European Commission

Artificial Intelligence in Science:

Promises or Perils for Creativity?

R&I PAPER SERIES
WORKING PAPER 2025/03

Stefano BIANCHINI, Valentina DI GIROLAMO, Julien RAVET, David ARRANZ

Research and Innovation

Artificial Intelligence in Science: Promises or Perils for Creativity?

European Commission
Directorate-General for Research and Innovation
Directorate G — Common Policy Centre
Unit G.1— Common R&I Strategy & Foresight Service
Contact Erik Canton
Valentina Di Girolamo
Email Erik.CANTON@ec.europa.eu
Valentina.DI-GIROLAMO@ec.europa.eu
RTD-ECONOMIC-ANALYSIS@ec.europa.eu
RTD-PUBLICATIONS@ec.europa.eu

European Commission
B-1049 Brussels

Manuscript completed in February 2025

PDF	ISBN 978-92-68-25300-7	doi:10.2777/6693925	KI-01-25-056-EN-N

Luxembourg: Publications Office of the European Union, 2025

EUROPEAN COMMISSION

Artificial Intelligence in Science:

Promises or Perils for Creativity?

Working Paper

Authors:

Stefano BIANCHINI, Valentina DI GIROLAMO, Julien RAVET, David ARRANZ

2025 Directorate-General for Research and Innovation EN

Contents

ABSTRACT

The use of AI for scientific discovery has advanced at pace in the last decades. While this technology holds great potential to transform research, concerns have been voiced about its adverse, often unintended consequences. Can AI actually boost scientific creativity and lead to more innovative and impactful discoveries? Thus far, answers to this question remain largely anecdotal and confined to a handful of disciplines. In this paper, we study the diffusion of AI across 80 scientific fields from 2000 to 2022 and its impact on creativity – measured through novelty and impact. We find that AI adoption has accelerated in nearly all disciplines since the early 2010s, with research activity becoming increasingly concentrated in three major regions: the EU, the US, and China. Our analysis confirms an overall positive effect of AI on scientific creativity, though with considerable variation across fields: while most have benefited, some have seen little to no gains, and a few have even experienced negative returns. We propose that the structural organisation of knowledge within a field – and, by extension, the patterns of knowledge production – may moderate the influence of AI on scientific discovery. Specifically, we show that AI has greater transformative potential in "rough" knowledge spaces, where ideas are more fragmented and disconnected, and human cognition struggle to cope with complexity. These findings contribute to the ongoing debate on the role of AI in science and are contextualised within recent policy initiatives designed to promote AI-powered science.

1. Introduction

Science is, at its core, about seeking and applying new knowledge to better understand the natural and social world. Over the past century, scientific progress has been a cornerstone in addressing global challenges and promoting economic growth. Yet, empirical evidence points to worrying trends: the productivity of scientific research has declined significantly in recent decades (see, e.g., Pammolli et al., 2011; Scannell et al., 2012; Boeing and Hünermund; 2020; Aghion et al., 2021), ideas are becoming increasingly harder to find (Bloom et al., 2020), and scientific papers and patents are becoming less disruptive over time (Park et al., 2023).

Against this backdrop, the fast-paced proliferation of Artificial Intelligence (AI) and the broad accessibility of AI/ML-powered tools have sparked both excitement and concern about its role in science. AI brings unprecedented capabilities to accelerate discovery – whether by processing vast amounts of data more efficiently to detect relationships, trends, or anomalies that might elude human researchers (saving time and resources) or by supporting research tasks like synthesising literature, brainstorming ideas, writing code, and more (see, e.g., Krenn et al., 2022; Peng et al., 2023; Van Noorden and Perkel, 2023; Musslick et al., 2025). Some have even suggested, as highlighted in a recent DeepMind report (Griffin et al., 2024), that we may be entering "*a new golden age of discovery*". Yet, alongside the epistemic benefits, AI carries epistemic risks when scientists trust it as knowledge-production partners. Indeed, the over-reliance on AI raises several concerns about its implications for scientific creativity, credibility, research integrity, and, more fundamentally, scientific understanding (Birhane et al., 2023; Messeri and Crockett, 2024).

A growing body of literature has explored the penetration of AI in the sciences and the factors driving its diffusion (Arranz et al., 2024; Bianchini et al., 2024; Gao and Wang, 2024; Schmallenbach et al., 2024). However, the empirical evidence on how AI actually affects scientific productivity and creativity remains limited (Bianchini et al., 2022; Noy and Zhang, 2023; Yu, 2024; Toner-Rodger, 2024). Defining and measuring productivity and creativity in science presents an additional challenge. Metrics such as the number of scientific publications, patents, or experiments completed within a given timeframe are among the most commonly used proxies for analysing research productivity. But volume-based indicators alone may not suffice to fully grasp the multifaceted nature of scientific performance; other dimensions, such as novelty and impact of scientific output, must also be considered. Novelty reflects the originality and fresh insights a piece of research brings to its field, while impact assesses its influence on future studies and developments. We believe that both dimensions are particularly relevant when assessing the role of AI in scientific discovery.

This paper investigates how the diffusion of AI is affecting scientific *creativity*, defined as the *novelty* and *impact* of scientific outcomes, using a large sample of OpenAlex publications spanning the period 2000-2022, and covering 80 scientific fields. In doing so, this research contributes to the emerging literature on "AI in science" (or AI4Science) in several ways.

First, unlike existing studies that often focus narrowly on specific AI technologies (e.g., AlphaFold, Chat-GPT) or application domains (e.g., health sciences), our study takes a broader view by defining AI more inclusively and assessing its impact across a wide range of research fields. Second, we move beyond "traditional" citation-based metrics and also consider up-to-date novelty indicators. Third, we propose that some inherent characteristics of a field – particularly the "roughness" or combinatorial complexity of its knowledge space – may mediate the effects of AI on discovery. Finally, we contribute to the literature by conducting our analysis at both the global level and across major economic regions, with a focus on comparing the effect of AI on science in the EU, the US and China.

To summarise the main takeaways of our work, our analysis reveals an overall positive effect of AI on scientific creativity, though such effects depend not just on the technology itself, but also on how, and where, it is used. We show that AI holds strong transformative potential in "rough" knowledge spaces, where the complexity and fragmentation of knowledge pose significant challenges to human cognition. By way of example, AI exerts strong influence in areas such as

Genetics Diagnosis or *Drug Target*, while its potential remains untapped in areas like *Industrial Engineering* or *Design Automation*. Furthermore, the overall effect of AI on both novelty and impact is stronger in China, followed by the US, and then the EU.

The reminder of the paper is structured as follows. Section 2 provides an overview of the relevant literature and the conceptual framework guiding our analysis. Section 3 describes the data and the methodology. Section 4 presents the empirical results, while Section 5 concludes with implications for policy.

2. Background

2.1. Why is science slowing down?

The production of new knowledge is central to the economic growth and societal well-being. As Romer (1993) observed, "*the potential for continued economic growth comes from the vast search space that we can explore*" (p.68-9), where the discovery of new ideas transforms limited physical resources into more valuable goods and services.

However, since the early 2000s, economists have diagnosed a global productivity slowdown across multiple countries and industries (Goldin et al., 2024). This deceleration in productivity growth has occurred despite the advent of important technological advancements, especially in the digital domain, that were expected to drive economic development (Aghion et al., 2017; Brynjolfsson et al., 2018). Evidence on declining productivity growth rate has been accompanied by a significant slowdown in the rate of scientific discoveries and technological breakthroughs, even as the volume of new scientific and technological knowledge has grown exponentially. That implies that the productivity of scientific research has been decreasing over time, a trend observed across various economic sectors and countries (see, among others, Scannell et al., 2012; Miyagawa and Ishikawa, 2019; Bloom et al., 2020; Boeing and Hünermund, 2020). Additionally, recent research confirms slowing rates of disruptiveness of papers and patents, indicating that they are less likely to break with the past and push science and technology in new directions (Park et al., 2023). Bloom et al. (2020) succinctly summarize the issue: "*Ideas – and in particular the exponential growth they imply – are getting harder and harder to find*" (p.1104).

Table 1: Why has scientific progress slowed down?

Hypothesis	Explanation	How AI can help
Exhaustion of "low-hanging fruits"	Most accessible and impactful discoveries have already been made, leaving future research with diminishing returns	AI can identify patterns and opportunities in complex data, uncovering "hidden" discoveries beyond the reach of traditional methods.
"Burden of knowledge"	As knowledge expands, it becomes increasingly difficult for individual researchers to explore the growing space of ideas.	AI can process vast and fragmented knowledge spaces, enabling researchers to integrate diverse fields and navigate complexity. It can also help reduce cognitive overload by synthesising and connecting information.
Inefficiencies in modern science	Heavy administrative burdens, biases against novelty, replication crises, stratification, and other systemic issues hinder productivity.	AI can streamline research workflows, reduce administrative overhead, and improve reproducibility through automation and better data management.

Notes: Details on how AI can boost science are discussed in the following section.

As is common in times of slowdown, commentators have focused on what might have gone wrong (see Table 1 for a summary). One of the simplest arguments is the fishing-out hypothesis: there is a finite pool of ideas and we are fishing the easiest first. According to this view, the “low-hanging fruits” have already been picked, leaving future research with diminishing returns (Cowen, 2011; Gordon, 2017). Another explanation rests on the “burden of knowledge” hypothesis, which suggests that as the knowledge frontier expands, the space of ideas becomes harder to explore (Jones, 2009).[1] As many has advocated, we are drawing in information but starved for knowledge. In support of this theory, empirical evidence confirms an increase in the share of scientific papers and patents authored by teams of multiple collaborators, suggesting that science increasingly requires increasing effort and specialisation to achieve ground-breaking discoveries (Wuchty et al. 2007). Not least, inefficiencies of the modern science system exacerbate the problem: heavy administrative workloads, biases against novelty, the replication crisis, and stratification are just a few of the challenges amplifying the slowdown in scientific progress (Jones et al., 2008; Stephan, 2012; Azoulay and Li, 2020; Franzoni et al., 2022).

2.2. Science in the age of AI

Progress in information and communication technologies (ICTs) has made knowledge more accessible and science more automatable (King et al., 2009; Waltz and Buchanan, 2009). Mokyr et al. (2015) argue that the tools developed during the ICT revolution substantially improved scientists’ ability to store vast amounts of data, search across information silos, and analyse them at a fraction of the cost compared to just a few decades ago. Yet, as discussed earlier, these advancements did not translate into a proportional increase in scientific productivity and truly innovative research.

AI represents the next frontier in this evolution and may offer a promising turnaround to make scientific research both more productive and creative. Due to its transformative and pervasive nature, indeed, some scholars have described AI as a “*general method of invention*” – a framework that positions AI as a technology broadly applicable across diverse fields to enhance problem-solving and facilitate the creation of new ideas, technologies, and innovations (Cockburn et al., 2019; Crafts, 2021; Bianchini et al., 2022).

How can AI systems contribute concretely to science? Distinct yet complementary AI interventions are being proposed that span the entire research pipeline (Table 2). Krenn et al. (2022) introduce three fundamental dimensions of impact for AI-assisted science. First, AI can function as a *computational microscope*, enhancing a laboratory’s measurement capabilities and uncovering insights that are currently beyond the reach of experimental methods. One implication of this is an increase in the complexity and accuracy of experiments and computer simulations that scientists can conduct. Also, with AI/ML tools, human scientists can expand their “spectrum of senses”, that is, their ability to sense structures and recognise underlying patterns in highly complex data.[2]

---

[1] Romer (1993), cited at the opening of this Section, continued along these lines: “*The curse of dimensionality is […] a remarkable blessing. To appreciate the potential for discovery, one need only consider the possibility that an extremely small fraction of the large number of possible mixtures may be valuable*” (p.69).

[2] This dimension – AI as a *computational microscope* – is the one most readily captured through bibliometric data, as scientists often explicitly acknowledge the use of AI/ML tools for specific research purposes (e.g., “*we use AlphaFold to reveal the 3D protein structure…*”). It is, in fact, the primary effect that we aim to capture in our empirical analysis. Note that the other dimensions of AI contribution are equally important, though they are very difficult, if not impossible, to quantify using publication data.

Table 2. Science in the age of AI: Potential benefits and risks

Dimension	Explanation	Example of application	Risks for scientific creativity
Computational microscope	AI analyses large and more complex systems. Uncovers insights not yet attainable by experiments. Enhances simulation accuracy and reduces timescale. Generates, extracts, and annotates large scientific datasets. Expands scientists' "spectrum of senses".	Finding the best architecture for a computer chip: https://doi.org/10.1038/s41586-021-03544-w Controlling the nuclear fusion plasma in a tokamak: https://doi.org/10.1038/s41586-021-04301-9 Generating a catalogue of genetic mutations: https://doi.org/10.1126/science.adg7492 Improving gravitational wave detectors: https://doi.org/10.48550/arXiv.2301.06221	Novelty: Risk of paradigm lock-in, where AI limits exploration to patterns found in historical data. Focus on "pseudo-novelty", where surprising patterns may lack scientific grounding. Lack of serendipity and intimations on phenomena that deviate from expectations. Illusion of explanatory depth and objectivity.
Source of inspiration	AI identifies unexpected regularities in data and literature. Synthesizes literature and highlights unexplored knowledge areas. Proposes interpretable solutions. Assists in brainstorming, writing, and coding.	Identifying unexpected phase of crystal structures in high-pressure physics: https://doi.org/10.1088/0953-8984/23/5/053201 Predicting future research trends in quantum physics: https://doi.org/10.1073/pnas.1914370116 Explaining facial appearance by deconstructing a neural network: https://doi.org/10.1016/j.dsp.2017.10.011 Re-discovering Newton's law of gravitation: https://doi.org/10.48550/arXiv.2202.02306	Impact: Over-reliance on AI outputs without validation. Loss of trust in results due to model opacity and lack of interpretability. Outputs are too narrowly focused and fail to address broader scientific questions or interdisciplinary contexts. Insights may be perceived as lacking causal reasoning or theoretical grounding.
Agent of understanding	AI autonomously generalizes observations, transfers scientific concepts to new phenomena, and achieves independent scientific understanding.	-	
Manager	AI as a manager of humans who perform research tasks. Five management functions: Task division/allocation, direction, coordination, motivation, supporting learning.	Allocating tasks to study mosquitos: https://doi.org/10.3390/insects13080675 Coordinating time and location of aurora sightings: https://doi.org/10.1002/2015SW001214 Sending motivational messages for crowd science on galaxy shapes: https://dl.acm.org/doi/10.5555/3061053.3061159	

Notes: This table, of our own elaboration, builds on the three dimensions of computer-assisted scientific understanding proposed in Krenn et al. (2022) and the survey of the literature presented in Section 2.2. The identified risks for scientific creativity (novelty and impact) apply to all dimensions.

Second, AI can act as *source of inspiration* for new concepts and ideas, expanding the boundaries of human imagination and creativity. This includes, for instance, identifying unexpected regularities in experiments or simulations that may surprise human scientists and suggesting novel directions for further investigation. One of the biggest challenges facing scientists is how to keep up to date with the vast and ever-growing body of scientific literature within and beyond their field, to determine what is already known and discovered. Here, AI can help alleviate this "knowledge burden" by assisting researchers in navigating the growing volume of papers, synthetising key insights, and spotting unexplored regions of the knowledge space, while uncovering surprising patterns in the scientific literature (for concrete examples see Hastings 2023, Ch.2). AI can also generate innovative concepts through model inspection, for example, by inverting a neural network to understand the internal representations it has learned. For a broader discussion on the potential of AI to generate novel research ideas, see Si et al. (2024), and the preliminary study by Microsoft Research AI4Science (2023) spanning a wide range of scientific areas including drug discovery, biology, computational chemistry, materials design, and partial differential equations.

Third, AI may act as an *agent of understanding*, taking on the role of generalising observations and transferring scientific concepts to new phenomena. While in the first two dimensions AI supports humans in gaining understanding, in this final dimension the machine would autonomously achieve new scientific understanding.[3]

In addition to the potential of AI in performing functional research tasks, Koehler and Sauermann (2024) provide evidence on the role of AI as a *manager* of scientists who perform such tasks. They identify five core "management" functions that are important in scientific research and that are particularly well-suited for intelligent machines: (i) task division and allocation, (ii) direction (providing guidance on how to perform specific tasks), (iii) coordination (integrating efforts and outputs from different scientists), (iv) motivation, and (v) supporting learning (tracking performance, identifying causes of problems, providing feedback).

The potential benefits of AI are worth taking seriously, but while AI may represent a game-changer for science, a growing number of scholars have raised concerns about potential, often unintended, consequences of its (mis)use (see, e.g., Birhane et al., 2023; Grimes et al., 2023; Messeri and Crockett, 2024). These concerns extend to both "traditional" AI/ML techniques and the newer generation of large language models (LLMs), as well as broader categories like foundation models and generative AI (GenAI).

Regarding the former, a long-standing debate surrounds the opacity and transparency of AI/ML models, especially deep neural networks. The lack of interpretability of these "black box" models brings up critical questions about trust and reliability, and thus confidence in the validity of AI-generated inferences. As Rudin (2019) put it: "*Let us stop calling approximations to black box model predictions 'explanations'* " (p.208). Issues also arise from data quality, particularly when data points are incomplete, erroneous, or inappropriate. Indeed, since AI learns from data, it may develop a skewed "tip-of-the-iceberg" view of the world or, worse, an entirely incorrect one that ultimately produces poor decisions (Budach et al., 2022). "*Garbage in, garbage out*," as Hanson et al. (2023) aptly put it.

AI could also have pervasive effects on scientific novelty, and the reason is simple: as AI models generate insights by identifying patterns based on past research, they risk engendering a paradigm lock-in and tamping down possibilities for new scientific directions (Birhane et al., 2023). This creates what Messeri and Crockett (2024) describe as an "illusion of explanatory breadth", where researchers may falsely believe they are exploring the full space of testable hypotheses, while in reality, they are constrained to a narrower subset of hypotheses that are testable with AI tools. So, the efficiency offered by AI could inadvertently foster the growth of scientific monocultures and homogeneity, in which certain forms of knowledge production – those best suited for AI assistance – come to dominate all the rest, thereby stifling novelty. A related concern is the potential loss of serendipity, a cornerstone of scientific breakthroughs, since machines do not (yet) have intimations of something that operates differently than expected. Science may therefore become too structured, too predictable, and too focused on what is computationally convenient for an AI.

Furthermore, scientific discoveries made through AI may struggle to gain acceptance within the scientific community and ultimately lack impact. This is because AI-generated insights are often regarded as both interpreter-dependent and theory-laden, lacking the causal principles needed to provide genuine understanding – i.e., an issue rooted in the long-standing debate on correlations vs. causation (Mullainathan and Spiess, 2017; Pearl and Mackenzie, 2018). Recent research in strategy science, indeed, has confirmed that when searching for the solution to a problem, being guided by a theory leads to better decisions and more innovative solutions (Felin and Holweg, 2024; Sorenson, 2024). Progress in AI may have given us

[3] We use "would" here because this capability remains a theoretical prospect rather than a current reality. Krenn et al. (2023) propose two sufficient conditions for an AI system – referred to as "scientific androids" in their terminology – to be considered an *agent of understanding*. *Condition 1*: An android gained scientific understanding if it can recognize qualitatively characteristic consequences of a theory without performing exact computations and use them in a new context. *Condition 2*: An android gained scientific understanding if it can transfer its understanding to a human expert. By strictly adhering to these two conditions, it becomes clear that we are by no means close to achieving true artificial autonomous scientists. More on the topic of "scientific understanding" in general and how this can be mastered by intelligent machines can be found in De Regt (2017).

machines with truly impressive abilities but no real “scientific intelligence” (i.e., the capacity to generate, understand, and apply causal explanations grounded in theory) and, as such, we may end up producing more, but understanding less.

With LLMs and GenAI, additional concerns arise about factual accuracy, biases, and the potential erosion of scientific rigor. For instance, these models are well-known to generate non-existent or false content, a phenomenon referred to as “hallucination”, which is clearly problematic in the realm of scientific research (Athaluri et al., 2023; Beutel et al., 2023). They also pose risks to research integrity and deontological obligations, for example, in cases of plagiarism or excessive reliance on writing or coding assistance. In short, whether involving LLMs or other AI/ML techniques, these technologies can fundamentally challenge the ethos of science and undermine some core values such as objectivity, rigor, and accountability. As four experts in artificial intelligence ethics and policy stated in a recent interview (Birhane et al., 2023): “*Researchers must proceed with caution, engaging the affordances provided by these technologies with the same kinds of epistemic humility, deflationary scepticism and disciplined adherence to the scientific method that have functioned as preconditions of modern scientific advancement since the dawn of the seventeenth-century Baconian and Newtonian revolutions*” (p.277).

From what has been discussed so far, it is clear that AI presents both opportunities and challenges for scientific discovery. The ultimate effect of AI on science, therefore, remains an open empirical question. We now turn to a review of the existing literature on the impact of AI in research. In doing so, we will also position our research within the current state of the art, outlining how our study builds on and extends prior work.

2.3. Empirical and theoretical evidence

A growing body of literature has focused on assessing the degree of penetration of AI technology in the sciences (see, among others, Xu et al., 2021; Gargiulo et al., 2023; Hajkowicz et al., 2023; Duede et al., 2024) and its adoption across different geographies (Arranz et al., 2023; AlShebli et al., 2024; Schmallenbach et al., 2024). We know collectively that the use of AI/ML in research is becoming pervasive across disciplines, fields, and geographic areas, with sharp growth in recent years. Fewer studies, however, have investigated the direct effects of AI on scientific research and R&D activities. Moreover, we identify two main limitations in the existing literature: first, an overemphasis on selected research fields or specific AI technologies; and second, a disproportionate focus on a handful of ground-breaking models, often resulting in anecdotal evidence drawn primarily from “success stories”.

For instance, Furman and Teodoridis (2020) study the effects of the Microsoft Kinect gaming system, powered by AI pattern recognition software, on knowledge production in computer science and electrical and electronics engineering. Their findings suggest that integrating AI technology leads to increased research output and greater research diversity. Bianchini et al. (2022) focus instead on neural-network-based technology in the health sciences, reporting a positive relationship between the adoption of such technologies and the likelihood of a scientific contribution becoming influential (i.e., highly cited), though not necessarily novel. Yu (2024) investigates the impact of AlphaFold on structural biology research, finding no significant effect on the number of publications but a positive impact on citation counts.

In the context of R&D and innovation, Rammer et al. (2022) analyse German firms and find that the use of AI technologies is associated with significantly higher rates of product and process innovations. In the same vein, Toner-Rodgers (2024) examines the impact of an AI-driven tool (i.e., graph neural networks) on materials discovery in a U.S. R&D lab, showing that AI-assisted scientists were able to discover 44% more materials, which led to an increase in patent filings and a rise in product innovation. Similar effects on AI patenting of climate-related inventions are found in Verendel (2023).[4]

Closer to our research is the recent work by Gao and Wang (2024), which examines the diffusion and impact of AI on citation counts across 19 disciplines and 292 fields using the MAG database, covering the period from 1960 to 2019. They find that publications that use AI – proxied by mentions of AI-related terms in publication titles and abstracts – tend to enjoy a citation premium, being more frequently cited both within and outside their disciplines. Notably, the authors highlight substantial heterogeneity in the direct use and potential benefits of AI across different disciplines. In fact, they conclude that almost every macro-discipline includes some subfields that experience high citation benefits from AI. For instance, they show that while ‘medicine’ as an aggregated discipline does not rank among the highest in terms of AI benefits, some of its subfields (e.g., ‘nuclear medicine’, ‘optometry’, and ‘medical physics’) exhibit substantial returns from AI integration.[5]

[4] For an in-depth discussion of the influence of AI on corporate innovation and other dimensions of innovation management, see, e.g., Bahoo et al. (2023) and Tekic and Füller (2023).
[5] Other studies on “AI in science” focus not directly on the impact of AI on knowledge production but rather on the structure of scientific teams and the challenges of interdisciplinary collaborations in AI research (Thu et al., 2022; Abbonato et al., 2024), on the growing influence of the private sector in AI research (Ahmed and Wahed, 2020;

Our study builds on and complement the above empirical contributions in several respects. First, we extend the analysis to more recent years (up to 2022), which is particularly important given the exponential growth in AI adoption, as documented in later sections. Second, we broaden the scope of impact metrics beyond citation counts to include exceptionally highly cited papers and additional measures of scientific novelty. Third, while Gao and Wang (2024) identify cross-field heterogeneity in the benefits of AI, they do not explore the underlying reasons for these differences. Here, we propose explanations related to the inherent characteristics of the knowledge space that scientists must navigate within their respective fields. In other words, we seek to understand *why* the potential impact of AI may vary from one field to another. Fourth, it is well known that countries such as China have recently been catching up in AI research, but we are not aware of any other studies that investigate the effects of AI in science at the geographical level. To be clear, we are not claiming that no studies compare the relative positions of countries worldwide in AI research; indeed, there is abundant literature on this topic (see, among others, Klinger et al. 2021; AlShebli et al., 2024; Schmallenbach et al., 2024). However, to our knowledge, no prior research has estimated the returns of AI in terms of scientific creativity while accounting for geographical differences. Our research addresses this gap as well.

Complementing the scarce empirical literature, a few theoretical studies show that, under certain assumptions, AI-enhanced R&D makes scientists and engineers more productive and, in turn, accelerates the production of new ideas (Besiroglu et al., 2024). This is especially true for scientific challenges involving combinatorial-type research problems (Agrawal et al., 2018). In fact, the notion that knowledge production is fundamentally recombinant in nature has deep intellectual roots (see, e.g., Arthur, 2009). The generation of new ideas hinges on the ability to combine existing knowledge into novel configurations, with an almost infinite number of ways in which different combinations can be put together (Weitzman, 1998; Fleming, 2001). According to this recombinant approach, thus, knowledge creation is inherently a process of *searching* and *combining* elements within a complex knowledge space. Yet, it is clear that as the knowledge frontier expands, both its morphology and complexity evolve: some regions become densely connected, while others grow increasingly isolated. In short, transforming the ever-expanding body of knowledge and information into valuable new ideas and innovations becomes progressively more challenging (Jones, 2009; Uzzi et al., 2013).

Some theoretical studies propose that AI may alleviate this burden. In their theoretical model, Agrawal et al. (2018) show that AI can support researchers to explore sparse, uncharted (and hence complex) territories within the theoretical search space, facilitate access to relevant knowledge, and enhance their ability to find new, useful combinations. The sequential search over a vast combinatorial knowledge space would ultimately result in "better, faster, cheaper" science (Agrawal et al., 2024). Along similar lines, Chen et al. (2024) suggest that AI may guide researchers to break away from longstanding, domain-specific mindsets and embrace new research paths via two mechanisms: knowledge *hybridization* across fields (exemplified by the development of MRI from medicine and physics) and knowledge *mutation* within a field (e.g., mRNA from well-understood principles of RNA biology), all while minimizing search costs.

One key takeaway from these theoretical models is that fields – and research problems therein – differ in how knowledge is structured: some knowledge spaces are smooth and well-connected, while others are rough and fragmented, requiring greater effort and new tools to navigate. As a result, we expect that the potential impact of AI is likely to vary across fields. In the second part of our empirical analysis, we will test how the effects of AI on creativity depend on the structure of field-specific knowledge spaces.

3. Data and Methods

3.1. The sample

We collected data from OpenAlex for the period 2000-2022, restricting the sample to peer-reviewed journal articles, conference proceedings, and preprint collections. The hierarchical structure of the OpenAlex "concept" taxonomy – which includes over 65,000 unique concepts at varying levels of granularity – allowed us to assign publications to different scientific fields. Each document may be associated with multiple concepts, and thus multiple fields, with each concept accompanied by a score indicating the confidence level of its classification. We classified a paper as an "AI paper" if it was associated with the (level-1) concepts *Artificial Intelligence* or *Machine Learning*, or with at least one of their 400+ sub-concepts.[6]

Analysing the entirety of OpenAlex is admittedly beyond our reach. The computational burden required to characterise the knowledge space of different fields (more on this in Section 3.3) and to estimate the effects of AI on novelty and impact through econometric models (details in Section 3.4) constrained our ability to

Ahmed et al., 2023), and the mobility of AI talent from academia to industry (Gofman and Jin, 2024; Jurowetzki et al., 2025).

[6] For readers unfamiliar with the concept taxonomy, an overview is available here: OpenAlex Concept Taxonomy.

work with the full dataset. We implemented a protocol for data retrieval that ensures coverage of a sufficiently large number of fields where AI has been widely adopted (e.g., *Remotely Operated Vehicles*) and fields with high volumes of scientific output (e.g., *Cancer Therapy*), even if they do not heavily involve AI. The final sample consists of approximately 3 million documents associated with 80 fields, identified at levels 2 and 3 of the concept hierarchy, of which ~35% were flagged as AI papers. Technical details on the protocol for data retrieval are provided in Appendix A1 (see also Figure A1), while the list of selected fields, along with their descriptions and examples of AI applications, is reported in Table A2.

For the sake of validation, we carried out a manual inspection of the abstracts of several papers in our sample, confirming the versatility of AI as a tool to support scientific research. By way of example, in some fields related to drug discovery (e.g., *Druggability*, *Drug Target*) AI was explicitly mentioned as being used to identify potential drug-binding sites on protein structures or to predict drug targets in pathogens. In cancer-related fields (e.g., *Cancer Therapy*, *Cancer Detection*), for instance, AI could assist physicians in tailoring cancer treatments based on patient profiles or classify cancer types from biopsy samples. Other examples include AI analysing real-time production data to adjust machine speeds (*Digital Manufacturing*), creating innovative designs for product development (*Ideation*), predicting timelines and costs of new projects (*Project Estimation*), optimising schedules and resource allocation based on historical project data (*Project Management*), and modelling customer behaviour in e-commerce (*Behavioural Modelling*), among many others.

As a cautionary note, we do not claim that the results reported in this manuscript fully capture the dynamics of the entire scientific landscape. However, we are confident that they are robust enough to offer meaningful insights and support some speculative conclusions on a larger scale. Further research is, of course, encouraged to expand and refine the scope of the analyses presented here.

3.2. Metrics for scientific creativity

Scientific creativity emerges when an individual or small group of individuals working together generates contributions to science that are both novel and useful. Novelty, or originality, involves breaking new ground or departing from the established status quo. Usefulness, on the other hand, pertains to the practical or theoretical value of an idea, and thus its influence within a scientific field or, more broadly, across science. This conceptual framework is largely inspired by the seminal works of Teresa Amabile and Margaret Boden. Amabile and Pratt (2016), for instance, define creativity as "*the production of novel and useful ideas*" (p.158); similarly, Boden (2004) describes it as "*the ability to come up with ideas or artefacts that are new, surprising and valuable*" (p.1). Guided by these definitions, we built a set of metrics to reflect the *novelty* and *impact* of each focal paper in our sample.

Traditionally, the novelty of scientific papers has been measured using citation patterns (e.g., atypical combination of cited references). Yet citation-based metrics often fall short in identifying novel scientific ideas at the time of publication and in capturing their true intellectual contribution to scientific progress (Fontana et al., 2020). More recently, Art et al. (2025) have suggested that scientific ideas are more effectively embedded in the text of scientific literature, with shifts in language being the main criterion for identifying novel research.

In our study, we follow this text-based approach to operationalise novelty. Hence, novelty was measured using: (1) the first appearance of new words or (2) noun phrases in the title or abstract of the focal paper; (3) the first appearance of unique pairwise combinations of words or (4) noun phrases in the titles or abstracts; and (5) the semantic distance of the focal paper from its most similar prior work.

Each of these variables were constructed as follows:[7]

- *New Words*: A binary variable equal to 1 if the focal paper introduces at least one new unigram (single word) in its title or abstract, and that word is subsequently reused in at least one other paper.
- *New Phrases*: A binary variable equal to 1 if the focal paper introduces at least one new noun phrase in its title or abstract, and that phrase is subsequently reused in at least one other paper. Noun phrases consist of one or more words with a noun as their head (e.g., "polymerase chain reaction").
- *New Words Combinations*: A binary variable equal to 1 if the focal paper introduces at least one novel pair of words (used together), and this pair is reused in at least one subsequent paper. The individual words themselves do not need to be new.
- *New Phrases Combinations*: Similar to *New Word Combinations*, but applied to phrases instead of single words.

[7] The original raw metrics for novelty were retrieved from here (Art et al., 2025) and matched to our sample using the unique OpenAlex publication ID. Except for the semantic distance, all variables in our analysis are from our own elaboration.

- *Semantic Distance*: Calculated as 1 minus the maximum cosine similarity between the focal paper and all papers published in the previous five years, providing a measure of how semantically different the focal paper is from prior work.

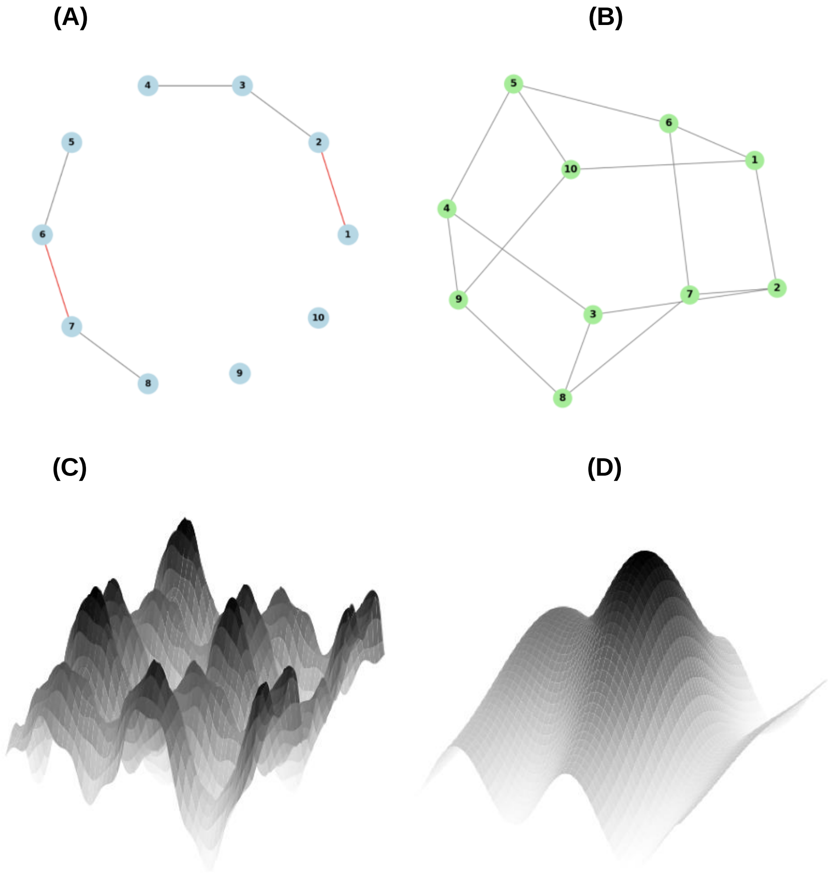
The second component of scientific creativity concerns impact, which is related to, but different from, novelty; if a paper provides novelty, that novelty must be adopted by the scientific community in order for its impact to be felt. Impact is measured by the number of citations received by the focal paper. Furthermore, we identified so-called "big hit" contributions – i.e., exceptionally cited papers. Specifically, we built the following metrics:

- *Weighted Nb. Citations*: The yearly count of citations the focal paper has received from its year of publication up to August 2024 (the time of data extraction).
- *Top Cited*: Three binary variables indicating whether the focal paper belongs to the top 10%, 5%, or 1% most-cited papers, with reference to other papers published in the same year and field.

3.3. Measuring the combinatorial complexity of a knowledge space

New knowledge somehow must come into being as fresh combinations of what already exists and is known – as principle known as "combinatorial evolution", backed by the rich literature of the evolution of science and technology (Kuhn, 1962; Kauffman, 1993 – ch.2; Weitzman, 1998; Arthur, 2009 – especially ch.2 and ch.9). Coming up with something new typically requires a rich repository of knowledge (a knowledge space) and the ability to explore, transform, and evaluate links across different elements. Of course, most combinations are entirely sterile, but a rare few prove to be extraordinary fruitful. Seen this from perspective, knowledge creation is inherently a process of *searching* and (re)*combining* elements within a knowledge space.

Figure 1. Topography of knowledge spaces



Notes: A simplified representation of a knowledge space as a network graph (Top), where nodes represent knowledge elements and edges denote connections between them, and a corresponding topographical surface (Bottom). Panel A: A rugged space, shown as a scattered and sparse network, illustrating a knowledge space with unconventional and novel combinations of ideas (highlighted in red). Ideas are fragmented and there are few connections between knowledge elements. For example, a scientist familiar with element 1 may find it difficult to access element 8 due to the lack of direct or intermediary connections. This corresponds to navigating a jagged surface, as shown in Panel C. Panel B: A smooth, well-trodden knowledge space, depicted as a dense, well-connected network where most nodes follow established pathways. Ideas build incrementally on each other and there are strong and frequent connections between knowledge elements. For example, a scientist familiar with element 1 can easily access element 9 through element 10 or other paths. This corresponds to navigating a smooth surface, as shown in Panel D. The illustration is intended solely for conceptual simplification.

Scientists from different fields must navigate knowledge spaces that can vary substantially in their structure and complexity. To measure the *structural organization of knowledge* – and, by extension, the process of knowledge production – within a scientific field, we used the indicator proposed by Lee et al. (2015), which identifies unusual or unprecedented combinations of references cited in a focal paper. This indicator was calculated for each paper in a field, and the values were then averaged across all papers within the field to produce a field-level measure. Appendix B provides technical details on the calculation of the indicator.

The field-level measure thus represents the "*roughness*" – or *combinatorial complexity* – of the knowledge space within a scientific field.[8] As illustrated in Figure 1 (top), fields with a higher score are characterized by knowledge space where connections between ideas are sparse or uneven, reflecting a less established and more fragmented knowledge structure. In these fields, scientists must first spot and then venture into new and unconventional combinations of ideas, akin to navigating a rugged landscape. On the contrary, fields with lower scores tend to be more incremental, with scientists predominantly building on well-established knowledge paths. In these cases, the knowledge landscape would resemble a smooth and well-trodden terrain.

Our metric can be meaningfully linked to concepts derived from the NK model (Kauffman, 1993) for interpretative clarity, as illustrated in Figure 1 (bottom). A useful analogy is to visualise a knowledge space as three-dimensional surface, where the horizontal dimensions represent possible combinations of knowledge elements, and the vertical dimension represents the value of a particular combination. In fields with high combinatorial complexity, akin to rugged NK landscapes with high K, the search processes are more complex and uncertain, but potentially rich with opportunities for unconventional and novel breakthroughs. In fields with low combinatorial complexity, ideas are more densely connected, and thus forming a smooth knowledge landscape that is easier and more predictable to explore.

In line with the theoretical models on AI for search and discovery discussed in Section 2.3, we expect AI to have greater potential for combinatorial creativity and thus to be particularly transformative in rough knowledge spaces, where human cognition may struggle to navigate complexity. Indeed, AI has been shown to excel at finding unconventional patterns and discovering hidden connections in vast amounts of data – hence the analogy of "AI as computational microscope". Moreover, AI can identify gaps and missing links between unrelated fields, literature, data, etc., facilitating connections that are often beyond human foresight – see its role as "source of inspiration"; and since scientific principles often transcend individual fields, and phenomena echo across fields, AI can borrow some principles from one domain of use and set to work in new ones. In addition to these capabilities, new AI-based techniques and methodological frameworks can be invented specifically to tackle complex problems that were previously intractable.

However, AI is inherently dependent on regularities in its training data, meaning that in sparse knowledge spaces, where such regularities are absent or weak, its performance may degrade. Still, the potential of AI should not be underestimated in smoother knowledge spaces. In these settings, AI can capitalise on abundant, often well-structured data to achieve high accuracy and efficiency, validate existing frameworks, and propose incremental yet meaningful advancements.

Ultimately, the extent to which the combinatorial complexity of a field mediates the effect of AI on scientific creativity remains an empirical question, one that we address in this study.

3.4. Empirical setting

The paper-level analysis compares AI vs. non-AI papers using a standard econometric framework. Hence, our main dependent variables are various measures of scientific novelty and impact. The main explanatory variable (*AI*) is a binary indicator that takes the value of 1 if the paper involves the use of AI and 0 otherwise. We considered a set of control variables to capture various characteristics of a focal paper, including the team size (*Nb. Authors*), the total count of references cited in the paper (*Nb. References*), a binary variable equal to 1 if the paper results from collaborations across multiple countries (*International Collab.*), and a binary variable indicating whether the paper provides a review or survey of extant literature (*Survey/Review*).[9]

[8] Note that the term "complexity" is often used to describe systems with many components and intricate interactions. However, in this study, we use complexity to describe the structural organisation of a knowledge space, specifically its level of fragmentation and sparsity. Under this definition, a knowledge space is considered more complex when its elements are more disconnected and scattered, making it harder to understand, navigate and recombine.

[9] The variable *International Collab.* takes value of 1 if there are at least two different countries in the authors' affiliations, 0 otherwise. *Survey/Review* takes a value of 1 if we detect in the title of the paper the terms 'survey', 'review', or 'overview', 0 otherwise.

For the geographic and field-level analyses, we used the same variables but examined how AI influences novelty and impact across different regions (China, the EU, and the US) and across fields characterised by varying degrees of "roughness" in the knowledge space.

All binary dependent variables are modelled with a Logit regression, whereas continuous dependent variables using Ordinary Least Squares (OLS). In all estimates, we included fixed-effects for scientific fields (concept level-2) and time to account for field-specific dynamics and cohort effects. Descriptive statistics for all variables are reported in Appendix C.

4. Results

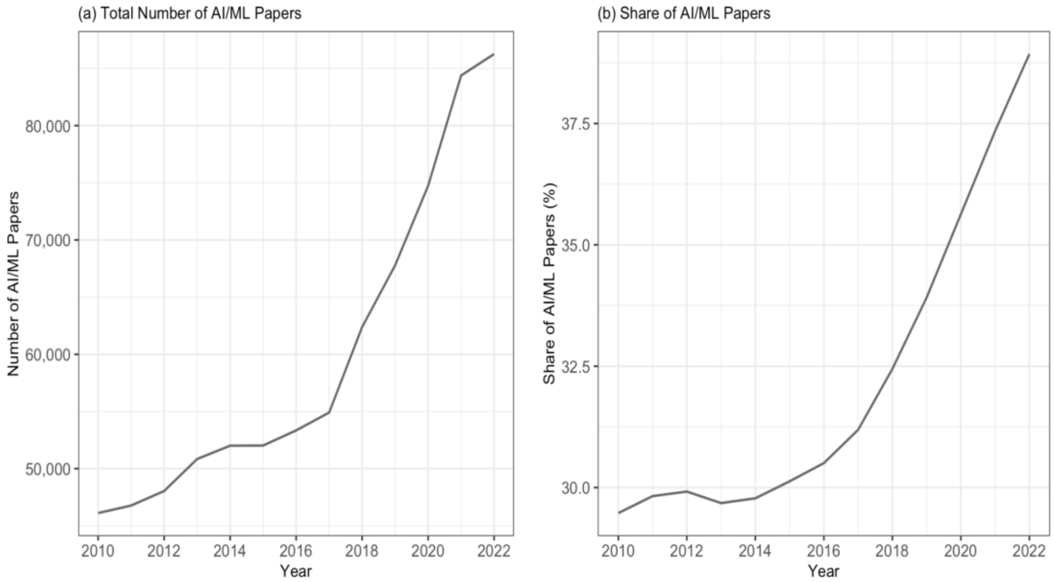
4.1. General trends

4.1.1. Growing AI adoption in science

The first statistic we present is the volume of scientific activity associated with AI (Figure 2a). Our findings are closely in line with the evidence documented in the literature (see Section 2.2), showing a steady increase during the period 2010-2015, followed by a sharp acceleration thereafter.

The first take-off can be attributed to the "deep learning revolution" sparked by the breakthrough of AlexNet (2012) and the popularisation of deep convolutional neural networks (CNNs). The second surge is likely driven by the advent and widespread adoption of transformer architectures, following Vasani et al.'s seminal paper "*Attention is all you need*". A similar pattern can be observed for the share of AI/ML papers (Figure 2b), confirming that the role played by AI is not only expanding in absolute terms, but that these tools are taking up increasingly larger share of the academic research output.

Figure 2. Trend in the total number and share of AI papers



Notes: Panel A: Total number of papers involving AI published over time. Panel B: Share of AI papers as percentage of total research output over time. Figures start from 2010 to focus on recent trends and minimize noise caused by smaller sample sizes in earlier years. The trends shown here are specific to the selected fields and should not be interpreted as representative of the entire scientific landscape. Based on our own elaboration.

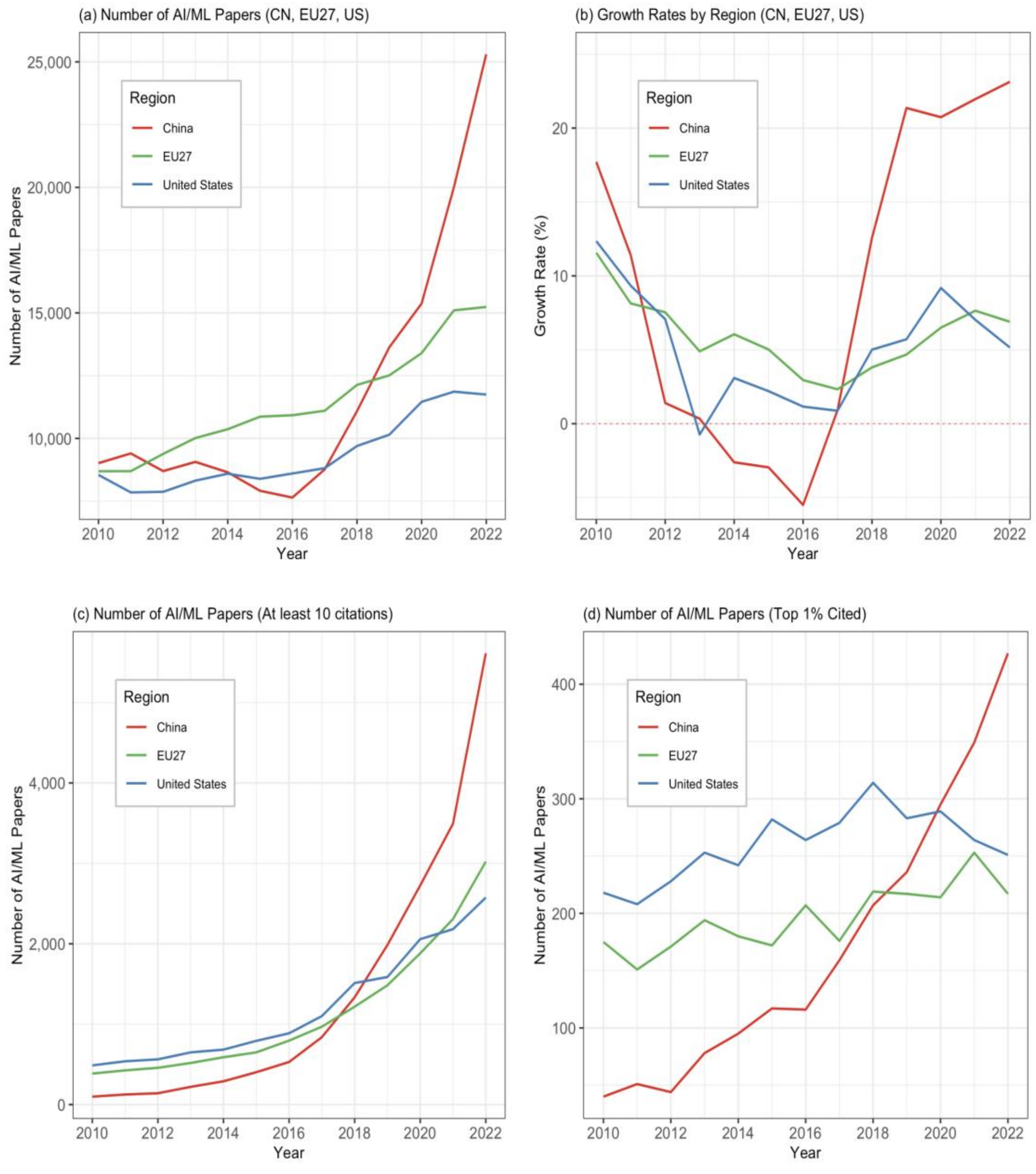
4.1.2. Global geography of AI research

Given its potential to shape global competitiveness, the race for world leadership in AI adoption in science and technology is intensifying globally. Countries are investing heavily in national AI strategies to guide and foster its development and deployment across various scientific domains and sectors of the economy. For example, China's *New Generation AI Development Plan* (2015-2030) aims to make the country as one of the world's leading AI powers by 2030; similar initiatives have been lunched in the US – the Biden administration released the updated *National Artificial Intelligence Research and Development Strategic*

Plan in May 2023 – and in Europe with its *Coordinated Plan on AI*, first published in 2018 and later updated in 2021.[10]

To compare the performance of these three global players, we analysed the number of AI papers (fractional count) published by each region. Figure 3a and 3b reveal a particularly striking performance by China: after a slight decline between 2010 and 2016, China's research output surged, surpassing both the EU and the US. In contrast, the EU and US followed a steady upward trend, though both regions experienced a slight slowdown in recent years.

Figure 3. Trend in the total number and growth rate of AI papers, by world region



Notes: Panel A: Fractional number of papers involving AI published over time by three regions (China, EU27, and the US). Panel B: Growth of AI papers calculated as 3-year rolling averages. Panel C: Same as Panel A but limited to papers with at least 10 citations. Panel D: Number of AI papers falling in the top 1% most-cited publications. Figures start from 2010 to focus on recent trends and minimize noise caused by smaller sample sizes in earlier years. The trends shown here are specific to the selected fields and should not be interpreted as representative of the entire scientific landscape. Based on our own elaboration.

The picture changes only slightly when filtering for papers that have received at least the arbitrary threshold of 10 citations (Figure 3c). Here, the US ranks ahead of the EU, except in recent years, but China still surpasses both in terms of volume after 2018. When considering the top-cited papers (Figure 3d), defined

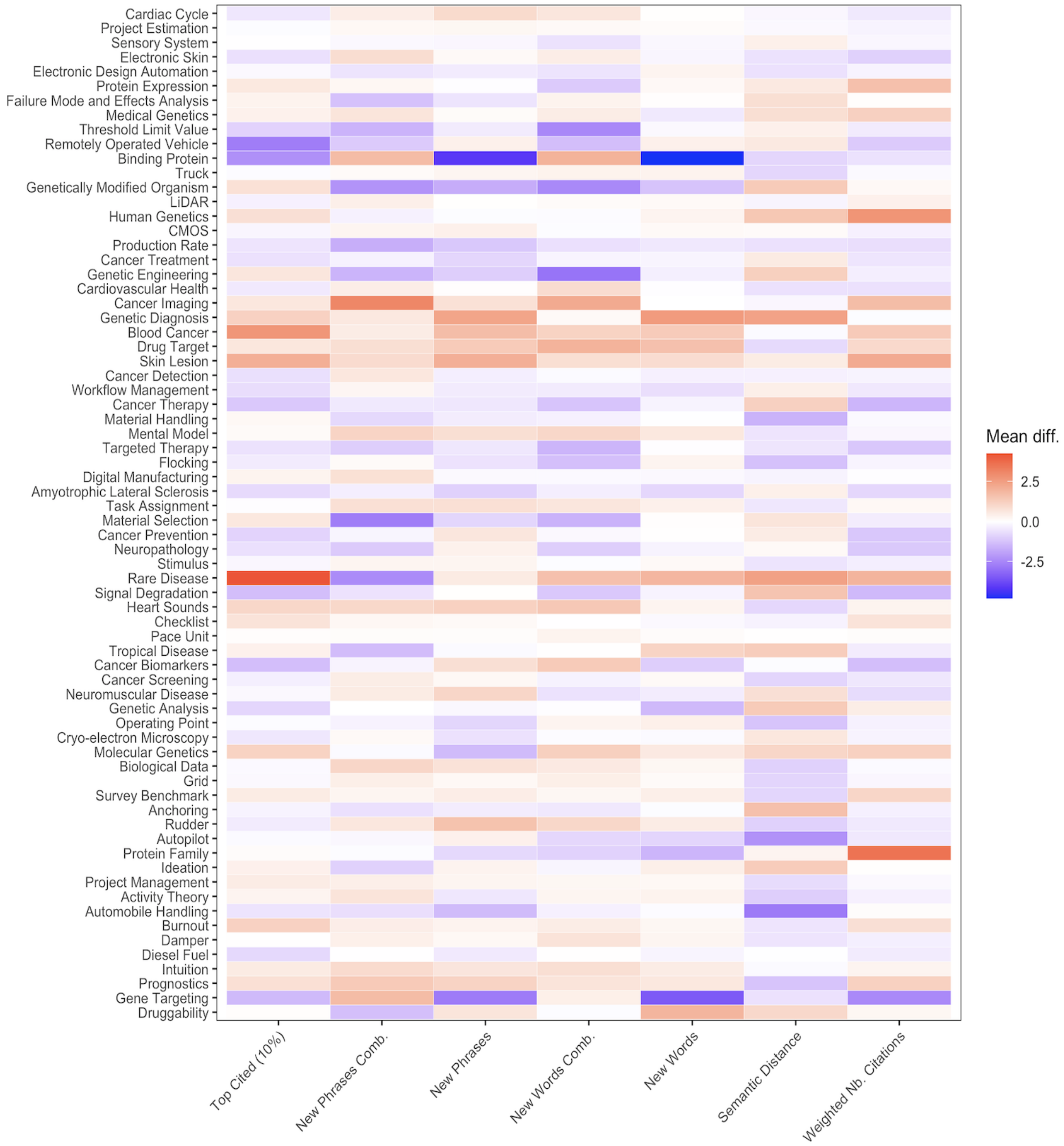
[10] On the EU Coordinated Plan on AI, see here. For more details on China's development plan, refer to Wu et al. (2020). For an overview of the over 1,000 AI policy initiatives launched by various countries, territories, and the EU, visit the OECD AI Policy Observatory (OECD, 2021). It is worth noting that "AI in science" has generally received only peripheral attention in these plans and policies, as the focus has mainly (so far) been on broader economic and social applications of AI technologies.

as those falling in the top 1% of the citation distribution, China takes the lead starting in 2020. In this case, the US systematically outperforms the EU.

4.1.3. Heterogeneity across fields

Figure 4 shows the mean difference between AI papers and other scientific publications across our indicators of novelty and impact, analysed by scientific fields. A clear takeaway from this figure is the substantial heterogeneity in how AI influences research output.

Figure 4. Novelty and impact for AI and non-AI papers across fields



Notes: The plot compares the mean difference between AI papers and other scientific publications across various indicators of novelty and impact (x-axis) and scientific fields (y-axis). All variables are standardised to facilitate comparison across indicators. Darker shades of red indicate that AI/ML papers exhibit a higher mean for a given indicator compared to non-AI papers, while blue shades indicate the opposite. Based on our own elaboration.

At a more aggregate level – concepts level-1 (see Figure C2 in Appendix C) – we identify areas where AI has strong influence (e.g., *Pathology*) and others with untapped potential (e.g., *Industrial Engineering*). More interesting, though puzzling, is the degree of heterogeneity that emerges at the finer-grained level within the domains, as illustrated in Figure 4. Several health-related fields (e.g., *Cancer Imaging*, *Blood Cancer*, *Drug Target*) benefit from balanced positive effects on both novelty and impact. But also, there are fields where AI positively affects impact but has limited effects on novelty (e.g., *Protein Family*, *Human Genetics*); and others seem to experience minimal or even negative effects (e.g., *Project Estimation*, *Production Rate*).

To be clear, different factors could explain the different effects across disciplines: some fields may lag in adopting AI/ML tools due to cultural barriers, insufficient data, or inadequate computing infrastructures; others may face misalignment between AI tools and field-specific research needs or differ in their dataintensiveness. We recognise the importance all these factors, yet we are unable to measure them directly and hope future research will explore these differences in greater depth. That said, our primary focus is on whether the combinatorial complexity of a field's knowledge space can influence the effects of AI on scientific creativity.

4.1.4. AI in different knowledge landscapes

Indeed, an important aspect is whether the strength of AI's contribution to scientific novelty and impact varies according to the topology of the knowledge space that characterises a scientific field.

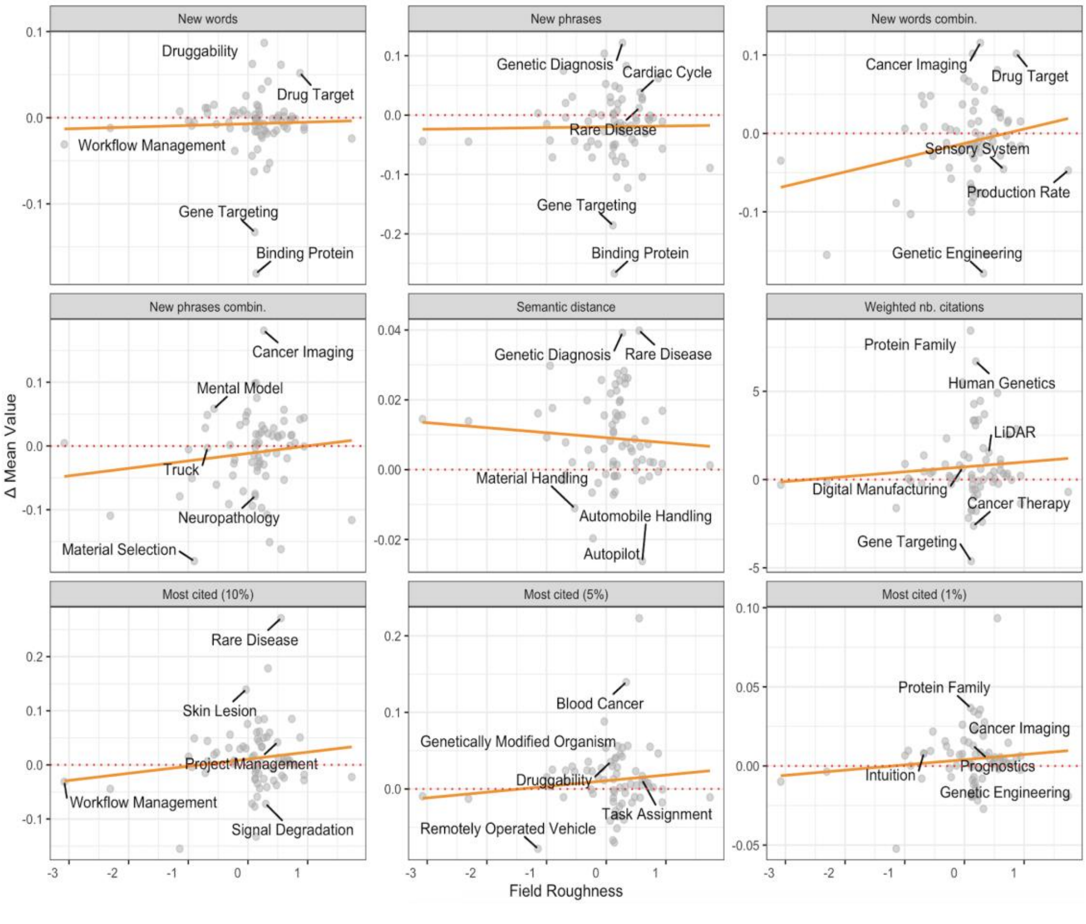
Figure 5 shows the relationship between the mean "roughness", or combinatorial complexity, of a field (x-axis) and the mean difference between AI and non-AI papers for various indicators (y-axis). Overall, we observe that the influence of AI on different metrics of novelty (apart from semantic distance) and impact tends to intensify as the roughness of the knowledge space increases.

A closer look at selected fields reveals that, e.g., *Cancer Imaging*, *Drug Target*, or *Rare Diseases* exhibit the highest positive returns from AI, particularly in generating novel ideas and achieving high citation impacts. Research in these fields relies heavily on different areas of expertise, and the knowledge space is, as a result, complex and fragmented: as an example, cancer imaging combines principles from medical imaging, molecular biology, and data analytics, whereas drug targeting relies on chemistry, genomics, and pharmacological modelling. As we see from the trends in Figure 5, AI seems particularly well suited to handle this complexity.

Other areas such as *Workflow Management* or *Remotely Operated Vehicles* may operate within clear boundaries, with research often progressing through incremental improvements. Here, AI's contribution appears more limited. While it can still provide gains in terms of efficiency in low-complexity fields, its transformative potential is comparatively weaker, as its use may be tied to optimising existing processes or models rather than reshaping the fundamental knowledge structure of the field.

It is important to note that the discussion above is univariate and does not account for potential confounding factors. Further evidence, some of which partially contrasts the patterns observed in Figure 5, will be presented in the following sections.

Figure 5. AI contribution to novelty and impact, by the combinatorial complexity of fields



Notes: The plot represents the relationship between the mean combinatorial complexity, or "roughness", of scientific fields (x-axis) and the mean difference in a specific metric between AI and non-AI papers (y-axis). All variables are standardized to facilitate comparison across indicators. A positive delta indicates that, on average, AI papers contributes more to a given metric, whereas negative values suggest a weaker contribution. The orange line represents a linear fit to capture the overall trend. Based on our own elaboration.

4.2. Econometric results

4.2.1. AI on novelty and impact

Table 3 reports the estimated coefficients for the effect of AI on different indicators of scientific novelty. Taken together, we find a positive correlation between AI adoption and all metrics of novelty. More specifically, AI increases the likelihood of introducing new words, new phrases, and new words/phrases combinations (Columns 1–4), an important signal of conceptual novelty; moreover, the integration of AI seems to lead to scientific outputs that are semantically more distant from past research, hence more novel (Column 5).

To further validate these findings, we conducted a manual inspection of some AI-related papers to spot newly introduced terms that were later widely re-used in future research. For example, in domains where the effect on novelty is particularly strong (see Figure 5) – e.g., *Drug Target*, *Druggability*, *Genetic Diagnosis* – we found new terms associated with the introduction of cutting-edge techniques for predicting drug-target binding affinity (*DeepDTA*), automated chemical compound classification (*ClassyFire*), or genetic newborn screening (*Screen4Care*). We therefore confirm that AI brings with it the capabilities to cope with the vast combinatorial space characteristic of fields such as chemistry and genetics, where the sheer number of possible interactions and configurations is a major challenge. We also came across new noun phrases such as '*drug similarity assessment*', '*predict drug repositioning*', or '*protein concavity*'. The appearance of these expressions reflects methodological refinements, practical applications and reframing of scientific questions, all reinforcing the idea that AI expands the scope of scientific inquiry by providing new tools and opening up new research possibilities – e.g., the identification of new therapeutic uses for existing or investigational drugs.[11]

Table 3. The effect of AI on novelty

	New Words	*New Words Comb.*	*New Phrases*	*New Phrases Comb.*	*Semantic Distance*
	(1)	(2)	(3)	(4)	(5)
AI	**0.079***** **(0.009)**	**0.110***** **(0.004)**	**0.112***** **(0.005)**	**0.201***** **(0.005)**	**0.005***** **(0.0001)**
Nb. Authors	0.029*** (0.001)	0.061*** (0.001)	0.036*** (0.001)	0.067*** (0.001)	-0.0004*** (0.00001)
Nb. References	0.016*** (0.003)	0.194*** (0.001)	0.099*** (0.001)	0.252*** (0.001)	-0.001*** (0.00002)
International Collab.	0.096*** (0.010)	0.123*** (0.006)	0.072*** (0.006)	0.113*** (0.006)	0.001*** (0.0001)
Survey/Review	-0.270*** (0.042)	-0.348*** (0.016)	-0.255*** (0.021)	-0.284*** (0.017)	0.001*** (0.0003)
Adjusted R^2					0.160
Log Likelihood	-287,885	-824,636	-741,794	-726,505	
AIC	575,985	1,649,488	1,483,803	1,453,225	
# Observations	1,397,173	1,397,173	1,397,173	1,397,173	1,397,173

Notes: The econometric models for evaluating the effect of AI on various indicators of novelty. Parameters in Columns 1–4 are estimated via Logit models, while those in Column 1 via OLS regression. All specifications include fixed effects for time and scientific fields. The number of observations falls due to the presence of missing values (NA) in the novelty metrics. The asterisks ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 4 presents the estimates of the effect of AI on citation counts. Here again, we can see a positive correlation between AI and all metrics of impact. As shown in Column 1, AI papers receive, on average, ~3% more citations compared to non-AI papers. And, interestingly, the magnitude of the effect strengthens

[11] Note that these are just a few examples, mainly for an internal check on the validity of our metrics. An in-depth analysis of newly introduced terms is beyond the scope of this study and, admittedly, outside our expertise. However, the dataset we are releasing with the publication (*add link later*) is an open-source resource for future research that may aim to explore this phenomenon in more depth.

as we move toward higher citation thresholds (Column 2–5) – see also Figure C2 in Appendix C. This suggests that AI disproportionately contributes to highly influential research. To get a sense of magnitude, the odd ratio from the Logit model of Column 5 tells us that AI papers are ~40% more likely to rank among the top 1% of most cited papers.

Table 4. The effect of AI on impact

	Weighted Nb. Citations	*Top Cited (10%)*	*Top Cited (5%)*	*Top Cited (1%)*
	(1)	(2)	(3)	(4)
AI	**0.030***** **(0.001)**	**0.168***** **(0.005)**	**0.220***** **(0.007)**	**0.338***** **(0.014)**
Nb. Authors	0.032*** (0.0001)	0.088*** (0.001)	0.077*** (0.001)	0.057*** (0.001)
Nb. References	0.282*** (0.0003)	1.122*** (0.003)	1.140*** (0.004)	1.170*** (0.007)
International Collab.	0.191*** (0.001)	0.391*** (0.005)	0.412*** (0.007)	0.483*** (0.014)
Survey/Review	0.134*** (0.004)	0.394*** (0.017)	0.490*** (0.021)	0.847*** (0.033)
Adjusted R^2	0.409			
Log Likelihood		-735,378	-464,142	-139,458
AIC		1,470,971	928,498	279,131
# Observations	2,889,302	2,889,302	2,889,302	2,889,302

Notes: The econometric models for evaluating the effect of AI on various indicators of impact. Parameters in Column 1 are estimated via OLS regression, while those in Columns 2–4 via Logit models. All specifications include fixed effects for time and scientific fields. The asterisks ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

The signs for control variables conform with our expectations: papers with more authors, international collaborations, higher reference counts, or surveys tend to receive a citation premium. It goes without saying that our models show correlations rather than definitive proof of causation – e.g., there could be endogeneity if, say, high-novelty teams might be more inclined to adopt AI.

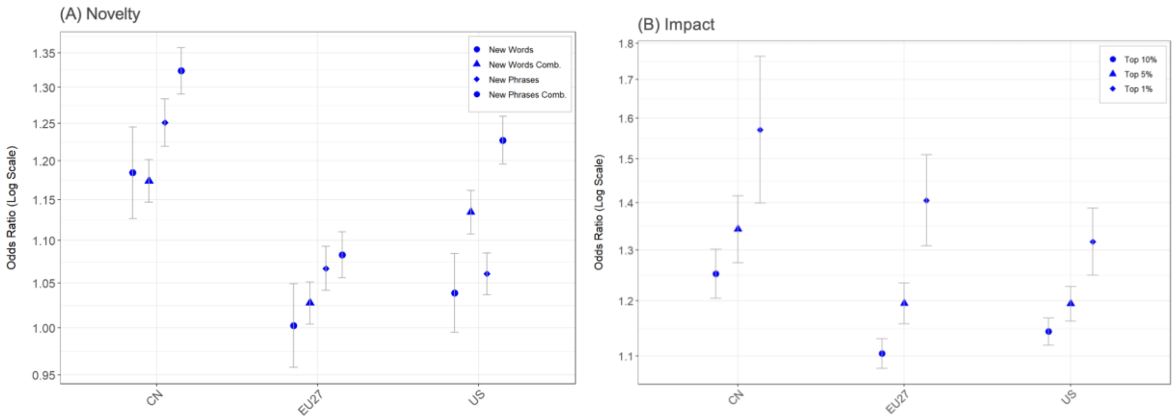
4.2.2. Differences across geographies

Do the benefits of AI vary by geographical area? The analysis shown in Figure 6 suggests that this is the case. Across novelty indicators (Panel A) and citation thresholds (Panel B), the contribution of AI to scientific outcome appears strongest in China, while the EU and the US exhibit similar trends.

So, China has not only caught up in terms of scientific production – a trend already documented in Section 4.1.2 – but the substantial government investment in AI research and large-scale AI infrastructures aimed at accelerating technological progress (Wu et al., 2020) has positioned the country as a global leader also in terms of AI-driven novelty and impact.[12]

[12] Our findings come with two main caveats. First, we report average effects across scientific fields; thus, we cannot exclude the possibility that AI has a more pronounced impact in specific domains within certain geographical areas, where a stronger domain-specific knowledge base may exist. Second, further research is needed to assess whether China's gains in novelty and impact are primarily a reflection of the sheer volume of research output or reflect truly disruptive breakthroughs. For this, future research could investigate the qualitative nature of AI-enhanced discoveries across different regions.

Figure 6. The effect of AI on novelty and impact, by geographical areas



Notes: The plot represents the odd ratios from Logit models estimating the effects of AI on novelty (Panel A) and impact (Panel B), separately for China, the UE, and the US. The econometric models are equivalent to those presented in Tables 3 and 4. Vertical bars indicate 95% confidence intervals.

4.2.3. Differences across fields

Do the benefits of AI vary by the degree of combinatorial complexity within a field? The estimates presented in Tables 5 and 6 provide partial support for this hypothesis.

We conducted two complementary analyses. First, we estimated the same models as above but introduced an interaction term between AI and the yearly mean field roughness (*AI* × *Field Roughness*); hence, the interaction term allows us to assess whether the effect of AI on novelty and impact is amplified or diminished in fields with higher combinatorial complexity. Second, we estimated separated models on subsamples classified as "low roughness" and "high roughness", where a field is categorized based on whether its roughness score falls below or above the yearly median, respectively. The results of this second exercise are reported in Appendix C.

The estimates in Table 5 confirm a reinforcing effect of AI on the introduction of new words and phrases (Columns 1 and 2), as well as their combinations (Columns 3 and 4), in fields with a high degree of combinatorial complexity, suggesting that AI plays a stronger role in fostering novelty in more fragmented knowledge spaces. To our knowledge, this is the first empirical evidence supporting theoretical predictions from recent models of search and discovery with intelligent machines over vast and complex combinatorial design spaces (see, e.g., Agrawal et al., 2018; Agrawal et al., 2024; Besiroglu et al., 2024). The robustness checks reported in Appendix C (Table C2) further support these findings, showing that AI has little to no effects on novelty in dense, well-connected knowledge space, whereas its effects are highly significant and positive in fields where knowledge elements are sparse and more disconnected.

Table 5. The effect of AI on novelty, by field "roughness" (interaction term)

	New Words	*New Words Comb.*	*New Phrases*	*New Phrases Comb.*	*Semantic Distance*
	(1)	(2)	(3)	(4)	(5)
AI	**0.061***** **(0.011)**	**0.085***** **(0.006)**	**0.080***** **(0.006)**	**0.166***** **(0.006)**	**0.006***** **(0.0001)**
Field Roughness	0.030 (0.018)	-0.040*** (0.010)	-0.016* (0.009)	-0.052*** (0.010)	0.001*** (0.0002)
AI × Field Roughness	**0.059***** **(0.017)**	**0.072***** **(0.009)**	**0.083***** **(0.009)**	**0.093***** **(0.009)**	**-0.002***** **(0.0002)**
Nb. Authors	0.029*** (0.001)	0.036*** (0.001)	0.061*** (0.001)	0.068*** (0.001)	-0.0003*** (0.00001)
Nb. References	0.015*** (0.003)	0.099*** (0.001)	0.192*** (0.001)	0.250*** (0.001)	-0.001*** (0.00002)
International Collab.	0.097*** (0.010)	0.073*** (0.006)	0.123*** (0.006)	0.114*** (0.006)	0.001*** (0.0001)
Survey/Review	-0.289*** (0.043)	-0.256*** (0.021)	-0.352*** (0.016)	-0.287*** (0.017)	0.001*** (0.0003)
Adjusted R^2					0.160
Log Likelihood	-280,795	-727,025	-812,721	-715,745	
AIC	561,806	1,454,268	1,625,659	1,431,706	
# Observations	1,374,533	1,374,533	1,374,533	1,374,533	1,374,533

Notes: The econometric models for evaluating the interaction between AI and field roughness on various indicators of novelty. Parameters in Columns 1–4 are estimated via Logit models, while those in Column 5 via OLS regression. All specifications include fixed effects for time and scientific fields. The asterisks ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

The negative sign of the interaction term for semantic distance (Column 5) might seem ambiguous, but in our judgment, it is particularly interesting. Recall that semantic distance primarily captures textual similarity, which can be different from conceptual novelty. Indeed, a lower semantic distance does not necessarily mean that the content is less novel but may instead indicate that AI-research tends to preserve some linguistic conventions (writing styles, framing of ideas, etc.) which are closer to the existing literature. Seen from this perspective, one possible explanation is that for novel research to be understandable and communicable within the scientific community, it must first be anchored in established scientific language. In other words, the balance between novelty and readability could explain why AI papers in complex knowledge spaces introduce new terminology but exhibit lower semantic distance.

Moving to the effects on citation counts, as shown in Table 6, we find no interaction effects. If anything, there appears to be a negative effect, albeit small in magnitude, on weighted number of citations (Column 1). These results partially contradict the estimates presented in Appendix C (Table C3), where the coefficients on AI are systematically higher in the sample comprising fields with a higher combinatorial complexity. Yet, given these discrepancies, we refrain from drawing robust conclusions on the relationship between AI adoption and citation-based impact across different knowledge landscapes.

Table 6. AI on impact, by field "roughness" (interaction term)

	Weighted Nb. Citations	*Top Cited (10%)*	*Top Cited (5%)*	*Top Cited (1%)*
	(1)	(2)	(3)	(4)
AI	**0.035*****	**0.165*****	**0.216*****	**0.337*****
	(0.001)	**(0.006)**	**(0.008)**	**(0.016)**
Field Roughness	0.140***	-0.304***	-0.303***	-0.302***
	(0.002)	(0.010)	(0.014)	(0.029)
AI × Field Roughness	**-0.011*****	**0.009**	**0.017**	**0.012**
	(0.002)	**(0.009)**	**(0.012)**	**(0.024)**
Nb. Authors	0.032***	0.088***	0.077***	0.057***
	(0.0001)	(0.001)	(0.001)	(0.001)
Nb. References	0.282***	1.134***	1.152***	1.181***
	(0.0003)	(0.003)	(0.004)	(0.008)
International Collab.	0.192***	0.390***	0.412***	0.485***
	(0.001)	(0.005)	(0.007)	(0.014)
Survey/Review	0.135***	0.387***	0.484***	0.847***
	(0.004)	(0.017)	(0.021)	(0.033)
Adjusted R^2	0.410			
Log Likelihood		-723,147	-456,482	-137,085
AIC		1,446,512	913,180	274,386
# Observations	2,846,014	2,846,014	2,846,014	2,846,014

Notes: The econometric models for evaluating the interaction between AI and field roughness on various indicators of impact. Parameters in Column 1 are estimated via OLS regression, while those in Columns 2–4 via Logit models. All specifications include fixed effects for time and scientific fields. The asterisks ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

5. Conclusion

AI is no longer confined to niche applications; from A for Archaeology to Z for Zoology, it is now (re)shaping research across disciplines and throughout the scientific workflow. Its transformative and far-reaching nature makes it a powerful “method of invention”, one that could help pull science out of the productivity slump of recent decades. But with big promises come big challenges, and as we have discussed in this paper, AI could just as easily hinder scientific progress as accelerate it.

Overall, our results offer an optimistic perspective on the benefits AI can bring to scientific creativity – acknowledging, of course, the inherent limitations of measuring such effects through bibliometric indicators of novelty and impact, as done in our study. Importantly, we also find that the influence of AI does not depend just on the technology itself, but also on how, and where, it is used. In fact, two other main findings are worth noting. First, AI seems to have the biggest impact in fields where the knowledge space is sparse and complex, meaning more fragmented and disconnected ideas. Second, regional differences are striking: in recent years, China has taken the lead in AI-driven research, outpacing both the US and the EU, not just in sheer output, but also in terms of scientific novelty and impact.

These findings have implications for policy, on the one hand, and for deeper philosophical and epistemological questions about what it means to “*be a scientist*” in the age of intelligent machines, on the other.

Let us start with the first point. While securing strategic advantage in AI research is a shared priority for most governments – and the application of AI for R&D is implicit in policy agendas aimed at strengthening research and industrial competitiveness – most national AI strategies pay limited *explicit* attention to the role of AI in science. A review of 32 national AI strategies found that very few included concrete measures to support AI in scientific research (OECD AI Policy Observatory 2024). Beyond this gap, the uneven distribution of AI’s benefits across disciplines documented in this study raises important policy considerations. Indeed, it may suggest that some disciplines may be underutilising AI’s potential, hence calling for targeted interventions to encourage both adoption and effective use. Policies could support these fields through financial incentives, improved infrastructure, and specialised training programs tailored to their unique needs. However, there is little to no systematic evidence on what funding streams are most effective to support the immediate deployment of existing AI capabilities and longer-term core AI research; what the actual infrastructures needs are across different application domains (e.g., computing power and data access); and how best to design advanced training programs to enhance AI literacy and skills among researchers. This gap is clearly a call for future research, which could extend and complement our analysis in several ways: first, by incorporating more recent data and thus capturing the latest technological development (*in primis* GenAI and LLMs, but also very recent reasoning models); second, by broadening the scope to additional scientific fields, ideally with a different degree of granularity as we have done here; third, by exploring alternative metrics of scientific outcomes and productivity; finally, by employing causal inference approaches to assess the effects of AI-oriented grants, access to computing facilities, or other institutional interventions aimed at fostering the adoption of AI in science. Of course, this list is far from exhaustive.

As discussed in this paper, the perils of AI are real and significant, and cannot be overlooked. The impressive acceleration in both the capabilities and popularity of AI systems has been accompanied by increasing fears regarding human ability to keep this fast-evolving technology under control. Policymaking, therefore, must remain agile and responsive to the emerging challenges, striking a delicate balance between incentives for AI-driven science and containment of risks linked to, among others, the black-box nature of complex AI/ML models, as well as potential misuses and biases, and research governance. Indeed, wider adoption of AI for science brings with it new challenges for research governance, including concerns about AI’s impact on research integrity and ethics (Resnik and Hosseini, 2024). In 2024, the European Commission published some guidelines on the responsible use of AI in science, outlining four basic principles: *reliability* (to ensure research quality), *honesty* (in reporting where and how AI has been used), *respect* (for the range of stakeholders that might be affected by AI in research), and *accountability* (for the use of AI and its outputs). But there is a need for practical guidance, which need to be domain-specific to gain traction. While high-level principles are clearly valuable in setting a direction of travel, their practical implementation will require translation to specific scientific contexts.

For the first time in history, scientists are confronted with a form of intelligence that, in many ways, mirrors our own in complexity and capability. This unprecedented interaction offers, therefore, an opportunity for introspection and self-examination: *what makes us, as human scientists, different from intelligent machines?* True, intelligence may be ever-increasing among machines, but genuinely creative intuitive thinking requires making mistakes, abandoning logic from one moment to the next, and learning through unpredictability. This is because our brains are a complex mix of determinism, chaos, and randomness, and current AI, despite its sophistication, still falls short in this respect.

Another aspect is vision, not to be confused with mere sight. Machines lacks (at least so far) vision: they do not independently decide to explore distant galaxies, though they excel at processing astronomical data once directed to do so; machines are better than most scientists at solving complex problems in calculus and quantum mechanics, yet they do not have the vision to see the need for such constructs in the first place.

Bottom line: to solve the great mysteries of nature, we will probably need intelligences other than human. The use of AI in science, thus, will move from a luxury to a necessity. But science is not just about answering questions, it is also about asking the right ones.

6. References

Abbonato, D., Bianchini, S., Gargiulo, F., & Venturini, T. (2024). Interdisciplinary research in artificial intelligence: Lessons from COVID-19. *Quantitative Science Studies*, 1-14. https://doi.org/10.1162/qss_a_00329

Aghion, P., Antonin, C., & Bunel, S. (2021). *The power of creative destruction: Economic upheaval and the wealth of nations*. Harvard University Press. https://doi.org/10.4159/9780674258686

Aghion, P., Jones, B. F., & Jones, C. I. (2017). Artificial intelligence and economic growth (Vol. 23928). Cambridge, MA: NBER. https://doi.org/10.7208/9780226613475-011

Agrawal, A., McHale, J., & Oettl, A. (2018). Finding needles in haystacks: Artificial intelligence and recombinant growth. In *The economics of artificial intelligence: An agenda* (pp. 149-174). University of Chicago Press. http://www.nber.org/chapters/c14024

Agrawal, A., McHale, J., & Oettl, A. (2024). Artificial intelligence and scientific discovery: A model of prioritized search. *Research Policy*, 53(5), 104989. https://doi.org/10.1016/j.respol.2024.104989

Ahmed, N., & Wahed, M. (2020). The De-democratization of AI: Deep learning and the compute divide in artificial intelligence research. *arXiv preprint arXiv:2010.15581*. https://arxiv.org/abs/2010.15581

Ahmed, N., Wahed, M., & Thompson, N. C. (2023). The growing influence of industry in AI research. *Science*, 379(6635), 884-886. https://doi.org/10.1126/science.ade2420

AI4Science, M. R., & Quantum, M. A. (2023). The impact of large language models on scientific discovery: A preliminary study using GPT-4. *arXiv preprint arXiv:2311.07361*. https://arxiv.org/abs/2311.07361

AlShebli, B., Cheng, E., Waniek, M., Jagannathan, R., Hernández-Lagos, P., & Rahwan, T. (2022). Beijing's central role in global artificial intelligence research. *Scientific Reports*, 12(1), 21461. https://doi.org/10.1038/s41598-022-25714-0

AlShebli, B., Memon, S. A., Evans, J. A., & Rahwan, T. (2024). China and the US produce more impactful AI research when collaborating together. *Scientific Reports*, 14(1), 28576. https://doi.org/10.1038/s41598-024-79863-5

Amabile, T. M., & Pratt, M. G. (2016). The dynamic componential model of creativity and innovation in organizations: Making progress, making meaning. *Research in Organizational Behavior*, 36, 157-183. https://doi.org/10.1016/j.riob.2016.10.001

Arranz, D., Bianchini, S., Di Girolamo, V. & Ravet, J. (2023). *Trends in the use of AI in science: A bibliometric analysis*. European Commission, Directorate-General for Research and Innovation, Publications Office of the European Union. https://data.europa.eu/doi/10.2777/418191

Arthur, W. B. (2009). *The nature of technology: What it is and how it evolves*. Simon and Schuster.

Arts, S., Melluso, N., & Veugelers, R. (2025). Beyond citations: Measuring novel scientific ideas and their impact in publication text. *The Review of Economics and Statistics* (forthcoming). https://arxiv.org/abs/2309.16437

Athaluri, S. A., Manthena, S. V., Kesapragada, V. K. M., Yarlagadda, V., Dave, T., & Duddumpudi, R. T. S. (2023). Exploring the boundaries of reality: Investigating the phenomenon of artificial intelligence hallucination in scientific writing through ChatGPT references. *Cureus*, 15(4). https://doi.org/10.7759/cureus.37432

Azoulay, P., & Li, D. (2020). *Scientific grant funding* (Vol. 26889). Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.7208/chicago/9780226805597-008

Bahoo, S., Cucculelli, M., & Qamar, D. (2023). Artificial intelligence and corporate innovation: A review and research agenda. *Technological Forecasting and Social Change*, 188, 122264. https://doi.org/10.1016/j.techfore.2022.122264

Besiroglu, T., Emery-Xu, N., & Thompson, N. (2024). Economic impacts of AI-augmented R&D. *Research Policy*, 53(7), 105037. https://doi.org/10.1016/j.respol.2024.105037

Beutel, G., Geerits, E., & Kielstein, J. T. (2023). Artificial hallucination: GPT on LSD?. *Critical Care*, 27(1), 148. https://doi.org/10.1186/s13054-023-04425-6

Bianchini, S., Müller, M., & Pelletier, P. (2022). Artificial intelligence in science: An emerging general method of invention. *Research Policy*, 51(10), 104604. https://doi.org/10.1016/j.respol.2022.104604

Bianchini, S., Müller, M., & Pelletier, P. (2024). Drivers and barriers of AI adoption and use in scientific research. arXiv preprint arXiv:2312.09843. https://arxiv.org/abs/2312.09843

Birhane, A., Kasirzadeh, A., Leslie, D., & Wachter, S. (2023). Science in the age of large language models. *Nature Reviews Physics*, 5(5), 277-280. https://doi.org/10.1038/s42254-023-00581-4

Bloom, N., Jones, C. I., Van Reenen, J., & Webb, M. (2020). Are ideas getting harder to find?. *American Economic Review*, 110(4), 1104-1144. https://doi.org/10.1257/aer.20180338

Boden, M. A. (2004). *The creative mind: Myths and mechanisms*. Routledge.

Boeing, P., & Hünermund, P. (2020). A global decline in research productivity? Evidence from China and Germany. *Economics Letters*, 197, 109646. https://doi.org/10.1016/j.econlet.2020.109646

Brynjolfsson, E., Rock, D., & Syverson, C. (2019). Artificial intelligence and the modern productivity paradox. *The economics of artificial intelligence: An agenda*, 23, 23-57. https://doi.org/10.7208/9780226613475-003

Budach, L., Feuerpfeil, M., Ihde, N., Nathansen, A., Noack, N., Patzlaff, H., ... & Harmouch, H. (2022). The effects of data quality on machine learning performance. *arXiv preprint arXiv:2207.14529*. https://arxiv.org/abs/2207.14529

Chen, Q., Ho, Y. J. I., Sun, P., & Wang, D. (2024). The philosopher's stone for science–The catalyst change of AI for scientific creativity. (March 5, 2024). https://ssrn.com/abstract=4749118

Cockburn, I. M., Henderson, R., & Stern, S. (2018). The impact of artificial intelligence on innovation (Vol. 24449). Cambridge, MA, USA: NBER. https://doi.org/10.3386/w24449

Cowen, T. (2011). *The great stagnation: How America ate all the low-hanging fruit of modern history, got sick, and will (eventually) feel better*: Penguin.

Crafts, N. (2021). Artificial intelligence as a general-purpose technology: An historical perspective. *Oxford Review of Economic Policy*, 37(3):521-536. https://doi.org/10.1093/oxrep/grab012

De Regt, H. W. (2017). *Understanding scientific understanding*. Oxford University Press.

Duede, E., Dolan, W., Bauer, A., Foster, I., & Lakhani, K. (2024). Oil & water? Diffusion of AI within and across scientific fields. *arXiv preprint arXiv:2405.15828*. https://arxiv.org/abs/2405.15828

Felin, T., & Holweg, M. (2024). Theory is all you need: AI, human cognition, and causal reasoning. *Strategy Science*, 9(4), 346-371. https://doi.org/10.1287/stsc.2024.0189

Fleming, L. (2001). Recombinant uncertainty in technological search. *Management Science*, 47(1), 117-132. https://doi.org/10.1287/mnsc.47.1.117.10671

Fontana, M., Iori, M., Montobbio, F., & Sinatra, R. (2020). New and atypical combinations: An assessment of novelty and interdisciplinarity. *Research Policy*, 49(7), 104063. https://doi.org/10.1016/j.respol.2020.104063

Franzoni, C., Stephan, P., & Veugelers, R. (2022). Funding risky research. *Entrepreneurship and Innovation Policy and the Economy*, 1(1), 103-133. https://doi.org/10.1086/719252

Furman, J. L., & Teodoridis, F. (2020). Automation, research technology, and researchers' trajectories: Evidence from computer science and electrical engineering. *Organization Science*, 31(2), 330-354. https://doi.org/10.1287/orsc.2019.1308

Gao, J., & Wang, D. (2024). Quantifying the use and potential benefits of artificial intelligence in scientific research. *Nature Human Behaviour*, 1-12. https://doi.org/10.1038/s41562-024-02020-5

Gargiulo, F., Fontaine, S., Dubois, M., & Tubaro, P. (2023). A meso-scale cartography of the AI ecosystem. *Quantitative Science Studies*, 4(3), 574-593. https://doi.org/10.1162/qss_a_00267

Gofman, M., & Jin, Z. (2024). Artificial intelligence, education, and entrepreneurship. *The Journal of Finance*, 79(1), 631-667. https://doi.org/10.1111/jofi.13302

Goldin, I., Koutroumpis, P., Lafond, F., & Winkler, J. (2024). Why is productivity slowing down?. *Journal of Economic Literature*, 62(1), 196-268. https://doi.org/10.1257/jel.20221543

Gordon, R. (2017). *The rise and fall of American growth: The US standard of living since the civil war*. Princeton university press. https://doi.org/10.1515/9781400888955

Griffin, C., Wallace, D., Mateos-Garcia, J., Schieve, H., & Kohli, P. (2024). A new golden age of discovery: Seizing the AI for science opportunity. *Google DeepMind, AI Policy Perspectives*. https://deepmind.google/public-policy/ai-for-science/

Grimes, M., Von Krogh, G., Feuerriegel, S., Rink, F., & Gruber, M. (2023). From scarcity to abundance: Scholars and scholarship in an age of generative artificial intelligence. *Academy of Management Journal*, 66(6), 1617-1624. https://doi.org/10.5465/amj.2023.4006

Hajkowicz, S., Sanderson, C., Karimi, S., Bratanova, A., & Naughtin, C. (2023). Artificial intelligence adoption in the physical sciences, natural sciences, life sciences, social sciences and the arts and humanities: A bibliometric analysis of research publications from 1960-2021. *Technology in Society*, 74, 102260. https://doi.org/10.1016/j.techsoc.2023.102260

Hanson, B., Stall, S., Cutcher-Gershenfeld, J., Vrouwenvelder, K., Wirz, C., Rao, Y., & Peng, G. (2023). Garbage in, garbage out: Mitigating risks and maximizing benefits of AI in research. *Nature*, *623*(7985), 28-31. https://doi.org/10.1038/d41586-023-03316-8

Hastings, J. (2023). *AI for scientific discovery*. CRC Press. https://doi.org/10.1201/9781003226642

Jones, B. F. (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder?. *The Review of Economic Studies*, 76(1), 283-317. https://doi.org/10.1111/j.1467-937X.2008.00531.x

Jones, B. F., Wuchty, S., & Uzzi, B. (2008). Multi-university research teams: Shifting impact, geography, and stratification in science. *Science*, 322(5905), 1259-1262. https://doi.org/10.1126/science.1158357

Jurowetzki, R., Hain, D. S., Wirtz, K., & Bianchini, S. (2025). The private sector is hoarding AI researchers: What implications for science? *AI & Society*, https://doi.org/10.1007/s00146-024-02171-z

Kauffman, S. A. (1993). The Structure of Rugged Fitness Landscapes. In *The Origins of Order: Self-Organization and Selection in Evolution* (New York, NY, 1993; Oxford Academic, 31 Oct. 2023), https://doi.org/10.1093/oso/9780195079517.003.0002

King, R. D., Rowland, J., Oliver, S. G., Young, M., Aubrey, W., Byrne, E., ... & Clare, A. (2009). The automation of science. *Science*, 324(5923), 85-89. https://doi.org/10.1126/science.1165620

Klinger, J., Mateos-Garcia, J., & Stathoulopoulos, K. (2021). Deep learning, deep change? Mapping the evolution and geography of a general-purpose technology. *Scientometrics*, *126*, 5589-5621. https://doi.org/10.1007/s11192-021-03936-9

Koehler, M., & Sauermann, H. (2024). Algorithmic management in scientific research. *Research Policy*, 53(4), 104985.

Krenn, M., Pollice, R., Guo, S. Y., Aldeghi, M., Cervera-Lierta, A., Friederich, P., ... & Aspuru-Guzik, A. (2022). On scientific understanding with artificial intelligence. *Nature Reviews Physics*, 4(12), 761-769. https://doi.org/10.1038/s42254-022-00518-3

Kuhn, T. S. (1962). *The structure of scientific revolutions*. University of Chicago Press.

Messeri, L., & Crockett, M. J. (2024). Artificial intelligence and illusions of understanding in scientific research. *Nature*, 627(8002), 49-58. https://doi.org/10.1038/s41586-024-07146-0

Miyagawa, T., & Ishikawa. T. (2019). On the Decline of R&D Efficiency, *RIETI Discussion Paper Series, No. 9-E-052*, Research Institute of Economy, Trade and Industry, Tokyo. Retrieved from: https://www.rieti.go.jp/jp/publications/dp/19e052.pdf

Mokyr, J., Vickers, C., & Ziebarth, N. L. (2015). The history of technological anxiety and the future of economic growth: Is this time different?. *Journal of Economic Perspectives*, 29(3), 31-50. https://doi.org/10.1257/jep.29.3.31

Mullainathan, S., & Spiess, J. (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31(2), 87-106. https://doi.org/10.1257/jep.31.2.87

Musslick, S., Bartlett, L. K., Chandramouli, S. H., Dubova, M., Gobet, F., Griffiths, T. L., ... & Holmes, W. R. (2025). Automating the practice of science: Opportunities, challenges, and implications. *Proceedings of the National Academy of Sciences*, 122(5), e2401238121. https://doi.org/10.1073/pnas.2401238121

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654), 187-192. https://doi.org/10.1126/science.adh2586

OECD AI Policy Observatory (2021). Database of national AI policies, powered by EC/OECD. https://oecd.ai

OECD AI Policy Observatory (2024). An overview of national AI strategies and policies, *OECD Going Digital Toolkit Notes*, No. 14, OECD Publishing, Paris, https://doi.org/10.1787/c05140d9-en

Pammolli, F., Magazzini, L., & Riccaboni, M. (2011). The productivity crisis in pharmaceutical R&D. *Nature Reviews Drug Discovery*, 10(6), 428-438. https://doi.org/10.1038/nrd3405

Park, M., Leahey, E., & Funk, R. J. (2023). Papers and patents are becoming less disruptive over time. *Nature*, 613(7942), 138-144. https://doi.org/10.1038/s41586-022-05543-x

Pearl, J., & Mackenzie, D. (2018). *The book of why: The new science of cause and effect*. Basic books.

Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023). The impact of AI on developer productivity: Evidence from GitHub copilot. *arXiv preprint arXiv:2302.06590*. https://arxiv.org/abs/2302.06590

Rammer, C., Fernández, G. P., & Czarnitzki, D. (2022). Artificial intelligence and industrial innovation: Evidence from German firm-level data. *Research Policy*, 51(7), 104555. https://doi.org/10.1016/j.respol.2022.104555

Resnik, D. B., & Hosseini, M. (2024). The ethics of using artificial intelligence in scientific research: new guidance needed for a new tool. *AI and Ethics*, 1-23. https://doi.org/10.1007/s43681-024-00493-8

Romer, P. (1993). Two strategies for economic development: Using ideas and producing ideas. *Proceedings of the World Bank Annual Conference on Development Economics*, Retrieved from: https://documents1.worldbank.org/curated/pt/353181468739245145/pdf/Proceedings-of-the-World-Bank-annual-conference-on-development-economics-1992.pdf

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206-215. https://doi.org/10.1038/s42256-019-0048-x

Scannell, J. W., Blanckley, A., Boldon, H., & Warrington, B. (2012). Diagnosing the decline in pharmaceutical R&D efficiency. *Nature Reviews Drug Discovery*, 11(3), 191-200. https://doi.org/10.1038/nrd3681

Schmallenbach, L., Bärnighausen, T. W., & Lerchenmueller, M. J. (2024). The global geography of artificial intelligence in life science research. *Nature Communications*, 15(1), 7527. https://doi.org/10.1038/s41467-024-51714-x

Si, C., Yang, D., & Hashimoto, T. (2024). Can LLMs generate novel research ideas? A large-scale human study with 100+ NLP researchers. *arXiv preprint arXiv:2409.04109*. https://arxiv.org/abs/2409.04109

Sorenson, O. (2024). Theory, Search, and Learning. *Strategy Science*, 9(4), 372-381. https://doi.org/10.1287/stsc.2024.0179

Stephan, P. (2012). *How economics shapes science*. Harvard University Press. https://doi.org/10.4159/harvard.9780674062757

Tekic, Z., & Füller, J. (2023). Managing innovation in the era of AI. *Technology in Society*, 73, 102254. https://doi.org/10.1016/j.techsoc.2023.102254

Thu, M. K., Beppu, S., Yarime, M., & Shibayama, S. (2022). Role of machine and organizational structure in science. PLOS One, 17(8), e0272280. https://doi.org/10.1371/journal.pone.0272280

Toner-Rodgers, A. (2024). Artificial intelligence, scientific discovery, and product innovation. *arXiv preprint arXiv:2412.17866*. https://arxiv.org/abs/2412.17866

Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical combinations and scientific impact. *Science*, 342(6157), 468-472. https://doi.org/10.1126/science.1240474

Van Noorden, R., & Perkel, J. M. (2023). AI and science: What 1,600 researchers think. *Nature*, 621(7980), 672-675. https://doi.org/10.1038/d41586-023-02980-0

Verendel, V. (2023). Tracking artificial intelligence in climate inventions with patent data. *Nature Climate Change*, 13(1), 40-47. https://doi.org/10.1038/s41558-022-01536-w

Waltz, D., & Buchanan, B. G. (2009). Automating science. *Science*, 324(5923), 43-44. https://doi.org/10.1126/science.1172781

Weitzman, M. L. (1998). Recombinant growth. *The Quarterly Journal of Economics*, 113(2), 331-360. https://doi.org/10.1162/003355398555595

Wu, F., Lu, C., Zhu, M., Chen, H., Zhu, J., Yu, K., ... & Pan, Y. (2020). Towards a new generation of artificial intelligence in China. *Nature Machine Intelligence*, 2(6), 312-316. https://doi.org/10.1038/s42256-020-0183-4

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, 316(5827), 1036-1039. https://doi.org/10.1126/science.1136099

Xu, Y., Liu, X., Cao, X., Huang, C., Liu, E., Qian, S., ... & Zhang, J. (2021). Artificial intelligence: A powerful paradigm for scientific research. *The Innovation*, 2(4). https://doi.org/10.1016/j.xinn.2021.100179

Yu, Z. (2024). The impacts of prediction AI on scientists: Evidence from AlphaFold. *Available at SSRN 4711334*. https://ssrn.com/abstract=4711334

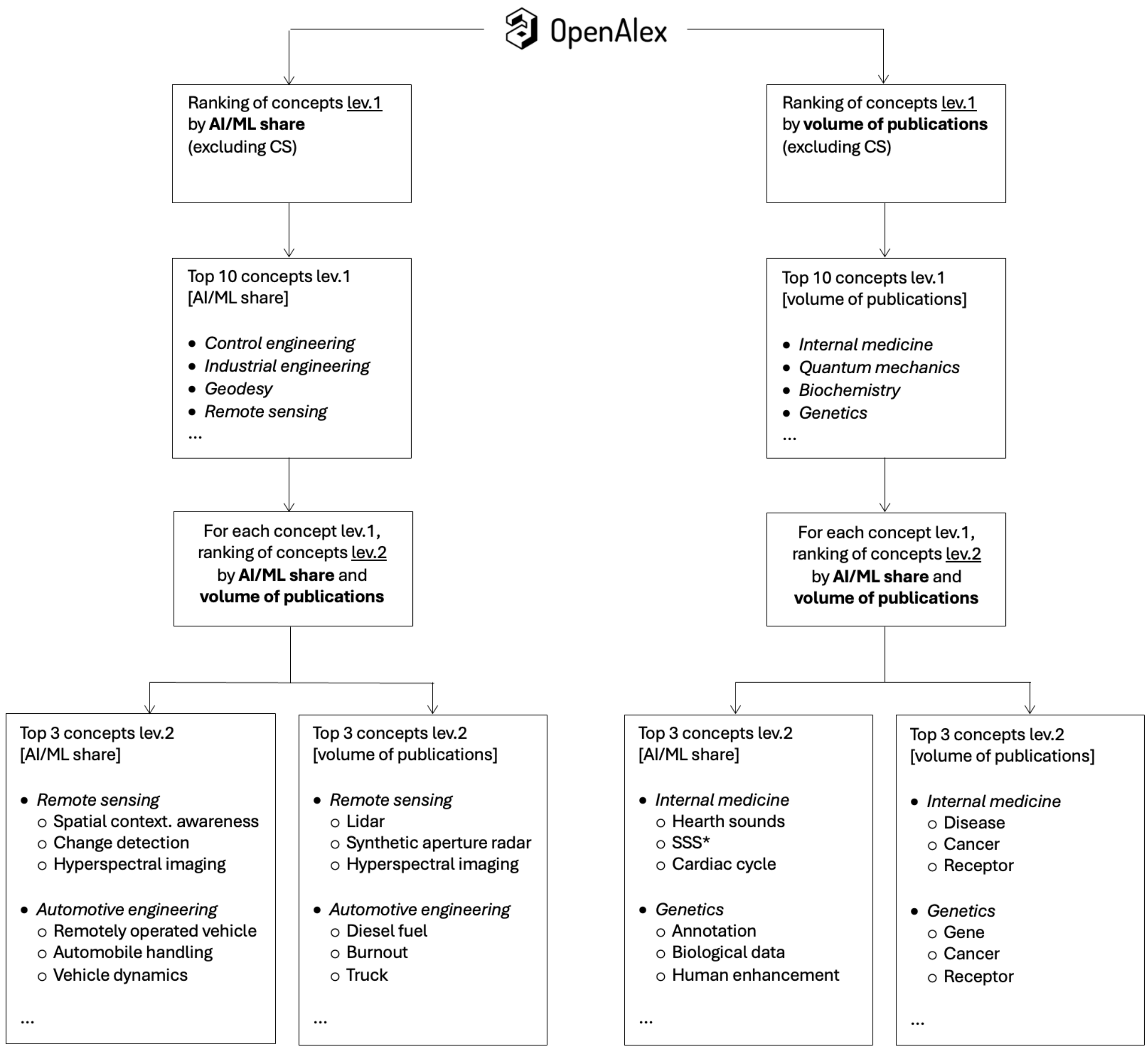
7. Appendix

7.1. Appendix A. Data processing

In this research, we aim to study the diffusion and impact of AI at a more granular level, moving beyond broad macro-disciplines (e.g., Chemistry, Medicine, Physics). However, this comes with a trade-off: the deeper we go, the more complex it becomes to manage field taxonomy. As of August 2024, OpenAlex lists 21,455 concepts at level-2 and 24,749 at level-3, making it impractical for the type of analysis and reporting conducted in this study.

We implemented a protocol to retrieve data by first ranking scientific fields (concept level-1) based on two criteria: (1) *AI penetration*, measured as the share of publications associated with AI (i.e., those classified under level-1 concepts 'Artificial Intelligence' or 'Machine Learning', or containing at least one of their 400+ sub-concepts], and (2) the *total volume of scientific publications* in each field.

Table A1. Data pipeline for the selection of scientific fields



From these rankings, we selected the top fields in each category – those with the highest AI penetration (left arm of Figure A1) and those with the highest publication volume (right arm). Thus, this approach ensures that our dataset contains both fields where AI is already widely integrated and those with intense research activity, even if AI adoption is still emerging.

Next, for each selected level-1 concept, we identified and ranked subfields (level-2 concepts) using the same criteria: AI share and publication volume. Since each document may be associated with multiple concepts, some subfields overlap across different rankings.

A final note concerns the fields '*Cancer*', '*Disease*', and '*Gene*'. Despite being classified as level-2, these fields cover between 2 to 4 million articles and are in fact too broad for our analysis. To refine our selection, we moved to level-3 and replicate the same approach, namely ranking subfields based on AI penetration and publication volume.

Recognising that there is no single best way to sample the data from OpenAlex, we tested the robustness of our findings using stricter criteria. For example, we experimented with a more conservative classification of AI papers by requiring their AI or ML concepts and sub-concepts score to exceed the arbitrary threshold of 0.5. We also replicated the analysis while excluding concepts with exceptionally high AI share or total publication volume, and explored alternative ranking methods. All these results are available upon request.

Table A2. Description of selected scientific fields

OpenAlex Code	Label	Description / Example of AI applications	Count [% AI/ML]
C6856738	Protein Expression†	Producing specific proteins for research, therapeutic, or industrial use *AI predicts optimal conditions for protein synthesis in bioreactors*	15,695 [2.85]
C10679952	Druggability†	Identifying biological targets that can bind effectively with drugs *AI identifies potential drug-binding sites on protein structures*	4,039 [4.46]
C15952604	Project Management	Achieving project objectives through structured processes *AI optimizes schedules and resource allocation based on historical project data*	58,914 [28.12]
C18020424	Autopilot	Automating vehicle navigation without manual control *Self-driving cars use AI to maintain lane and avoid obstacles*	5,203 [77.07]
C18483071	Anchoring	Understanding decision-making biases from initial information *AI analyzes how initial product reviews or ratings influence purchasing decisions*	9,637 [15.02]
C20702342	Cryo-electron Microscopy	Imaging molecular structures at near-atomic resolution under cryogenic conditions *AI reconstructs 3D molecular structures from 2D microscopy images*	3,174 [35.07]
C22762622	Operating Point	Optimizing device performance by identifying critical operational conditions *AI adjusts the speed and energy consumption of industrial robots on an assembly line*	4,726 [84.36]
C23085057	Genetic Analysis†	Detecting DNA mutations to predict risks or optimize treatments *AI identifies disease-related genetic mutations from DNA sequences*	6,287 [5.15]
C40506919	Sequence Learning	Exploring how humans learn and recall sequential information *AI analyzes sequences of spoken words to simulate how children acquire grammar rules*	2,919 [100.00]
C46362747	CMOS	Creating efficient integrated circuits for electronics	299,019

		AI designs optimized transistor layouts in circuit simulations	[22.80]
C47042493	Human Genetics†	Understanding genes, heredity, and human genetic variations *AI predicts genetic predisposition to diseases based on genome data*	16,951 [3.89]
C51399673	LiDAR	Mapping environments using laser-based imaging technology *AI processes LiDAR data for autonomous vehicle navigation*	57,382 [42.85]
C51456166	Genetically Modified Organism (GMO)†	Altering organisms' genes for desired traits *ML suggests genetic edits to improve crop yield*	11,173 [5.19]
C52121051	Truck	Optimizing the design and operation of commercial vehicles *AI-powered fleet management optimizes delivery routes*	58,954 [19.88]
C60627051	Body Shape	Studying human form based on skeletal and muscular structure *AI estimates body composition from medical imaging*	1,908 [100.00]
C64228939	Remotely Operated Vehicle (ROV)	Developing or using ROVs to explore underwater environments via remote-controlled devices *AI assists in underwater object recognition*	1,853 [89.85]
C64413873	Threshold Limit Value (TLV)	Defining safe exposure levels to hazardous substances *AI analyzes historical workplace exposure data and real-time readings from air quality sensors*	2,819 [64.14]
C64474127	Medical Genetics†	Diagnosing and managing hereditary diseases *AI identifies mutations linked to familial breast cancer*	6,188 [4.09]
C64754055	Spatial Contextual Awareness	Leveraging location-based data for computing and decision-making *AI enhances AR applications with location-based data*	4,839 [100.00]
C66283442	Failure Mode and Effects Analysis (FMEA)	Preventing system failures by identifying potential issues *AI monitors vibration and temperature data from industrial machinery to predict failures*	24,223 [18.91]
C74172769	Electronic Design Automation (EDA)	Streamlining electronic circuit design processes *AI analyzes circuit designs to automatically detect errors and optimize layouts*	177,430 [26.06]
C78639753	Behavioral Modeling	Analyzing systems through behavioral patterns *AI models customer behavior in e-commerce*	3,011 [100.00]
C87360688	Synthetic Aperture Radar (SAR)	Capturing high-resolution radar imagery *AI interprets radar images for environmental monitoring*	67,258 [100.00]

C90130585	Electronic Skin	Developing skin-mimicking technologies for sensors *AI decodes tactile data from e-skin sensors for prosthetics*	1,096 [46.62]
C94487597	Sensory System	Processing sensory input in living organisms *AI helps a robotic hand adjust its grip based on the weight of an object it is holding*	87,997 [24.65]
C96250715	Project Estimation	Forecasting resources and time for project completion *AI predicts timelines and costs of new projects*	166,415 [42.47]
C99398487	Cardiac Cycle	Understanding the phases of heart function during a heartbeat *AI detects arrhythmias in ECG data*	7,085 [27.30]
C111829913	Gene Targeting†	Modifying specific genes via homologous recombination *AI suggests gene-editing targets for CRISPR applications*	3,192 [0.75]
C129364497	Prognostics	Predicting system failures or maintenance needs *AI forecasts machine part failures for predictive maintenance*	5,731 [57.93]
C132010649	Intuition	Gaining insights without recourse to conscious reasoning *AI mimics human intuitive decision-making in games like Go*	24,958 [34.13]
C138171918	Diesel Fuel	Powering engines using efficient liquid fuels *AI optimizes fuel injection rates and air intake in a diesel engine to reduce emissions*	98,845 [10.57]
C139489369	Structural Similarity	Measuring image or video quality based on human perception *AI evaluates image quality in video streaming*	3,152 [100.00]
C140096630	Damper	Controlling airflow or energy dissipation in systems *AI adjusts damper settings for improved HVAC efficiency*	37,740 [50.85]
C143916079	Burnout (Automotive)	Analyzing tire behavior, traction, and heat generation during burnouts *AI systems detect burnout through tire and traction analysis to study material durability*	54,864 [14.44]
C148381915	Automobile Handling	Assessing how vehicles respond to driver inputs *AI optimizes suspension settings for performance vehicles*	2,838 [86.47]
C148699463	SSS*	Optimizing search efficiency in algorithms *AI enhances heuristic search algorithms for large-scale scheduling problems in airline operations*	4,235 [100.00]
C159334719	Activity Theory	Examining human activities in societal and systemic contexts	1,877 [39.53]

		AI analyzes user interactions in educational software	
C170477896	Ideation	Generating and refining creative ideas *Generative AI creates innovative designs for product development*	6,046 [30.57]
C171897839	Protein Family†	Understanding evolutionary relationships among proteins *AI classifies protein families based on sequence similarities*	3,039 [13.92]
C183776436	Rudder	Controlling direction in air or water vehicles *AI stabilizes rudder control in autonomous ships*	4,752 [56.92]
C185798385	Survey Benchmark	Providing reference points for geographic measurements *AI improves geospatial mapping using drone data*	163,763 [68.03]
C187691185	Grid	Dividing surfaces for mapping or data indexing *AI optimizes grid layouts for efficient urban planning*	270,622 [41.71]
C201797286	Biological Data	Collecting and analyzing data for biological insights *AI identifies patterns in omics datasets for biomedical research*	3,584 [68.69]
C203357204	Chunking	Studying how people group related information to enhance memory *AI powers adaptive learning platforms that group related concepts based on a student's learning pace*	2,005 [100.00]
C204315192†	Molecular Genetics	Studying genes at the molecular level *AI predicts genetic expressions from molecular data*	2,708 [3.58]
C553089730†	Binding Protein	Understanding DNA-protein interactions *AI analyzes DNA sequences and protein structures to predict binding sites and affinities*	2,388 [0.25]
C2776356578	Neuromuscular Disease†	Researching disorders affecting muscle-controlling nerves *AI detects early signs of ALS from movement data*	3,385 [2.27]
C2776463041	Cancer Screening†	Detecting cancer early in asymptomatic individuals *AI identifies early cancer markers in imaging scans*	9,140 [14.07]
C2777002142	Cancer Biomarkers†	Identifying substances indicating cancer presence *AI discovers new biomarkers using genomic data*	1,848 [3.35]
C2777474118	Tropical Disease†	Studying diseases common in tropical regions *AI maps disease outbreaks based on environmental data*	3,426 [3.82]
C2777522853	Digital Pathology	Managing pathology data digitally for better diagnostics *AI processes histology slides to identify cancerous cells*	2,910 [100.00]
C2777526511	Pace Unit	Measuring length in human steps	64,171

		AI improves wearable step-counting and heart rate accuracy	[20.64]
C2779356329	Checklist	Studying workflows and human behavior to design and optimize standardized procedures *AI analyzes past aviation data to suggest improvements to checklists to reduce errors*	64,066 [12.57]
C2779435589	Heart Sounds	Analyzing acoustic signals of heart function *AI analyzes heart sounds to detect murmurs*	2,247 [62.66]
C2779679103	Signal Degradation	Mitigating quality loss in electronic signals *AI enhances degraded audio or video signals*	169,908 [6.56]
C2779701055	Rare Disease†	Researching diseases affecting small populations *AI identifies rare disease cases from sparse datasets*	13,025 [1.04]
C2779706800	Human Enhancement	Improving physical or mental abilities using technology *AI aids cognitive enhancement through neurofeedback*	1,380 [100.00]
C2779918689	Stimulus	Studying how organisms respond to environmental triggers *AI models neural responses to visual stimuli*	96,159 [36.83]
C2780130745	Neuropathology†	Diagnosing diseases in nervous system tissues *AI identifies brain abnormalities in imaging data*	10,382 [3.33]
C2780234812	Cancer Prevention†	Reducing cancer risk through proactive measures *AI designs lifestyle interventions to reduce cancer risk*	8,418 [7.23]
C2780443751	Material Selection	Choosing materials for specific applications *AI recommends optimal materials for 3D printing*	4,035 [55.86]
C2780451532	Task Assignment	Allocating tasks effectively to achieve goals *AI distributes workloads for efficient team performance*	524,246 [56.72]
C2780596555	Amyotrophic Lateral Sclerosis (ALS)†	Researching and managing neurodegenerative diseases *AI predicts ALS progression from patient data*	25,175 [4.37]
C2780841897	Digital Manufacturing	Using computer technology to optimize manufacturing processes *AI analyzes real-time production data to adjust machine speeds*	1,450 [28.96]
C2781220375	Flocking	Studying the coordinated movement of individual agents *AI optimizes flocking behaviors in drone swarms*	3,310 [68.28]
C2781230642	Targeted Therapy†	Developing drugs targeting specific disease pathways *ML identifies drug targets specific to cancer cell pathways*	16,484 [1.56]

C2982912361	Mental Model	Simulating scenarios mentally to anticipate outcomes *AI creates virtual simulations to train mental models for pilots*	4,174 [37.99]
C2983137510	Material Handling	Managing materials efficiently during production and distribution *AI automates warehouse logistics for inventory management*	1,968 [37.30]
C2983331546	Cancer Therapy†	Treating cancer using various therapeutic methods *AI personalizes cancer treatments based on patient profiles*	14,111 [2.93]
C2985179714	Workflow Management	Streamlining processes and resource allocation *AI prioritizes tickets based on urgency and automates responses for common queries*	3,792 [24.84]
C2985322473	Cancer Detection†	Diagnosing cancer in symptomatic individuals or individuals with an elevated risk *AI classifies cancer types from biopsy samples*	5,609 [31.41]
C2988168687	Skin Lesion	Studying skin disorders or abnormalities *AI detects melanoma from skin lesion images*	4,244 [33.98]
C2989108626	Drug Target	Investigating molecular targets for therapeutic intervention *AI predicts potential drug targets in pathogens*	1,931 [33.20]
C2992972558	Blood Cancer†	Understanding and treating cancers of blood or bone marrow *AI aids in diagnosing leukemia from blood samples*	1,581 [7.78]
C2993153387	Genetic Diagnosis†	Identifying genetic causes of diseases *AI analyzes genetic data for precise diagnoses*	1,600 [5.81]
C2994114330	Cancer Imaging†	Visualizing cancer using medical imaging tools *AI enhances tumor visualization in MRI scans*	1,463 [18.93]
C3018284874	Cardiovascular Health†	Promoting heart and blood vessel health *AI predicts cardiovascular events from wearable data*	5,079 [8.92]
C3019111730	Genetic Engineering†	Manipulating genomes for research or applications *AI designs synthetic genes for biotech applications*	3,958 [3.39]
C3019816032	Cancer Treatment†	Exploring methods to treat cancer effectively *AI predicts cancer response to immunotherapy*	7,587 [5.71]
C3020597237	Production Rate	Measuring the movement of inputs and outputs in manufacturing systems *AI forecasts production bottlenecks in real time*	4,506 [12.52]

Total			2,889,302 [36.55]

Note: † Concept at level 3

7.2. Appendix B. Combinatorial complexity of a scientific field

We used the indicator proposed by Lee et al. (2015), which relies on comparing the *observed frequency* of knowledge combinations with their *expected frequency* in a given year. The methodology involves the following steps:

1. Construct the co-citation network
    - Each scientific paper cites a set of references. Let *FP* be a focal paper citing references *A*, *B*, *C*, etc.
    - Each pair of cited references (*A, B*), (*A, C*), (*B, C*), etc. forms a knowledge combination.
2. Compute observed frequency
    - The observed frequency $w_t(V_i, V_j)$ of a combination (V_i, V_j) in year *t* is the actual number of times this pair has been cited together in papers from the same year.
3. Compute expected frequency
    - The expected frequency of a combination is based on a null model: $\frac{k_i \times k_j}{N_t}$ where k_i and k_j are the degrees (citation counts) of references *i* and *j*, respectively, and N_t is the total number of reference pairs in year *t*.
    - The formula assumes a random combination model, meaning that if references were cited randomly, their expected co-occurrence should be proportional to their overall citation frequency.
4. Calculate the commonness score for each combination
    - The commonness score for a pair (i, j) is given by: $\text{commonness}_{ijt} = \frac{w_t(V_i,V_j) \times N_t}{k_i \times k_j}$
    - A high value means the combination is frequently used (common), while a low value means the combination is rarely used (uncommon).
5. Aggregate at the paper level
    - For each focal paper *FP*, the 10th percentile of the commonness distribution of all its reference pairs is computed: $\text{commonness}_{FP} = -\log\left(P_{10}(C_{FP})\right)$
    - The negative log transformation ensures that lower commonness values (i.e., rarer combinations) receive higher scores
6. Aggregate at the field level
    - To measure the combinatorial complexity, or “roughness”, of a scientific field, we compute the mean commonness score across all papers in that field.
    - Higher values indicate a more fragmented knowledge structure.

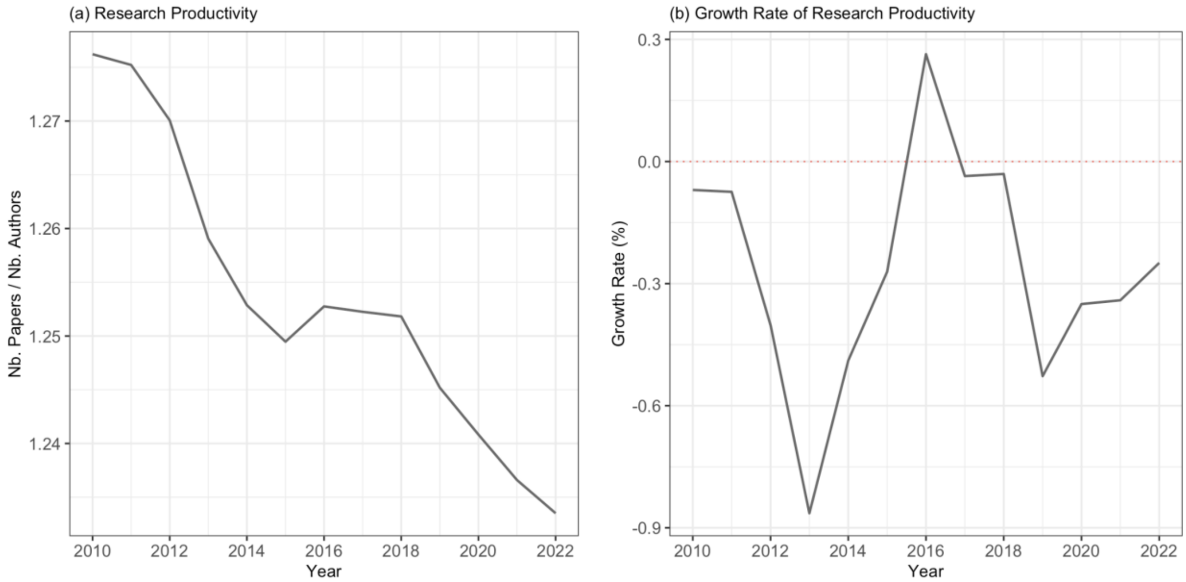
7.3. Appendix C. Supplementary statistics and analysis

Table C1. Descriptive statistics of the variables used in this study

	Mean	Min	Med	Max
New Words	0.10	0	0	1
New Words Combinations	0.70	0	1	1
New Phrases	0.20	0	0	1
New Phrases Combinations	0.70	0	1	1
Semantic Distance	0.10	0	0.20	0.50
Weighted Nb. Citations	2.48	0	0.33	996.45
Top Cited (10%)	0.10	0	0	1
Top Cited (5%)	0.05	0	0	1
Top Cited (1%)	0.01	0	0	1
AI	0.36	0	0	1
Field Roughness	0.30	-4.15	0.44	3.04
Nb. Authors	3.47	1	3	100
Nb. References	18.34	0	10.00	5,499
International Collab.	0.12	0	0	1
Survey/Review	0.01	0	0	1

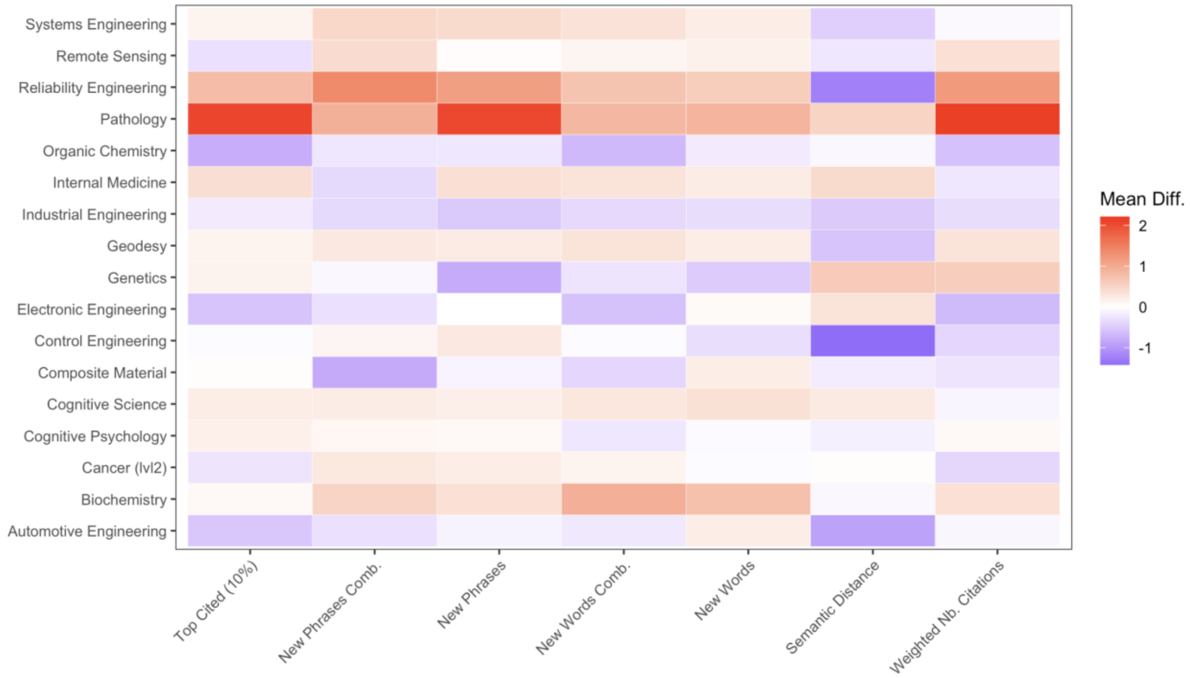
Notes: Statistics are calculated on the full sample.

Figure C1. Trend and growth rate in research productivity



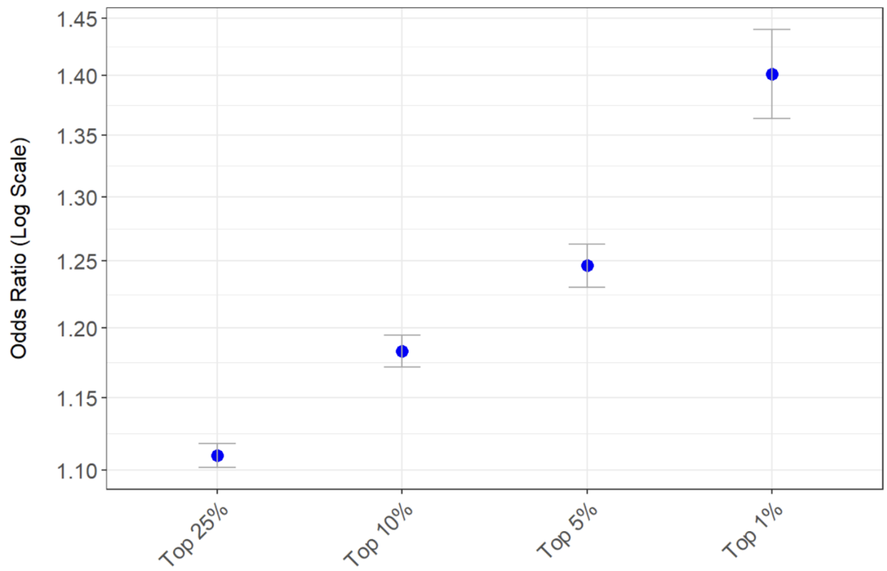
Notes: Panel A: Research productivity is measured as the ratio of the number of scientific publications to the number of unique authors per year. Panel B: Growth of research productivity is calculated as 3-year rolling averages. Figures start from 2010 to focus on recent trends and minimize noise caused by smaller sample sizes in earlier years. Based on our own elaboration.

Figure C2. Mean differences between indicators of novelty and impact for AI and other papers, lvl.1



Notes: The plot compares the mean difference between AI papers and other scientific publications across various indicators of novelty and impact (x-axis) and scientific fields (y-axis). All variables are standardized to facilitate comparison across indicators. Darker shades of red indicate that AI papers exhibit a higher mean for a given indicator compared to non-AI papers, while blue shades indicate the opposite. Based on our own elaboration.

Figure C3. Odds ratio for impact indicators



Notes: The plot represents the odd ratios from Logit models estimating the effects of AI on impact for different citation thresholds – see Table 4 in the main text. Vertical bars indicate 95% confidence intervals.

Table C2. The effect of AI on novelty, by field "roughness" (low vs. high)

	New Words		*New Words Comb.*		*New Phrases*		*New Phrases Comb.*		*Semantic Distance*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Low	High	Low	High	Low	High	Low	High	Low	High
AI	0.042**	0.096***	0.021**	0.139***	0.036***	0.141***	0.090***	0.235***	0.006***	0.004***
	(0.016)	(0.011)	(0.009)	(0.005)	(0.009)	(0.006)	(0.010)	(0.006)	(0.0002)	(0.0001)
Nb. Authors	0.023***	0.035***	0.052***	0.065***	0.028***	0.042***	0.054***	0.075***	-0.0002***	-0.0005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.00002)	(0.00001)
Nb. References	-0.008*	0.030***	0.167***	0.204***	0.080***	0.108***	0.220***	0.263***	-0.002***	-0.001***
	(0.004)	(0.009)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.00004)	(0.00003)
International Collab.	0.103***	0.089***	0.140***	0.115***	0.094***	0.059***	0.142***	0.101***	0.0001	0.001***
	(0.019)	(0.014)	(0.011)	(0.007)	(0.010)	(0.007)	(0.012)	(0.007)	(0.0002)	(0.0001)
Survey/Review	-0.133***	-0.404***	-0.332***	-0.364***	-0.216***	-0.280***	-0.264***	-0.301***	-0.005	0.002***
	(0.064)	(0.057)	(0.029)	(0.020)	(0.036)	(0.027)	(0.030)	(0.021)	(0.001)	(0.0004)
Adjusted R^2									0.160	0.156
Log Likelihood	-100,760	-179,881	-245,788	-566,513	-240,506	-486,144	-214,523	-500,677		
AIC	201,696	359,939	491,752	1,133,203	481,188	972,465	429,222	1,001,531		
# Observations	428,107	946,426	428,107	946,426	428,107	946,426	428,107	946,426	428,107	946,426

Notes: The econometric models for evaluating the effect of AI on various indicators of novelty across two levels of field roughness: low and high. Parameters in Columns 1–8 are estimated via OLS regression, while those in Columns 9–10 via Logit models. All specifications include fixed effects for time and scientific fields. The asterisks ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table C3. The effect of AI on impact, by field "roughness" (low vs. high)

	Weighted Nb. Citations		*Top Cited (10%)*		*Top Cited (5%)*		*Top Cited (1%)*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Low	High	Low	High	Low	High	Low	High
AI	0.030***	0.027***	0.132***	0.183***	0.155***	0.250***	0.254***	0.381***
	(0.002)	(0.001)	(0.009)	(0.006)	(0.012)	(0.008)	(0.025)	(0.017)
Nb. Authors	0.028***	0.035***	0.080***	0.094***	0.068***	0.083***	0.049***	0.066***
	(0.0002)	(0.0002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Nb. References	0.263***	0.290***	1.107***	1.149***	1.120***	1.170***	1.179***	1.184***
	(0.0005)	(0.0003)	(0.004)	(0.003)	(0.006)	(0.006)	(0.012)	(0.010)
International Collab.	0.183***	0.195***	0.393***	0.387***	0.420***	0.406***	0.487***	0.479***
	(0.002)	(0.002)	(0.010)	(0.006)	(0.012)	(0.008)	(0.024)	(0.017)
Survey/Review	0.096***	0.152***	0.300***	0.424***	0.440***	0.502***	0.852***	0.847***
	(0.006)	(0.004)	(0.032)	(0.424)	(0.038)	(0.025)	(0.061)	(0.040)
Adjusted R^2	0.388	0.414						
Log Likelihood			-236,457	-486,499	-148,970	-307,395	-44,873	-92,143
AIC			473,090	973,174	298,116	614,967	89,922	184,463
# Observations	928,952	1,917,062	928,952	1,917,062	928,952	1,917,062	928,952	1,917,062

Notes: The econometric models for evaluating the effect of AI on various indicators of impact across two levels of field roughness: low and high. Parameters in Columns 1–2 are estimated via OLS regression, while those in Columns 3–8 via Logit models. All specifications include fixed effects for time and scientific fields. The asterisks ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct centres. You can find the address of the centre nearest you online (european-union.europa.eu/contact-eu/meet-us_en).

On the phone or in writing

Europe Direct is a service that answers your questions about the European Union.
You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696,
- via the following form: european-union.europa.eu/contact-eu/write-us_en.

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website (european-union.europa.eu).

EU publications

You can view or order EU publications at op.europa.eu/en/publications. Multiple copies of free publications can be obtained by contacting Europe Direct or your local documentation centre (european-union.europa.eu/contact-eu/meet-us_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex (eur-lex.europa.eu).

EU open data

The portal data.europa.eu provides access to open datasets from the EU institutions, bodies and agencies. These can be downloaded and reused for free, for both commercial and non-commercial purposes. The portal also provides access to a wealth of datasets from European countries.

The paper explores the impact of AI on scientific creativity, examining its use across 80 fields from 2000 to 2022. AI adoption has surged in nearly all areas since the early 2010s, although striking regional differences emerge. In recent years, China has taken the lead in AI-driven research, outpacing both the US and the EU, not just in sheer output, but also in terms of scientific novelty and impact. The study concludes that AI generally enhances scientific creativity, measured by novelty and impact, though the effects vary by field. Most fields benefit from AI applications, although great heterogeneity is observed with some fields seeing little to no improvement, and a few experiencing negative impacts. The influence of AI is moderated by the structural organisation of knowledge within fields, with greater potential in "rough" knowledge spaces where ideas are fragmented. These findings contribute to discussions on AI's role in science and are relevant to policy initiatives promoting AI-driven research.

Studies and reports

Publications Office
of the European Union