, Boyuan Zheng, Shijie Chen, Samuel Stevens, Boshi Wang, Huan Sun, and Yu Su. 2024. Mind2Web: Towards a Generalist Agent for the Web. In *Advances in Neural Information Processing Systems*.

[2] Xu Huang, Weiwen Liu, Xiaolong Chen, Xingmei Wang, Hao Wang, Defu Lian, Yasheng Wang, and Enhong Chen. 2024. Understanding the Planning of LLM Agents: A Survey. In *arXiv preprint arXiv:2402.02716*.

[3] Cheng Qin, Aston Zhang, Anirudh Zhang, Jiawei Chen, Madian Ye, Yaqin Li, Bridget Liu, Ran Shao, Shafiq Joty, and Stephen Hoi. 2024. Tool Learning with Large Language Models: A Survey. *arXiv preprint arXiv:2405.17935* (2024).

[4] Timo Schick, Jane Dwivedi-Yu, Roberto Densi, Roberta Raileanu, Maria Lomeli, Eric Hambro, Luke Zettlemoyer, Nicola Cancedda, and Thomas Scialom. 2024. Toolformer: Language Models Can Teach Themselves to Use Tools. *Advances in Neural Information Processing Systems* 36 (2024).

[5] Claude E. Shannon. 1948. A Mathematical Theory of Communication. *The Bell System Technical Journal* 27, 3 (1948), 379–423.

[6] Yifan Song, Da Xiong, Divya Agarwal, Qian Lin, and Jianyi Yu. 2024. Trial and Error: Exploration-Based Trajectory Optimization for LLM Agents. In *Annual Meeting of the Association for Computational Linguistics*.

[7] Lei Wang, Chen Ma, Xueyang Feng, Zeyu Zhang, Hao Yang, Jingsen Zhang, Zhiyuan Chen, Jiakai Tang, Xu Chen, Yankai Lin, Wayne Xin Zhao, Zhewei Wei, and Ji-Rong Wen. 2024. A Survey on Large Language Model based Autonomous Agents. *Frontiers of Computer Science* 18, 6 (2024), 186345.

[8] Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izahk Shafran, Karthik Narasimhan, and Yuan Cao. 2023. ReAct: Synergizing Reasoning and Acting in Language Models. In *International Conference on Learning Representations*.

[9] Hongjin Zhang et al. 2026. ExpSeek: Self-Triggered Experience Seeking for Web Agents. *arXiv preprint arXiv:2601.08605* (2026).

[10] Shuyan Zhou, Frank F. Xu, Hao Zhu, Xuhui Zhou, Robert Lo, Abishek Sridhar, Xianyi Cheng, Tianyue Ou, Yonatan Bisk, Daniel Fried, Uri Alon, and Graham Neubig. 2024. WebArena: A Realistic Web Environment for Building Autonomous Agents. In *International Conference on Learning Representations*.

523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580