Truth Matters: Generative AI as Muse or Tool in the Research Process

Marcus Birkenkrahe
Division of Math & Science, Lyon College, Batesville, USA
Email: birkenkrahe@lyon.edu

Abstract: **The integration of generative artificial intelligence (AI) into research marks both a provocation and an inflection point. While gains in productivity and access draw attention, a deeper transformation is underway: how knowledge is conceptualized, mediated, and validated amid systems that simulate understanding without possessing it. If current trends hold, AI will amplify existing dynamics in scholarly communication. Publication volume may rise, but trust could decline. Conventional markers of originality and rigor may destabilize—not through automation alone, but through shifting norms around authorship, evaluation, and epistemic authority. This paper argues that AI is neither just a tool nor merely a muse, but a structural participant in research — shaping inquiry through fluent simulation but without understanding. A cognitive map is introduced to model how researchers interact with AI across phases of the research process, alternating between instrumental and generative uses. Generative systems can assist and accelerate scholarly work, but their role must be framed within a broader account of intellectual labor and meaning-making. Confusing fluency for insight risks eroding core scholarly practices. Implications extend to pedagogy, authorship policy, and the design of AI-aware research infrastructure. Ultimately, scholarship in the age of AI will depend as much on critical literacy as on technical fluency. AI is not merely a tool of transformation but a mirror—reflecting the values and assumptions of the communities that create and use it.**

Keywords: **Generative Artificial Intelligence, Research Methodology, Cognitive Mapping, Epistemology of AI, Scholarly Communication.**

I. INTRODUCTION

The advent of generative artificial intelligence (AI) signals a significant shift in research activities [1]. This paper examines the dual role of generative AI as both a muse and a tool in scholarly practice, proposing a cognitive framework intended to guide reflective use and preserve the integrity of human-driven inquiry.

This analysis takes place against the backdrop of a growing erosion of trust in information. The World Wide Web—originally conceived as a global library of knowledge—has increasingly commodified information, obscuring sources and fostering widespread skepticism [2]. Information is often consumed via anonymous social media platforms and filtered through opaque, algorithm-driven search engines. The credibility of science itself faces significant challenges: in 2023, over 10,000 published research papers—the highest annual number on record—were retracted worldwide [3]. Although retraction rates remain low relative to total research output (which exceeded 3 million publications in 2023, an increase of nearly 60% since 2012) [4], error rates in some disciplines—whether due to fraud or other issues—are estimated at 20–30% [5]. While these trends cannot be attributed solely to AI, it remains unclear how AI adoption might improve the situation, particularly given its potential to generate convincingly fabricated scientific articles [6].

This inquiry originated from an effort in early 2025 to employ AI in the development of a research proposal on modeling hallucinations as emergent phenomena using methods from mathematical physics. Prior to this, AI had been used primarily for enhancing coding workflows and teaching efficiency. However, in this context, AI appeared to function as a reflective instrument—raising the question of its value as a muse or source of inspiration.

Initial consideration was given to teacher-scholars at small colleges and resident-scholars in medical training programs as a representative audience. These researchers often work under considerable constraints, balancing intensive teaching or clinical duties with scholarly expectations. For such individuals, AI offers a potential lifeline by supporting tasks such as data analysis, literature reviews, and manuscript drafting, while also potentially facilitating creative insight and novel inquiry.

Subsequent reflection expanded the scope of this work to a broader audience. Anyone who engages in systematic search processes is, in effect, performing research. Many users have already replaced traditional search engines with conversational AI interfaces, and most major search platforms now incorporate generative AI capabilities [7]. Accordingly, this article aims to provide a principled foundation for the reflective and ethical use of AI in research—not merely as an efficiency enhancer, but as a tool for rethinking how knowledge is sought, generated, and communicated.

The discussion begins with a brief overview of recent technological developments and evolving usage patterns over the past two years. It then addresses the persistent mismatch between the demand for and supply of high-quality data, and considers the potential for AI augmentation to address this gap. The central contribution is a cognitive map outlining four scenarios that align distinct phases of the research process with the muse/tool dichotomy in AI-supported inquiry. The paper concludes with a concise overview of emergent AI tools, a set of forward-looking predictions, and a brief reflective commentary on future directions.

This paper originated as a faculty colloquium presentation delivered at Lyon College in April 2025.

II. EXPLORATION AND SPECIALIZATION

In August 2023, as it became increasingly evident that generative AI would persist as a transformative force, the author delivered a public lecture titled "How Machine Learning Can Help Your Publishing Career" at White River Medical Hospital in Batesville AR, USA [8]. Since that time, both the underlying technologies and their modes of use have evolved significantly.

DOI: 10.53375/ijecer.2025.472

By the end of 2023, approximately 33% of organizations had adopted generative AI in at least one business function [9]. At this stage, emphasis was placed on general-purpose tools (e.g., ChatGPT-4) used for tasks such as text generation, code synthesis, and summarization. In October 2023, studies reported productivity improvements of up to 25% among consultants with access to AI-assisted workflows [10]. Concurrently, discussions emerged around ethical concerns, including bias, factual reliability, and potential misuse of generative models [11]. The overarching approach in this early period was characterized by rapid experimentation and open-ended exploration of novel capabilities.

By the spring of 2025, the field had progressed through the early stages of the innovation adoption curve and entered a phase of heightened public discourse. Notably, Dennis Hassabis, the creator of the AlphaFold model for protein structure prediction (developed in 2018), was featured in TIME magazine with the assertion that artificial general intelligence (AGI) may soon be within reach [12]. Bill Gates, co-founder of Microsoft, suggested that key professional roles—such as those of doctors and teachers—could become largely replaceable by AI within the next decade [13].

As of 2025, adoption rates have risen to approximately 71% of organizations, with technology firms, professional services, and media companies leading in generative AI integration [14]. A noticeable shift has occurred from general-purpose models to domain-specific, multi-modal models—capable of ingesting and producing information across formats including text, images, audio, video, and symbolic data. These specialized models are increasingly applied to research-intensive tasks in areas such as marketing, product development, and services. Organizations report significant productivity enhancements, with 58% citing "exponential" benefits [15]. In parallel, regulatory scrutiny has intensified, prompting many institutions to implement structured governance protocols, including output review procedures and internal AI policies. For example, Lyon College has introduced institutional AI guidelines for faculty, staff, and students. Specialized AI systems are now broadly deployed in sectors such as healthcare, finance, and design, where domain adaptation is essential.

Two important caveats warrant attention when evaluating these favorable assessments. First, based on established patterns in technology adoption, benefit analyses—especially those related to tools involving complex human interaction—frequently overstate potential returns. Such projections are often used to justify strategic investments and may later be revised in light of practical limitations. Second, while multi-modal models appear capable of integrating diverse data types, their operational foundation remains rooted in language modeling. These models function by generating probabilistic token sequences based on large-scale training data. Their multi-modal competencies emerge from aligned datasets (e.g., image–caption or audio–transcript pairs), which allow them to simulate non-linguistic reasoning through semantically consistent outputs. However, this ability should not be mistaken for genuine multi-modal understanding.

III. LIMITS AND PROMISES

Modern generative AI models operate by estimating the most probable continuation of a given sequence of tokens, based on statistical patterns learned from extensive corpora of text. Despite significant advancements in scale and performance, the underlying mechanism remains essentially unchanged since 2022: these models function as stochastic pattern-matching engines. Their results are not based on the meaning or comprehension [16].

Large Language Models (LLMs) do not perform symbolic reasoning, which is governed by explicitly defined logical rules and typically yields high reliability [17]. Nor do they exhibit consciousness — an attribute whose nature remains scientifically elusive. Critically, LLMs lack semantic processing capabilities: they do not comprehend the physical world, social nuance, or idiomatic expressions. For example, a phrase like “The glass is half full” evokes layered meanings for human readers but is treated by the model as a statistical configuration of tokens, devoid of conceptual understanding. The model recognizes neither the semantic significance of the sentence as a whole nor of its constituent parts.

Questions regarding how much context is “enough” for coherent output depend on both the prompt’s specificity and the model’s contextual window, which may encompass up to 100,000 tokens in state-of-the-art systems. However, regardless of length, the model’s output is always bounded by what it has encountered during training and the information provided in the current prompt. While human performance may even improve when a human is asked ambiguous or paradoxical questions, the AI's output quality degrades. In fact, the reason why AI output inevitably degrades over time, a process that is sometimes euphemistically called "data drift" or “content drift”, is not understood at all [18].

The integration of web search capabilities aims to supplement the model’s limitations by injecting real-time context. Nevertheless, even “deep search” functions as a hybrid process: it combines traditional information retrieval (i.e., fetching content from web pages) with the model’s native statistical pattern completion. This does not constitute genuine comprehension. Rather, it is more akin to augmenting auto-complete with live data — a process that increases relevance but does not confer understanding. Complicating matters further, much of the modern web is dynamic and interactive, making web scraping technically challenging and prone to error. Consequently, the reliability of such retrieved content varies considerably.

What generative AI systems produce is best described as syntactic fluency combined with statistical mimicry. There is no semantic depth whatsoever. The apparent wisdom of these systems is a byproduct of their exposure to vast and diverse textual data, including materials that contain insightful or well-articulated ideas. Through pattern recombination, these models can reproduce formulations that appear profound. However, this effect is analogous to an actor delivering lines from a script: the performance may seem authentic, but it does not imply identity or understanding. An actor who recites the words of Julius Caesar in Shakespeare's play does not become a statesman.

Some have argued that denying models the label of “creative” or “intelligent” is problematic, especially in light of the conceptual ambiguity surrounding those terms [19]. Yet the absence of clear definitions does not justify retrofitting human-centered constructs to accommodate machine capabilities. The inability to fully define creativity or intelligence in human contexts does not necessitate redefining these terms to align with statistical language models.

When LLMs are described as operating in a "thinking mode," this often reflects only rather the appearance of reasoning. Internally, the model selects the next token by assigning probabilities to many potential outcomes based on learned patterns encoded in its neural network weights. This process is fundamentally non-reflective and non-referential. Unlike humans, who may reason about causality, abstract relationships, or counterfactuals, LLMs merely produce the statistically most likely continuation of a prompt.

Although it is difficult to articulate what reasoning is in humans, we possess an intuitive sense when someone is genuinely reasoning with us versus merely echoing familiar phrases. While this distinction is challenging to formalize, it is essential to how we evaluate credibility and establish trust—both of which are foundational to effective secondary research. Once trust is undermined by what is perceived as unreasonableness, it is not simply diminished; it is often withdrawn altogether.

IV. DATA AND DISPLACEMENT

A useful metaphor for understanding our relationship with data and artificial intelligence today is Plato's Allegory of the Cave, first articulated in "The Republic" around 375 AD [20]. In this allegory, prisoners are confined in a cave, chained in such a way that they can only see shadows projected on a wall by objects behind them—mere representations of reality. When one prisoner escapes and experiences the world outside, he returns to share the truth, only to be rejected by those who remain shackled. The story illustrates that grasping and facing the truth requires effort, reflection, and courage.

The dilemma we face today as researchers and digital citizens interacting with generative large language models and vast information infrastructures has several dimensions. Three stand out as particularly relevant to the question of whether AI can aid or distort our pursuit of knowledge:

1. *Fragmentation of knowledge vs. integration of insight:* Contemporary research problems—whether scientific, medical, environmental, or social—require synthesis across domains. Yet academic labor has become increasingly specialized, with experts knowing more and more about narrower fields. Ideally, the World Wide Web would facilitate bridging these silos, enabling shared access to distributed knowledge. This was its original vision: a global network of linked ideas rather than duplicated effort. But that ideal has proven elusive—not due to technical constraints, but to cultural, economic, and epistemic ones [21].

2. *Transparency and trust have eroded despite (or because of) information abundance.* Hyperlinking made information accessible, but not always interpretable. In commodifying information, the Web separated it from its human source and context. Today, digital representations of knowledge often obscure more than they reveal. We can access more than ever before, but with less confidence in the accuracy, provenance, or intent of what we find. This paradox — visibility without reliability — contributes to widespread epistemic skepticism [22]. If the Web is a hall of mirrors, then the analog, embodied, and socially situated aspects of knowledge are more valuable than ever.

3. *Digital mediation distorts rather than reveals reality.* Most of our digital encounters with the world — through social media, search engines, or institutional platforms — are mediated. We rarely engage directly with information or experience; instead, we receive curated outputs shaped by commercial, algorithmic, and institutional forces. In this sense, we resemble Plato's prisoners: our access to reality is limited, our perceptions are mediated, and our trust is increasingly precarious. These conditions create vulnerabilities to manipulation and misinformation that extend beyond any one platform or model—suggesting a kind of "digital metempsychosis," whereby our experience of reality and even our sense of self is shaped and displaced by the digital dwellings we ourselves have constructed [23].

V. AUTOMATION AND EMANCIPATION

In response to the mounting cognitive and procedural demands of modern research, generative artificial intelligence—particularly LLMs—has been widely promoted as a transformative solution [24]. The prevailing promise is clear: both mundane and complex intellectual tasks can now be delegated to machines. This vision of liberation permeates institutional bulletins, academic blogs, opinion pieces, and online forums. The message is consistent and optimistic: the machine will unchain the human mind.

This supposed liberation encompasses all levels of academic labor. Generative AI is expected to handle repetitive, time-consuming tasks such as summarization, formatting, and proofreading, while also assuming responsibility for more sophisticated activities, including conceptual analysis, drafting, and synthesis. These systems are said to perform such tasks more efficiently, more rapidly, and—according to advocates—with outcomes that often exceed expectations. The result is frequently described as "satisfying," implying a seamless match between algorithmic output and human standards of coherence or utility.

Yet this is a formidable claim. Even in the early stages of widespread AI deployment, as in 2023, the pressures on knowledge workers were acutely visible. Medical residents, for example, found themselves constrained by systems that demanded not only excellence in patient care but also continuous scholarly output as a condition of advancement. Similarly, faculty at teaching-intensive institutions are expected to deliver high-quality instruction, engage in service, and simultaneously maintain active research agendas. These overlapping obligations reflect structural tensions that cannot be alleviated by efficiency tools alone.

The appeal of technological relief is neither new nor unreasonable. Indeed, the history of technology is frequently narrated as a sequence of innovations intended to reduce human labor and stress. Generative AI is but the latest installment in this trajectory, framed not only as a productivity enhancer but also as a cognitive liberator.

What remains uncertain is whether this promise will lead to genuine emancipation or merely to the further automation of already fragmented intellectual life. The significance of generative AI lies not only in what it can accomplish, but in how its role is interpreted: as replacement, assistant, inspiration, or constraint. This interpretive question is crucial to any ethical and reflective engagement with AI in research settings.

VI. TRUST AND DEGRADATION

One of the most widely discussed liabilities of generative artificial intelligence is its tendency to produce outputs that are convincingly fluent but factually incorrect. These phenomena are frequently referred to as *hallucinations* or *confabulations*. The terms have gained traction across technical and journalistic domains, but their use warrants scrutiny: both originate in clinical and cognitive science and imply subjective mental states or intentional agency. Their application to language models risks anthropomorphizing systems that possess neither consciousness nor internal representation. More appropriate terminology would describe these behaviors as *glitches*, or simply *errors* — misalignment between the model's outputs and objective or verifiable reality.

These errors are not merely occasional flaws but structural features of probabilistic text generation. Trust is central to any working relationship, and generative AI systems, by their nature, cannot be trusted in the same way that human assistants can. Their failures are not detectable in advance, nor is there any mechanism by which they can self-correct with intention or reflection. As such, while they may perform helpful tasks, they cannot yet be considered collaborators in any meaningful epistemic sense.

Recent research reinforces this concern. Contrary to optimistic projections, larger models with enhanced reasoning capabilities appear to produce more, not fewer, hallucinations. In April 2025, OpenAI acknowledged that newer models — including the o3 and o4-mini reasoning systems — exhibited significantly higher hallucination rates on tasks requiring factual accuracy. On the PersonQA benchmark, o3 hallucinated in 33% of test cases, while o4-mini reached 48%—compared to only 16% in earlier models [25]. The underlying causes remain unclear, and the findings suggest that increasing model scale does not straightforwardly reduce error rates.

An illustrative example of model persistence can be seen in visual generation tools integrated with conversational agents. In one case, an image generator began repeatedly including the Twitter bird icon in unrelated illustrations. The bird had been requested for a prior image, but subsequent prompts—lacking any reference to it—still yielded visual outputs containing the motif. Despite efforts to correct this, the model persisted. This behavior — akin to a form of fixation — offers a visual analogue to linguistic hallucination. It reflects the entanglement of prompt history, token prediction, and model memory. More strikingly, it demonstrates how human users may adapt to these persistent errors, accepting or incorporating them rather than resisting the system's outputs.

Such examples raise a deeper, more speculative concern about AI's impact on human learning and reasoning: Could Generative AI subtly diminish their intellectual agency? The use of AI in scholarly work introduces the risk of cognitive dependency: The illusion of insight may replace genuine inquiry. Users may feel informed or competent without acquiring new knowledge or deepening their understanding. Worse still, this over-reliance may degrade existing knowledge through contamination or misapplication.

This worry, while difficult to empirically verify, is conceptually serious. If generative AI tools produce outputs that feel authoritative but lack epistemic reliability, users may unknowingly internalize distorted representations of their subject matter. The resulting feedback loop — false confidence, reduced engagement, degraded learning — poses a structural challenge to the use of AI in both educational and research contexts.

Whether generative AI will ultimately make scholars more efficient or less effective remains an open question. It is, fundamentally, a problem for the learning sciences. But as with many questions about human cognition and development, consensus is elusive. In the absence of definitive evidence, researchers and educators must rely on reflective judgment, incremental experimentation, and ethical caution as they integrate AI into their work.

VII. MODELS AND METAPHORS

Although this paper primarily explores the meta-question of how generative AI may support or distort the research process, its origins lie in a concrete attempt to use AI as a co-developer of a highly technical research proposal. The hypothesis in question arose from a speculative analogy: that hallucinations in generative AI might resemble phase transitions in physical systems. Much like boiling water undergoes a sudden transformation from liquid to gas, the model might transition from coherence to incoherence — from accurate representation to confabulated response.

This line of inquiry reflects a broader curiosity about whether deep learning systems exhibit emergent behavior that could be analyzed using mathematical physics. Specifically, the idea was to apply the Renormalization Group (RG) — a conceptual framework from quantum field theory used to study scale-dependent behavior in physical systems — to the multi-scale information compression performed by neural networks. Just as RG techniques discard microscopic details to uncover stable macroscopic patterns, deep networks compress and abstract from raw data to yield semantically useful features. In both domains, the process involves hierarchical organization, information loss, and the emergence of invariant patterns. Prior work has shown that restricted Boltzmann machines trained on data from the Ising model can reconstruct RG transformations [26].

Further speculative steps in this project included mapping AI hallucinations using category theory and representing information flow using Feynman diagrams — techniques drawn from earlier work in lattice gauge theory [27]. In this framework, hallucinations might be described as transitions in internal symmetry, expressible as morphisms between categorical states. Whether these mappings are formal or metaphorical remains an open question, but the exercise illustrates the kind of creative conceptual bridging that human researchers bring to complex domains.

Testing this research idea with the aid of generative AI proved both illuminating and frustrating. During spring 2025, a sustained effort was made to use AI tools to shape the research proposal. The result was uneven. In some cases, AI provided helpful scaffolding and stylistic feedback, generating prompts that led to conceptual clarification. In others, it produced misleading analogies or syntactically plausible but substantively vacuous responses. While the experience did not yield a completed proposal, it did clarify the dual nature of AI: at times a useful tool, at times a provocative muse.

This episode underscores a central theme of the paper: the value of AI is not just technical but *epistemological*. Its usefulness does not rest solely on its capacity to generate coherent output, but on its capacity to engage human intuition in unexpected ways. Yet the boundary between inspiration and illusion remains precarious. As AI systems simulate understanding with increasing fluency, researchers risk mistaking verbal plausibility for conceptual clarity. The true challenge lies not only in the model's capacity, but also in the researcher's discernment.

VIII. INSPIRATION AND EXECUTION

The expanding presence of AI in scholarly practice invites a conceptual distinction between two modes of engagement: *AI as muse* and *AI as tool.* This distinction informs not only practical usage patterns but also the epistemological and ethical implications of AI-supported inquiry.

A *muse*, in classical and contemporary usage, denotes a source of inspiration—one that prompts creativity, facilitates imaginative leaps, and engages the researcher as a cognitive partner. Muses are traditionally invoked, not constructed; they are felt rather than controlled. The concept presumes a subjective experience of being "touched" by insight, a phenomenon foundational to the emergence of original thought in the research process.

By contrast, a *tool* is an instrument of execution. Tools follow instructions, enable task completion, and extend human capabilities through mechanized reliability. Their value lies not in provoking thought but in structuring, organizing, or accelerating it. In computational contexts, tools are typically designed to be optimized, integrated, and assessed on the basis of performance and reproducibility.

The muse/tool dichotomy provides a useful frame for understanding both the potential and limitations of generative AI in research. As a tool, AI supports well-bounded tasks such as summarization, formatting, citation generation, and code synthesis. As a muse, it may provoke conceptual reframing, suggest metaphors, or reveal previously unconsidered directions—albeit without intention or comprehension.

Empirical insights, though preliminary, support this dual framing. In a small classroom survey (N = 15), approximately 53% of students reported using generative AI primarily as a tool to accomplish concrete tasks, while 47% described it as a muse supporting creative ideation. This distribution suggests that users experience AI both as a procedural asset and as a cognitive stimulant, depending on context and familiarity.

A parallel audience poll conducted in an academic setting (N = 49) yielded similar ambiguity. When asked whether generative AI is best understood as a muse or a tool, 49% selected "both equally." Another 38.8% identified AI primarily as a tool, while 6.1% viewed it chiefly as a muse. An additional 6.1% expressed uncertainty. These results suggest that for many academic users, AI occupises a hybrid space—neither fully instrumental nor purely inspirational. A second poll addressed the widely circulated prediction asserting that AI will replace doctors and teachers within ten years [13]. Responses were overwhelmingly skeptical: 51% strongly disagreed, 34.7% somewhat disagreed, 12.2% somewhat agreed, and only 2% strongly agreed. The results reflect continuing doubt about the capacity of AI to replicate roles grounded in interpretive judgment, interpersonal responsiveness, and trust-based relationships.

Together, these findings support a scenario-based understanding of AI's role in research, wherein the same system may operate alternately as muse or tool, contingent on user intention, task complexity, and disciplinary context. The next section formalizes this view through a cognitive map that aligns these roles with discrete phases of the research process.

IX. SCENARIOS AND SUPPORT

A scenario-based framework was developed to illustrate how AI may function as a muse or a tool across distinct phases of the research process. Drawing on the scenario-building tradition of the 1970s, particularly as practiced in strategic foresight and systems analysis [28], this framework locates the evolving relationship between user and system along two intersecting dimensions: from research *idea* to research *result* (horizontal axis), and from AI as *tool* to AI as *muse* (vertical axis).

At the center of this coordinate system lies the activity of research *review*—the reflective movement between inspiration and implementation that marks the transition from isolated ideation to communicable outcome. The model suggests that research activity rarely proceeds in a linear fashion. Instead, it evolves through iterative navigation between conceptual, procedural, creative, and mechanistic modes.

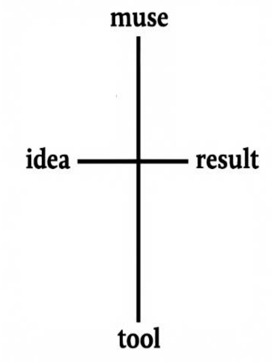

Figure 1: Four quadrants of AI-augmented Research

The proposed quadrants serve not as fixed categories but as heuristic markers that encourage researchers to locate themselves and their AI usage patterns within a broader cognitive landscape. This map can be used to identify one's position in the research workflow, clarify the type of support AI may offer at that stage, encourage reflective practice rather than blind automation, and bridge abstract thinking and practical implementation.

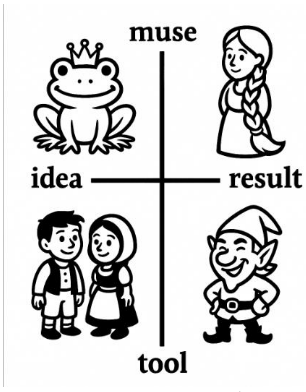

Figure 2: Four scenarios of AI-augmented research

To illustrate the four quadrants, a set of metaphorical scenarios — drawn from familiar European fairy tales — was employed [29]. These narratives offer an intuitive and culturally universally recognizable entry point into otherwise abstract research behaviors. Each quadrant emphasizes a different combination of input (idea/result) and orientation (muse/tool):

1. *Idea–Muse: "The Frog Prince"*. AI reveals the unexpected by introducing conceptual metaphors or surprising analogies. For example: when asking the AI for metaphors about hallucinations, the answers revealed a deeper context than I had anticipated and helped me to better define my research goal.

2. *Result–Muse: "Rapunzel"*. AI gently unlocks new interpretive paths during the validation or presentation of a result. For example: the AI suggested to test a research result with a

series of sample cases and provided a list of specific experimental setups.

3. *Idea–Tool: "Hänsel and Gretel".* AI helps to navigate early-stage ideation by producing language variations, structural outlines, or reformulations of loosely defined concepts—much like dropping and following breadcrumbs through a conceptual forest. For example: when playing with different ways to phrase ideas, the AI offered a sequence of rewordings that revealed underlying assumptions.

4. *Result–Tool: "Rumpelstiltskin".* AI assists in transforming completed work into publishable or presentable form. For example: after drafting a methods section the AI helped rephrase technical language to suit an interdisciplinary audience.

While metaphorical, these scenarios highlight practical and cognitive orientations that are observable in actual user behavior. They emphasize that generative AI systems do not function monolithically. Instead, they shift roles depending on the stage of inquiry, the user's objectives, and the interpretive or technical challenges at hand. Understanding these roles not only enhances critical engagement but may help prevent the epistemic and procedural drift associated with unreflective reliance on AI-generated output.

The cognitive map can thus serve as both a typology and a guide—a prompt for inquiry into how, when, and why AI might augment (rather than distort) the research process.

X. EVIDENCE AND EXTENSION

Empirical data and sector-specific commentary provide early insight into the evolving integration of generative artificial intelligence across educational, disciplinary, and public domains. This section synthesizes these developments and identifies a parallel expansion in the design of AI tools, with growing specialization and adaptive memory capabilities that shape the trajectory of research augmentation.

A 2025 survey conducted by the Higher Education Policy Institute reported that 92% of UK undergraduate students now use generative AI tools such as ChatGPT, a sharp increase from 66% in the previous year. Primary motivations included perceived gains in time efficiency and improved work quality [30]. A separate 2024 study by Anthropic, surveying 5,000 university students, found widespread use of Claude for ideation, outlining, and formative feedback—particularly in the early stages of the research and writing process [31].

In the STEM disciplines, the integration of AI has been positioned as a productivity enhancer rather than a source of original thought. In a 2024 lecture "The Potential for AI in Science and Mathematics", mathematician Terence Tao characterized AI as a "co-pilot" capable of identifying patterns and executing routine procedures. However, Tao stressed that foundational ideas and novel proofs must continue to emerge from human intuition. In this vision, generative AI systems function as auxiliary engines that automate the mundane while preserving space for creative insight [32].

The reception of generative AI in the humanities has been notably ambivalent. While recent studies document its growing use in digital humanities for tasks such as automation and augmentation, scholars remain divided over its epistemic implications and the potential disruption of interpretive methodologies [33].

In broader public usage, a 2024 report by Common Sense Media indicated that 70% of U.S. teenagers had engaged with generative AI tools, primarily for school-related tasks such as brainstorming and homework assistance. Notably, only 37% of parents were aware of their children's AI use. The study also reported significant disparities in access and perception along demographic lines, reinforcing the need for targeted education and institutional guidance on ethical and effective AI use [34].

In parallel with these usage patterns, 2025 has witnessed a notable diversification of generative AI tools designed for specific research and educational contexts. Among the most consequential innovations is the emergence of multi-context prompting (MCP), now supported by systems such as ChatGPT’s “Memory” and Claude’s document-based context. These tools retain conversational state and user-defined goals across sessions, facilitating continuity and iteration in research workflows.

A related development is the growing presence of autonomous AI agents, exemplified by systems like AutoGPT and Devin. These agents are designed to execute multistep tasks—including web search, filtering, and basic decision-making—without ongoing user intervention. While still prone to error and difficult to audit, such systems represent a shift from responsive to delegated computation.

Specialized platforms have also emerged. Tools such as Google’s NotebookLM offer document-aware assistance for literature review and information synthesis. Other domain-specific copilots have been fine-tuned for targeted applications: SciSpace (science publishing), CaseText and CoCounsel (legal reasoning), Cursor (software development), and Diffit (education). These systems reduce hallucination rates by constraining output to disciplinary language models and structured corpora. More traditional platforms, such as Google Colab, GitHub, and Datalab, have likewise integrated generative assistants into their environments, increasing their utility without replacing expert oversight.

Together, these trends suggest both a broadening and a deepening of AI engagement in scholarly practice. The tools are becoming more personalized, the tasks more intricate, and the user roles more variable. If generative AI is to serve as either muse or tool—or both—it will do so increasingly through systems that are not only generative, but also context-sensitive and domain-adaptive.

XI. CONCLUSIONS AND OUTLOOK

The increasing entanglement of generative artificial intelligence with research practice presents both a provocation and an inflection point. While measurable effects—such as increased productivity or expanded access—continue to draw attention, the more consequential shift lies in how knowledge is conceptualized, mediated, and validated in the presence of systems that can simulate understanding without possessing it.

If current trajectories persist, generative AI will likely amplify existing patterns in scholarly communication: publication volumes will rise, the credibility of output will face renewed scrutiny, and traditional indicators of originality or value may be destabilized. Yet these outcomes are neither fixed nor uniform. They depend not on technological capacity alone, but on institutional adaptation, scholarly discernment, and the evolution of norms surrounding authorship, trust, and interpretation.

This paper has argued that generative AI should not be viewed solely as a tool for efficiency nor merely as a muse for inspiration. Instead, it should be recognized, and carefully used, as a *structural participant* in the research process — shaping inquiry by offering synthetic fluency that mimics human reasoning while remaining fundamentally distinct from it. The cognitive map developed here suggests that researchers already navigate a shifting terrain in which AI occupies multiple, unstable roles. These roles must be understood not only functionally but also epistemologically.

The most important outcome of this inquiry is not a definitive claim about AI's utility, but a call for *reflective integration*. Generative systems can assist, provoke, and accelerate research — but their contributions must be framed within a broader understanding of what constitutes intellectual labor, methodological rigor, and disciplinary meaning. Misinterpreting convenience for comprehension, or automation for insight, risks eroding the very foundations of scholarly knowledge.

Looking forward, applications of this work may extend to curriculum design, authorship policy, or the development of AI-aware research infrastructures. More abstractly, the findings suggest that scholarship in the age of generative AI will require not just new tools, but new habits of mind. Critical literacy—about systems, signals, and sources—will become as essential as technical fluency.

In this light, generative AI is not merely an instrument of change but a mirror: reflecting back the priorities, assumptions, and aspirations of those who build, deploy, and rely upon it.

ACKNOWLEDGMENT

The author gratefully acknowledges the support and engagement of students, staff and faculty at Lyon College. Valuable discussions and collaboration with Wesley Beal, Cristian Del Gobbo, Pietro Dall'Olio, Harald Kjellin, and Dave Sonnier also informed the development of this paper. Generative AI tools were used during the drafting process.

REFERENCES

[1] A. Ganguly, A. Johri, A. Ali, and N. McDonald, "Generative artificial intelligence for academic research: Evidence from guidance issued for researchers by higher education institutions in the United States," *AI and Ethics*, vol. 5, pp. 3917-3933, 2025. DOI: 10.1007/s43681-025-00688-7.

[2] S. Zuboff, *The Age of Surveillance Capitalism*. New York, NY: Public Affairs, 2019.

[3] R. Van Noorden, "More than 10,000 research papers were retracted in 2023 — a new record," *Nature*, 12 Dec., 2023. [Online]. DOI: 10.1038/d41586-023-03974-8

[4] National Science Board, *Publications output: U.S. trends and international comparisons* (NSB-2023-33), National Science Foundation, 2023. [Online]. Available: https://ncses.nsf.gov/pubs/nsb202333

[5] O. C. Robertson, L.-E. Becerra-Garcia, C. J. Vorland, and D. B. Allison, "When scientists don't correct errors, misinformation and deadly consequences can follow," *Scientific American*, 25 Mar., 2025. [Online]. Available: https://www.scientificamerican.com/article/when-scientists-dont-correct-errors-misinformation-and-deadly-consequences/.

[6] M. Májovský, M. Černý, M. Kasal, M. Komarc, and D. Netuka, "Artificial Intelligence Can Generate Fraudulent but Authentic-Looking Scientific Medical Articles: Pandora's Box Has Been Opened," *J. Med. Internet Res.*, vol. 25, art. no. e46924, 2023. DOI: 10.2196/46924

[7] GeeksforGeeks, "Use of AI in Search Engines," Dec. 16, 2024. [Online]. Available: https://www.geeksforgeeks.org/use-of-ai-in-search-engines/.

[8] M. Birkenkrahe, "Medical AI - How Machine Learning Can Help Your Publishing Career," White River Medical Center, Batesville, AR, 8 Aug., 2023.

[9] National Science Board, "Publication output by region, country, or economy, and by scientific field," *Science and Engineering Indicators*, National Science Foundation, 2024. [Online]. Available: https://ncses.nsf.gov.

[10] F. Dell'Acqua et al., "Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity," Working paper no. 24-013, Harvard Business School, 2023. [Online]. Available: https://www.hbs.edu/ris/Publication%20Files/24-013_d9b45b68-9e74-42d6-a1c6-c72fb70c7282.pdf.

[11] E. Ferrara, "Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies," *Science*, vol. 6, no. 1, p. 3, 2023. DOI: 10.3390/sci6010003.

[12] B. Perrigo, "Google DeepMind CEO Demis Hassabis on AGI and AI in the Military," *Time*, 27 Apr., 2025. [Online]. Available: https://time.com/7280740/demis-hassabis-interview/.

[13] OpenDataScience, "Bill Gates predicts AI will replace doctors and teachers within a decade," 28 Mar., 2025. [Online]. Available: https://opendatascience.com/bill-gates-predicts-ai-will-replace-doctors-and-teachers-within-a-decade/.

[14] McKinsey & Company, "The state of AI: Global survey," 2025. [Online]. Available: https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai.

[15] R. Bean and T. H. Davenport, "Five Trends in AI and Data Science for 2025," *MIT Sloan Management Review*, 2025. [Online]. Available: https://sloanreview.mit.edu/article/five-trends-in-ai-and-data-science-for-2025/.

[16] V. Havlík, "Meaning and Understanding in Large Language Models," arXiv preprint, Oct. 2023. [Online]. Available: https://arxiv.org/abs/2310.17407.

[17] R. Sullivan and N. Elsayed, "Can Large Language Models Act as Symbolic Reasoners?" arXiv preprint, Oct. 2024. [Online]. Available: https://arxiv.org/abs/2410.21490.

[18] A. Arrieta-Ibarra, G. Gopalakrishnan, A. Hernandez, J. Hernandez-Orallo, and S. Bryson, "Temporal quality degradation in AI models," *Scientific Reports*, vol. 12, art. no. 12143, 2022. DOI: 10.1038/s41598-022-15245-z.

[19] E. Creely, "Conceiving Creativity and Learning in a World of Artificial Intelligence: A Thinking Model," in: Henriksen, D., Mishra, P. (eds) *Creative Provocations: Speculations on the Future of Creativity, Technology & Learning*, Creativity Theory and Action in Education, vol 7., Springer, Cham, 2022, pp. 35-50. DOI: 10.1007/978-3-031-14549-0_3.

[20] Plato, *The Republic*, trans. B. Jowett, Oxford University Press, 2008.

[21] Matteo, D'Ambrosio,"Decoding and Coping with the Forms of Fragmentation of Knowledge," PhilArchive, 2023. [Online]. Available: https://philarchive.org/archive/MALDAC-4.

[22] B. Leiter, "Free Speech on the Internet: The Crisis of Epistemic Authority," *Daedalus*, vol. 153, no. 3, pp. 91-105, Summer 2024. DOI: 10.1162/daed_a_02091.

[23] A. De Cesaris, "Digital Metempsychosis: The Technological Mediation of Subjectivity," *Techné: Research in Philosophy and Technology*, vol. 27, no. 3, pp. 287-307, 2023. DOI: 10.5840/techne202387185.

[24] E. Brynjolfsson and A. McAfee, *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*, New York, NY, USA: W. W. Norton & Company, 2014.

[25] M. Zeff, "OpenAI's new reasoning AI models hallucinate more," *TechCrunch*, 18. Apr., 2025. [Online]. Available: https://techcrunch.com/2025/04/18/openais-new-reasoning-ai-models-hallucinate-more/.

[26] P. Mehta, M. Bukov, C.-H. Wang, A. G. R. Day, C. Richardson, C. K. Fisher, and D. J. Schwab, "A high-bias, low-variance introduction to Machine Learning for physicists," *Physics Reports*, vol. 810, pp. 1–124, 2019. DOI: 10.1016/j.physrep.2019.03.001.

[27] M. Bäker, T. Kalkreuter, G. Mack, and M. Speh, "Neural multigrid for gauge theories and other disordered systems," *Int. J. Mod. Phys. C*, vol. 4, no. 2, pp. 239-247, 1993. Presented at the 4th International Conference on Computational Physics PC '92, Prague, Aug. 1992.

[28] P. Wack, "Scenarios: Uncharted waters ahead," *Harvard Business Review*, vol. 63, no. 5, pp. 73-89, 1985.

[29] J. and W. Grimm, *Grimm's Fairy Tales*, trans. M. Hunt, New York, NY, USA: Penguin Classics, 2003.

[30] Higher Education Policy Institute, "Student generative AI survey 2025," 26 Feb., 2025. [Online]. Available:

https://www.hepi.ac.uk/2025/02/26/student-generative-ai-survey-2025/.

[31] Anthropic, "How university students use Claude: Findings from a 5,000-student survey," 10 Apr., 2024. [Online]. Available: https://www.anthropic.com/news/anthropic-education-report-how-university-students-use-claude.

[32] T. Tao, "The potential of AI in science and mathematics," YouTube, 13 Dec., 2024. [Online]. Available: https://www.youtube.com/watch?v=_sTDSO74D8Q.

[33] J. Liu, Z. Wang, J. Xie, and L. Pei, "From ChatGPT, DALL-E 3 to Sora: How has generative AI changed digital humanities research and services?" arXiv preprint, Apr. 2024. [Online]. Available: https://arxiv.org/abs/2404.18518.

[34] K. Knibbs, "Most US Teens Use Generative AI. Most of Their Parents Don't Know," *Wired Buisiness*, 18 Sept., 2024. [Online]. Available: https://www.wired.com/story/teens-generative-ai-use-schools-parents/.