The Philosopher's Stone for Science – The Catalyst Change of AI for Scientific Creativity

Qian Chen (quc20@psu.edu)
Smeal College of Business, Pennsylvania State University, University Park, PA

Yi-Jen (Ian) Ho (ian.ho@tulane.edu)
Freeman School of Business, Tulane University, New Orleans, LA

Pin Sun (pin.sun@ruc.edu.cn)
School of Applied Economics, Renmin University, Beijing, China

Dashun Wang (dashun.wang@kellogg.northwestern.edu)
Kellogg School of Management, Northwestern University, Evanston, IL

Abstract

Limited research studies the impact of AI on scientific creativity. The work investigates whether AI is a catalyst for scientific creativity and what theoretical explanations behind observed AI-supported creativity are. Employing the Logical Creative Thinking (LCT) framework, we conjecture that AI enhances scientific creativity by providing faster search algorithms and offering possibilities to explore new search paths for uncovered knowledge. AI is expected to facilitate creative knowledge *hybridization* (i.e., *recombination* in LCT) across fields and serve as a stimulus for knowledge mutation (i.e., *replacement* in LCT) within a field. Accordingly, we consider two measures of scientific creativity: novelty (as hybridization) and disruption (as mutation). To quantify the AI impact, we analyze the publications from 2000 to 2021 and their citation networks. Our findings first inform that AI increases the novelty of mediocre (medium-level) and the top (90th-percentile) papers while enhancing the disruption of the mediocre papers only. Second, we identify nuanced variations in the impact on creativity across fields. Specifically, AI has the strongest and least impacts on basic science and humanity, respectively. Third, our citation-network analyses further uncover the direct and indirect effects of AI. Citing AI-related papers from other fields fosters novelty due to the hybridization of diverse techniques. Yet, citation concentration within specific or own fields leads to an indirect negative impact on novelty. As for disruption, we observe a similar pattern applied to mediocre papers only. For the most disruptive papers, over-emphasizing specific fields and references still hurts scientific mutation, whereas the increase in within-field citations deepens the understanding of prior work and illuminates new reasoning paths. Overall, the theoretical mapping and comprehensive analyses shed light on the intricate dynamics between AI and the evolution of creative research landscapes.

Keywords: artificial intelligence, scientific creativity, logical creative thinking (LCT), citation network

Electronic copy available at: https://ssrn.com/abstract=4749118

"*Imagination is more important than knowledge.*
Knowledge is limited. Imagination encircles the world."

– Albert Einstein

1. Introduction

Creativity is at the heart of the imaginative progress in science, technology, and society, driving humanity forward with revolutionary innovations and game-changing ideas. Understanding where creativity comes from remains puzzling. To solve the puzzle, the work of analytical creativity (e.g., Altshuller 1956; Ding 2020) believes that a creative process is never random but learnable, practicable, and replicable to produce valuable products (or work) across fields. Specifically, the *Theory of Inventive Problem Solving* (TRIZ) by Altshuller (1956, 1984, 1996) focuses on technical creativity and argues that heuristics are systematically used to create technological innovation. Generally, the *Logical Creative Thinking* (LCT) theory by Ding (2020) interprets creative tasks as *search* problems, which take place in every domain. Thus, the objective of creativity then becomes to search for novel and valuable outcomes within a broad theoretical space while minimizing search costs. In other words, LCT regards innovation (or invention) as a search process, and innovators (or inventors) as explorers.

Scientific creativity is the application of analytical creativity using scientific methods to search for new findings. Scientific creativity is undoubtedly significant, expanding the frontier of knowledge and thus contributing to societal betterment through continual exploration and enlightenment. It involves a dynamic and inventive process, wherein scientists formulate novel ideas, theories, and solutions for scientific inquiry. Scientific creativity requires departing from established conventions, challenging longstanding assumptions, and searching unexplored theoretical spaces. Prior research studies scientific creativity via four dimensions: creative *product* (*work*), *people*, *process*, and *situation* (*environment*), as Stumpf (1995) summarizes.

We apply the LCT framework to study scientific creativity with a focus on the product and process dimensions. From the product perspective, creative work (scientific publications) should be innovative and impactful (Ochse 1990). The community objectively measures a paper's creativity using its citations (or references). Beyond oversimplified citation counts, *novelty* (Uzzi et al. 2013) and *disruption* (Funk &

Electronic copy available at: https://ssrn.com/abstract=4749118

Owen-Smith 2017) are developed in the Science of Science domain. The novelty score means the extent to which the existing knowledge is utilized differently by a focal publication. It considers how atypical the focal paper's reference composition is, compared with those expected by random chance. The disruption score captures the extent to which a focal paper disrupts or develops the existing knowledge. It examines the likelihood that subsequent studies cite the focal paper rather than its references. To exemplify the two measures, the *hybridization* and *mutation* mechanisms in the *Theory of Evolution* serve as vivid analogies for novelty and disruption, respectively. Novel scientific findings originate from hybridizing seemingly unrelated knowledge (e.g., MRI from medicine and physics) across domains, making unusual combinations of references. The birth of disruptive work (e.g., mRNA) would require mutating the established scientific genes, altering longstanding assumptions, and trying different search paths.

From the process perspective, the LCT framework theorizes that the objective of scientific research is to search for new breakthroughs, which is achieved through the *replacement* and *recombination* methods for knowledge mutation and hybridization, respectively. Ding (2020) metaphorizes replacement as a single-point gene mutation and recombination as chromosome crossovers (in a relatively large scope). In other words, a scientist could find a new replacement for a critical component of her research via *abstraction* (i.e., similar) or *contrarian* searches. Alternatively, she may go broad and combine the existing knowledge across context-theoretical space using *sharing* (i.e., combining different solutions), *arbitrage* (i.e., learning from successful solutions in different contexts), and the other two methods.

Powerful, logical search methods are always needed to boost scientific creativity. Indeed, artificial intelligence (AI) is reshaping the search process of scientific research by offering faster, more precise tools and algorithms, which are more efficient than humans for specific inquiries.[1] Specifically, AI algorithms (e.g., generative AI and text-mining) accelerate the process by navigating scientists to identify (1) similar (or competing) theories and techniques to replace prior assumptions and formulations (i.e., mutation), and (2) successful (or enhancing) solutions across contexts to recombine prior knowledge to have explore new

[1] AI is broadly defined in this study, including but not limited to machine learning, convolutional neural network, deep learning, natural language processing, and image processing, along with other popular technologies.

Electronic copy available at: https://ssrn.com/abstract=4749118

solution paths (i.e., hybridization). With the theorization of LCT, we are interested in investigating the role of AI in scientific creativity by asking the following questions: (1) *Would artificial intelligence improve scientific creativity?* (2) *If so, would the improvements vary with the distinctive natures of discipline fields?* (3) *Would artificial intelligence alter the scientific creation process regarding knowledge hybridization and mutation?*

To answer the above questions, we collect data from SciSciNet (Lin et al., 2023), which is a large-scale open data set encompassing over 134 million academic publications.[2] Given our interest in AI, our analysis focuses on the 93 million papers published between 2000 and 2021 and their corresponding citation networks. A keyword-based approach (Gao and Wang 2023) is applied to divide the 93 million publications into the AI and non-AI-related sections using general AI keywords.[3] We construct the citation networks of the two sections across 19 research fields on an annual basis.[4] An annual citation network consists of filed nodes and directed citation edges among them (including self-loops as self-citations). The network structure captures the knowledge flow (i.e., mutation and hybridization patterns) across fields. As for the outcomes of interest, we consider each paper's characteristics (e.g., its novelty score) and aggregate them to the field level as the node attributes (e.g., the distribution of the novelty scores). To be comprehensive, we append other field-level information (e.g., the number of papers and funding). The data operationalization results in 794 observations at the field-year level.

Our empirical strategies center on fixed-effect regression models and network analyses. First, we study the impact of AI on scientific creativity by comparing AI- and non-AI-related papers. Since the *top* creative work may have a distinct nature from its *mediocre* counterpart, we conduct subsample analyses by comparing the publications at the 90^{th} and 50^{th}-percentiles of each creativity score. As a result, we have our

[2] SciSciNet is built upon the Microsoft Academic Graph (MAG) dataset, renowned as one of the largest bibliometric datasets and widely utilized for the Science of Science research.

[3] The general AI keywords used are the Top 25 most frequently mentioned AI terms in scientific research in 2019. The complete list of these keywords is provided in Online Appendix A.

[4] The MAG dataset considers 19 major fields: Art, Biology, Business, Chemistry, Computer Science, Economics, Engineering, Environmental Science, Geography, Geology, History, Materials Science, Mathematics, Medicine, Philosophy, Physics, Political Science, Psychology, and Sociology.

Electronic copy available at: https://ssrn.com/abstract=4749118

models in a 2x2 framework (i.e., the two creativity dimensions at the two different percentiles). Second, we consider the variation in the AI impact due to field-specific search needs across fields. To have a big picture, we group 19 fields into four disciplines – Basic Science, Social Science, Applied Science, and Humanities. As AI is more efficient in logic searches, we expect the moderating effects of disciplines. Last, we utilize the network structure to understand how AI alters knowledge creation using mediating analyses, wherein the observed impact is attributed to the direct effect (AI → creativity) and the indirect effect (AI → citation networks → creativity).

We show an overall positive impact of AI on scientific creativity. First, the novelty of both top and mediocre papers (i.e., papers at the 90th and 50th-percentile novelty in their field) in the AI-related section is significantly higher than in the non-AI-related section. Using AI technologies increases the novelty scores of mediocre papers by 0.69 standard deviations and that of the top novel paper by 0.76 standard deviations. In addition, AI makes mediocre work (i.e., the paper at the 50th-percentile disruption) more disruptive but not for the top (i.e., the paper at the 90th-percentile disruption), as shown a 0.40 standard deviation increase in the former's disruption index.

Second, our investigation discovers the heterogeneous impacts of AI across research disciplines. In specific, basic science enjoys the most substantial augmentation, followed by applied and social sciences (with moderate and insignificant impacts, respectively). However, humanity subjects experience the least (or even negative) AI influence. These nuanced variations underscore valuable insights into the interplay between AI and distinct academic domains. In short, the more analytical creativity a discipline emphasizes, the more creativity gains AI delivers.

Last, our citation-network-based mediating analyses reveal significant direct and indirect AI effects. On the one side, the direct effect is the key route through which AI catalyzes unconventional recombination of prior knowledge and replacement of established thinking. In this regard, AI serves as a search algorithm, lowering search costs and *reconfiguring* the current research designs to obtain new insights.[5] On the other

[5] Reconfiguration (besides replacement and recombination) is one of the three search methods in the LCT framework (Ding 2020), wherein researchers decompose a concept or solution into small pieces and reform them to create new variants.

Electronic copy available at: https://ssrn.com/abstract=4749118

side, AI indirectly guides researchers to move away from longstanding, domain-specific mindsets and embrace new research paths. Such an indirect effect significantly alters the knowledge flows among fields, leading to a paradigm shift of knowledge hybridization and mutation. We summarize the indirect effect on citation composition (i.e., mediators), novelty, and disruption.

First, AI-related publications cite more cross-field and within-field references by 53% and 102%, respectively, compared with non-AI-related ones. The former enhances creativity in general, whereas the latter deters it. In addition, additional cross-field references are concentrated on fewer specific fields. Such cross-field reference concentration lowers knowledge hybridization and hurts creativity. Second, AI has a negative indirect effect on novelty, overall. While AI promotes cross-field hybridization for novel research, the negative impacts of overly concentrated cross-field references and exaggerated within-field references are overwhelming. Last, a similar pattern repeats when we study mediocre disruptive research. By contrast, we find a neutral indirect effect for the top disruptive research. On the one hand, as LCT theorizes earlier, the most disruptive work relies on replacement (knowledge mutation) rather than recombination. Citing more within-field work is critical to staying focused on breakthroughs by replacing a point gene of the focal research. On the other hand, additional cross-field references distract disruptive scientists, and the negative distraction is worsened by the associated concentration.

Our literature contributions have three layers. First, we enrich the analytical creativity literature by applying the LCT theory to empirically investigate scientific creativity. Grounded in the LCT theory, our investigation explores how AI functions as powerful search algorithms and stimuli to accelerate logical search processes for creation. Through our explorations, we deepen our understanding of the LCT theory. Specifically, we map its replacement and recombination to the roles of scientific mutation and hybridization in creative research. Second, we advance the scientific creativity literature by providing novel empirical analyses powered by large-scale citation networks, beyond descriptive analyses (e.g., Gao and Wang 2023). Our network analysis illustrates how the knowledge flows (within and across fields) catalyze scientific mutation and hybridization, ultimately boosting disruption and novelty for academic publications. Last, we contribute to the literature on AI and creativity. While extant research has focused on AI-enhanced artistic

Electronic copy available at: https://ssrn.com/abstract=4749118

creativity (e.g., Ali Elfa & Dawood 2003), our study stands out as one of the pioneers in exploring the impact of AI on scientific creativity. We consider the moderating effect of discipline heterogeneity and the mediating effects of information flow by comparing AI-related populations and their non-AI-related counterparts. We uncover how AI functions as a general-purpose tool, catalyzing the novelty and disruption of research across fields. This comprehensive analysis sheds light on the intricate dynamics between AI and scientific creativity. Overall, we answer our research questions by hybridizing the knowledge of the analytical, scientific, and AI creativity literature (i.e., the creative process and methods of LTC, creative product measures, and AI-related publications).

Our study also provides valuable implications for several stakeholders. First, for scientists, utilizing AI algorithms and techniques does produce more creative work. AI enables scientists to implement LCT's reconfiguration, replacement, and recombination methods effectively. Indeed, scientists can easily leverage AI to efficiently (1) collect, process, and visualize big data, (2) operationalize new constructs, (3) screen much more related work to frame hypotheses, (4) test various experiment designs (in simulations) and tune parameters with a minimal cost, and (5) validate models and results. In addition, scientists should be aware that AI influences novelty and disruption differently, resulting in two scenarios. On the one side, AI brings novel research by promoting recombining cross-filed knowledge. Thus, scientists should be more proactive in "borrowing" knowledge from other fields to create novel aspects. On the other side, AI supports the most disruptive work by pointing out critical components of within-field knowledge as the replacement of old mindsets. As a result, scientists can continue focusing on deepening the understanding of a specific topic, eventually developing disruptive work. Such a use case is particularly practical for basic science.

Second, our empirics provide guidance for developers to design new AI tools to boost scientific creativity. Specifically, general AI could be further refined to support novel and disruptive research. For instance, novelty-enhancing AI emphasizes LTC's recombination methods by searching for and combining different solutions (sharing) in different contexts (arbitrage). In this case, a large language model (LLM) for biologists trained by extra texts in adjacent fields (e.g., chemistry and engineering) to learn solutions not uncommon in biology. By contrast, disruption-boosting AI implements LTC's replacement methods by

Electronic copy available at: https://ssrn.com/abstract=4749118

finding similar or opposite concepts (abstraction and contrarian) to change mindsets. Thus, an LLM should provide more within-field contrasts for physicists to go deep and disruptive (rather than broad and novel). However, we worry that AI-related references are overly concentrated on certain domains (i.e., Computer Science, Engineering, and Psychology), perhaps hurting the overall creativity in the long run. To promote AI dissemination and adoption, developers may create more stylized algorithms to fit heterogeneous needs. By decentralizing AI, developers will help scientific creativity grow exponentially.

Last, policymakers should take the initiative to build a healthy ecosystem whereby AI-developing and AI-using scientists benefit each other. To incentivize AI research, officers may provide extra funding opportunities for AI research and development in academic institutions, facilitating the creation of AI tool, algorithm, and interdisciplinary infrastructure. Establishing funding programs tailored for interdisciplinary AI research not only drives innovation but also enlarges the academic audience. To encourage innovative AI use cases, investing in AI education programs equips researchers, educators, and students with essential skills for AI utilization in academic research and teaching. In addition, it will be useful to promote open data (and model) initials that foster cross-field collaboration and within-field reconfiguration.

The rest of this paper is organized as follows. Section 2 reviews related literature. We introduce our empirical setting and describe our data. Sections 4 and 5 cover the AI impacts on creativity and the related mediating effects, respectively. We conclude this research by providing executable implications for stakeholders in Section 6.

2. Literature

We review the related creativity literature in a three-layer fashion. We start with the understanding of the analytical creativity theories. Then, the discussion transits from a general framework of analytical creativity to a valuable domain of scientific creativity. Last, we specifically look at the research gap of the AI and scientific creativity intersection.

2.1. Analytical Creativity

Analytical creativity is the specific form of creativity that generates reproducible creative results by learning

Electronic copy available at: https://ssrn.com/abstract=4749118

the innovative process. The advocates of analytical creativity believe that innovation is never a random process but a learnable and replicable one. The most presentative work in this literature stream consists of Altshuller's *Theory of Inventive Problem Solving* (1970) and Ding's *Logical Creativity Thinking* (2020). The two illustrate how creativity is perceived and formulated as follows.

The Theory of Inventive Problem Solving (Romanized from *Teoriya Resheniya Izobretatelskikh Zadach*) is a philosophical and methodological combination, including a set of problem definitions and solving tools and strategies (Altshuller 1996). The originality of TRIZ dates back to 1946 (Altshuller and Shapiro 1956) with an attempt to characterize technical creativity. Focusing on pattern invention, Altshuller exerts multi-year efforts to analyze over two hundred thousand significant patents worldwide. Altshuller argues that technical innovation is not a random process of uncontrolled (against analytical) creativity but involves learning from prior knowledge and leveraging it to come up with new, successful solutions (Souchkov 2015). After years of development, TRIZ debuted in 1970 eventually. Altshuller (1985) further extends TRIZ to ARIZ-85C, a comprehensive algorithm of 32 steps for solving inventive problems.

The Logical Creative Thinking (LCT) paradigm provides a unique perspective on the innovative nature of research. It reinterprets creativity by seeing innovation as a search process and innovators as *explorers*. Under LCT, forthcoming innovations are theoretically possible but practically undiscovered. Thus, the challenge of innovation centers on employing and developing search algorithms to accelerate knowledge discovery. The LCT paradigm underscores the interplay of *logical* and *creative* thinking for shaping the trajectory of research endeavors. In particular, logical thinking serves as a guiding force, directing researchers to adhere to existing rules, facilitating precise inferences, and guiding subsequent steps. By contrast, creative thinking encourages a departure from established rules, empowering researchers to explore new routes that may lead to unexpected and groundbreaking findings. LCT posits that innovation combines both types of thinking as a logical process to produce unexpectedly valuable outcomes. LCT derives three tiers of methods to implement in the innovation process. The first-tier methods (including reconfiguration, replacement, and recombination) are derived from fundamental innovation principles. The second and third-tier methods specify innovation principles and implementable applications, respectively.

Electronic copy available at: https://ssrn.com/abstract=4749118

Our paper attempts to contribute to the analytical creativity literature by applying conceptual LTC to concrete scientific creativity. To the best of our knowledge, we are among the first to provide empirical insights to support the LCT arguments. Our empirics speicifically show that replacement and recombination significantly promote scientific creativity via knowledge mutation and hybridization. Our work highlights the versatility of the LCT framework and underscores its relevance in the domains of advanced technology.

2.2. Scientific Creativity

What drives creative work remains a puzzle. Different approaches and viewpoints are applied to get close to the answer. As Stumpf (1995) surveys, prior literature studies scientific creativity from four perspectives: the creative *product*, *person*, *process*, and *situation* (or *environment*). In the following discussion, we briefly go over each of them for completeness while focusing on the work and process dimensions more related to our work.

Creative Product (Work)

Ochse (1990) synthesizes prior definitions of creative products, stating that scientific work must be original (new, unusual, novel, and unexpected) and valuable (useful, good, adaptive, and appropriate). Yet, under such an abstract definition, judges evaluate scientific work subjectively. To alleviate potential subjectivity, the field develops two classes of objective methods, including *expert rating* and *citation-based* metrics. On the one side, the former relies on domain experts' ratings to evaluate publication quality. Domain experts rate a study based on predefined creative standards, such as the Creative Product Analysis Matrix (Besemer & Treffinger 1981) and the Creative Product Semantic Scale (Besemer & O'Quinn 1986, 1989, and 1999). Note that this method remains subject to experts' subjectivity.

On the other side, the latter infers the creativity of a study using its citation-based information. The most straightforward index is *citation counts*. However, it may overlook cross-discipline heterogeneities. For example, a study's citation counts highly depend on the size of the community to which it belongs. It is unfair to compare the citation counts of a pure-mathematics paper and a machine-learning paper. Other factors, such as writing languages and topic popularity, also affect citation counts. Two neat measures are developed to better identify creative work via citation-based data: *Novelty* (Uzzi et al. 2013) and *Disruption*

Electronic copy available at: https://ssrn.com/abstract=4749118

(Owen-Smith 2017).

The novelty measure examines how uniquely a focal paper combines prior work (from the theoretic parameter space) by checking its reference pairs. The more unconventional its reference pairs are, the more novel the work is. The disruption measure considers how much a focal paper advances the existing literature (e.g., developing a new theoretical framework or creating a new research path). More disruptive work has a higher ratio of its subsequent work citing it exclusively to its subsequent work citing both itself and its references. This research considers novelty and disruption measures to objectively evaluate two distinct dimensions of scientific creativity.

Creative Person

The literature also analyzes the personality traits of scientists and associates them with scientific creativity (e.g., Ochse 1990; Rushton et al. 1987). Rushton et al. (1987) find that researchers with low sociability, aggressiveness, dominance, and introversion bring more creative work. Similarly, Feist (1998, 2008) uses the *big five* personality traits to show that creative scientists are more conscientious and open to experience, dominant and assertive, and achievement-driven and ambitious. In addition, Grosul and Feist (2014) buster the myth about the link between mental illness (or craziness) and creativity. Creative scientists exhibit lower levels of psychoticism, hostility, impulsivity, and coldness, echoing that creative work is built on a focused, thoughtful, and deliberate approach.

The people-related discussion extends to research teams beyond individuals. For team sizes, Wu et al. (2019) find that smaller teams are most interested in generating disruptive ideas in the fields of science and technology, whereas larger teams are formed to advance the existing knowledge. For team formation, Zheng et al. (2022) show the positive relationship between a team's expertise diversity and its scientific creativity. Thus, while authors' personality traits cannot be accessed in this study, we incorporate research team size as a control variable to measure the characteristics of creative people.

Creative Process

A creative process is a sequence of thoughts and actions (in stages), resulting in innovative and adaptive scientific products (Lubart 2001). Wallas (1926) first proposes the classic four stages of a creative process.

Electronic copy available at: https://ssrn.com/abstract=4749118

First, *preparation* involves screening, defining, and setting up a research problem. Second, *incubation* (as a search process) refers to working on the problem without an agenda and forming associations among (or combinations of) ideas and concepts. Third, *illumination* describes sudden enlightenment or a flash of a solution. Last, *verification* is developing, evaluating, and refining the enlightening solution. The following work extends the classic four-stage model by adding more phases. For instance, some think that there should be a question-identifying (or -formulating) phase before the preparation stage (e.g., Amabile 2011; Getzels & Csikszentmihalyi 1976; Osborn 1953) and a frustration phase after (e.g., Goleman et al. 1992; Hutchinson 1949). Amabile (1996) and Stein (1974) append a deployment and communication phase as the final stage.

Besides, other studies look at specific subprocesses of scientific creation. For example, Getzels & Csikszentmihalyi (1976) study problem-finding, formulation, and redefinition. They note that individuals who start painting without a specific plan create more creative artwork than those with predetermined ideas. The literature also notices the significance of divergent thinking that generates numerous alternative ideas (e.g., Guilford 1967; Khandwalla 1993; Runco 1991).

We quantitatively study creative processes using LTC methods as the theoretical foundation. Prior research qualitatively investigates creative mechanisms using information processing, such as analogy and metaphor (Ward et al. 1997; Weisberg 1986) and remote associations (Mednick 1962). We supplement the literature by looking at the process dimension across academic disciplines from an ecosystemic perspective, whereas prior research focuses on specific mechanisms from an application-level perspective.

Creative Situation (Environment)

Environmental factors – social, cultural, political, and historical background – affect scientific creativity. Ochse (1990) shows how creativity varies across different religious groups. Others investigate the impacts of political factors (e.g., internal war, external threat, political instability, and civil disturbances) on societal creativity (e.g., Simonton 1990). Science historians consider that particular discoveries or the evolution of some theories naturally take place at certain junctures in history. While Merton (1961) argues that identical discoveries are frequently made by independent scientists simultaneously (i.e., *multiples*), Simonton (1988, 2004) shows that creation is more like a stochastic process using Monte Carlo simulations. More recently,

Electronic copy available at: https://ssrn.com/abstract=4749118

Heinze et al. (2009) note the importance of organizational and institutional drivers on scientific creativity, including technical skills, research sponsorships, access to extramural skills and resources, and leadership.

In this paper, our focus centers on both the creative product and the creative process. Specifically, we investigate how AI accelerates scientific creativity, gauged through measures of novelty and disruption. To gain deeper insights into the creative process, we leverage large-scale citation networks and conduct network analysis. Our study delves into how the flow of information and knowledge induced by AI, both within and across fields, facilitates scientific hybridization and mutation, thereby influencing the levels of novelty and disruption observed in academic publications. Our research enriches theoretical discussions of scientific creativity and provides practical implications for fostering innovation in scientific research.

2.3. AI and Creativity

AI is driving a revolutionary era. As AI continues evolving, its impact on creativity grows exponentially, attracting more devoted researchers. While AI-supported creativity disrupts scientific (quantitative) and artistic (qualitative) fields, the ongoing literature is exploring AI's role in the latter (e.g., Miller 2019; Figoli et al. 2022). AI is generally seen as creativity enhancement when humans collaborate with it. As Bonnardel and Zenasni (2010) argue, AI (similar to other support systems) remains a computational tool for designers. It potentially helps idea-divergent and -convergent processes of creativity. In addition, Liapsis et al. (2016) show that AI sometimes offers humans random stimuli in creative processes. Such offerings might break the designers' calcified reasoning patterns (or even biases) and inspire designers to explore new creative paths. Recently, Ali Elfa and Dawood (2023) discover that AI helps artists generate new compositions from existing work by analyzing enormous images efficiently. Also, Tigre Moura et al. (2023) suggest that integrating AI into the production process promotes the audience's perception of innovation. However, the audience may discredit a piece of artwork's creativity if told that it is made by AI rather than an artist. Similarly, people feel that artists spend less effort in creation with the help of generative AI. Last, Wu et al. (2021) demonstrate that collaborating with AI is more beneficial than competing against AI. It complements scientists' intelligence by assembling all of the existing knowledge in the humanity discipline. A Human-AI Co-Creation model is proposed to explain the creative dynamics, wherein great human-AI integration

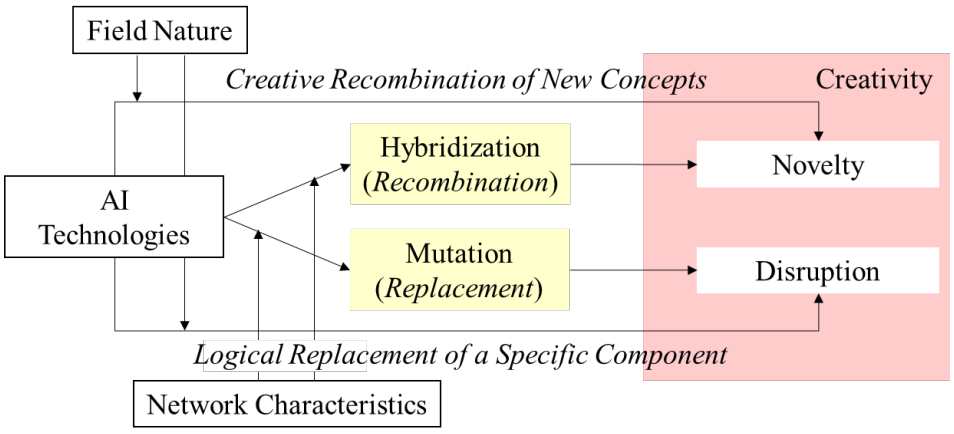
Electronic copy available at: https://ssrn.com/abstract=4749118

makes creative processes more accessible and inclusive.

The impact of AI on scientific creativity receives relatively limited attention than the above artistic counterpart. Ideally, Colton and Steel (1999) feel that AI can support the following scientific creation: (1) generating new descriptions of phenomena, (2) innovating new concepts and categorizations, identifying empirical trends, and formulating new hypotheses, (3) identifying examples of a phenomenon, (4) designing experiments, testing hypotheses, and engaging in the closed-loop discovery to elucidate the advancement of a theory, and (5) making explicit the unquestioned assumptions. Due to AI winters, the field has had no significant advancement until recently. Gao and Wang (2023) describe the surging adoption of AI across research fields and highlight AI's significance in scientific publications regarding citation counts within and beyond their respective disciplines. In addition, Kureha (2023) explores the potential of AI in making creative discoveries and the feasibility of automating the process. While there is no backlash against AI-enabled creative discovery theoretically, state-of-the-art AI does not achieve the goal practically.

We probe into whether AI substantially enhances different types of scientific creativity and explore potential mechanisms by employing comprehensive LCT as the theoretical foundation, well-established creativity measures (i.e., novelty and disruption scores), and rigorous analyses. As Figure 1 illustrates, we compare the publications involving AI technologies and those not. Our novel investigation provides insights into the complex interplay between AI and scientific creativity, enriching the literature in this are of AI.

Figure 1. Conceptual Framework



Electronic copy available at: https://ssrn.com/abstract=4749118

3. Data and Method

We collect our data from SciSciNet (Lin et al., 2023), a comprehensive open data repository tailored for the Science of Science research. Its foundation is rooted in the Microsoft Academic Graph (MAG) dataset, recognized as one of the largest and most expansive bibliometric datasets worldwide. Encompassing over 134 million scientific publications since 1950 and millions of external links to funding and public usage, SciSciNet provides comprehensive information on publication records, including authors, citation records among publications, and upstream funding support and downstream uses in the public domain. Given our interest in AI, our analyses focus on the 93 million papers published between 2000 and 2021, along with their corresponding citation networks.

3.2. Data and Variables

We utilize the keyword-based methodology outlined by Gao and Wang (2023) to identify AI-related papers. Specifically, we compile the Top 25 most frequently mentioned AI terms in scientific research in 2019, as documented by Gao and Wang (2023). The list of these top 25 AI terms is provided in Online Appendix A. In our dataset, any paper containing at least one of these keywords in its abstract is categorized as AI-related, while the remaining papers are considered non-AI-related.

For each paper, we employ the variables *Novelty* and *Disruption* to measure creativity. *Novelty* measures how uncommon a paper's reference pairs are as compared to random chance. Following Uzzi et al.'s paper published in *Science* in 2013, we calculate *Novelty* as follows. For a given paper, we determine the *Novelty* score by calculating the Z-score for every pairing of its referenced journals. Then, we assess the distribution of Z-scores for journal pairs by examining all potential pairs of references cited in this given paper. Out of this distribution, the 10th percentile Z-score represents this focal paper's novelty score. A higher novelty score signifies a greater degree of novelty. Using genetics as an analogy. *Novelty* measures a paper's level of scientific hybridization. *Disruption* describes the degree to which a paper disrupts or advances the existing literature. Following Funk and Owen-Smith (2017) and Wu et al. (2019) published in *Nature*, *Disruption* of a focal paper is calculated through citation networks. Among all subsequent papers, it compares the ratio that exclusively cites the focal paper with the ratio of works that cite both the focal

Electronic copy available at: https://ssrn.com/abstract=4749118

paper and its references. The difference between the abovementioned two ratios is regarded as the disruption score. A higher score indicates a greater degree of disruption, which describes a paper's level of scientific mutation. We also incorporate other paper characteristics. *TeamSize* refers to the number of authors of the paper, while *Institution* quantifies the number of institutions affiliated with the author team. *Citation* denotes the number of subsequent works citing the focal paper, whereas *Reference* is the number of references cited in the focal paper. *Twitter* indicates the number of Twitter mentions of the paper, while NSF or NIH refers to the number of grants associated with the paper.

Table 1 presents the descriptive statistics for paper-level variables. Our dataset comprises a total of 93,916,411 papers published between 2000 and 2021, with 3% being AI-related. The average number of references per paper is 14.39, accompanied by a standard deviation of 27.89. Papers, on average, receive 10.25 citations, with a standard deviation of 66.84. The typical author team size is 3.35, with a standard deviation of 13.18, and the average number of institutions to which the author team belongs is 1.60, with a standard deviation of 1.84. Papers, on average, receive 0.01 NSF grant and 0.05 NIH grant, with standard deviations of 0.17 and 0.50, respectively. The average number of Twitter mentions per paper is 0.59, with a standard deviation of 38.43. Regarding our outcome variables of interest, the average *Novelty* of papers is -45.91, with a standard deviation of 338.30. Meanwhile, the average *Disruption* of papers is 0.00, with a standard deviation of 0.05.

Table 1. Descriptive Statistics of Paper-Level Variables

	N	Mean	Std. Dev	Min	Max
Novelty	93,916,411	-45.91	338.30	-102,810.92	445.27
Disruption	93,916,411	0.00	0.05	-1.00	1.00
AI	93,916,411	0.03	0.17	0	1
Reference	93,916,411	14.39	27.89	0	11,278
Citation	93,916,411	10.25	66.84	0	127,777
TeamSize	93,916,411	3.35	13.18	1	6,760
Institution	93,916,411	1.60	1.84	1	777
NSF	93,916,411	0.01	0.17	0	75
NIH	93,916,411	0.05	0.50	0	192
Twitter	93,916,411	0.59	38.43	0	159,156

Focusing on AI-related papers, we establish annual citation networks across 19 research fields spanning from 2000 to 2021. In these networks, individual nodes represent fields, while the edges are

Electronic copy available at: https://ssrn.com/abstract=4749118

directed and weighted. *Size* of a node represents the number of AI papers published in year *t* in the specific field. The edge weight signifies the number of citations from AI-related papers published in the citing field in year *t* during the time period from the year 2000 to 2021, towards AI-related papers published in the cited field between the years 1950 and *t-1*. Meanwhile, *SelfCitation* represents the number of references citing the papers within the AI papers' own field. Additionally, we formulate node characteristics to signify publication attributes within each field. Field-level data is derived through the consolidation of paper-level variables. In light of the distributions observed in paper-level variables, we choose to compute the medians and 90th percentiles of these paper-level variable values within each field as the field-level variables. This approach ensures that the resulting field-level data effectively captures the overall distributions. Simultaneously, we craft annual citation networks for non-AI papers, mirroring the structure and node characteristics of the AI citation networks. The key distinction lies in the focus on non-AI papers, identified by the absence of AI-related keywords in their abstracts. This differentiation enables a comprehensive analysis of both AI and non-AI paper citation networks. We present the correlation among field-level variables in Online Appendix B.

3.3. Visualization

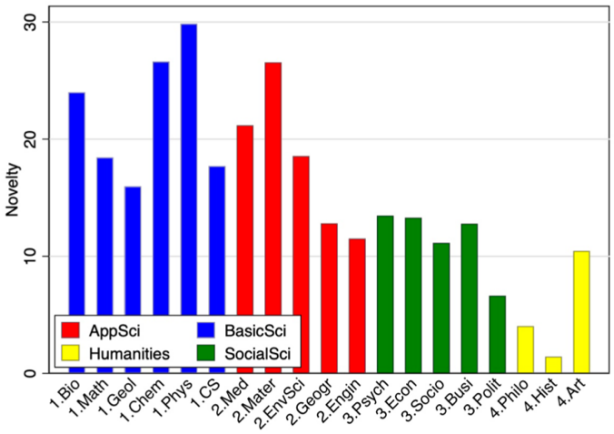
Figure 2 illustrates the novelty and disruption of the papers under examination categorized by field. We present both the median and the 90th percentile novelty (disruption) for AI papers versus non-AI papers in each of the 19 fields. We further classify all fields into distinct discipline areas – Basic Science, Social Science, Applied Science, and Humanities.[6] As shown in Figure 2(a), both mediocre (i.e., 50th percentile) and top (i.e., 90th percentile) novel papers show that the novelty is highest in the fields of Basic Science and Applied Science, while it is lowest in Humanities for both AI and non-AI papers. AI papers consistently exhibit higher novelty scores across the majority of fields compared to non-AI papers. However, the extent of the difference between AI papers and non-AI papers varies by field. Notably, the difference in novelty

[6] We group the 19 fields into 4 discipline areas as follows. (1) Basic Sciences include Biology, Chemistry, Computer Science, Geology, Mathematics, and Physics. (2) Applied Sciences consist of Medicine, Material Science, Environmental Science, Geography, and Engineering. (3) Social Sciences cover Business, Economics, Sociology, and Political Science. (4) Hummanities have Art, History, and Philosophy.

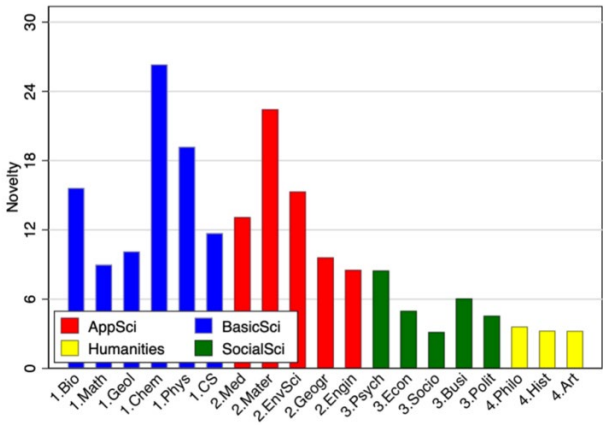
Electronic copy available at: https://ssrn.com/abstract=4749118

between AI and non-AI papers is most pronounced in Basic Science and smallest or even negative in Humanities.

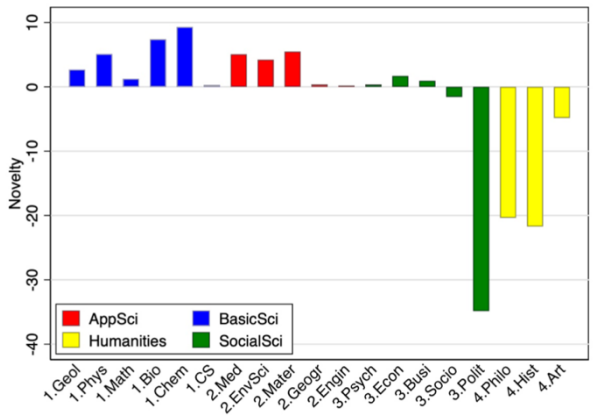
Figure 2a. Novelty of AI and Non-AI papers (by Fields)



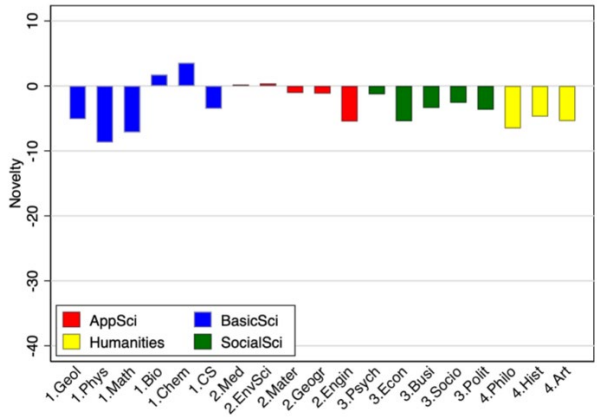
90th-Percentile Novelty of AI Papers



90th-Percentile Novelty of Non-AI Papers



50th-Percentile Novelty of AI Papers

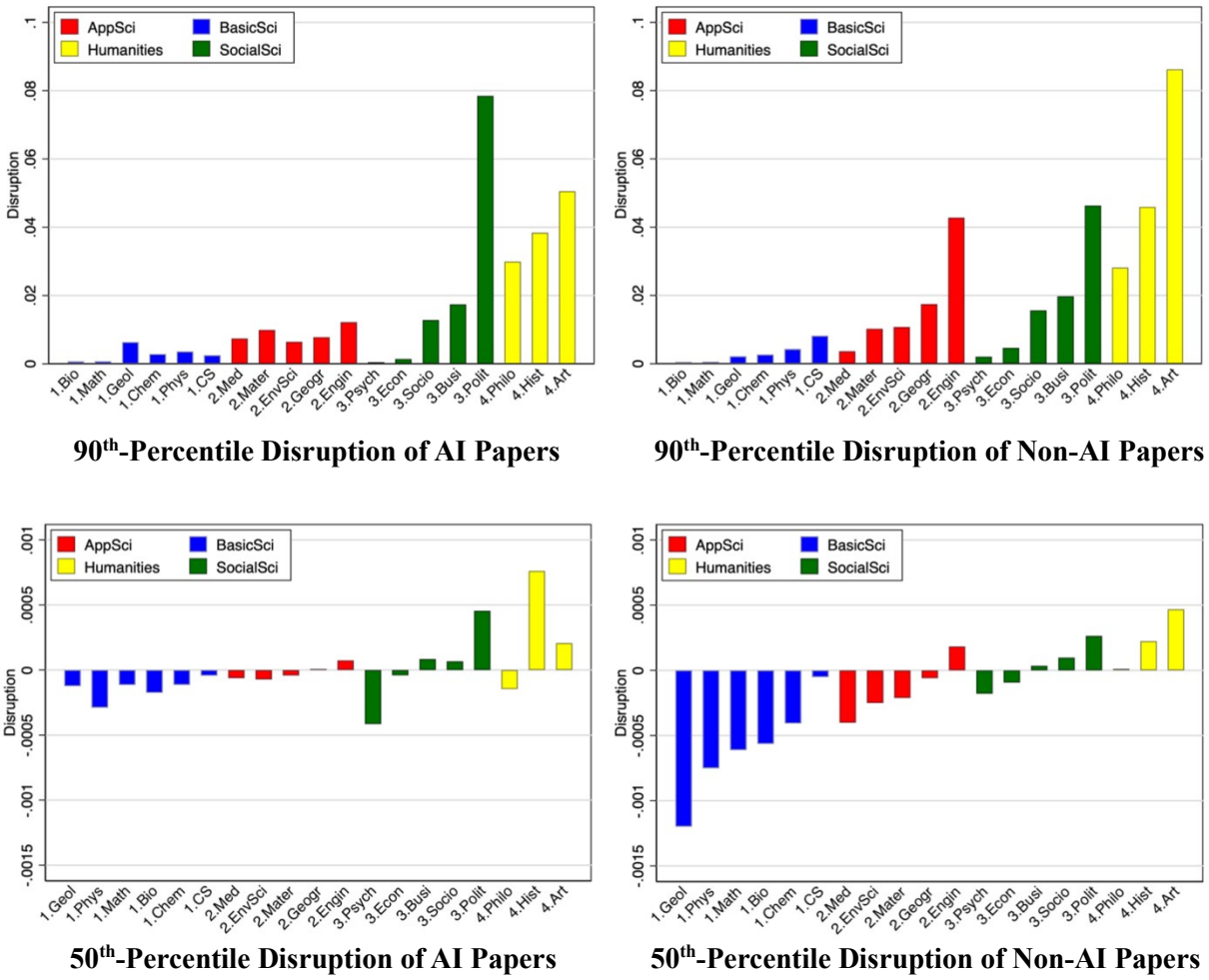


50th-Percentile Novelty of Non-AI Papers

As illustrated in Figure 2(b), both mediocre and top disruptive papers reveal that disruption is predominantly highest in Humanities, followed by Social Sciences and Applied Science, while it is lowest in Basic Science for both AI and non-AI papers. Remarkably, for top disruptive papers, there is no significant difference in disruption between AI and non-AI papers. In the case of mediocre disruptive papers, AI papers demonstrate higher disruption across most fields compared to non-AI papers. The extent of this difference is most pronounced in Basic Science.

Electronic copy available at: https://ssrn.com/abstract=4749118

Figure 2b. Disruption of AI and Non-AI papers (by Fields)



90th-Percentile Disruption of AI Papers

90th-Percentile Disruption of Non-AI Papers

50th-Percentile Disruption of AI Papers

50th-Percentile Disruption of Non-AI Papers

4. Analysis: AI's Impacts on Scientific Creativity

This section covers two sets of empirical analyses, starting with quantifying the impact of AI on scientific novelty and disruption. Given the significance of the impact, we then examine how it varies across distinct academic discipline areas using moderation analyses.

4.1. AI's Overall Impacts on Novelty and Disruption

The main goal of this study is to examine if AI functions as a general-purpose tool to accelerate scientific creativity. To address this question, we examine the impact of AI on scientific creativity by comparing papers related to AI with those that are not. Since highly creative work may exhibit distinct characteristics from mediocre work, we conduct subsample analyses for papers at the 90th and 50th percentiles in terms

Electronic copy available at: https://ssrn.com/abstract=4749118

of creativity indices. Thus, we estimate the models within a 2x2 framework. We employ the following linear regression models:

$$Y_{ijt} = \beta_0 + \beta_1 AI_{jt} + \gamma' X_{ijt} + F_{it} + \varepsilon_{it}, \quad (2)$$

where Y_{ijt} represents the outcome of our interest (*Novelty* or *Disruption*), measuring the degree of creativity in Type j papers (AI papers or non-AI papers) within Field i published in Year t. AI_{jt} is a binary variable indicating whether or not Type j papers published in Year t are AI-related based upon our definition. X_{ijt} includes the attributes of Type j papers within Field i as controlled variables, such as the log(*Size*), log(*TeamSize*), log(*SelfCitation*), *NIH*, *NSF*, and *Twitter*. F_{it} represents *field-year* fixed effects controlling for fields' time trends. Ordinary least squares (OLS) are used to estimate the regression, with the estimated coefficient β_1 helping to identify AI's impact on scientific creativity. The results of the models are listed in Table 2. The dependent variable in Model 1 is a field's 50th-percentile novelty, while in Model 2, it is the 90th-percentile novelty. Similarly, the dependent variable in Model 3 is a field's median disruption, and in Model 4, it is the 90th percentile disruption. We discuss the results of all models together.

Table 2. AI's Impact on Field Creativity

	Novelty		Disruption	
	(1) 50th percentile	(2) 90th percentile	(3) 50th percentile	(4) 90th percentile
Constant	-45.341** (21.849)	-14.796*** (5.334)	-0.636 (0.631)	0.575 (0.658)
AI	14.243** (5.930)	9.963*** (1.900)	0.397*** (0.151)	0.111 (0.183)
log(*Size*)	5.774* (3.162)	4.140*** (0.965)	0.199* (0.115)	0.227* (0.128)
log(*TeamSize*)	1.966** (0.842)	1.363*** (0.440)	-0.142** (0.072)	-0.143*** (0.042)
log(*SelfCitation*)	-2.582* (1.405)	-2.434*** (0.451)	-0.101 (0.069)	-0.151** (0.075)
Year-by-Field Fixed Effect	Yes	Yes	Yes	Yes
N	794	794	794	794
Adjusted R^2	0.098	0.636	0.246	0.493

Notes. Standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01. We have also included *Institution*, *NIH*, *NSF*, and *Twitter* as additional controlled variables.

The *AI* variable has a positive sign, and it is statistically significant in Models 1-3. This suggests that, on average, AI technologies contribute to increased novelty in both mediocre (i.e., 50th percentile) and

Electronic copy available at: https://ssrn.com/abstract=4749118

top (i.e., 90th percentile) novel papers across various fields. Using AI technologies increases the novelty of mediocre papers by 0.69 standard deviations and that of the top novel paper by 0.76 standard deviations. The integration of AI tools introduces a fresh perspective on reevaluating existing knowledge and fostering scientific hybridization. Acting as a novel search algorithm, AI explores uncharted territory within the theoretical search space, thereby enhancing novelty for both top and mediocre novel papers. However, while AI enhances disruption in mediocre disruptive papers by 0.40 standard deviation, it does not significantly impact the top disruptive papers. AI assists researchers in exploring alternative pathways, enabling deviations from conventional scientific trajectories, and stimulating scientific mutation. Consequently, it enhances disruption in mediocre disruptive papers. However, the highly disruptive work requires researchers to deviate entirely from established paths. Given that AI technologies are primarily trained on existing data, it is difficult to incubate entirely novel work that is independent of existing research. This explains why AI does not increase disruption in the most groundbreaking papers.

The control variable log(*Size*) has a positive sign, and it is significant for Models 1–4, indicating that, on average, larger fields with more papers tend to have a higher degree of novelty and disruption. The control variable log(*TeamSize*) has a positive and statistically significant sign in Models 1–2, reflecting that larger author teams, on average, increase novelty in both mediocre and top novel papers. Nevertheless, it exhibits an expected negative and statistically significant sign in Models 3–4, suggesting that smaller author teams' work tends to be more disruptive in both mediocre and top disruptive papers. The role of team size in disruption is consistent with the findings of Wu et al. (2019). The control variable log(*SelfCitation*) has a negative sign, and it is significant in Models 1, 2, and 4, reflecting that citing more works from own fields on average reduces novelty in both mediocre and top novel papers and disruption in top disruptive papers. We also conducted the same analysis at the subfield level. Subfields refer to the sub-areas within the academic field, and our dataset includes 292 subfields. The results are presented in Online Appendix C.

4.2. Moderating Effects of Discipline Areas

To explore how AI technologies influence scientific creativity differently across fields, we classify all fields into four distinct discipline areas – Basic Sciences, Social Sciences, Applied Sciences, and Humanities, as

Electronic copy available at: https://ssrn.com/abstract=4749118

illustrated in Figure 2. We then compare the impacts of AI on *Novelty* and *Disruption* across these discipline areas. Specifically, we treat discipline areas as moderators and employ the following specifications to analyze their moderating effects:

$$Y_{ijt} = \beta_0 + \beta_1 AI_{jt} \times C_{g(i)} + \beta_2 C_{g(i)} + \gamma' X_{ijt} + F_t + \varepsilon_{it}, \quad (2)$$

where Y_{ijt} represents the outcome of our interest (*Novelty* or *Disruption*), measuring the overall creativity of Field *i*'s Type *j* papers published in Year *t*. $AI_{jt} \times C_{g(i)}$ is an interaction term between the binary variable AI_{jt}, indicating whether Type *j* papers are AI-related. The categorical variable $C_{g(i)}$ capturing to which discipline area Field *i* belongs (i.e., *BasicScience, AppliedScience, Social Science, and Humanities,* with the baseline as *BasicScience*). $C_{g(i)}$ functions as the moderator variable. The estimated coefficients of the interaction term β_1 measure the magnitude of moderating effects. We incorporate control variables X_{ijt} which include Type *j* papers within Field *i*, such as the log(*Size*), log(*TeamSize*), log(*SelfCitation*), *NIH*, *NSF* and *Twitter*. In addition, we control for year fixed effects F_{it}.

The results of the models (Equation (2)) are presented in Table 3. On average, the variable *AI* shows a positive and statically significant coefficient for Models 1–4, suggesting AI's positive impacts on novelty and disruption of both mediocre (i.e., 50th percentile) and top (i.e., 90th percentile) novel and disruptive papers in basic science. Specifically, for papers in basic science, AI increases the novelty for mediocre and top novel papers, while it increases the disruption for mediocre and top disruptive papers. Notably, the interaction terms are of interest. The coefficients for the three interaction terms ($AppliedScience \times AI$, $SocialScience \times AI$, and $Humanity \times AI$) are negative and significant, implying that AI has the most substantial positive impacts on novelty and disruption in Basic Science. AI's effects vary across the other three discipline areas. For mediocre novel papers, AI moderately increases the novelty of papers in Applied Science and has minimal effects on the novelty of papers in Social Science, while it decreases the novelty of papers in Humanities. In the case of top novel papers, AI's positive impacts on novelty are similar between Applied Science, Social Science and Humanities. For mediocre disruptive papers, AI increases the disruption of papers in Applied Science while slightly decreasing the disruption of papers in Social Science

Electronic copy available at: https://ssrn.com/abstract=4749118

and Humanities. For top disruptive papers, AI marginally increases the disruption in Applied Science, has minimal impacts on Social Science papers, and notably decreases the disruption of papers in Humanities. To summarize, our analysis reveals differences in the influence of AI on novelty and disruption across academic disciplines. Basic Science sees the most significant enhancement, Humanities subjects show the least or even negative impact, while Applied and Social Sciences demonstrate moderate or minimal effects. As anticipated, these findings are in line with our expectations. Basic Science, commonly associated with natural or physical sciences, heavily relies on advanced technologies and instruments for observation, experimentation, and data analysis. Contemporary scientific research, particularly in disciplines such as Physics, Chemistry, and Biology, often utilizes advanced technology. The advent of AI technologies introduces a new tool, serving as a novel search algorithm, empowering researchers in these fields to delve into uncharted territories in their quest for undiscovered insights. In contrast, research in the disciplines of Humanities such as Philosophy, History, and Art, does not rely as heavily on advanced technologies. While welcoming technological progress, the Humanities prioritize imagination and interpretation to understand the complexity of human experiences and culture.

Table 3. Moderating Effects of Discipline Areas on AI's Impact on Field Creativity

	Novelty		Disruption	
	(1) 50th percentile	(2) 90th percentile	(1) 50th percentile	(2) 90th percentile
AI	8.810** (3.460)	6.211*** (1.091)	1.107*** (0.117)	0.444*** (0.111)
AppliedScience × AI	-4.092*** (0.758)	-3.510*** (0.993)	-0.810*** (0.100)	-0.353*** (0.064)
SocialScience × AI	-8.424*** (2.716)	-3.116*** (1.206)	-1.273*** (0.099)	-0.433*** (0.104)
Humanities × AI	-12.291*** (3.306)	-3.441* (1.894)	-1.385*** (0.289)	-1.504*** (0.196)
log(Size)	1.816 (1.376)	0.743 (0.492)	0.415*** (0.051)	0.471*** (0.061)
log(TeamSize)	3.584*** (0.614)	3.042*** (0.364)	0.018 (0.031)	-0.010 (0.022)
log(SelfCitation)	-1.060 (0.647)	-0.816** (0.322)	-0.259*** (0.035)	-0.349*** (0.037)
Year FE	Yes	Yes	Yes	Yes
N	815	815	815	815
Adjusted R^2	0.115	0.558	0.357	0.474

Notes. Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Electronic copy available at: https://ssrn.com/abstract=4749118

5. Empirical Analysis: AI's Impacts via Citation Networks

In this section, we attempt to explore the underlying mechanism by which AI accelerates scientific activity. We closely follow the LCT framework, mapping knowledge recombination and replacement to citation flows across fields. Our mediation analyses separate the direct and indirect effects that are mediated through changes in citation flows.

5.1. Citation Patterns of AI and Non-AI Papers

Besides the field variables derived from paper-level data, we construct three network characteristics from AI papers' citation network and non-AI papers' to capture the citation patterns across fields. First, we calculate the weighted outdegree centrality (*WOutdegree*) to measure a field's tendency to cite works from other fields. Based on a given citation network, for a given field *i*, we calculate the variable *WOutdegree* as the sum of weights associated with all its outgoing edges to other nodes:

$$WOutdegree_i = \sum_{k \in \boldsymbol{S}(i)} w_{ik}\,, \tag{3}$$

where $\boldsymbol{S}(i)$ represents the set of nodes connected to node *i* and w_{ik} denotes the respective edge weight. A higher *WOutdegree* indicates a stronger propensity to cite papers from external fields. Secondly, we calculate the betweenness centrality (*Betweenness*), a measure that quantifies the extent to which a node lies on the shortest paths between other nodes in the network. In our setting, it reflects the field's importance in facilitating communication or the flow of information between different fields of the network. For a given Node *i*, *Betweenness* is calculated as the number of shortest paths that pass through that node. Thirdly, we use entropy to assess the diversity of a field's references. Variable *Entropy* of Field i is calculated as follows:

$$Entropy_i = -\sum_{k \neq i} P_k log_2 P_k, \tag{4}$$

where P_k is the normalized weight from node i to node k. A lower entropy implies that citations to external fields are more concentrated in a few specific areas. Conversely, a higher entropy suggests a more even distribution of citations across external fields. Table 4 summarizes the statistics of the network measures.

Electronic copy available at: https://ssrn.com/abstract=4749118

Table 4. Descriptive Statistics of Field-Level Network Variables

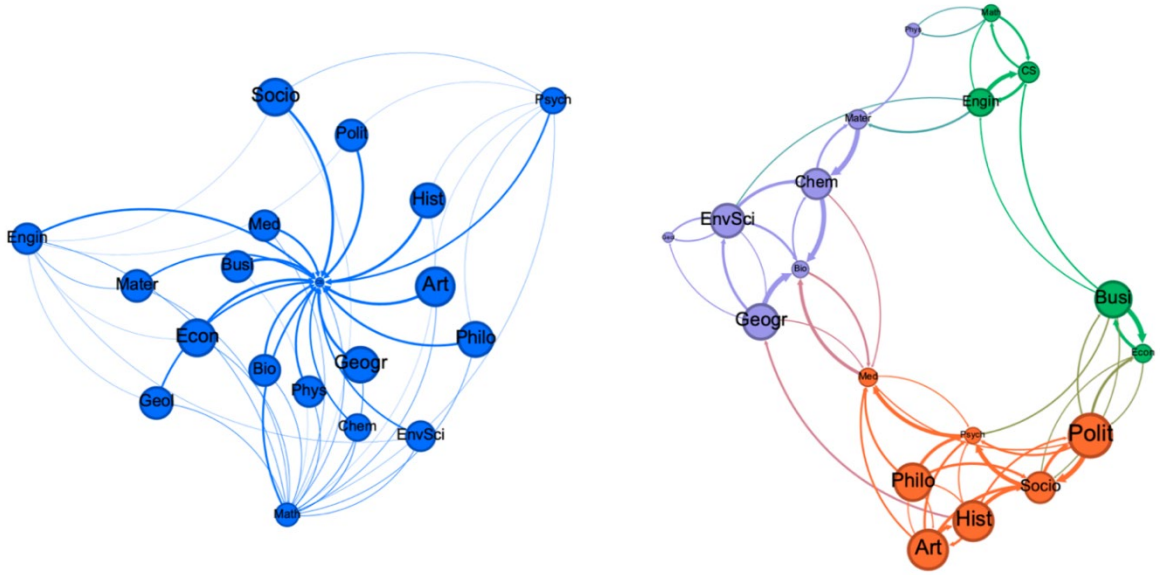
		AI				Non-AI			
	N	Mean	Std. Dev	Min	Max	Mean	Std. Dev	Min	Max
WOutdegree	836	17,962.89	57,366.86	0	534,977	762,602.18	1,167,363.22	1,238	7,627,389
Entropy	836	1.45	0.38	0.00	2.16	1.97	0.66	0.84	3.17
Betweenness	836	11,304.16	132,928.41	0	2,359,286	39.32	49.32	0	168

Figures 3a, 3b, and 3c present a visual representation of the three network metrics, WOutdegree, Betweenness, and *Entropy*, for AI networks and non-AI networks, respectively. In these figures, the edge width represents the citations from an out-node (citing field) to an in-node (cited field). Fileds are color-coded using community detection, a method in network analysis that identifies groups of nodes exhibiting denser connections among themselves. Thus, fields sharing the same color are more likely to have a higher number of citations from one another. In Figure 3a, the size of a node indicates the weighted outdegree centrality of the specific field. As anticipated, within the AI papers' citation network, Computer Science is the most cited field, evident from the numerous and prominently directed edges pointing towards it. It, the birthplace of AI, serves as the primary hub of AI, providing information and technologies to other fields. Yet, its weighted outdegree centrality is relatively low, indicating that AI papers in the field of Computer Science tend to exhibit a higher degree of self-interest. By contrast, per the non-AI paper citation network, multiple communities emerge, and fields are more inclined to recombine with each other in a community. We imply that information and knowledge sharing shall be more frequent within the same community. Non-AI-related knowledge originates from various fields, whereas AI-related technologies are diffused from Computer Science only.

Following the same representation, we consider betweenness centrality in Figure 3b. Interestingly, we notice that Engineering and Psychology play a significant role in disseminating AI knowledge to their communities. Engineering (Psychology) facilitates another route to learn, access, and adopt AI for Basic and Applied Science (Social Science and Humanities) areas. Again, different from the knowledge-sharing of AI, non-AI-related knowledge-sharing seems more balanced and collaborative, which reflects a general pattern from Figure 3a.

Electronic copy available at: https://ssrn.com/abstract=4749118

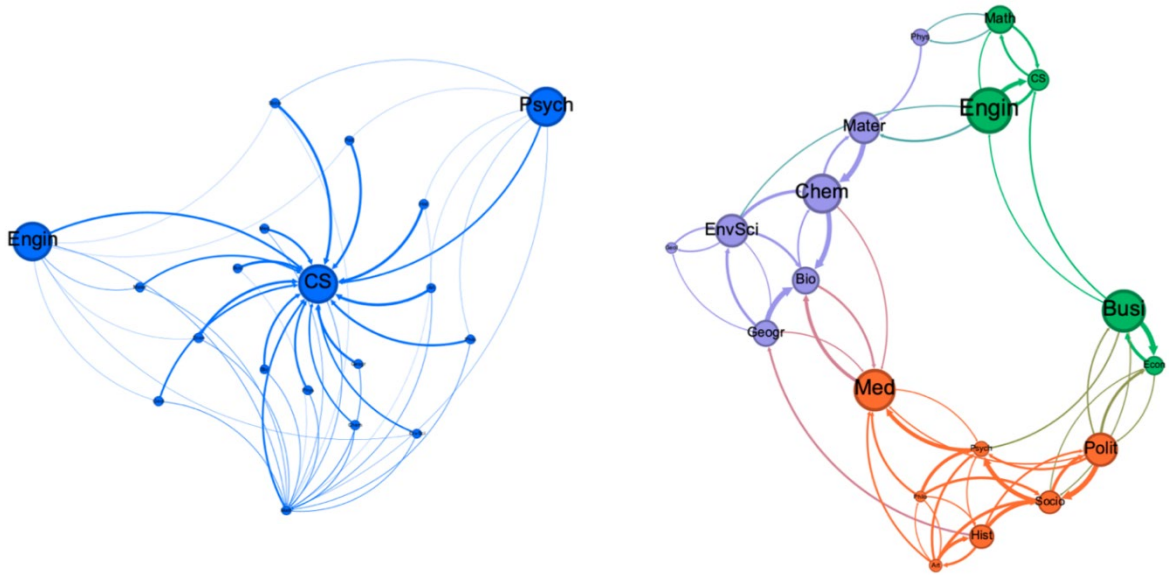
Figure 3a. Weighted Outdegree Centrality (by Fields)



AI Citation Network **Non-AI Citation Network**

Notes. The size of a node indicates the weighted outdegree centrality of the specific field. The edge width represents the citations from the out-node (citing) to the in-node (cited). Nodes are color-coded through community detection, a process in network analysis that identifies groups of nodes exhibiting denser connections among themselves compared to connections with nodes outside the group.

Figure 3b. Betweenness Centrality (by Fields)



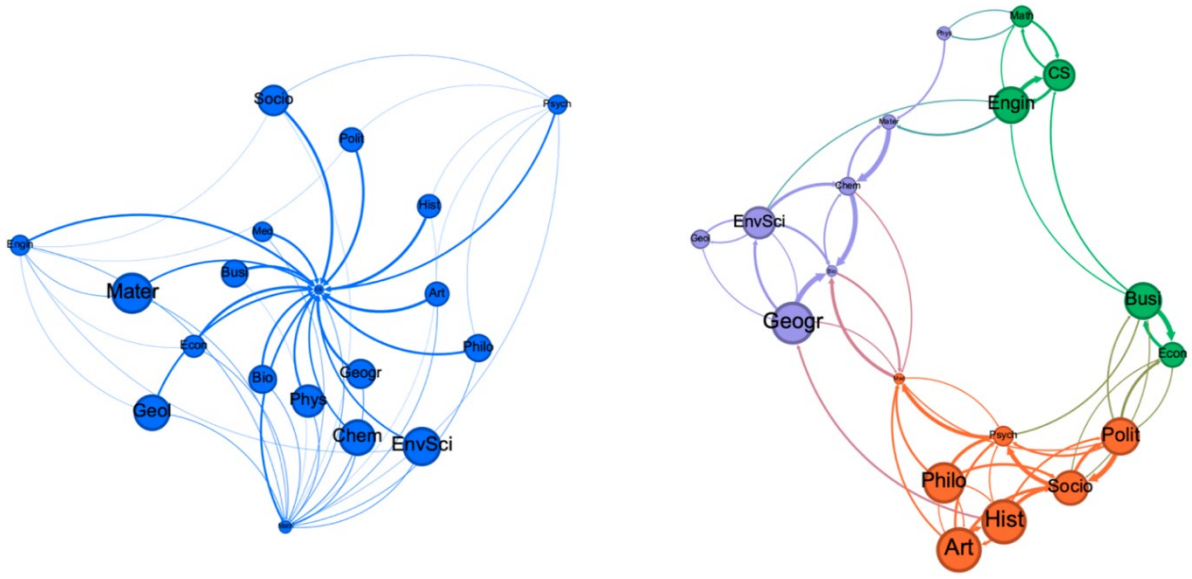
AI Citation Network **Non-AI Citation Network**

Notes. The size of a node indicates the betweenness centrality of the specific field. The edge width represents the citations from the out-node (citing) to the in-node (cited). Nodes are color-coded through community detection, a process in network analysis that identifies groups of nodes exhibiting denser connections among themselves compared to connections with nodes outside the group.

Electronic copy available at: https://ssrn.com/abstract=4749118

In Figure 3c, we use the size of a node to depict its entropy. In the AI papers' citation network, Computer Science has the smallest entropy, meaning that citations to external fields are super concentrated in very few areas, echoing what we find in Figures 3a and 3b. Materials Science and Environmental Science exhibit the highest entropy, followed by Chemistry, Geology, Physics, and Sociology, indicating that these fields focus on more recombination while conducting AI-related research. As for the network of non-AI papers, the diversity of citation sources demonstrates a contrasting pattern. Geography, Art, History, and Philosophy rely on the highest entropy knowledge-crossover.

Figure 3c. Entropy (by Fields)



AI Citation Network **Non-AI Citation Network**

Notes. The size of a node indicates the entropy of the specific field. The edge width represents the citations from the out-node (citing) to the in-node (cited). Nodes are color-coded through community detection, a process in network analysis that identifies groups of nodes exhibiting denser connections among themselves compared to connections with nodes outside the group.

5.2. Mediation Analyses of Network Measures

To investigate the underlying mechanism, we delve into the citation networks to disentangle the direct and indirect effects of AI on novelty and disruption. To explore the indirect effect, we perform mediation analysis, treating the three network measures – *WOutdegree*, *Entropy*, and *SelfCitation* – as mediators.[7]

[7] *SelfCitaion* refers to the self-loop in our citation networks as an important network measure symetically.

Electronic copy available at: https://ssrn.com/abstract=4749118

This analysis solves the following structural equations to estimate the mediation effects of thee three network measures:

$$Y_{ijt} = \beta_0 + \beta_1 AI_{jt} + \pi' N_{ijt} + \gamma' X_{ijt} + F_{it} + \varepsilon_{it}, \quad (5)$$

$$\log(Outdegree_{ijt}) = \beta_{01} + \beta_{11} AI_{jt} + \gamma_1' X_{ijt} + F_{it} + \varepsilon_{it}, \quad (6)$$

$$Entropy_{ijt} = \beta_{02} + \beta_{12} AI_{jt} + \gamma_2' X_{ijt} + F_{it} + \varepsilon_{it}, \quad (7)$$

$$\log(SelfCitation_{ijt}) = \beta_{03} + \beta_{13} AI_{jt} + \gamma_3' X_{ijt} + F_{it} + \varepsilon_{it}, \quad (8)$$

Where again Y_{ijt} represents the outcome of our interest, *Novelty or Disruption*, measuring the overall creativity of Field *i*'s type-*j* papers published in Year *t* and AI_{jt} is a binary variable indicating whether Type j papers are AI-related. N_{ijt} Represents the three network characteristics, log(*Outdegree*), *Entropy*, and log(*SelfCitation*). The controlled variable X_{ijt} includes log(*Size*) and log(*TeamSize*). We also employ the field-year fixed effects F_{it} to control for fields' time trends. In the structural equation system, the first equation regresses the creativity measures on the *AI* indicator, network measures, and controlled covariates. The subsequent three equations involve the regression analysis of the AI indicator and controlled covariates on the three network measures, respectively. To estimate the mediation effects within this structural equation system, we employ the Generalized Structural Equation Modeling (GSEM) approach.

Direct Effects

Table 5 reports the estimated direct and indirect effects. Detailed results of the above models, Equations (5)–(8), are presented in Online Appendix D. Figure 4 provides a visual presentation of the signs of both the direct and indirect effects. As shown in Figure 4, *AI* has direct effects on both *Novelty* and *Disruption* for both mediocre and top papers, which are positive and statistically significant. The ratios of AI's direct effect on *Novelty* out of the total are 1.35 for mediocre novel papers and 1.46 for top novel papers. The ratios of AI's direct effect on *Disruption* out of the total is 1.73 for mediocre disruptive papers and 4.38 for top disruptive papers. The substantial magnitudes of the direct effects indicate that the direct pathway serves as a primary channel through which AI technologies accelerate unusual combinations of existing knowledge and the disruptive departure from established knowledge. Namely, AI itself severs as the new search algorithms and stimulus of the search paths for exploring new and valuable knowledge in the theoretical

Electronic copy available at: https://ssrn.com/abstract=4749118

search space.

Table 5. Mediating Effects of Network Measures on AI's Impact on Field Creativity

	Novelty		Disruption	
	(1) 50th percentile	(2) 90th percentile	(1) 50th percentile	(2) 90th percentile
Direct effect of *AI*	17.15***	10.43***	0.78***	0.47***
Indirect eff via log(*WOutdegree*)	2.52***	0.66*	0.19***	-0.30***
Indirect eff via *Entropy*	-1.76***	-0.52**	-0.24***	-0.21***
Indirect eff via log(*SelfCitation*)	-5.18***	-3.52***	-0.28***	0.15***
Ratio of Indirect Effect	-0.35**	-0.46***	-0.73*	-3.38
N	836	836	836	836

Notes. Standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Indirect Effects

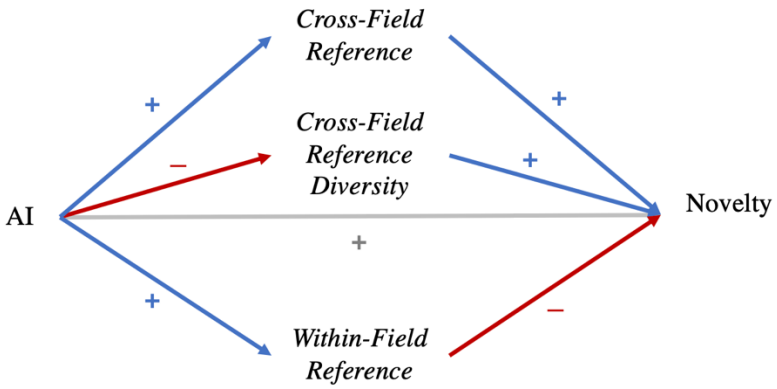
Additionally, *AI*'s effects on *Novelty* and *Disruption* are mediated by the variables log(*WOutdegree*), *Entropy,* and log(*SelfCitation*), offering further evidence of AI's role in altering the citation pattern. Namely, AI may significantly change the information flow within the academic sphere, thus accelerating scientific hybridization and mutation. As depicted in Figures 4a and 4b, on average, *AI* increases log(*WOutdegree*) and log(*SelfCitation*) significantly, indicating that AI papers tend to cite more from outside fields and within their own fields. However, *AI* significantly decreases *Entropy*, indicating that AI papers' references tend to concentrate within a few specific fields, which is consistent with the findings in the previous section.

With the mediating effects, the overall indirect effects of AI on novelty are negative for both mediocre and top novel papers. As shown in Figure 4a, the impacts of log(*WOutdegree*), *Entropy*, and log(*SelfCitation*) share the same direction concerning *Novelty* of both mediocre and top novel papers. Both log(*WOutdegree*) and *Entropy*, on average, contribute positively, indicating that a higher degree of citing references from external fields and a more diverse range of citation sources tend to foster scientific hybridization. By contrast, log(*SelfCitation*) exerts a negative influence on *Novelty*, suggesting that increased citations from the paper's own field somewhat restrict the potential for uncommon combinations of references. Consequently, three types of indirect effects from *AI* to *Novelty* emerge. Firstly, *AI* positively impacts *Novelty* via increasing citations from external fields (increased log(*WOutdegree*)). Conversely, AI negatively influences *Novelty* via reducing diversity in external citations (decreased *Entropy*) and increasing citations within the same field (increased log(*SelfCitation*)). Combining all three indirect effects,

Electronic copy available at: https://ssrn.com/abstract=4749118

the total indirect effects are negative and statistically significant. The ratios of the indirect effects are –0.35 and –0.46 respectively for mediocre novel papers and top novel papers.

Figure 4a. Mediating Effects on Novelty

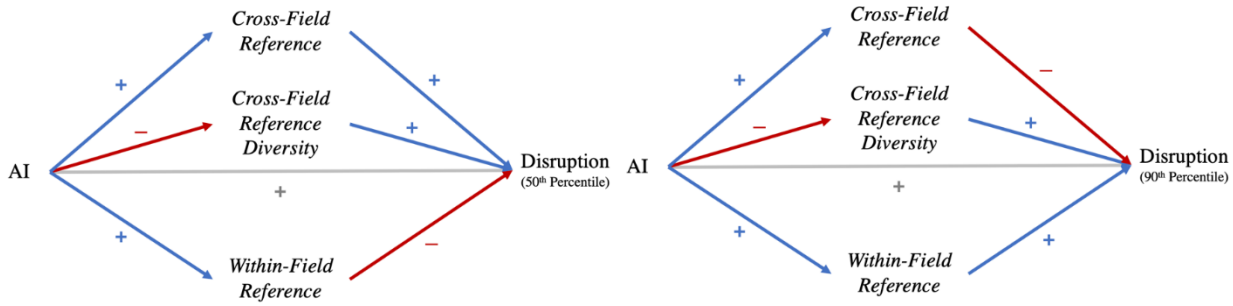


For mediocre disruptive papers, the overall indirect effect of AI on disruption is negative and significant. As depicted in Figure 4b, similar indirect paths of *AI* on 50th percentile *Disruption* are observed for these moderately disruptive papers as seen in Novelty. The ratio of the indirect effects is –0.73. However, for the top disruptive papers, AI's indirect effects are insignificant. As shown in Figure 4b, it takes different indirect paths. For the most groundbreaking works, on average, Log(*WOutdegree*) contributes negatively to *Disruption*, suggesting that the increased citations from outside fields may impede scientific mutation. This could be attributed to the heightened reliance on external references, potentially limiting researchers' imaginative exploration and reducing the likelihood of deviating from established search paths in the pursuit of new knowledge. By contrast, *Entropy* and log(*SelfCitation*) have positive effects on *Disruption*, meaning that increased diversity in external citation sources and elevated citations from the paper's own field foster the potential for scientific mutation. Citing more works from a specific field could be likened to genetic inbreeding. While inbreeding does not directly induce mutations, it does influence the distribution and expression of existing genetic variations within a population. In our context, the concentration of citations within one's own field may explain the heightened level of discourse observed in disruptive papers, as this practice may impact the prevalence of certain traits within the scholarly landscape. As a result, three types of indirect effects from *AI* to 90th Percentile *Disruption* emerge. Firstly, *AI* negatively impacts *Disruption*

Electronic copy available at: https://ssrn.com/abstract=4749118

by increasing citations from external fields (increased log(*WOutdegree*)) and reducing diversity in external citations (decreased *Entropy*). Conversely, AI positively affects *Disruption* via increasing citations within the same field (increased log(*SelfCitation*)). Mitigated by the negative indirect effects, the overall indirect effect on *Disruption* is insignificant for top disruptive papers.

Figure 4b. Mediating Effects on Disruption



6. Conclusions

In this paper, we have examined the role of AI technologies in advancing scientific creativity. Our study is framed within the context of *Logical Creative Thinking* (LCT), which considers innovation as a search process for undiscovered valuable knowledge. In this metaphor, researchers act as explorers employing various search algorithms to navigate the landscape of scientific exploration. Specifically, our exploration centered on the impact of AI on the *novelty* and *disruption* of papers. Novelty measures how atypical a paper's reference pairs are, while disruption measures the extent to which a paper either disrupts or improves the existing literature. We conceptualize the generation of scientific work through creative combinations of existing knowledge as scientific hybridization. In contrast, the birth of new work that disrupts existing paradigms is likened to scientific mutation. We have investigated both the overarching impact of AI across diverse fields and how this impact is moderated by discipline areas. In addition, we have explored the underlying mechanisms by disentangling their direct impacts and indirect effects through performing mediation analysis with citation networks.

In our analysis, we find an overall enhancement of novelty in both mediocre and top novel papers due to the adoption of AI technologies. The use of AI tools provides a new angle for reevaluating existing

Electronic copy available at: https://ssrn.com/abstract=4749118

knowledge, fostering scientific hybridization, and thereby promoting novelty. Also, AI contributes to the disruption of mediocre disruptive papers, yet this effect is not evident in top disruptive papers. AI enables researchers to explore new paths and deviate from conventional scientific trajectories, stimulating scientific mutation and enhancing disruption in mediocre disruptive papers. However, achieving groundbreaking disruption requires researchers to diverge entirely from established paths, which poses a challenge given that AI technologies are primarily trained on existing data. As a result, AI may not effectively enhance disruption in the most disruptive papers.

Our findings also indicate that AI has the most substantial positive influences in the fields of Basic Science, exhibits marginal impact in Social Science, and demonstrates minimal effects in Applied Science. Furthermore, the least, or even negative, influence is observed in Humanities. Basic science heavily relies on advanced technologies and instruments for observation, experimentation, and data analysis. The introduction of AI technologies serves as a search algorithm, motivating researchers to venture into uncharted paths in pursuit of undiscovered insights. In contrast, the Humanities prioritize imagination and interpretation in understanding human experiences and culture, relying less on advanced technologies.

We use network measures to explain how citations are made across different fields. We have also employed mediation analysis to disentangle the direct and indirect effects of AI on novelty and disruption, with network measures as mediators. AI technologies have direct positive effects both novelty and disruption for mediocre and top papers. The significant positive effects suggest that the direct pathway is a primary means by which AI technologies expedite novel combinations of existing knowledge and departure from established norms, paving the way for new search algorithms and paths to uncover valuable insights. The overall indirect effects of AI on novelty are negative for both mediocre and top novel papers. AI boosts novelty by increasing citations from external sources, fostering scientific hybridization. However, it also hampers novelty by reducing diversity in external citation sources and increasing citations within the same field, limiting the potential for uncommon reference combinations. The negative impact of the latter two paths outweighs the positive impact of the first one, resulting in an overall negative effect. The cumulative indirect effects of AI on disruption are negative for both mediocre and top disruptive papers. For top

Electronic copy available at: https://ssrn.com/abstract=4749118

disruptive papers, the indirect effects are insignificant, and the paths differ. AI diminishes disruption by boosting citations from external fields and narrowing the variety of external sources cited. This could potentially limit researchers' imaginative exploration and hinder scientific innovation, thereby restricting scientific mutation. However, AI increases disruption by fostering citations from the paper's own field, akin to genetic inbreeding, potentially influencing the distribution and expression of existing variations of scholarly traits. The negative and positive impacts offset each other.

Our contribution to the literature is three-fold. Firstly, we expand the literature on analytical creativity by extending the empirical application of the LCT theory to scientific creativity, specifically in the context of AI-accelerated creativity. Secondly, we advance the literature on scientific creativity by leveraging the structure of large-scale citation networks. Through network analysis, we illustrate how information flow drives scientific hybridization and mutation, enhancing novelty and disruption in academic publications. Lastly, we introduce a novel perspective to the literature on AI's influence on human creativity. Our study pioneers the exploration of AI's impact on scientific creativity.

The insights uncovered in this paper also hold practical implications for innovation, science, and advanced technologies. Our analysis confirms the concept of analytical creativity, which proposes that the creative process is teachable, trainable, and applicable. Further endeavors in this direction are necessary to enhance the replicability of innovation. Second, Our findings emphasize the importance of supporting AI endeavors, as AI adoption and interdisciplinary collaboration can yield novel tools and methodologies for scientific study, thereby enhancing scientific creativity. Finally, our insights can be applied to other general-purpose technologies, as well as critical and emerging technologies. As indicated by our findings, there is a clear necessity to promote further advancement and wider adoption of these advanced tools, thereby accelerating innovation across diverse fields.

Addressing limitations and suggesting directions for future research, our analyses primarily operate at the field level, constrained by computational limitations. This constraint hinders the granularity of our insights and implications. A more granular understanding could be achieved by future research through an investigation at the paper level. Additionally, our current analyses concentrate on 50th - and 90th-percentile

Electronic copy available at: https://ssrn.com/abstract=4749118

novelty and disruption scores, offering only snapshots of the distributions. A more comprehensive examination of the entire distribution could yield a more nuanced perspective. This avenue could be explored in future research endeavors. Furthermore, future research could delve into the impact of AI technologies or other advanced technologies on broader outcomes, such as patent citations that signify the rate of knowledge transformation from science to market applications. Exploring associations with news coverage and media mentions could also provide valuable insights. In summary, this study pioneers the examination of AI's role in scientific creativity by analyzing publication records and citation networks, contributing valuable insights to academic communities and society at large.

"Logic will get you from A to B. Imagination will take you everywhere."

—Albert Einstein

Electronic copy available at: https://ssrn.com/abstract=4749118

References

Ali Elfa, M. A., & Dawood, M. E. T. (2023). Using Artificial Intelligence for enhancing Human Creativity. *Journal of Art, Design and Music*, *2*(2), 3.

Altshuller, G. S., & Shapiro, R. B. (1956). Psychology of inventive creativity. *Issues of Psychology*, 6, 37-49.

Altshuller, G. (1984). *Creativity as an Exact Science*. New York, NY: Gordon & Breach.

Altshuller, G. S. (1985). *Algorithm for solving inventive problems ARIZ-85C*. Retrieved July, 8, 2014.

Altschuller, G., Zlotin, B., Zusman, A., & Filatov, V. I. (1989). *Search for new ideas: from insight to technology.*

Altshuller, G. S. (1996). *And suddenly the inventor appeared: TRIZ, the theory of inventive problem solving*. Technical Innovation Center, Inc.

Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of personality and social psychology*, *43*(5), 997.

Amabile, T. (2011). *Componential theory of creativity* (pp. 538-559). Boston, MA: Harvard Business School.

Besemer, S. P., & O'Quin, K. (1986). Analyzing creative products: Refinement and test of a judging instrument. *The Journal of Creative Behavior*.

Besemer, S. P., & O'Quin, K. (1987). Creative product analysis: Testing a model by developing a judging instrument. *Frontiers of creativity research: Beyond the basics*, 367-389.

Besemer, S. P., & Treffinger, D. J. (1981). Analysis of creative products: review and synthesis. *The Journal of Creative Behavior*.

Besemer, S. P., & O'Quin, K. (1999). Confirming the three-factor creative product analysis matrix model in an American sample. *Creativity Research Journal*, *12*(4), 287-296.

Bonnardel, N., & Zenasni, F. (2010). The impact of technology on creativity in design: an enhancement?. Creativity and innovation management, 19(2), 180-191.

Busse, T. V., & Mansfield, R. S. (1980). Theories of the creative process: A review and a perspective. *The Journal of Creative Behavior*.

Colton, S., & Steel, G. (1999). Artificial intelligence and scientific creativity. *Artificial Intelligence and the Study of Behaviour Quarterly*, *102*.

Ding, M. (2020). *Logical creative thinking methods*. Routledge.

Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and social psychology review*, *2*(4), 290-309

Feist, G. J. (2008). *The psychology of science and the origins of the scientific mind*. Yale University Press.

Electronic copy available at: https://ssrn.com/abstract=4749118

Figoli, F. A., Mattioli, F., & Rampino, L. (2022). *Artificial intelligence in the design process: The Impact on Creativity and Team Collaboration* (p. 106). FrancoAngeli.

Funk, R. J., & Owen-Smith, J. (2017). A dynamic network measure of technological change. *Management science*, *63*(3), 791-817.

Gao, J. and Wang, D. (2023). Quantifying the Benefit of Artificial Intelligence for Scientific Research. https://arxiv.org/abs/2304.10578

Getzels, J. W., & Csikszentmihalyi, M. (1976). The creative vision: A longitudinal study of problem finding in art.

Grosul, M., & Feist, G. J. (2014). The creative person in science. *Psychology of Aesthetics, Creativity, and the Arts*, *8*(1), 30.

Goleman, D., Kaufman, P., & Ray, M. L. (1992). *The creative spirit*. New York: Dutton.

Guilford, J. P. (1967). The nature of human intelligence.

Heinze, T., Shapira, P., Rogers, J. D., & Senker, J. M. (2009). Organizational and institutional influences on creativity in scientific research. *Research Policy*, *38*(4), 610-623.

Hutchinson, E. D. (1949). How to think creatively.

Khandwalla, P. N. (1993). An exploratory investigation of divergent thinking through protocol analysis. *Creativity Research Journal*, *6*(3), 241-259.

Kureha, M. (2023). Implications of Automating Science: The Possibility of Artificial Creativity and the Future of Science.

Liapis, A., Smith, G., & Shaker, N. (2016). Mixed-initiative content creation. Procedural content generation in games, 195-214.

Lin, Z., Yin, Y., Liu, L., & Wang, D. (2023). SciSciNet: A Large-Scale Open Data Lake for the Science of Science Research.

Liu, L., Wang, Y., Sinatra, R., Giles, C. L., Song, C., & Wang, D. (2018). Hot streaks in artistic, cultural, and scientific careers. *Nature*, *559*(7714), 396-399.

Lubart, T. I. (2001). Models of the creative process: Past, present and future. *Creativity research journal*, *13*(3-4), 295-308.

Magni, F., Park, J., & Chao, M. M. (2023). Humans as Creativity Gatekeepers: Are We Biased Against AI Creativity?. *Journal of Business and Psychology*, 1-14.

Mann, D. (2001). An introduction to TRIZ: The theory of inventive problem solving. *Creativity and Innovation Manage.*

Mednick, S. (1962). The associative basis of the creative process. *Psychological review*, *69*(3), 220.

Merton, R. K. (1961). Singletons and multiples in scientific discovery: A chapter in the sociology of science. *Proceedings of the American Philosophical Society*, *105*(5), 470-486.

Electronic copy available at: https://ssrn.com/abstract=4749118

Miller, A. I. (2019). *The artist in the machine: The world of AI-powered creativity*. Mit Press.

Ochse, R. (1990). Before the gates of excellence: The determinants of creative genius. CUP Archive.

Osborn, A. F. (1953). Applied imagination.

Runco, M. A. (1991). *Divergent thinking*. Norwood, NJ: Ablex Publishing Corporation.

Rushton, J. P., Murray, H. G., & Paunonen, S. V. (1987). Personality characteristics associated with high research productivity.

Simonton, D. (1986). Multiple discovery: Some Monte Carlo simulations and Gedanken experiments. *Scientometrics*, *9*(5-6), 269-280.

Simonton, D. K. (1990). Political pathology and societal creativity. *Creativity Research Journal*, *3*(2), 85-99.

Simonton, D. K. (2004). *Creativity in science: Chance, logic, genius, and zeitgeist*. Cambridge University Press.

Stein, M. I. (2014). Stimulating creativity: Individual procedures. Academic Press.

Souchkov, V. (2015), A brief history of TRIZ. https://www.xtriz.com/BriefHistoryOfTRIZ.pdf

Stumpf, H. (1995). Scientific creativity: A short overview. *Educational Psychology Review*, *7*(3), 225-241.

Tigre Moura, F., Castrucci, C., & Hindley, C. (2023). Artificial Intelligence Creates Art? An Experimental Investigation of Value and Creativity Perceptions. *The Journal of Creative Behavior*.

Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical combinations and scientific impact. *Science*, *342*(6157), 468-472.

Wallas, G. (1926). *The art of thought* (Vol. 10). Harcourt, Brace.

Ward, T. B., Smith, S. M., & Vaid, J. E. (1997). *Creative thought: An investigation of conceptual structures and processes* (pp. xv-567). American Psychological Association.

Weisberg, R. (1986). *Creativity: Genius and other myths*. WH Freeman/Times Books/Henry Holt & Co.

Wu, Z., Ji, D., Yu, K., Zeng, X., Wu, D., & Shidujaman, M. (2021). AI creativity and the human-AI co-creation model. In *Human-Computer Interaction. Theory, Methods and Tools: Thematic Area, HCI 2021, Held as Part of the 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part I 23* (pp. 171-190). Springer International Publishing.

Wu, L., Wang, D., & Evans, J. A. (2019). Large teams develop and small teams disrupt science and technology. *Nature*, *566*(7744), 378-382.

Zheng, H., Li, W., & Wang, D. (2022). Expertise diversity of teams predicts originality and long-term impact in science and technology. *arXiv preprint arXiv:2210.04422*.

Electronic copy available at: https://ssrn.com/abstract=4749118

Online Appendices

Appendix A. Top 25 Most Frequently Mentioned AI Terms in Scientific Research

A paper containing any of the following keywords in the abstract will be considered AI-related. These keywords have been chosen as they represent the Top 25 most frequently mentioned AI terms in scientific research in 2019. (Gao and Wang, 2023).

- Machine learning
- Convolutional neural network
- Deep learning
- Artificial intelligence
- Deep neural network
- Feature selection
- Feature extraction
- Object detection
- Image segmentation
- Artificial neural network
- Support vector machine
- Reinforcement learning
- Facial recognition system
- Hyperspectral imaging
- Image processing
- Sentiment analysis
- Generative adversarial network
- Learning model
- Recurrent neural network
- Wavelet transform
- Image retrieval
- Point cloud
- Natural language
- Action recognition
- Genetic algorithm

Electronic copy available at: https://ssrn.com/abstract=4749118

Appendix B. Correlations among Field-Level Variables

	Novelty	Disruption	WOutdegree	Betweenness	Entropy	SelfCitation	Size	Institution	TeamSize
Novelty	1								
Disruption	-0.19***	1							
WOutdegree	0.07**	-0.14***	1						
Betweenness	-0.00	0.01	-0.03	1					
Entropy	-0.04	0.31***	-0.08**	-0.06*	1				
SelfCitation	0.06*	-0.18***	0.88***	-0.02	-0.26***	1			
Size	0.04	-0.15***	0.83***	-0.03	-0.11***	0.87***	1		
Institution	0.11***	-0.06*	0.18***	-0.01	-0.07**	0.17***	0.03	1	
TeamSize	0.33***	-0.35***	0.41***	-0.04	-0.38***	0.45***	0.32***	0.37***	1

Electronic copy available at: https://ssrn.com/abstract=4749118

Appendix C. AI's Impact on Sub-Field Creativity

	Novelty		Disruption	
	(1)	(2)	(3)	(4)
	50th percentile	90th percentile	50th percentile	90th percentile
Constant	-69.419***	-48.221*	0.407***	-0.280**
	(31.653)	(27.669)	(0.095)	(0.109)
AI	9.041***	7.616*	-0.278***	-0.001
	(3.005)	(4.052)	(0.026)	(0.022)
log(*Size*)	13.186***	8.591***	0.071***	0.207***
	(5.045)	(3.265)	(0.024)	(0.023)
log(*TeamSize*)	6.790	5.063	-0.093***	-0.019
	(5.776)	(3.712)	(0.017)	(0.016)
log(*SelfCitation*)	-8.318***	-6.014***	-0.108***	-0.172***
	(3.069)	(1.851)	(0.017)	(0.017)
Year-by-Field Fixed Effect	YES	YES	YES	YES
N	6,576	6,576	6,616	6,616
Adjusted R^2	0.050	0.037	0.0147	0.549

Notes. Standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Electronic copy available at: https://ssrn.com/abstract=4749118

Appendix D. Model Results of Mediation Analysis

	Novelty		Disruption	
	(1) 50th percentile	(2) 90th percentile	(1) 50th percentile	(2) 90th percentile
DV: Novelty/Disruption				
Constant	-70.83***	-32.19***	-1.67***	-0.04
	(10.85)	(4.99)	(0.49)	(0.39)
AI	17.14***	10.43***	0.78***	0.47***
	(2.75)	(1.18)	(0.12)	(0.09)
log(Outdegree)	4.73***	1.07**	0.35***	-0.48***
	(1.22)	(0.55)	(0.06)	(0.04)
Entropy	4.83***	1.18**	0.66***	0.59***
	(0.93)	(0.50)	(0.04)	(0.04)
log(SelfCitation)	-5.06***	-3.11***	-0.27***	0.13***
	(1.04)	(0.47)	(0.05)	(0.04)
log(Size)	4.69***	3.88***	0.11**	0.33***
	(1.23)	(0.55)	(0.06)	(0.04)
log(TeamSize)	3.39***	2.01***	0.05	0.12***
	(1.30)	(0.42)	(0.06)	(0.03)
DV: *log(Outdegree)*				
Constant	-3.20***	-3.20***	-3.20***	-3.20***
	(0.35)	(0.35)	(0.35)	(0.35)
AI	0.53***	0.62***	0.53***	0.62***
	(0.11)	(0.10)	(0.11)	(0.10)
log(Size)	1.08***	1.09***	1.08***	1.09***
	(0.02)	(0.02)	(0.02)	(0.02)
log(TeamSize)	0.13**	0.01	0.13**	0.01
	(0.05)	(0.03)	(0.05)	(0.03)
DV: *Entropy*				
Constant	2.72***	3.64***	2.72***	3.64***
	(0.34)	(0.29)	(0.34)	(0.29)
AI	-0.36***	-0.36***	-0.36***	-0.36***
	(0.10)	(0.08)	(0.10)	(0.08)
log(Size)	0.01	0.01	0.01	0.02
	(0.02)	(0.02)	(0.02)	(0.02)
log(TeamSize)	-0.31***	-0.52***	-0.31***	-0.52***
	(0.05)	(0.03)	(0.05)	(0.03)
DV: *log(SelfCitation)*				
Constant	-8.32***	-8.35***	-8.32***	-8.35***
	(0.40)	(0.41)	(0.40)	(0.41)
AI	1.02***	1.13***	1.02***	1.13***
	(0.12)	(0.12)	(0.12)	(0.12)
log(Size)	1.53***	1.54***	1.53***	1.54***
	(0.02)	(0.02)	(0.02)	(0.02)
log(TeamSize)	0.17***	0.02	0.17***	0.02
	(0.06)	(0.04)	(0.06)	(0.04)

Notes. Standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

Electronic copy available at: https://ssrn.com/abstract=4749118

Electronic copy available at: https://ssrn.com/abstract=4749118