Friction or Breakdown? Rethinking the Timing of GenAI Intervention in Knowledge Work

Zhitong Guan[1,*]

[1]University of Texas at Austin, School of Information

Abstract

Generative AI assistants are increasingly moving from reactive tools toward proactive collaborators that intervene during ongoing knowledge work. Yet timely intervention is not simply a matter of detecting user difficulty. In knowledge work, difficulty, hesitation, uncertainty, and reformulation are often the conditions through which users reflect, evaluate, synthesize, and create. This position paper argues that proactive GenAI assistance should be designed and evaluated through the distinction between productive cognitive friction and unproductive cognitive breakdown. Drawing on research in cognitive augmentation, information seeking and use, and proactive AI assistance, I argue that common timing signals, such as pauses, reformulations, low output volume, and rapid acceptance, are ambiguous unless interpreted in relation to task stage, artifact development, and users' evolving cognitive trajectories. Intervention timing should therefore be evaluated through both immediate interaction effects and downstream cognitive consequences. This perspective reframes timely AI assistance as the problem of knowing not only when to help, but when not to help.

Keywords

proactive AI assistance, intervention timing, knowledge work, cognitive augmentation, human-AI interaction

1. Introduction

Current GenAI assistants are increasingly imagined as proactive collaborators in knowledge work. Rather than waiting for users to issue explicit prompts, future assistants may monitor ongoing activity, infer user needs, and offer timely support. This shift creates an important design challenge: when should an assistant intervene, and when should it remain silent?

A common way to approach this problem is to treat intervention timing as a matter of detecting user difficulty. If the user pauses for a long time, reformulates repeatedly, produces few outputs, or appears frustrated, the system may infer that help is needed. While such signals are useful, they are insufficient for knowledge work. Knowledge work often requires uncertainty, delay, and struggle. A researcher may pause because they are integrating evidence. A writer may repeatedly revise because they are clarifying an argument. A student may hesitate because they are evaluating a source rather than passively accepting an answer. If a GenAI assistant intervenes too early in these moments, it may reduce opportunities for reflection, self-regulation, and learning.

This paper argues that the timing of GenAI intervention should be examined through a distinction between productive cognitive friction and unproductive cognitive breakdown, between difficulty that advances cognitive engagement and difficulty that signals it has stalled. From this follows the central claim of the paper: timely GenAI intervention in knowledge work should be evaluated not by whether assistance reduces visible difficulty, but by whether it occurs at moments that preserve or restore productive cognitive engagement.

This perspective builds on research treating information seeking and use as iterative, evolving cognitive processes in which users interpret information, reformulate goals, and make situated judgments

Proceedings of TimelyAI: When Should Generative AI Assistants Intervene? A Workshop Co-located with the 5th Symposium on Human-Computer Interaction for Work (CHIWORK 2026), Linz, Austria, June 22 to 25, 2026

*Corresponding author.
klarazt@utexas.edu (Z. Guan)
https://klaraguan.github.io/ (Z. Guan)
0009-0008-6950-2638 (Z. Guan)

about relevance and quality; these processes extend beyond search interactions to reading, writing, analysis, and deliberation in knowledge work more broadly [1, 2, 3, 4, 5, 6]. The perspective also connects to recent work on proactive AI assistance, where systems anticipate needs rather than waiting for explicit input, raising timing questions closely related to those addressed here. As GenAI increasingly mediates knowledge work by synthesizing, explaining, and generating content [7, 8], it can both support human cognition and encourage premature closure, overreliance, and shallow engagement [9, 10, 11, 12]. To design proactive GenAI assistants responsibly, we need timing models and evaluation methods that ask not only whether intervention improves immediate task performance, but whether it preserves the user's role as an active cognitive agent.

The paper makes three contributions. First, it proposes productive friction and cognitive breakdown as a conceptual distinction for reasoning about intervention timing in knowledge work. Second, it argues that common timing signals are ambiguous and must be interpreted in relation to cognitive trajectory rather than treated as direct indicators of need. Third, it proposes a layered approach to evaluating intervention timing, combining immediate interaction outcomes with cognitive outcomes that capture the quality of users' thinking and work products.

2. Ambiguous Timing Signals in Knowledge Work

Designing proactive systems for knowledge work requires user modeling that infers when users need help, often through behavioral signals. But knowledge work is rarely a linear process of problem input and answer output. It involves forming and reforming questions, gathering and evaluating information, comparing sources, constructing interpretations, drafting and revising artifacts, and deciding how to proceed. Across these activities, users frequently begin with incomplete or poorly articulated needs that develop only through engagement with the problem [13, 14], and their criteria for what counts as relevant or sufficient evolve as understanding deepens [15, 16]. Behaviors such as reformulating questions, comparing sources, taking and reorganizing notes, and revising drafts are not merely operational steps toward a fixed goal; they are traces of ongoing cognition.

This iterative nature makes behavioral signals difficult to interpret. A long pause could mean confusion, but it could also mean reflection. Repeated reformulation could mean failure, but it could also mean deeper exploration [17, 18]. Low output volume could mean disengagement, but it could also mean careful evaluation. Rapid acceptance of AI-generated content could mean that the system provided useful support, but it could also indicate overreliance or premature closure. The same observable behavior can indicate either productive friction or cognitive breakdown depending on contexts such as task stage, cognitive state, and user goals.

GenAI intensifies this ambiguity. On one hand, GenAI assistants can help users articulate vague needs, synthesize large bodies of information, compare perspectives, generate alternatives, and reflect on their reasoning [19, 20]. On the other hand, fluent AI-generated responses may compress or obscure the cognitive steps through which users build understanding. When an assistant summarizes, evaluates, or recommends too readily, it may make knowledge work more efficient while reducing users' opportunities to practice judgment and learn [21, 22]. Proactive assistance therefore creates a difficult timing problem: the assistant must infer whether difficulty is a sign that help is needed, or whether it is part of the cognitive effort that should be preserved.

The challenge is not to minimize cognitive effort. It is to support the right kind of effort. A timely intervention is not necessarily the earliest possible intervention. It is an intervention that occurs when assistance can help sustain cognitive engagement without replacing the user's own interpretive work.

3. Productive Friction or Cognitive Breakdown

This paper proposes a distinction between two cognitive conditions that share much of their behavioral surface but differ in what intervention should do. I treat productive cognitive friction and cognitive breakdown not as strictly binary states, but as conceptual poles on a trajectory of cognitive engagement.

Table 1
Three Timing Judgments for Proactive GenAI Intervention

Timing judgment	Description	Possible consequence
Too early	The assistant intervenes while the user's difficulty still shows signs of productive cognitive movement.	The intervention may reduce reflection, agency, learning, or creative exploration.
Too late	The assistant waits until the user has already disengaged, narrowed the problem, or accepted a weak answer.	The intervention may fail to recover meaningful cognitive engagement.
Appropriate	The assistant intervenes in the transitional risk zone, where the user remains engaged but recent activity shows weakening conceptual movement.	The intervention may restore orientation, broaden thinking, or support continued evaluation.

Between them is a transitional risk zone in which the user remains engaged, but signs of productive movement begin to weaken. This framing is important for intervention timing: the goal is not to eliminate friction, but to detect when friction is no longer supporting progress and is beginning to turn into breakdown.

Productive cognitive friction is defined here as effortful activity that contributes to the development of understanding, judgment, or ideas. In knowledge work, not all resistance should be smoothed away, because some forms of effort create the conditions for reflection, evaluation, and creativity. Productive friction may involve uncertainty, hesitation, ambiguity, and even temporary frustration, but the user remains cognitively engaged. Importantly, productivity is not defined retrospectively by whether the final outcome is successful. Productive friction can be characterized during the process by indicators of continued cognitive movement, such as increasingly specific reformulations, emerging structure in notes or artifacts, active comparison of sources, articulation of criteria, generation of alternatives, or explicit evaluation of evidence. These indicators, elaborated in Table 2, provide an independent basis for distinguishing productive friction from breakdown. Poorly timed intervention during productive friction risks short-circuiting the very cognitive processes the user is engaged in, replacing user-owned understanding with system-supplied conclusions.

This idea aligns with theories of sensemaking, which describe understanding as an iterative process of fitting data to frames, revising frames when inconsistencies emerge, and constructing representations that support action [5, 6, 23]. It aligns with work on critical thinking and metacognition, which emphasize that careful evaluation, the questioning of assumptions, and reflective self-regulation depend on a state of perplexity or productive uncertainty that fluent answers tend to dissolve [24, 25, 26]. It also aligns with work on learning-oriented inquiry and exploratory information use, where uncertainty and reformulation are expected parts of the process rather than simply symptoms of poor system performance [27, 28, 29].

By contrast, cognitive breakdown refers to the point at which effort no longer contributes to progress. Breakdown may appear as looping without conceptual change, fixation on a narrow framing, inability to synthesize across materials, repeated reliance on AI outputs without verification, or premature closure around a weak answer. In these moments, intervention may be valuable because it can help users regain orientation, broaden their perspective, or re-engage in evaluation before the task is abandoned or prematurely resolved.

The most important timing challenge lies in the transitional risk zone between productive friction and breakdown. A process may be considered at risk but recoverable when the user still shows signs of engagement, such as continued interaction with sources, artifacts, or the assistant, but the trajectory of that engagement shows weakening cognitive movement. For example, reformulations may continue but stop becoming more specific; notes may accumulate without becoming more structured; source consultation may continue without comparison or integration; or generated outputs may be accepted with decreasing evidence checking. In such cases, the issue is not the presence of difficulty itself, but the loss of development across the user's recent activity. This is why timing judgments should rely on trajectories of signals rather than isolated behavioral cues.

Because the behavioral surface underdetermines the cognitive condition, distinguishing productive friction, transitional risk, and breakdown from individual signals examined in isolation is rarely possible. If assistants treat all signs of difficulty as triggers for intervention, they risk interrupting productive friction. If they wait only for obvious failure, they may intervene too late. A useful model of timing must therefore account for at least three timing judgments, summarized in Table 1.

A concrete example clarifies the distinction. A student comparing conflicting sources may benefit from continued silence while they examine the evidence; an assistant that prematurely synthesizes the sources into a single coherent answer would intervene too early. The same student, having stopped engaging with the contradiction and copied an AI-generated answer, would no longer be in a recoverable state; an intervention here arrives too late. An appropriately timed intervention might occur when the student still consults sources and revises notes, but no longer compares the contradiction or develops the interpretation. At that point, the assistant could prompt them to articulate which evidence supports each interpretation, restoring productive engagement without replacing their judgment.

4. Signals as Evidence, Not Triggers

Building on this ambiguity, timing judgments should treat signals as evidence within a trajectory rather than as direct triggers. Table 2 illustrates how candidate signals may be interpreted differently depending on whether the user's cognitive trajectory is still developing or has stalled, organized by the higher-order cognitive process they index.

The table should be read relationally rather than categorically. For instance, "repeated reformulation" is not inherently productive or unproductive. In exploratory information seeking, reformulation may reflect a user's evolving understanding of the problem [17]. In other contexts such as look-up search, repeated reformulation without clicks or changes in sources consulted may indicate a failure. Similarly, convergence is not inherently a problem. In creative work, convergence may be a productive stage of selecting and refining ideas after a period of divergent exploration [30]. This differs from declining diversity as a breakdown signal. Productive convergence narrows the idea space by applying emerging criteria, comparing alternatives, and elaborating selected directions; breakdown narrows the idea space without such evaluative movement. Thus, the same surface pattern, fewer ideas over time, can indicate either refinement or premature closure depending on whether narrowing is accompanied by explicit criteria, comparison, and elaboration.

This suggests that proactive GenAI assistants should not treat observable signals as simple triggers. Instead, signals should be interpreted as evidence within a trajectory. Relevant questions include: Is the user's understanding becoming more structured? Are criteria becoming clearer? Are sources being evaluated rather than merely collected? Are ideas becoming more diverse or more refined? Is the user still exercising judgment, or delegating judgment to the system?

The data sources used to answer these questions should be proportionate to the task. Interaction logs can capture queries, clicks, dwell time, revisions, scrolling, and transitions [31, 32]. Artifact traces can capture changes in notes, outlines, annotations, diagrams, or drafts, which are especially important in sensemaking and writing tasks [33, 34]. Lightweight self-reports or reflection prompts can help clarify whether users experience their effort as productive or stuck [35, 36]. More intrusive data sources, such as eye tracking or physiological sensing, may improve detection of cognitive load or attention but also raise concerns about proportionality, consent, and privacy [37]. The methodological challenge is not only how to detect signals accurately, but how to ensure that detection respects the cognitive and ethical character of knowledge work.

Table 2
Illustrative Signals That May Indicate Productive Friction or Cognitive Breakdown Across Cognitive Processes

Cognitive process	Productive friction may involve...	Breakdown may involve...
Sensemaking	Comparing sources, restructuring notes, creating schema, moving between evidence and interpretation.	Revisiting the same materials without conceptual change, accumulating fragmented notes without synthesis, ignoring unresolved contradictions.
Decision-making	Refining criteria, comparing trade-offs, seeking missing information, calibrating uncertainty.	Premature closure, overconfidence, reliance on one criterion, failure to consider alternatives, accepting a recommendation without understanding assumptions.
Critical thinking	Checking sources, questioning assumptions, generating counterarguments, comparing evidence quality.	Passive acceptance of fluent AI output, lack of verification, shallow agreement, failure to notice unsupported claims.
Creativity	Divergent exploration, recombination of ideas, movement across domains, iterative refinement.	Fixation, repetitive prompts, declining diversity of ideas, overreliance on generic AI suggestions.

5. Evaluating Intervention Timing

Evaluating whether an intervention was timely in knowledge work requires more than asking whether the user accepted it. Acceptance may reflect convenience rather than cognitive benefit. A user may welcome an intervention that gives them an answer quickly, even if it reduces reflection or weakens learning. Conversely, a user may ignore an intervention that is beneficial but poorly framed. Timing evaluation therefore needs to examine both immediate interaction effects and cognitive consequences.

At the immediate interaction layer, researchers can assess whether the intervention disrupted flow, whether the user responded to it, whether it helped the user continue the task, and whether it changed subsequent behavior. Useful indicators may include acceptance or dismissal, task resumption, reduced unproductive looping, renewed source exploration, or changes in revision patterns. Subjective measures such as perceived helpfulness, interruption cost, confidence, and control can provide additional insight [32, 4]. These measures are important, but they are not sufficient.

At the cognitive outcome layer, evaluation should ask whether the timing of the intervention improved or weakened the quality of the user's thinking or work product. For example, in a synthesis task, researchers might examine whether the intervention improved the structure, coverage, and integration of evidence in the final artifact. In a decision task, they might evaluate whether the user articulated clearer criteria, considered alternatives, and justified trade-offs. In a critical evaluation task, they might examine whether the user checked sources, identified unsupported claims, or considered counterarguments. In creative work, they might assess whether the intervention preserved idea diversity rather than narrowing users toward generic AI-suggested solutions [38, 39, 34].

This layered evaluation perspective allows researchers to distinguish different timing failures. Too-early interventions may produce short-term efficiency gains but weaker cognitive outcomes. Too-late interventions may be rated as helpful but fail to recover task progress. Appropriately timed interventions should support both immediate continuation and higher-quality cognitive engagement.

A useful study design may compare intervention timing conditions rather than intervention content alone. For example, a system could offer the same reflective prompt at different moments: during early exploration, after repeated reformulation without conceptual change, or after rapid acceptance of an AI-generated answer. Researchers could then examine not only which condition users prefer, but which condition leads to better synthesis, deeper evaluation, or more justified decisions.

Longitudinal studies may be especially important because the cognitive costs of premature assistance may accumulate over time. Recent work suggests several possible mechanisms: GenAI support can reduce learners' metacognitive engagement and motivation during learning tasks [9]; repeated reliance on AI assistance may produce cognitive debt, where users complete tasks but show weaker independent recall, ownership, or subsequent performance [10]; and GenAI systems can shift metacognitive demands from doing the task to monitoring and regulating the assistant, creating risks when users do not actively maintain that oversight [12]. These findings suggest that timing should be evaluated not only within a single interaction, but also in terms of whether repeated intervention affects users' capacity or willingness to reason independently.

6. Discussion and Research Agenda

This position invites further work along several directions.

Operationalizing cognitive state transitions. The productive-to-breakdown transition is presented here as a conceptual distinction, but operationalizing it for real-time inference will require empirical study of which trajectory signatures distinguish the two states across cognitive processes and tasks.

Individual differences. The same surface behaviors carry different meanings across users with different expertise, working styles, and cognitive preferences. A general timing model risks systematic misjudgment of users at the tails of these distributions. Adaptive or user-modeled approaches deserve exploration.

Longitudinal evaluation. The cognitive outcome layer is most meaningful over time. Single-session studies may not detect the gradual erosion of higher-order capacities that motivates concern about over-intervention. Longitudinal designs, though resource-intensive, may be necessary for the evaluation of timing systems claiming cognitive augmentation.

The status of beneficial silence. If silence is a timing decision, the design space of proactive GenAI includes not only when to speak but how to remain present without intervening, through ambient cues, deferred suggestions, or other forms of light-touch availability. This design space is largely unexplored.

The framing offered here does not prescribe a complete solution to the timing problem. It argues that the problem itself is more demanding than current framings acknowledge: timely GenAI intervention in knowledge work requires not the detection of difficulty but the discrimination of cognitive states that share its appearance, and the evaluation not only of help given but of help withheld.

7. Conclusion

In knowledge work, difficulty is often the condition through which users reflect, evaluate, synthesize, and create. This paper has argued that proactive GenAI assistants should be designed and evaluated through the distinction between productive cognitive friction and unproductive cognitive breakdown, two states that may share similar behavioral signals but call for different timing judgments. Timely intervention therefore depends on operationalizing users' cognitive states and needs within knowledge work tasks. Evaluation should consider not only whether assistance is useful, but whether its timing preserves human cognitive engagement. The future of proactive GenAI depends as much on beneficial silence as on helpful intervention.

Declaration on Generative AI

During the preparation of this work, the author used Claude Opus 4.6 for grammar and spelling checks and for adapting the LaTeX file to the required submission format. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the publication's content.

References

[1] N. J. Belkin, Interaction with texts: Information retrieval as information seeking behavior, Information retrieval 93 (1993).

[2] P. Ingwersen, K. Järvelin, The Turn: Integration of Information Seeking and Retrieval in Context, Springer Netherlands, Dordrecht, NETHERLANDS, THE, 2005.

[3] T. Saracevic, The stratified model of information retrieval interaction: Extension and applications, in: Proceedings of the Annual Meeting-American Society for Information Science, volume 34, Learned Information (Europe) Ltd, 1997, pp. 313–327.

[4] R. W. White, Interactions with Search Systems, Cambridge University Press, 2016.

[5] G. Klein, B. Moon, R. Hoffman, Making Sense of Sensemaking 1: Alternative Perspectives, IEEE Intelligent Systems 21 (2006) 70–73. doi:10.1109/MIS.2006.75.

[6] P. Pirolli, S. Card, The Sensemaking Process and Leverage Points for Analyst Technology as Identified Through Cognitive Task Analysis (2005).

[7] R. W. White, C. Shah (Eds.), Information Access in the Era of Generative AI, The Information Retrieval Series, 51, 1st ed. 2025. ed., Springer Nature Switzerland, Cham, 2025. doi:10.1007/978-3-031-73147-1.

[8] C. Zhai, Large Language Models and Future of Information Retrieval: Opportunities and Challenges, in: Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval, ACM, Washington DC USA, 2024, pp. 481–490. doi:10.1145/3626772.3657848.

[9] Y. Fan, L. Tang, H. Le, K. Shen, S. Tan, Y. Zhao, Y. Shen, X. Li, D. Gašević, Beware of Metacognitive Laziness: Effects of Generative Artificial Intelligence on Learning Motivation, Processes, and Performance, British Journal of Educational Technology 56 (2025) 489–530. doi:10.1111/bjet.13544. arXiv:2412.09315.

[10] N. Kosmyna, E. Hauptmann, Y. T. Yuan, J. Situ, X.-H. Liao, A. V. Beresnitzky, I. Braunstein, P. Maes, Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task, 2025. doi:10.48550/arXiv.2506.08872. arXiv:2506.08872.

[11] N. Sharma, Q. V. Liao, Z. Xiao, Generative Echo Chamber? Effects of LLM-Powered Search Systems on Diverse Information Seeking, 2024. arXiv:2402.05880.

[12] L. Tankelevitch, V. Kewenig, A. Simkute, A. E. Scott, A. Sarkar, A. Sellen, S. Rintel, The Metacognitive Demands and Opportunities of Generative AI, 2024. doi:10.48550/arXiv.2312.10893. arXiv:2312.10893.

[13] N. J. Belkin, Anomalous states of knowledge as a basis for information retrieval, Canadian journal of information science 5 (1980) 133–143.

[14] R. S. Taylor, Question-negotiation and information seeking in libraries., Technical Report, 1967.

[15] C. C. Kuhlthau, Inside the search process: Information seeking from the user's perspective, Journal of the American Society for Information Science 42 (1991) 361–371. doi:10.1002/(SICI)1097-4571(199106)42:5<361::AID-ASI6>3.0.CO;2-\#.

[16] P. Vakkari, A theory of the task-based information retrieval process: A summary and generalisation of a longitudinal study, Journal of documentation 57 (2001) 44–60.

[17] S. Y. Rieh, H. I. Xie, Analysis of multiple query reformulations on the web: The interactive information retrieval context, Information Processing & Management 42 (2006) 751–768. doi:10.1016/j.ipm.2005.05.005.

[18] A. Spink, M. Park, B. J. Jansen, J. Pedersen, Multitasking during web search sessions, Information Processing & Management 42 (2006) 264–275.

[19] M. Aliannejadi, J. Gwizdka, H. Zamani, Interactions with Generative Information Retrieval Systems, 2024. doi:10.48550/arXiv.2407.11605. arXiv:2407.11605.

[20] Y. Liu, V. N. Shah, S. Suh, P. Siangliulue, T. August, Y. Huang, Perspectra: Choosing Your Experts Enhances Critical Thinking in Multi-Agent Research Ideation, in: Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems, CHI '26, Association for Computing Machinery, New York, NY, USA, 2026, pp. 1–27. doi:10.1145/3772318.3791560.

[21] C. Shah, E. M. Bender, Envisioning Information Access Systems: What Makes for Good Tools and a Healthy Web?, ACM Trans. Web 18 (2024) 33:1–33:24. doi:10.1145/3649468.
[22] A. Singh, K. Taneja, Z. Guan, A. Ghosh, Protecting Human Cognition in the Age of AI, 2025. doi:10.48550/arXiv.2502.12447. arXiv:2502.12447.
[23] D. M. Russell, M. J. Stefik, P. Pirolli, S. K. Card, The cost structure of sensemaking, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '93, ACM Press, Amsterdam, The Netherlands, 1993, pp. 269–276. doi:10.1145/169059.169209.
[24] D. F. Halpern, Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring., American psychologist 53 (1998) 449.
[25] D. Kuhn, A developmental model of critical thinking, Educational researcher 28 (1999) 16–46.
[26] M. E. Martinez, What is metacognition?, Phi delta kappan 87 (2006) 696–699.
[27] G. Marchionini, Exploratory search: From finding to understanding, Communications of the ACM 49 (2006) 41–46.
[28] S. Y. Rieh, K. Collins-Thompson, P. Hansen, H.-J. Lee, Towards searching as a learning process: A review of current perspectives and future directions, Journal of Information Science 42 (2016) 19–34. doi:10.1177/0165551515615841.
[29] R. W. White, R. A. Roth, Exploratory Search: Beyond the Query—Response Paradigm, Synthesis Lectures on Information Concepts, Retrieval, and Services, Springer International Publishing, Cham, 2009. doi:10.1007/978-3-031-02260-9.
[30] M. D. Mumford, M. I. Mobley, R. Reiter-Palmon, C. E. Uhlman, L. M. Doares, Process analytic models of creative capacities, Creativity Research Journal 4 (1991) 91–122. doi:10.1080/10400419109534380.
[31] B. J. Jansen, A. Spink, How are we searching the World Wide Web? A comparison of nine search engine transaction logs, Information processing & management 42 (2006) 248–263.
[32] D. Kelly, Methods for Evaluating Interactive Information Retrieval Systems with Users, Foundations and Trends® in Information Retrieval 3 (2009) 1–224. doi:10.1561/1500000012.
[33] P. H. Nguyen, K. Xu, A. Wheat, B. W. Wong, S. Attfield, B. Fields, SensePath: Understanding the Sensemaking Process Through Analytic Provenance, IEEE Transactions on Visualization and Computer Graphics 22 (2016) 41–50. doi:10.1109/TVCG.2015.2467611.
[34] M. J. Wilson, M. L. Wilson, A comparison of techniques for measuring sensemaking and learning within participant-generated summaries, Journal of the American Society for Information Science and Technology 64 (2013) 291–306.
[35] A. Cole, H. O'Brien, Using data-prompted interviews in interactive information retrieval research: A reflection on the study of self-efficacy when learning using search, in: Proceedings of the 2023 Conference on Human Information Interaction and Retrieval, 2023, pp. 406–411.
[36] A. Singh, Z. Guan, S. Y. Rieh, Enhancing Critical Thinking in Generative AI Search with Metacognitive Prompts, 2025. doi:10.48550/arXiv.2505.24014. arXiv:2505.24014.
[37] J. Mostafa, J. Gwizdka, Deepening the role of the user: Neuro-physiological evidence as a basis for studying and improving search, in: Proceedings of the 2016 Acm on Conference on Human Information Interaction and Retrieval, 2016, pp. 63–70.
[38] T. M. Amabile, Social psychology of creativity: A consensual assessment technique, Journal of Personality and Social Psychology 43 (1982) 997–1013. doi:10.1037/0022-3514.43.5.997.
[39] D. H. Jonassen, Designing for decision making, Educational technology research and development 60 (2012) 341–359.