UNIVERSITÉ DE STRASBOURG

École doctorale Augustin Cournot ED 221

Bureau d'Économie Théorique et Appliquée UMR 7522

THÈSE

pour l'obtention du titre de Docteur en Sciences Économiques

Présentée et soutenue le 23 juin 2023 par

Pierre Pelletier

ARTIFICIAL INTELLIGENCE IN SCIENCE: DIFFUSION AND IMPACT

Préparée sous la direction de Stefano Bianchini

Membres du jury :

Aldo Geuna	Professeur des Universités, Université de Turin	Rapporteur
Michele Pezzoni	Maître de Conférences, HDR, Université Côte d'Azur	Rapporteur
Reinhilde Veugelers	Professeure des Universités, Université de Louvain	Examinatrice
Sandrine Wolff	Maîtresse de Conférences, HDR, Université de Strasbourg	Examinatrice
Jacques Mairesse	Professeur des Universités, Université de Maastricht	Membre invité
Stefano Bianchini	Maître de Conférences HDR, Université de Strasbourg	Directeur

À mon frère Guillaume,

Remerciements

Mes remerciements vont en premier lieu à Stefano Bianchini, mon directeur de thèse, qui m'a accompagné, soutenu et encouragé depuis mon mémoire de master. Il m'a rapidement considéré comme un collègue et m'a laissé une image de la relation directeur-doctorant exceptionnelle. Sa gentillesse et sa bonne humeur ont rendu ce travail de thèse beaucoup moins éprouvant ; la qualité de son encadrement a grandement facilité la réalisation de ce travail. C'est aujourd'hui bien plus qu'un collègue, mais bel et bien un ami avec qui travailler à été et sera toujours un réel plaisir. Je suis également reconnaissant envers sa famille, Angela, Lorenzo et Ariana pour m'avoir accueilli chaleureusement et avoir supporté nos nombreuses sessions de travail.

Je remercie vivement Michele Pezzoni et Julien Penin pour leur disponibilité et leurs précieux conseils, lors des comités de suivi. Michele fait partie des premiers chercheurs que j'ai rencontrés et ses travaux ont grandement contribué à mon intérêt pour l'économie de la science. Julien Penin m'a fait découvrir l'économie de l'innovation ; ce cours a grandement influencé ma conception de l'open-source.

Merci également à Aldo Geuna, Jacques Mairesse, Michele Pezzoni, Reinhilde Veugelers et Sandrine Wolff pour avoir accepté de compter parmi les membres de mon jury. Je suis extrêmement honoré d'être évalué par des chercheurs ayant tant contribué à l'économie de la science et de l'innovation.

Je remercie également les coauteurs ayant collaboré aux travaux de cette thèse. Merci à Stefano Bianchini, Moritz Müller et Kevin Wirtz pour toutes ces réunions et ces heures passées à délirer. Travailler avec eux a été extrêmement instructif et j'espère que nous deviendrons des "super ties".

Je tiens aussi à remercier celles et ceux qui ont peuplé au quotidien mon monde académique.

Merci au bureau 133, à Kenza, pour ces 4 ans côte à côte, donnant lieu à l'obtention d'une deuxième thèse en commérage et archéologie sous la direction de Louis et Théo. Vous êtes le 100 de la veine. Merci à Kéa, mais aussi à Laura et Sofiene pour leur patience.

Merci au bureau 126, à Anne-Gazelle àlias las papillas, tu as un don le sang. Merci aux meilleurs monchus, à Romane (Gemini) et à Jérôme pour nos escapades à vélo ou à 3842 mètres, pour nos risques et nos funs.

Merci au bureau 154, au trio infernal, Emilieng (Gemini), Guillaume, et Agathe (la + belle), et bien sûr merci à Kevin pour les balades à travers le Danemark et ces 9 dernières années passés dans ces locaux. Merci enfin à Diletta et Shengxhi et également à toutes les autres personnes qui font vivre cet établissement. Merci aussi aux personnes formidables que j'ai pu rencontrer hors de la fac lors de cette thèse, avec une mention spéciale pour Daniel Souza.

Un grand merci à ma famille, à ma mère Geneviève et mon père Nicolas. Vous êtes des parents formidables et la réalisation de ce document n'est que le fruit des conditions favorables que vous m'avez offertes. Merci également à mes frères et sœurs, à Laurence, à Guillaume, à Géraldine et à Étienne pour tout ce que vous avez pu me transmettre. Je vous aime fort.

Merci infiniment aux gens qui ont modellé la personne que je suis devenu. Plus spécialement merci à Alexis, Antoine, Camille, Carole, Chloé, Corentin, Florian, Frank, Jamel, Joséphine, Ludovic (Gemini), Martin (Gemini), Mathieu, Robin, Tina (Gemini), Tufan, Victor (Gemini) et Yassine. Vous etes les bests et je vous dois tout.

Finalement merci à Agathe. Je n'imagine ni ma thèse ni ma vie sans toi. Merci pour tout ce que l'on a pu vivre ces dernières années, trop hate que l'on arpente l'Irlande tout les deux durant celles qui viennent. Merci par extension aussi à Marvin (aka Bootsyboots), le garant de ma santé mentale en cette fin de thèse. Tu es le meilleur des chats, puisse cette thèse te servir à te poser confortablement.

Contents

General Introduction

Launched on March 14, 2023, the new version of ChatGPT, GPT-4, marks a turning point in Artificial Intelligence (AI). Due to its ability to combine text and vision, it can now successfully pass exams – with high grades – in nearly any domain. When I started this thesis in 2019, only seven years had passed since we could recognize a cat with 75% accuracy [Le, 2013], and only four years since Microsoft's AI had "surpassed" humans in their ability to recognize entities in images [He et al., 2015]. Four years later, we have seen an explosion in the development of AI, especially in natural language processing, and still, little is known about its potential implications and dangers (e.g. misinformation and fake news, cybersecurity risks, ethical concerns or psychological impact on individuals). In response to this growing development of AI, for which we do not fully comprehend and handle the consequences, an open letter calling on AI research to immediately pause for at least six months the training of other AI systems more powerful than GPT-4 was published the 22 March 2023. This petition was signed by more than 25,000 individuals including 1,000 AI researchers and experts[1]. The proliferation of AI and its abilities for ten years involves societal changes and might alter how science works. This thesis specifically addresses the relationship between science and artificial intelligence technology. Specifically, it aims to understand how various AI applications can affect the nature of research conducted in application domains. In doing so, we consider social factors to better understand how researchers adopt technology and produce new knowledge. In the following, we provide an overview of the role of AI in the new scientific paradigm.

Scientists depend on evolving technology to conduct experiments and validate theories. New technologies often enable scientists to explore the knowledge space differently and make new discoveries. As Derek de Solla Price states, "*The changes of paradigm that accompany great and revolutionary changes may sometimes be caused*

[1]Including Joshua Bengio, one of the three founding fathers of deep learning and winner of the Turing Prize. The petition can be found here: https://futureoflife.org/open-letter/pause-giant-ai-experiments/

by inspired thought, but much more commonly they seem due to the application of technology to science" [de Solla Price, 1963]. Science and technology have a bidirectional relationship: science provides fundamental principles for the development of new technologies, and technology, in turn, generates the instrumentation and technics needed to address novel and more challenging scientific questions more efficiently [Brooks, 1994]. Several technological advances have reshaped the scientific landscape throughout history. Think of the invention of the microscope which led to the discovery of cells or X-ray crystallography that facilitated the elucidation of three-dimensional molecular structures. These developments have deepened our understanding of biological mechanisms and provided a foundation for countless discoveries and innovations in various scientific disciplines. More recently, advances in informatics have enabled scientists to create complex mathematical models and solve problems previously considered intractable. For example, computers have been used to solve problems in number theory, cryptography, combinatorial optimization, simulate natural phenomena (climate systems, molecular dynamics, galaxy structures), and finally, enable the emergence of artificial intelligence.

AI lies at the core of the current technological paradigm, sharing several similarities in scale and scope with previous technological revolutions that have shaped and fueled long-term cycles of economic growth and structural change. The term "Artificial Intelligence" was coined by the computer scientist John McCarthy for the 1956 Dartmouth Summer Research Project on Artificial Intelligence, a seminal event for the field [McCarthy et al., 1955]. The goal of AI was to make machines use language, form abstractions, solve human problems, and improve themselves. Definitions of AI have varied but generally involve machines simulating intelligent behavior, performing complex tasks, and learning from experience. For instance, the European Commission refers to AI as "*machines or agents capable of observing their environment, learning, and taking intelligent action or proposing decisions*" [Annoni et al., 2018, p.19]. According to the OECD, AI systems are *machine-based systems that can make predictions, recommendations, or decisions for a given set of human-defined objectives* "[OECD, 2019, p.23]. WIPO defines AI systems as *learning systems that can improve at tasks typically performed by humans with limited or no human intervention* "[WIPO, 2019, p.19]. Terms like machine learning, deep learning, and artificial intelligence are often used interchangeably.

In the early days, AI focused on solving problems that formal mathematical rules could describe. These problems are intellectually challenging for humans but simpler for computers, as real-world knowledge can be hard-coded into formal languages, allowing logical inference rules to find solutions. This method, known as the 'knowledge-based' approach, involves a typical architecture with a knowledge base and an inference engine. The knowledge base stores real-world information, while the inference engine enables the machine to deduce insights from the stored information. This approach was dominant during the first few decades, with applications like "expert systems" introduced in the 1970s to simulate human judgment and behavior in specific fields. These systems were effective for certain problem types but not those requiring substantial subjective and intuitive knowledge or perceptual capabilities. Such problems are easy for humans to perform but difficult to articulate formally and mathematically [Nilsson, 2009].

In the same period, an alternative approach to machine intelligence began to take hold in the scientific community. This approach soon became known as "machine learning", which focused on designing intelligent systems that can acquire knowledge by extracting patterns from raw data. Unlike knowledge-based systems, machine learning methods construct hypotheses directly from the data through inductive inference, allowing machines to tackle problems involving real-world knowledge and achieve some human-like abilities, such as recognizing objects. Although machine learning proved to be a successful alternative to knowledge-based systems and became one of the most prominent branches of AI starting in the 1980s, particular challenges remained. Mainly, traditional machine learning methods encountered significant difficulties in extracting high-level abstract features from raw data due to factors of variation, such as different shapes, shadows, and viewing angles [Nilsson, 2009, Goodfellow et al., 2016]. All these attributes are known as factors of variations, essentially constructs in the human mind that can be thought of as high-level abstractions that help us make sense of the rich variability of the observed data.

The "deep learning" approach to machine intelligence emerged as an effective solution to the challenges faced by traditional machine learning methods. Deep learning (DL) systems learn from experience and comprehend the world through a hierarchy of abstract concepts, each defined in relation to simpler concepts [Schmidhuber, 2015, LeCun et al., 2015, Goodfellow et al., 2016]. This approach offers two significant advantages. First, like simpler machine learning algorithms, the machine acquires knowledge from past experiences, eliminating the need for humans to provide all the formal knowledge required to achieve a specific goal. Second, the complexity and abstraction of concepts are no longer barriers, as the machine can reconstruct and combine them on top of each other. This hierarchy of concepts makes the learning process that can be seen as structured into multiple layers, hence the term "deep". AI technics have been successfully used in diverse areas as predicting the 3D structure of proteins [Jumper et al., 2021], regulating nuclear fusion plasma in the tokamak configuration [Degrave et al., 2022], predicting the formation of the structure of the Universe [He et al., 2019], and creating a map of the brains of small insects [Winding et al., 2023] to name few. In 2017, AI witnessed another remarkable breakthrough with the emergence of Transformer models with self-attention mechanisms [Vaswani et al., 2017].

The impressive results of Generative Pre-trained Transformers (GPT), such as ChatGPT or GPT-4, now clearly illustrate artificial intelligence's general-purpose technology (GPT) nature, showcasing their adaptability and broad applicability across numerous domains. These models not only interact with users on various subjects but also support human thought processes by providing additional perspectives backed by near-expert knowledge on diverse topics. Thus, GPTs are GPTs. In March 2023, Eloundou et al. [2023] immortalized the wordplay and published the paper "GPTs are GPTs: An early look at the labor market impact potential of large language models", highlighting GPT-4's capabilities and its potential impact on the job market, suggesting that 15% of all worker tasks in the US could be completed significantly faster while maintaining the same quality level. Note that some of these tasks are also part of the scientific system, like programming and writing [Eloundou et al., 2023]. AI possesses the attributes of a general-purpose technology, with wide-ranging applications across numerous disciplines. GPTs, such as AI, stand out from other innovations due to their extensive application across various sectors, ability to catalyze further innovation in application sectors, and continuous rapid improvement [David, 1990, Bresnahan and Trajtenberg, 1995]. Classic examples of GPTs include the electric motor and the microprocessor, which have driven significant technological and organizational change across diverse sectors like manufacturing, agriculture, retail, and residential construction.

AI's role as a GPT in science is further exemplified by its function as an "Invention in the Method of Invention" (IMI). IMIs create or improve specific products and provide a new way of generating new products with broader applications. For example, double-cross hybrid in agriculture was an IMI that led to the development of numerous new crop varieties, profoundly impacting agricultural productivity [Griliches, 1957]. The economic impact of AI as a research tool extends beyond merely reducing the costs of specific innovation activities as it enables an entirely new approach to innovation itself. The pervasive nature of AI as a GPT and IMI positions it as a unique technology capable of driving organizational change across a wide range of scientific fields. Yet, it should be used responsibly. This is why institutions such as the European Commission are supporting a regulatory and investment-oriented approach to promoting AI uptake while addressing risks associated with the technology [European-Commission, 2020]. Similar initiatives have been implemented by the OECD, which launched a specific technology observatory to closely monitor technology evolution and provide evidence-based policy analysis on AI[2]. These initiatives are not limited to Europe, since 2019 institutions such as Stanford University have regularly reported yearly metrics on AI's evolution, labor market, skills, and automation to guide responsible and ethical decision-making [3].

Ongoing advancements in AI within scientific disciplines, particularly the remarkable accomplishments achieved through neural network technics, extend beyond providing impressive anecdotal discoveries. Instead, these advancements have resulted in increasing pressure to transition from hypothesis-driven to data-driven scientific exploration. The emerging scientific paradigm is founded upon data-intensive computing, facilitated by the widespread implementation of intelligent machines capable of discerning representations, rules, and patterns in an ever-growing volume of structured and unstructured data [King et al., 2009, Hey et al., 2009]. Scientific discovery can be viewed as the process or product of successful scientific inquiry. In its narrowest sense, the term discovery might refer to the so-called 'eureka moment' of gaining new insights. However, in this context, we adopt its broadest meaning—using the term discovery as synonymous with 'successful scientific endeavour' as a whole. Historically, the process of scientific inquiry has evolved through paradigms, i.e. symbolic generalizations, metaphysical commitments, values, and exemplars shared by a community of scientists that guide their research [Kuhn, 1962]. For most of human history, scientists have observed phenomena and postulated laws or principles to simplify the complexity of observations into more manageable concepts. Initially, there were only experimental and theoretical sciences. Hey et al. [2009] refer to

[2] https://oecd.ai/en/

[3] https://aiindex.stanford.edu/report/

empirical observation and logical (theory) formulation as the first and second scientific paradigms, respectively. However, by the mid-20th century, numerous problems became too complex for analytical solutions, leading researchers to adopt simulation methods. Science entered a third paradigm characterized by the development of computational models and simulations to understand complex phenomena. We are transitioning towards a fourth scientific paradigm where scientific investigation is rooted in data-intensive computing, enabled by the extensive deployment of intelligent machines capable of extracting representations, rules, and patterns from data [King et al., 2009, Hey et al., 2009].

The shift towards this fourth scientific paradigm implies that there may be a change in the way science progresses. Scientific advancements stem from individuals' abilities to balance exploration and exploitation of the knowledge space efficiently [Uzzi et al., 2013]. With the integration of data-intensive computing and intelligent machines, the process of knowledge creation could potentially be transformed or even accelerated as these new technologies help uncover representations, rules, and patterns from vast amounts of data. The theory of re-combinatorial knowledge creation posits that new knowledge primarily results from the recombination of existing pieces of knowledge [Weitzman, 1998, Uzzi et al., 2013, Wang et al., 2017]. Scientific progress is thus the outcome of individual and collective creativeness, where *creativity* is defined as the "*production of high-quality, original, and elegant solutions to complex, novel, ill-defined, or poorly structured problems*" [Hemlin et al., 2013]. Innovation relies on the exploration of the knowledge space. The way organizations manage the balance between exploring new ideas and exploiting existing ones can be easily transferred to the scientific community. This balance determines the trade-off that organizations face when attempting to innovate while preserving established routines and practices [March, 1991]. March argues that organizations must find a suitable compromise between these two aspects to survive and succeed in the long run. The same stand for science since scientists try to innovate within established paradigms, and survival in science can be considered through peer recognition. This perspective emphasizes two facets of creativity in science: the novelty and relevance of the research conducted. Measures of novelty or atypicality are based on the concept of knowledge recombination [Uzzi et al., 2013, Lee et al., 2015, Foster et al., 2015, Wang et al., 2017, Shibayama et al., 2021]. One can approximate the difficulties of combining pieces of knowledge within a scientific document (i.e.

an article). Although these measures may contain biases due to their reliance on citation networks, they remain crucially helpful. They provide evidence of possible biases in journals' peer-review processes [Wang et al., 2017] as well as in funding allocations towards too-novel research [Boudreau et al., 2016, Carayol et al., 2017, Franzoni et al., 2022]. Furthermore, these metrics also provide a basis for analyzing the influence of technology on research. They help to identify whether the technology offers a more cross-disciplinary viewpoint by helping navigation through the knowledge space or allows for exploiting this space by providing a sharper perspective on a well-defined problem. These questions of how AI would lead to further exploration or exploitation of the knowledge space and how it affects the associated recognition will be addressed in Chapter 1 of this thesis.

Since science is a social phenomenon [Fleck, 2012], factors related to individuals' social capital will determine how AI disseminates within the scientific community. This paradigm shift suggests that technology can impact the nature of the research conducted but also implies that a growing number of researchers in various application domains will focus on these technics. Not all individuals have equal access to AI-based technologies. A researcher's ability to adopt AI largely depends on their scientific and technical human capital, including cognitive skills, scientific and technical knowledge, and contextual skills. Resources can be divided into two broad categories: those that reside within the individual and those that are anchored in the relationships between the individual and their working environment [Bozeman et al., 2001, Bozeman and Corley, 2004]. Consequently, the adoption of AI in science is closely linked to how researchers mobilize and are limited by their resources. Although the objective is not for all researchers to necessarily use AI, it still seems important to understand the factors that promote its adoption to uncover mechanisms that enable a broader range of researchers to benefit from the advantages linked to AI. Chapter 2 will discuss the relationship between individuals' scientific and technical human capital and the adoption of artificial intelligence in application domains.

The third contribution of this thesis, presented in Chapters 3 and 4, focuses on the relation between the cognitive dimension within a research team and knowledge creation. Nahapiet and Ghoshal [1998] conceptualizes three dimensions of social capital that impact intellectual capital development: structural, relational, and cognitive. Structural capital examines the connections between individuals and their respective networks; relational capital represents the nature and intensity of the relationships between team members; and cognitive capital symbolises the shared background between individuals and their common language. In science, cognitive diversity is often promoted through interdisciplinary projects, as the intersection of different perspectives is commonly needed to solve complex scientific problems [Page, 2008]. Indeed, people from outside a domain may have an advantage in offering fresh ideas through their distinct knowledge [Jeppesen and Lakhani, 2010, Kuhn, 1962]. Chapters 3 and 4 enhance our understanding of novelty, scientific impact indicators, and their association with social dimensions. Chapter 3 introduces an open-source Python-based tool, "*Novelpy*", which allows the computing of various metrics of novelty and disruption. This chapter also formalises existing indicators mathematically in a common framework. It seems essential to consider the social dimension of the innovation process to better understand how to identify potentially innovative research without solely relying on measures based on citation networks. The final chapter of this thesis, chapter 4, takes a step back and analyzes the origin of these novelty indicators, considering the cognitive dimension of the team as a determining factor of its creativity.

Outline of the thesis

In the current context of artificial intelligence becoming a transformative force in research, there is a growing need to address questions surrounding its adoption and impact on the scientific process. The primary objective of this thesis is to shed light on three main questions:

- How does AI affect the knowledge production process in terms of novelty and scientific recognition?
- What are the factors that promote the adoption of this technology in scientific application domains?
- How individuals' ability to explore the knowledge space, and cognitive distances between team members influence their capacity to combine distant knowledge and the resulting recognition?

To answer these questions, Chapters 1 and 2 will offer insights into the two first queries, while Chapters 3 and 4 will concentrate on the third one. The subsequent sections will present a comprehensive overview of each chapter and the methodologies employed throughout this thesis.

Chapter 1

This initial chapter provides insights into the dissemination and impact of artificial intelligence, specifically neural networks, in science. Some recent studies have documented the diffusion of AI and deep learning in science [Cockburn et al., 2018, Klinger et al., 2021], but none have explored how its use influences scientific discovery. Our article addresses this gap by examining how the use of neural networks (NNs) affects combinatorial novelty and the scientific impact of articles in the health sciences.

To identify articles using AI, we employed a new and original method based on word embedding which allowed us to identify some 250,000 documents published between 1990 and 2018 from Web of Science. By analyzing these documents, we considered five key attributes that define a technology as 'emerging' – namely: (i) radical novelty, (ii) fast growth, (iii) coherence, (iv) prominent impact, and (v) uncertainty and ambiguity [Rotolo et al., 2015] – and demonstrated that NNs conform to these properties. We found that NN research activity has grown exponentially across nearly all sciences and globally, with the diffusion process following a double-boom cycle and a strong reconfiguration of global actors. The diffusion of NN methods into application domains began cross-disciplinary involving the computer sciences, breaking their way into 'pure' field-specific research within the various application domains.

We subsequently investigated the impact of technology adoption on scientific discovery, particularly focusing on health sciences. We found a negative correlation between adopting NN methods and combinatorial novelty by employing novelty indicators from Uzzi et al. [2013] and Wang et al. [2017]. At the same time, we observed a positive correlation with the expectation and dispersion of citations received, thus increasing a contribution's likelihood of becoming a 'big hit.'

Our findings prompt us to adopt a more moderate stance in the recent debate regarding AI's influence on knowledge development. We conclude that while NN methods do not yet function as an autopilot for navigating the sea of knowledge and connecting ideas, they represent a potent and versatile research tool that impacts knowledge creation in tangible ways. As such, we propose that AI be considered as an *emerging general method of invention.*

Chapter 2

This chapter aims to understand the factors that promote AI adoption by domain scientists. While most literature focuses on an article-level analysis, we propose studying the dynamics of AI adoption at the *individual level*.

We relied on the Scientific & Technical Human Capital (STHC) framework proposed in the seminal paper by Bozeman and Corley [2004] and, hence, divided the authors' STHC into three main dimensions: individual characteristics, social environment, and institutional context. We operationalized the three dimensions by blending OpenAlex data with information on the computational capabilities of institutions, social networks and exploratory profiles of individuals.

In this study, we show that the proportion of researchers adopting AI who will eventually use the technology again remains relatively stable at around 35%. Despite technological advancements and the increasing availability of resources to facilitate AI utilization, incorporating AI into a researcher's future work seems not determined by its progress and accessibility. Researchers adopting AI often apply this technology in a familiar field, with 62% publishing their first AI article in a journal sharing the same primary concept as their initial publication. On average, the number of researchers with computer science or AI skills is higher in AI-based papers involving domain scientists, suggesting that published AI articles demand specialized skills compared to researchers' prior publications.

Our results indicate that the STHC offers a valuable framework for understanding drivers of AI adoption in application domains. Some institutional dimensions, such as the degree of specialization, significantly affect individuals' ability to transition to AI usage and to appropriate it in the long term. Physical infrastructures (High-performance computing) appear beneficial only in some domains, emphasizing that the lack of local physical infrastructure may not be the most significant barrier to undertaking AI research contrary to the popular belief and evidence from macro-level studies [Ahmed and Wahed, 2020]. Also, the composition of the social environment (i.e. past collaboration network) is strongly related to AI integration in researchers' practices and its long-term adoption; both adopters and reusers belong to networks populated by computer scientists or individuals with AI backgrounds at the expense of domain scientists. Scholars with more diverse backgrounds are more likely to embrace and reuse AI in their research, indicating that individuals with a more exploratory profile are more prone to transitioning toward new technology. Lastly, we found that young researchers enhance AI adoption and reuse; many past collaborations with young researchers make AI adoption more straightforward, and their presence on the team during initial trials strongly influences the reuse of the technology.

Institutions must foster a culture that encourages knowledge sharing, promotes interaction among scientists, and identifies and supports 'boundary-spanning' individuals who can bridge the gap between AI expertise and other scientific domains. Additionally, reconsidering the allocation of resources towards more modest but widespread investments in data science or ML infrastructure could democratize AI, promoting its adoption across a broader range of scientific disciplines.

Chapter 3

Chapter 3 is a methodological chapter. It proposes *Novelpy*, an open-source Python package designed to compute novelty and disruption indicators for scientific articles and patents. This chapter also provides a comprehensive review of the various indicators available in *Novelpy* by formally describing these measures (both mathematically and graphically).

Novelty measures are based on the concept of knowledge combination, the indicators calculate the difficulty associated with the combinations realized in an article to determine whether it is based on distant or proximate knowledge within a certain knowledge space. As it is common in the literature, pieces of knowledge are represented by the journals or abstracts of an article's references or its keywords.

Disruption measures, on the other hand, analyze how a focal article acts as a bottleneck between future papers and the references of the focal papers. They capture whether a document consolidates a domain (i.e., future papers rely on the references used in the focal paper) or disrupts it (i.e., future papers only reference the focal paper).

Although there are several packages available in *R* and *Python* designed to study citation, co-authorship, or any coupling (e.g. *ScientoPy*, Ruiz-Rosero et al. [2019] ; *scientoText*, Uddin et al. [2016]; *Metaknowledge*, McLevey and McIlroy-Young [2017] or *bibliometrix*, Aria and Cuccurullo [2017]), yet libraries to compute novelty and disruptiveness indicators remain unavailable. Our effort aims to provide the scientometrics community with a tool that centralizes different measures of novelty and disruptiveness, facilitates their comparison, and promotes reproducibility.

Novelpy package incorporates novelty measures from Uzzi et al. [2013], Foster et al. [2015], Lee et al. [2015], Wang et al. [2017], and Shibayama et al. [2021], as well as disruptiveness measures from Wu et al. [2019], Bornmann et al. [2019a], and Bu et al. [2019]. To demonstrate the module's capabilities, we close the chapter by comparing the different measures on a random sample of 1.5M articles drawn from the PubMed Knowledge Graph.

Chapter 4

This final chapter focuses more closely on novelty in science. Only a few studies seek to explain the mechanisms that give rise to novelty. In this work, we develop a new indicator that allows us to measure the team's *cognitive diversity* and the propensity of its members to *explore* the knowledge space. The indicator is built using word embedding techniques on the publication history of team members. We test its relationship with novelty indicators and validate it using peer recommendations from Faculty Opinions, following Bornmann et al. [2019b].

We can think of our indicator as a measure of *potential novelty*, i.e., opportunities for new knowledge recombination available through the diversity of backgrounds in the team and the capacity of individuals to bridge the gap between other team members. In comparison, combinatorial novelty indicators would capture the *realized novelty*, i.e., the output of the research conducted by this team in terms of pieces of knowledge used. Finally, Faculty Opinion labelling and other external validation methods can describe the *perceived novelty*, i.e., the peers' perception of this study. Seen from this perspective, we investigate whether *potential* novelty contributes to *realized* and *perceived* novelty and its scientific recognition, measured with metrics of disruptiveness [Wu et al., 2019, Bornmann et al., 2019a, Bu et al., 2019]. To this end, we use the PubMed Knowledge Graph and examine approximately 1.8M articles from the 2000-2005 period, focusing on less recent publications to manage the fact that novel articles are more often "sleeping beauties" and accumulate citations in the long run [Lin et al., 2021].

Our findings emphasize the critical role of the cognitive dimensions in creativity, as it significantly influences originality and success. We show that cognitive diversity always seems beneficial to combine more distant knowledge. In contrast, the within-team average exploratory profile follows an inverse U-shaped relation with combinatorial novelty (i.e. there is a turning point where it is no longer beneficial). The same relation can be found when examining the impact in terms of citations. However, our study highlights the strong connection between the cognitive dimension and the nature of these citations. More specifically, teams with more *exploitative* profiles tend to consolidate science, while those with more *exploratory* individuals disrupt it when associated with *exploitative* ones.

In short, our research underscores the importance of team composition in terms of profiles for scientific creativity. We show that the joint presence of highly exploratory and exploitative individuals constitutes the most effective team compositions for disrupting science, yet a limited number of highly exploratory individuals is essential to maximize the relevance of knowledge combinations achieved.

Methodology

This thesis is largely based on quantitative analyses of science, employing methods from bibliometrics and scientometrics. Thus, we make use of massive databases represented as graphs, connecting scientific entities with one another, such as authors, articles, institutions, etc. A brief history of these domains is in order.

The term bibliometrics was first defined by Belgian author Paul Otlet in 1934 [Otlet, 1934] and reintroduced in its English version by Pritchard et al. [1969] in the paper “Statistical Bibliography or Bibliometrics?”. The discipline’s early focus was improving the classification and organization of books to avoid the flood of knowledge. More precisely, librarians used it to select relevant items for their collections [Sugimoto and Larivière, 2018]. In 1955, chemist and documentalist Eugene K. Garfield proposed the creation of a citation index to offer an analysis tool by studying the links between different scientific documents. The Institute for Scientific Information, founded by Garfield in 1960, developed the Science Citation Index (SCI), first launched in 1963 for researchers and librarians. However, the beginning of bibliometrics studies was in the 1960s, with one of the central figures being Derek J. de Solla Price [de Solla Price, 1965, Boyack et al., 2005]. Back then, the initial focus was to understand research as a system by examining the growth in publications and the outline of citation activity. Early on, the number of references, the density of citation count for papers, and the inequality in the citation process was of particular interest. The fundamental problems of keeping track of relevant pieces of knowledge and the progression of the science system remain at the core of the science of science research [Fortunato et al., 2018]. The number of databases used by scholars has increased in the last decades. New database structures like Knowledge Graphs (KG) emerged (e.g. Microsoft Academic Graph (MAG, replaced by OpenAlex in 2022), PubMed Knowledge Graph (PKG)). Although the name KG was first used in Schneider [1973], it was only popularized in 2012 when Google presented their own KG. MAG and PKG help match different knowledge units for a paper, a crucial task in scientometrics. These units can be an author profile, a journal, a reference, or even topics. KGs help us understand more deeply how science is structured and performed.

The increasing availability of data has led in turn to a diversification in the field of scientometrics. Scholars often use terms like Bibliometrics, Scientometrics, Informetrics, Webometrics, and Altmetrics interchangeably. Extensive literature traces the history of these fields and seeks to understand their inter-differences, as well as creating intra-field taxonomies. They all share a common goal: studying science as a system, using scientific methods. Yet some differences remains, Informetrics is a sub-discipline of information sciences and is defined as the application of mathematical methods to the content of information science [Chellappandi and Vijayakumar, 2018]. In other words, Informetrics is the highest level of abstraction, and every other field is a subset of Informetrics. Bibliometrics, as we have seen above, focuses on citations for collection management and document retrieval using specific aspects of the document without placing it in the overall context in which it was created. In contrast, Scientometrics is a “meta-science” that quantitatively analyzes the production, dissemination, and underlying system’s mechanisms [Sugimoto and Larivière, 2018, Chellappandi and Vijayakumar, 2018]. Finally, Webometrics and Altmetrics are both concerned with information available on the web, but Webometrics is document-focused, with the document being a web page, while Altmetrics is focused on the networking aspect and complements citations with the number of likes and retweets [Mingers and Leydesdorff, 2015].

This thesis is essentially built on methods from Scientometrics. But Scientometrics itself can be further divided into two macro types of analysis: performance and Science Mapping Analysis (SMA) [Moral Muñoz et al., 2020]. The goal of the former is to assess scientific actors’ activities and their impact. Its purpose is, therefore, to assign a value to the productivity and pervasiveness of the research conducted by a unit (article, author, institution). SMA “*is mostly directed at monitoring a scientific field to determine its (cognitive) structure, its evolution, and main actors within*” [Noyons et al., 1999]; it takes a snapshot of a part of the scientific system at a given moment to analyze its structure.

Inputs, outputs, and impacts of these scientific activities are the three perspectives used in performance analysis and SMA [Sugimoto and Larivière, 2018]. Input refers to human and financial resources and captures the different interactions of agents in the system at various levels (Authors/Institutional/Country levels); output is the end result of the research process, which is the composition of this document and the different entities that characterize it; and finally, impact measures study the repercussions of the outputs, the dissemination of knowledge that an article generates through citations, attention by the general public, or reutilization of the document's component.

This thesis addresses these three dimensions in the study of artificial intelligence in science. Indeed, in Chapter 1, as outlined earlier, we investigate the outputs and scientific impact of AI publications. In Chapter 2, we study the inputs of this research. Chapters 3 and 4 aim finally to understand the relationships between the inputs at the author level, the outputs of the conducted research and their scientific impact.

Introduction générale

Lancée le 14 mars 2023, la nouvelle version de ChatGPT, GPT-4, marque un tournant dans le domaine de l'intelligence artificielle (IA). Grâce à sa capacité à combiner le texte et la vision, cette entité peut désormais passer avec succès des examens – avec des notes élevées – dans presque tous les domaines. Lorsque j'ai commencé cette thèse en 2019, il ne s'était écoulé que sept ans depuis que nous pouvions reconnaître un chat avec une précision de 75% [Le, 2013], et seulement quatre ans depuis que l'IA de Microsoft avait "surpassé" les humains dans leur capacité à reconnaître des entités dans des images [He et al., 2015]. Quatre ans plus tard, nous avons assisté à une explosion du développement de l'IA, en particulier dans le domaine du traitement du langage naturel, et nous en savons encore peu sur ses implications et ses dangers potentiels (par exemple, la désinformation et les "fake news", les risques de cybersécurité, les préoccupations éthiques ou l'impact psychologique sur les individus). En réponse à ce développement croissant de l'IA, dont nous ne comprenons ni ne gérons pleinement les conséquences, une lettre ouverte appelant la recherche sur l'IA à faire une pause immédiate d'au moins six mois dans l'entrainement d'autres systèmes, plus puissants que GPT-4, a été publiée le 22 mars 2023. Cette pétition a été signée par plus de 25 000 personnes, dont 1 000 chercheurs et experts en IA[4]. La prolifération de l'IA et l'amélioration de ses capacités depuis dix ans impliquent des changements sociétaux et pourrait modifier la façon dont la science fonctionne. Cette thèse aborde spécifiquement la relation entre la science et la technologie de l'intelligence artificielle. Plus précisément, elle vise à comprendre comment les diverses applications de l'IA peuvent affecter la nature de la recherche menée dans les domaines d'application. Pour ce faire, nous prenons en compte les facteurs sociaux afin de mieux comprendre comment les chercheurs adoptent la technologie et produisent de nouvelles connaissances. Nous présentons ci-dessous une vue d'ensemble

[4]Parmi lesquels Joshua Bengio, l'un des trois pères fondateurs de l'apprentissage profond et lauréat du prix Turing. La pétition est disponible à l'adresse suivante : https://futureoflife.org/open-letter/pause-giant-ai-experiments/

du rôle de l'IA dans ce nouveau paradigme scientifique.

Les scientifiques dépendent de l'évolution de la technologie pour mener des experience et valider des théories. Grâce aux nouvelles technologies, les scientifiques peuvent explorer différemment l'espace de connaissances et faire de nouvelles découvertes. Comme le souligne Derek de Solla Price, "*The changes of paradigm that accompany great and revolutionary changes may sometimes be caused by inspired thought, but much more commonly they seem due to the application of technology to science*" [de Solla Price, 1963]. La science et la technologie entretiennent une relation bidirectionnelle : la science fournit des principes fondamentaux pour le développement de nouvelles technologies, et la technologie, à son tour, génère les instruments et les techniques nécessaires pour répondre plus efficacement à des questions scientifiques nouvelles et plus difficiles [Brooks, 1994]. Plusieurs avancées technologiques ont remodelé le paysage scientifique au cours de l'histoire. Il suffit de penser à l'invention du microscope, qui a conduit à la découverte des cellules, ou à la cristallographie aux rayons X, qui a facilité la compréhension des structures moléculaires et de l'ADN. Ces développements ont permis d'approfondir notre compréhension des mécanismes biologiques et ont servi de base à de nombreuses découvertes et innovations dans diverses disciplines scientifiques. Plus récemment, les progrès de l'informatique ont permis aux scientifiques de créer des modèles mathématiques complexes et de résoudre des problèmes auparavant considérés comme insolubles. Par exemple, les ordinateurs ont été utilisés pour résoudre des problèmes de théorie des nombres, de cryptographie, d'optimisation combinatoire, pour simuler des phénomènes naturels (systèmes climatiques, dynamique moléculaire, structures des galaxies) et enfin, ont permis l'émergence de l'intelligence artificielle.

L'IA est au cœur du paradigme technologique actuel, partageant plusieurs similitudes en termes d'échelle et de portée avec les révolutions technologiques précédentes. Le terme "intelligence artificielle" a été lancé par l'informaticien John McCarthy à l'occasion du projet de recherche estival de Dartmouth sur l'intelligence artificielle en 1956, un événement fondateur pour le domaine [McCarthy et al., 1955]. L'objectif de l'IA était de faire en sorte que les machines utilisent le langage, construisent des abstractions, résolvent des problèmes humains et s'améliorent d'elles-mêmes. Les définitions de l'IA varient, mais elles impliquent généralement que les machines simulent un comportement intelligent, exécutent des tâches complexes et tirent des enseignements de leurs expériences. Par exemple, la Commission européenne définit l'IA comme suit : "*machines or agents capable of observing their environment, lear-*

ning, and taking intelligent action or proposing decisions" [Annoni et al., 2018, p.19]. Selon l'OCDE, les systèmes d'IA sont des "*machine-based systems that can make predictions, recommendations, or decisions for a given set of human-defined objectives* "[OECD, 2019, p.23]. L'OMPI définit les systèmes d'IA comme des "*learning systems that can improve at tasks typically performed by humans with limited or no human intervention* "[WIPO, 2019, p.19]. Les termes "apprentissage automatique", "apprentissage profond" et "intelligence artificielle" sont souvent utilisés de manière interchangeable.

Au début, l'IA s'est concentrée sur la résolution de problèmes que des règles mathématiques formelles pouvaient décrire. Ces problèmes sont intellectuellement difficiles pour les humains, mais sont plus simples pour les ordinateurs car ils peuvent être codés formellement, ce qui permet de trouver des solutions à l'aide de règles d'inférence logiques. Cette méthode, connue sous le nom d'approche "basée sur la connaissance", implique une architecture avec une base de connaissances et un mécanisme d'inférence. La base de connaissances stocke les informations du monde réel, tandis que le mécanisme d'inférence permet à la machine de déduire des schemas à partir des informations stockées. Cette approche a été dominante au cours des premières décennies, avec des applications telles que les "systèmes experts" introduits dans les années 1970 pour simuler le jugement et le comportement humains dans des domaines spécifiques. Ces systèmes étaient efficaces pour certains types de problèmes, mais pas pour ceux qui nécessitaient des connaissances substantielles subjectives et intuitives ou des capacités de perception. Ces problèmes sont faciles à résoudre pour les humains, mais difficiles à formuler de manière formelle et mathématique [Nilsson, 2009]. Au cours de la même période, une autre approche à l'intelligence machine a commencé à s'imposer dans la communauté scientifique. Cette approche est rapidement devenue connue sous le nom d'"apprentissage machine". L'"apprentissage machine" se concentre sur la conception de systèmes intelligents capables d'acquérir des connaissances en extrayant des régularités à partir de données brutes. Contrairement aux systèmes basés sur la connaissance, les méthodes d'apprentissage automatique construisent des hypothèses directement à partir des données par inférence inductive, ce qui permet aux machines de s'attaquer à des problèmes impliquant des connaissances du monde réel et d'atteindre certaines capacités semblables à celles de l'homme, telles que la reconnaissance d'objets. Bien que l'apprentissage automatique se soit avéré être une alternative efficace aux systèmes basés sur la connaissance et qu'il soit devenu l'une des branches les plus importantes de l'IA à partir des années 1980, des défis particuliers subsistaient. Principalement, les méthodes traditionnelles d'apprentissage automatique ont rencontré d'importantes difficultés pour extraire des caractéristiques abstraites de haut niveau à partir de données brutes en raison de facteurs de variation, tels que les différentes formes, les ombres et les angles de vue. Ces facteurs de variation sont essentiellement des constructions de l'esprit humain qui peuvent être considérés comme des abstractions de haut niveau qui nous aident à donner un sens à la riche variabilité des données observées [Nilsson, 2009, Goodfellow et al., 2016]. L'approche 'apprentissage profond' de l'intelligence artificielle est apparue comme une solution efficace aux défis posés par les méthodes traditionnelles d'apprentissage automatique. Les systèmes d'apprentissage profond (DL) apprennent par l'expérience et appréhendent le monde à travers une hiérarchie de concepts abstraits, chacun défini en relation avec des concepts plus simples [Schmidhuber, 2015, LeCun et al., 2015, Goodfellow et al., 2016]. Cette approche présente deux avantages significatifs. Premièrement, à l'instar des algorithmes d'apprentissage automatique plus simples, la machine acquiert des connaissances à partir d'expériences passées, ce qui évite aux humains de devoir fournir toutes les connaissances formelles nécessaires pour atteindre un objectif spécifique. Deuxièmement, la complexité et l'abstraction des concepts ne sont plus des obstacles, car la machine peut les reconstruire et les combiner les uns avec les autres. Cette hiérarchie de concepts fait que le processus d'apprentissage peut être considéré comme structuré en plusieurs couches, d'où le terme "profond". Les techniques d'IA ont été utilisées avec succès dans des domaines aussi variés que la prédiction de la structure 3D des protéines [Jumper et al., 2021], la régulation du plasma de fusion nucléaire dans la configuration du tokamak [Degrave et al., 2022], la prédiction de la formation de la structure de l'Univers [He et al., 2019], ou la création d'une carte du cerveau des petits insectes [Winding et al., 2023], pour n'en citer que quelques-uns. En 2017, l'intelligence artificielle a connu une autre avancée remarquable avec l'émergence de modèles de Transformers dotés de mécanismes d'auto-attention [Vaswani et al., 2017].

Les résultats impressionnants des "Generative Pre-trained Transformers" (GPT), tels que ChatGPT ou GPT-4, illustrent désormais clairement la nature de technologie polyvalente de l'intelligence artificielle (General-Purpose Technology - GPT), en montrant leur adaptabilité et leur large applicabilité dans de nombreux domaines. Ces modèles ne se contentent pas d'interagir avec les utilisateurs sur divers sujets, aussi ils peuvent accompagner les humains dans un processus de pensée, en apportant des perspectives supplémentaires étayées par des connaissances quasi-expertes sur divers sujets. Les GPTs (Generative Pre-trained Transformers) sont donc des GPTs (General-Purpose Technologies). En mars 2023, Eloundou et al. [2023], ont immortalisé le jeu de mots en publiant l'article "GPTs are GPTs : An early look at the labor market impact potential of large language models", qui met en lumière les capacités du GPT-4 et son impact potentiel sur le marché du travail, suggérant que 15 % de toutes les tâches des travailleurs aux États-Unis pourraient être accomplies beaucoup plus rapidement tout en conservant le même niveau de qualité. À noter que certaines de ces tâches font également partie du monde scientifique, comme la programmation et l'écriture [Eloundou et al., 2023]. L'intelligence artificielle possède les attributs d'une technologie à usage universel, avec des applications très variées dans de nombreuses disciplines. Les GPT, comme l'IA, se distinguent des autres innovations par leur application étendue dans divers secteurs, leur capacité à catalyser d'autres innovations dans les secteurs d'application et leur amélioration rapide et continue [David, 1990, Bresnahan and Trajtenberg, 1995]. Parmi les exemples classiques de GPT, on peut citer le moteur électrique et le microprocesseur, qui ont entraîné d'importants changements technologiques et organisationnels dans divers secteurs tels que l'industrie manufacturière, l'agriculture, le commerce de détail et la construction de logements.

Le rôle de l'IA en tant que GPT dans la science est également illustré par sa position d'"invention dans les méthodes d'invention" (IMI). Les IMI créent ou améliorent des produits spécifiques et fournissent une nouvelle façon de générer de nouveaux produits avec des applications plus larges. Par exemple, le croisement en agriculture est une IMI qui a conduit au développement de nombreuses nouvelles variétés de plantes cultivées, ce qui a eu un impact profond sur la productivité agricole. L'impact économique de l'IA en tant qu'outil de recherche va au-delà de la simple réduction des coûts d'activités liés a des innovations spécifiques, car elle permet une approche entièrement nouvelle de l'innovation elle-même. L'omniprésence de l'IA en tant que GPT et IMI la positionne comme une technologie unique capable de conduire des changements organisationnels dans un large éventail de domaines scientifiques. Cependant, elle doit être utilisée de manière responsable. C'est pourquoi des institutions telles que la Commission européenne soutiennent une approche réglementaire et axée sur l'investissement pour promouvoir l'adoption de l'IA tout en abordant les risques associés à la technologie [European-Commission, 2020]. Des initiatives similaires ont été mises en œuvre par l'OCDE, qui a lancé un observatoire de la technologie spécifique pour suivre de près l'évolution de la technologie et fournir une analyse politique de l'IA fondée sur des données empiriques[5]. Ces initiatives ne se limitent pas à l'Europe ; depuis 2019, des institutions comme l'Université de Stanford publient régulièrement des données annuelles sur l'évolution de l'IA, le marché du travail, les compétences et l'automatisation, afin de guider la prise de décisions responsables et éthiques[6].

Les progrès continus de l'IA dans les disciplines scientifiques, en particulier les réalisations remarquables obtenues grâce aux techniques des réseaux neuronaux, ne se limitent pas à des découvertes anecdotiques impressionnantes. Au contraire, ces progrès ont entraîné une tendance croissante à passer d'une exploration scientifique fondée sur des hypothèse à une exploration scientifique fondée sur des données. Le paradigme scientifique émergent est fondé sur le recours intensif aux données, facilité par la généralisation de machines intelligentes capables de discerner des représentations, des règles et des schémas dans un volume sans cesse croissant de données structurées et non-structurées. La découverte scientifique peut être considérée à la fois comme le processus et le produit d'une étude scientifique aboutie. Dans son sens le plus étroit, le terme de découverte peut se référer à ce que l'on appelle le “moment d'eurêka”, qui consiste à acquérir de nouveaux points de vue. Toutefois, dans ce contexte, nous adoptons le sens le plus large de découverte, comme synonyme d'“ effort scientifique fructueux”.

Historiquement, le processus de recherche scientifique a évolué à travers des paradigmes, c'est à dire, des généralisations symboliques, des principes philosophiques, des valeurs et des références partagés par une communauté de scientifiques qui guident leurs recherches [Kuhn, 1962]. Durant la majeure partie de l'histoire de la science, les scientifiques ont observé des phénomènes et postulé des lois ou des principes pour simplifier la complexité des observations en concepts plus faciles à manipuler. Au départ, il n'existait que des sciences expérimentales et théoriques. Hey et al. [2009] considèrent l'observation empirique ainsi que la formulation logique (théorie) comme étant respectivement le premier et le deuxième paradigme scientifique. Toutefois, vers le milieu du 20e siècle, de nombreux problèmes sont devenus trop complexes pour être résolus de manière analytique, ce qui a conduit les chercheurs à adopter des méthodes de simulation. La science est entrée dans un troisième

[5] https://oecd.ai/en/
[6] https://aiindex.stanford.edu/report/

paradigme caractérisé par le développement de modèles informatiques et de simulations pour comprendre des phénomènes complexes. Nous sommes en train de passer à un quatrième paradigme scientifique dans lequel l'investigation scientifique est ancrée dans le traitement intensif des données, rendu possible par le déploiement à grande échelle de machines intelligentes capables d'extraire des représentations, des règles et des régularités à partir des données [King et al., 2009, Hey et al., 2009].

Le passage à ce quatrième paradigme scientifique implique qu'il pourrait y avoir un changement dans la façon dont la science progresse. Les progrès scientifiques découlent de la capacité des individus à équilibrer efficacement l'*exploration* et l'*exploitation* de l'espace de la connaissance [Uzzi et al., 2013]. Avec l'intégration de données et de machines intelligentes, le processus de création de connaissances pourrait potentiellement être transformé, voire accéléré. Ces nouvelles technologies permettent, de manière autonome, de découvrir des représentations, des règles et des modèles à partir de vastes quantités de données. La théorie de la création de connaissances par recombinaison postule que les nouvelles connaissances résultent principalement de la re-combinaison de connaissances existantes [Weitzman, 1998, Uzzi et al., 2013, Wang et al., 2017]. Le progrès scientifique est donc le résultat de la créativité individuelle et collective, *créativité* étant définie comme la "*production of high-quality, original, and elegant solutions to complex, novel, ill-defined, or poorly structured problems*". [Hemlin et al., 2013]. L'innovation repose sur l'exploration de l'espace des connaissances, en effet la manière dont les organisations gèrent cet équilibre entre l'exploration de nouvelles idées et l'exploitation des idées existantes peut être facilement transposée au monde scientifique. Cet équilibre détermine les compromis auxquels les organisations sont confrontées lorsqu'elles tentent d'innover tout en préservant les routines et les pratiques établies [March, 1991]. March soutient que les organisations doivent trouver un équilibre adéquat entre ces deux aspects pour survivre et réussir à long terme. Il en va de même pour la science, puisque les scientifiques tentent d'innover dans le cadre des paradigmes établis, et que la survie en science peut être envisagée par le biais de la reconnaissance par les pairs. Cette perspective met l'accent sur deux aspects de la créativité en science : la nouveauté et la pertinence de la recherche menée. Les mesures de nouveauté ou d'atypicité sont basées sur le concept de recombinaison des connaissances [Uzzi et al., 2013, Lee et al., 2015, Foster et al., 2015, Wang et al., 2017, Shibayama et al., 2021]. On peut estimer les difficultés à combiner des éléments de connaissance au sein d'un article scientifique. Bien que ces mesures puissent présenter des biais en raison de leur dépendance aux réseaux de citations, elles n'en demeurent pas moins d'une utilité capitale. Elles fournissent des preuves d'éventuels biais dans les processus d'évaluation par les pairs dans les revues [Wang et al., 2017] ainsi que dans les allocations de fonds en défaveur des recherches trop novatrices [Boudreau et al., 2016, Carayol et al., 2017, Franzoni et al., 2022]. Ces indicateurs fournissent également une base pour analyser l'influence de la technologie sur la recherche. Ils permettent de comprendre si la technologie offre un point de vue plus transversal, en aidant à naviguer dans l'espace des connaissances ou permet d'exploiter cet espace en offrant une perspective plus fine sur un problème bien défini. Ces questionnements sur la manière dont l'IA peut conduire à une exploration ou une exploitation de l'espace de connaissance et comment cela affecte la reconnaissance seront traités dans le chapitre 1 de cette thèse.

La science étant un phénomène social [Fleck, 2012], les facteurs liés au capital social des individus déterminent la manière dont l'IA se diffuse au sein de la communauté scientifique. Ce changement de paradigme suggère que la technologie peut avoir un impact sur la nature de la recherche menée, mais implique également qu'un nombre croissant de chercheurs dans divers domaines d'application se concentreront sur ces techniques. Toutefois, tous les individus n'ont pas le même accès aux technologies basées sur l'IA. La capacité d'un chercheur à adopter l'IA dépend largement de son capital humain scientifique et technique, notamment de ses compétences cognitives, de ses connaissances scientifiques et techniques et de ses compétences contextuelles. Les ressources d'un individu peuvent être divisées en deux grandes catégories : celles propres à l'individu et celles qui sont ancrées dans les relations entre l'individu et son environnement de travail [Bozeman et al., 2001, Bozeman and Corley, 2004]. L'adoption de l'IA dans le domaine scientifique est étroitement liée à la manière dont les chercheurs mobilisent leurs ressources et sont limités par celles-ci. Bien que l'objectif ne soit pas que tous les chercheurs utilisent nécessairement l'IA, il semble néanmoins important de comprendre les facteurs qui favorisent son adoption et son application afin de comprendre les mécanismes permettant à un plus grand nombre de chercheurs de bénéficier des avantages liés à l'IA. Le chapitre 2 examine la relation entre le capital humain scientifique et technique des individus et l'adoption de l'intelligence artificielle dans les domaines d'application.

La troisième contribution de cette thèse, présentée dans les chapitres 3 et 4, se concentre sur la relation entre la dimension cognitive dans une équipe de recherche et la création de connaissances. Nahapiet and Ghoshal [1998] conceptualise trois dimensions du capital social qui influencent le développement du capital intellectuel : le capital structurel, relationnel et cognitif. Le capital structurel examine les liens entre les individus et leurs réseaux respectifs ; le capital relationnel représente la nature et l'intensité des relations entre les membres de l'équipe ; et le capital cognitif symbolise les compétences partagées entre les individus et leur langage commun. En science, la diversité cognitive est souvent encouragée par des projets interdisciplinaires, car l'intersection de différents points de vue est généralement nécessaire pour résoudre des problèmes scientifiques complexes [Page, 2008]. En effet, les personnes extérieures à un domaine peuvent avoir l'avantage d'apporter des idées nouvelles grâce à leurs connaissances distinctes [Jeppesen and Lakhani, 2010, Kuhn, 1962]. Les chapitres 3 et 4 améliorent notre compréhension de la nouveauté, des indicateurs d'impact scientifique et de leur association avec les dimensions sociales. Le chapitre 3 présente "*Novelpy*", un outil open-source basé sur Python visant à calculer divers indicateurs de nouveauté et de disruption. Ce chapitre formalise également mathématiquement ces indicateurs. Pour mieux comprendre comment identifier des recherches potentiellement innovantes sans s'appuyer uniquement sur des mesures basées sur le réseau de citations, il semble important de considérer la dimension sociale de ce processus d'innovation. Le dernier chapitre de cette thèse, le chapitre 4, fait un pas en arrière et analyse la source de ces indicateurs de nouveauté, en considérant la dimension cognitive de l'équipe comme un facteur déterminant de sa créativité.

Description de la thèse

Dans un contexte où l'intelligence artificielle constitue une source de mutation de la science, il est de plus en plus nécessaire d'aborder les aspects liés à son adoption et son impact sur le processus scientifique. L'objectif principal de cette thèse est de faire la lumière sur trois questions principales :

- Comment l'IA affecte-t-elle le processus de production de connaissances en termes de nouveauté et de reconnaissance scientifique ?
- Quels sont les facteurs qui favorisent l'adoption de cette technologie dans les domaines d'application scientifiques ?
- Comment l'aptitude des individus à explorer l'espace de connaissances, ainsi que les distances cognitives entre les membres de l'équipe, influencent leurs capacités à combiner des connaissances éloignées et la reconnaissance scientifique ?

Pour répondre à ces questions, les chapitres 1 et 2 apporteront des éclaircissements sur les deux première questions, tandis que les chapitres 3 et 4 se concentreront sur la troisième. Les sections suivantes présenteront un aperçu de chaque chapitre et des méthodologies employées tout au long de cette thèse.

Chapitre 1

Ce premier chapitre donne un aperçu de la diffusion et de l'impact de l'intelligence artificielle, en particulier des réseaux neuronaux, dans la science. Certaines études récentes ont documenté la diffusion de l'IA et de l'apprentissage profond dans la science [Cockburn et al., 2018, Klinger et al., 2021], mais aucune n'a exploré la façon dont leur utilisation influence la découverte scientifique. Notre article répond à cette question en examinant comment l'utilisation des réseaux neuronaux (NN) affecte la nouveauté combinatoire et l'impact scientifique des articles dans les sciences de la santé.

Pour identifier les articles utilisant l'IA, nous avons employé une méthode nouvelle et originale basée sur la vectorisation de mots qui a permis d'identifier environ 250 000 documents publiés entre 1990 et 2018 provenant de la base de données Web of Science. En analysant ces documents, nous avons pris en compte cinq attributs clés qui définissent une technologie comme "émergente" - à savoir : (i) nouveauté radicale, (ii) croissance rapide, (iii) cohérence, (iv) impact important, et (v) incertitude et ambiguïté [Rotolo et al., 2015] - et avons démontré que les NNs se conforment à ces propriétés. Nous avons constaté que l'activité de recherche sur les NNs a connu une croissance exponentielle dans presque toutes les sciences et à l'échelle mondiale, le processus de diffusion suivant un cycle à deux phases et une forte recomposition des acteurs mondiaux. La diffusion des méthodes de NN dans les domaines d'application a démarré de manière transdisciplinaire en impliquant les sciences informatiques, puis s'est frayée un chemin dans la "pure" recherche propre a différents domaines d'application.

Nous avons ensuite étudié l'impact de l'adoption de la technologie sur la découverte scientifique, en nous concentrant particulièrement sur les sciences de la santé. Nous avons constaté une corrélation négative entre l'adoption des méthodes NN et la nouveauté combinatoire en utilisant les indicateurs de nouveauté de Uzzi et al. [2013] et Wang et al. [2017]. Dans le même temps, nous avons observé une corrélation positive avec la probabilité et la dispersion des citations reçues, augmentant ainsi la probabilité qu'une contribution devienne un "grand succès", mais aussi un article peu cité.

Nos résultats nous incitent à adopter une position plus modérée dans le récent débat concernant l'influence de l'IA sur le développement des connaissances. Nous concluons que si les méthodes de NN ne fonctionnent pas encore comme un pilote automatique pour naviguer dans l'espace des connaissances et relier les idées, elles représentent un outil de recherche puissant et polyvalent qui a un impact tangible sur la création de connaissances. À ce titre, nous proposons que l'IA soit considérée comme un *méthode générale émergente d'invention*.

Chapitre 2

Ce chapitre vise à comprendre les facteurs qui favorisent l'adoption de l'IA par les scientifiques de domaine d'application. La plupart des études se concentrent sur une analyse au niveau article, mais nous proposons d'étudier la dynamique de l'adoption de l'IA au niveau *individuel*.

Nous avons utilisé le concept du capital humain scientifique et technique (STHC) proposé dans l'article fondateur de Bozeman and Corley [2004] et, par conséquent, nous avons divisé le STHC des auteurs en trois dimensions principales : les caractéristiques individuelles, l'environnement social et le contexte institutionnel. Nous avons opérationnalisé ces trois dimensions en associant les données d'OpenAlex à des informations sur les capacités informatiques des institutions.

Dans cette étude, nous montrons que la proportion de chercheurs qui adoptent l'IA et qui finiront par utiliser à nouveau cette technologie reste relativement stable, autour de 35 %. Malgré les avancées technologiques et la disponibilité croissante de ressources pour faciliter l'utilisation de l'IA, l'intégration de l'IA dans le travail futur d'un chercheur ne semble pas déterminée par ses progrès et son accessibilité. Les chercheurs qui adoptent l'IA appliquent souvent cette technologie dans un domaine familier, 62 % d'entre eux publiant leur premier article sur l'IA dans une revue partageant le même concept principal que leur publication initiale. En moyenne, le nombre de chercheurs ayant des compétences en informatique ou en IA est plus élevé dans les articles basés sur l'IA impliquant des scientifiques du domaine, ce qui suggère que les articles publiés basés sur l'IA requièrent des compétences plus spécialisées que les publications antérieures du chercheur.

Nos résultats indiquent que le STHC offre un cadre pertinent pour comprendre les catalyseurs de l'adoption de l'IA dans les domaines d'application. Certaines dimensions institutionnelles, telles que le degré de spécialisation, affectent de manière significative la capacité des individus à faire la transition vers l'utilisation de l'IA et de se l'approprier à plus long terme. Les infrastructures physiques ("High performance computing") ne semblent bénéfiques que dans certains domaines, ce qui souligne que le manque d'infrastructures physiques locales n'est peut-être pas l'obstacle le plus important à l'adoption de l'IA, contrairement à ce que l'on croit généralement et à ce que montrent les études au niveau macroéconomique [Ahmed and Wahed, 2020]. En outre, la composition de l'environnement social (c'est-à-dire le réseau de collaborations antérieures) est étroitement liée à l'intégration de l'IA dans les pratiques des chercheurs et à son adoption à long terme ; les chercheurs qui essaient l'IA et ceux qui réutiliseront cette technologie par la suite appartiennent à des réseaux peuplés d'informaticiens ou d'individus ayant une expérience de l'IA, au détriment des scientifiques du domaine. Les chercheurs ayant des profils plus variés sont plus susceptibles d'adopter et de réutiliser l'IA dans leurs recherches, ce qui indique que les personnes ayant un profil plus exploratoire sont plus susceptibles de s'orienter vers de nouvelles technologies. Enfin, nous avons constaté que les jeunes chercheurs favorisent l'adoption et la réutilisation de l'IA ; de nombreuses collaborations passées avec de jeunes chercheurs rendent l'adoption de l'IA plus simple, et leur présence dans l'équipe lors des premiers essai influe fortement sur la réutilisation de la technologie.

Les institutions doivent encourager une culture favorisant le partage des connaissances, la promotion des interactions entre chercheurs et l'identification et le soutien des individus "interdisciplinaires" capables de combler le fossé entre l'expertise en IA et les autres domaines scientifiques. De plus, une réévaluation de l'allocation des ressources vers des investissements plus modestes mais généralisés dans les infrastructures de science des données ou d'apprentissage automatique pourrait faciliter la démocratisation de l'IA.

Chapitre 3

Le chapitre 3 est un chapitre méthodologique. Il propose *Novelpy*, un module Python open-source conçu pour calculer des indicateurs de nouveauté et de disruption d'articles scientifiques et de brevets. Ce chapitre fournit également un aperçu détaillé des différents indicateurs disponibles dans *Novelpy* en décrivant formellement ces mesures (à la fois mathématiquement et graphiquement).

Les mesures de nouveauté sont basées sur le concept de combinaison de connaissances, et les indicateurs calculent la difficulté associée aux combinaisons réalisées dans un article pour déterminer s'il est basé sur des connaissances éloignées ou proches dans un certain espace de connaissances. Comme couramment dans la littérature, les éléments de connaissance sont représentés par les revues ou les abstracts des références d'un article ou par ses mots-clés.

Les mesures de disruption, quant à elles, analysent la manière dont un article cible agit comme un goulot d'étranglement entre les futurs articles et les références de l'article cible. Elles déterminent si un document consolide un domaine (c'est-à-dire que les futurs articles s'appuient sur les références utilisées dans l'article cible) ou le bouleverse (c'est-à-dire que les futurs articles ne font référence qu'à l'article cible).

Bien qu'il existe plusieurs packages disponibles dans *R* et *Python* conçus pour étudier la citation, la coécriture ou tout autre couplage (par exemple *ScientoPy* de Ruiz-Rosero et al. [2019] ; *scientoText* de Uddin et al. [2016] ; *Metaknowledge* de McLevey and McIlroy-Young [2017] ou *bibliometrix* de Aria and Cuccurullo [2017]), les bibliothèques permettant de calculer les indicateurs de nouveauté et de perturbation restent inexistantes. Notre effort vise à fournir à la communauté de la scientométrie un outil qui centralise les différentes mesures de nouveauté et de disruption, facilitant leur comparaison et promouvant leur reproductibilité.

Le module *Novelpy* intègre les mesures de nouveauté de Uzzi et al. [2013], Foster et al. [2015], Lee et al. [2015], Wang et al. [2017], et Shibayama et al. [2021], ainsi que les mesures de disruption de Wu et al. [2019], Bornmann et al. [2019a], et Bu et al. [2019]. Pour démontrer les capacités du module, nous terminons le chapitre en comparant les différentes mesures sur un échantillon aléatoire de 1,5 million d'articles tirés du PubMed Knowledge Graph.

Chapitre 4

Ce dernier chapitre s'intéresse de plus près à la nouveauté en science. Seules quelques études ont cherché à expliquer les mécanismes à l'origine de la nouveauté. Dans ce travail, nous développons un nouvel indicateur qui nous permet de mesurer la *diversité cognitive* d'une équipe et la propension de ses membres à *explorer* l'espace des connaissances. L'indicateur est construit à l'aide de techniques de plongement de mots (word2vec, Mikolov et al. [2013b]) sur l'historique des publications des membres de l'équipe. Nous testons sa relation avec les indicateurs de nouveauté et nous le validons en utilisant les recommandations des pairs de la Faculty Opinions, en suivant Bornmann et al. [2019b].

Nous pouvons considérer notre indicateur comme une mesure de la *nouveauté potentielle*, c'est-à-dire des possibilités de nouvelles combinaisons de connaissances offertes par la diversité des profils d'une équipe et la capacité des individus à établir des passerelles entre les autres membres de l'équipe. En comparaison, les indicateurs de nouveauté combinatoire captureraient la *nouveauté réalisée*, c'est-à-dire le résultat de la recherche menée par cette équipe en termes d'éléments de connaissance utilisés. Enfin, la labellisation des membres de Faculty Opinions et d'autres méthodes de validation externe peuvent décrire la *nouveauté perçue*, c'est-à-dire la perception de cette recherche par les pairs. Dans cette perspective, nous cherchons à savoir si la nouveauté potentielle contribue à la nouveauté réalisée et perçue et à sa reconnaissance scientifique, mesurée à l'aide de métriques de disruption [Wu et al., 2019, Bornmann et al., 2019a, Bu et al., 2019]. Pour ce faire, nous utilisons PubMed Knowledge Graph et examinons environ 1,8 million d'articles de la période 2000-2005, en nous concentrant sur les publications moins récentes pour gérer le fait que les nouveaux articles sont plus souvent des "beautés endormies" (*sleeping beauties*) et accumulent davantage de citations sur le long terme [Lin et al., 2021].

Nos résultats soulignent le rôle critique de la dimension cognitive dans la créativité, car elle influence de manière significative l'originalité et le succès. Nous montrons que la diversité cognitive semble toujours bénéfique pour combiner des connaissances plus éloignées. En revanche, le profil exploratoire moyen au sein de l'équipe suit une relation en forme de U inversé avec la nouveauté combinatoire (c'est-à-dire qu'il existe un point d'inflexion où cela n'est plus bénéfique). La même relation peut être trouvée en examinant l'impact en termes de citations. Cependant, notre étude met en évidence le lien étroit entre la dimension cognitive et la nature de ces citations. Plus précisément, les équipes composées d'individus *exploitatifs* ont tendance à consolider la science, tandis que celles composées de profils plus *exploratifs*, lorsqu'ils sont associés à des profils *exploitatifs*, la "disruptent". En résumé, notre recherche souligne l'importance de la composition des équipes en termes de profil cognitif pour la créativité scientifique.

Nous montrons que la présence conjointe d'individus hautement exploratifs et exploitatifs constitue la composition d'équipe la plus efficace pour perturber la science ; cependant, un nombre limité d'individus hautement exploratifs est essentiel pour maximiser la pertinence des combinaisons de connaissances créées.

Méthodologie

Cette thèse repose en grande partie sur des analyses quantitatives de la science, employant des méthodes issues de la bibliométrie et de la scientométrie. Ainsi, nous utilisons des bases de données massives représentées sous forme de graphes, reliant les entités scientifiques entre elles, telles que les auteurs, les articles, les institutions, etc. Un bref historique de ces domaines s'impose.

Le terme bibliométrie a été défini pour la première fois par le Belge Paul Otlet en 1934 [Otlet, 1934] et réintroduit dans sa version anglaise par Pritchard et al. [1969] dans l'article "Statistical Bibliography or Bibliometrics ?" L'objectif initial de la discipline était d'améliorer la classification et l'organisation des livres pour contenir l'afflux de connaissances. Plus précisément, les bibliothécaires l'utilisaient pour sélectionner les articles pertinents pour leurs collections [Sugimoto and Larivière, 2018]. En 1955, le chimiste et documentaliste Eugene K. Garfield proposa la création d'un index des citations pour offrir un outil d'analyse en étudiant les liens entre différents documents scientifiques. L'Institute for Scientific Information, fondé par Garfield en 1960, a développé le Science Citation Index (SCI), lancé pour la première fois en 1963 à l'intention des chercheurs et des bibliothécaires. Toutefois, les études bibliométriques ont débuté dans les années 1960, l'une des figures centrales étant Derek J. de Solla Price [de Solla Price, 1965, Boyack et al., 2005]. À l'époque, l'objectif initial était de comprendre la recherche en tant que système en examinant la croissance des publications et les grandes lignes de l'activité de citation. Très tôt, le nombre de références, la densité du nombre de citations pour les articles et l'inégalité dans le processus de citation ont fait l'objet d'un intérêt particulier. Les problèmes fondamentaux liés à la détection des connaissances importantes et à la progression du système scientifique restent au cœur de la recherche en science des sciences [Fortunato et al., 2018]. Le nombre de bases de données utilisées par les chercheurs a augmenté au cours des dernières décennies. De nouvelles structures telles que les graphes de connaissances (KG) sont apparues[7]. Bien que le nom KG ait été utilisé pour la première fois dans Schneider [1973], il n'a été popularisé qu'en 2012 lorsque Google a présenté son propre KG. Ils permettent de faire des correspondances entre différentes entités liées un article, une tâche cruciale en scientométrie. Ces unités peuvent être un profil d'auteur, une revue, une référence ou même des sujets. Ces entités nous aident à mieux comprendre la manière dont la science est structurée et réalisée.

La disponibilité croissante des données a entraîné une diversification du domaine de la scientométrie. Les chercheurs utilisent souvent de manière interchangeable des termes tels que bibliométrie, scientométrie, informétrie, webométrie et altmétrie. Une

[7] Par exemple, Microsoft Academic Graph (MAG, qui a été remplacé par OpenAlex en 2022) ou bien PubMed Knowledge Graph (PKG)

littérature abondante retrace l'histoire de ces domaines et cherche à comprendre leurs différences, ainsi qu'à créer des taxonomies intra-champ. Ils partagent tous un objectif commun : étudier la science en tant que système, en utilisant des méthodes scientifiques. L'informétrie est une sous-discipline des sciences de l'information et se définit comme l'application de méthodes mathématiques au contenu des sciences de l'information [Chellappandi and Vijayakumar, 2018]. En d'autres termes, l'informétrie est le plus haut niveau d'abstraction, et tous les autres domaines sont un sous-ensemble de l'informétrie. La bibliométrie, comme nous l'avons vu plus haut, se concentre sur les citations pour la gestion des collections et la recherche de documents en utilisant des aspects spécifiques du document sans le placer dans le contexte global dans lequel il a été créé. En revanche, la scientométrie est une "méta-science" qui analyse quantitativement la production, la diffusion et les mécanismes du système sous-jacent [Sugimoto and Larivière, 2018, Chellappandi and Vijayakumar, 2018]. Enfin, Webometrics et Altmetrics s'intéressent tous deux aux informations disponibles sur le web, mais Webometrics est centré sur le document, le document étant une page web, tandis qu'Altmetrics est centré sur l'aspect réseau et complète les citations par le nombre de likes et de retweets [Mingers and Leydesdorff, 2015].

Cette thèse est essentiellement basée sur des méthodes issues de la scientométrie. Mais la scientométrie elle-même peut être divisée en deux macro-types d'analyse : la performance et l'analyse de la cartographie des sciences (SMA) [Moral Muñoz et al., 2020]. L'objectif de cette première est d'évaluer les activités des acteurs scientifiques et leur impact. Son but est donc d'attribuer une valeur à la productivité et à l'étendue de la recherche menée par une unité (article, auteur, institution). SMA "*is mostly directed at monitoring a scientific field to determine its (cognitive) structure, its evolution, and main actors within*" [Noyons et al., 1999] ; elle prend un cliché d'une partie du système scientifique à un moment donné pour analyser sa structure.

Les intrants, les extrants et les impacts de ces activités scientifiques sont les trois perspectives utilisées dans l'analyse de performance et la SMA [Sugimoto and Larivière, 2018]. L'input fait référence aux ressources humaines et financières et capture les différentes interactions des agents du système à différents niveaux (niveaux Auteur/Institutionnel/Pays) ; l'output est le résultat final du processus de recherche, c'est-à-dire la composition de ce document et les différentes entités qui le caractérisent ; et enfin, les mesures d'impact étudient les répercussions des outputs, la diffusion des connaissances qu'un article génère à travers les citations, l'attention du grand public, ou la réutilisation des composantes du document.

Cette thèse aborde ces trois dimensions dans le cadre de l'étude de l'intelligence artificielle en science. En effet, dans le chapitre 1, comme indiqué dans l'introduction, nous étudions les extrants et l'impact scientifique des publications utilisant l'IA. Dans le chapitre 2, nous étudions les intrants de la recherche en IA. Les chapitres 3 et 4 visent enfin à comprendre les relations entre les intrants au niveau de l'auteur et les extrants de la recherche menée et leur impact scientifique.

Chapitre 1

Artificial Intelligence in Science : An Emerging General Method of Invention

This chapter was co-authored with

Stefano BIANCHINI and Moritz MÜLLER

Summary of the chapter

This paper offers insights into the diffusion and impact of artificial intelligence in science. More specifically, we show that neural network-based technology meets the essential properties of emerging technologies in the scientific realm. It is novel, because it shows discontinuous innovations in the originating domain and is put to new uses in many application domains; it is quick growing, its dimensions being subject to rapid change; it is coherent, because it detaches from its technological parents, and integrates and is accepted in different scientific communities; and it has a prominent impact on scientific discovery, but a high degree of uncertainty and ambiguity associated with this impact. Our findings suggest that intelligent machines diffuse in the sciences, reshape the nature of the discovery process and potentially affect the organization of science. We propose a new conceptual framework that considers artificial intelligence as an *emerging general method of invention* and, on this basis, derive its policy implications.

Publication: This chapter was published in a similar version as:

BIANCHINI, Stefano, MÜLLER, Moritz, and PELLETIER, Pierre. Artificial intelligence in science: An emerging general method of invention. Research Policy, 2022, vol. 51, no 10, p. 104604. DOI:10.1016/j.respol.2022.104604

Acknowledgement. The research leading to the results of this chapter has received financial support from the CNRS through the MITI interdisciplinary programs [reference: Artificial intelligence in the science system (ARISE)] and the French National Research Agency [reference: DInnAMICS -ANR-18-CE26-0017-01]. Earlier versions of this paper were presented at the INNOVA MEASURE III Expert Workshop 'Brainstorming in Ispra'; 7th European Conference on Corporate R&D and Innovation, Seville (Spain); EMAEE 2019 'Economics, Governance and Management of AI, Robots and Digital Transformations', Brighton (UK); OLKC 2019 'The Human Side of Innovation – Understanding the Role of Interpersonal Relations in an Increasingly Digitised Workplace'; CAGE-NESTA Workshop on 'Data Science for the Economics of Science, Technology and Innovation', London (UK); and the Workshop Series on 'The Economics and Management of AI Technologies', Copenhagen (Denmark) and Strasbourg (France). The authors thank seminar participants and in particular Tommaso Ciarli, Giacomo Damioli, Mirko Draca, Daniel S. Hain, Björn Jindra, Bertrand Koebel, Roman Jurowetzki, Patrick Llerena, Juan Mateos-Garcia, Maria Savona, Simone Vannuccini, Daniel Vertesy, and Marco Vivarelli. We also are very grateful to two anonymous referees for their constructive criticism.

1.1 Introduction

Measurable research outputs such as papers, patents, and innovations have been subject to high enduring growth rates over the last century. Yet, recent empirical evidence suggests that research productivity is ever falling and new ideas are becoming increasingly harder to find [Gordon, 2016, Bloom et al., 2020]. A common narrative for this decline in productivity rests on the so-called 'knowledge burden'. Over the past few decades, data and information have begun to grow and accumulate on an unprecedented scale, and searching through an increasingly vast and complex knowledge space has become prohibitively expensive [Weitzman, 1998, Fleming, 2001, Jones, 2009].

Recent advances in artificial intelligence (AI) – in particular the rapid improvements in prediction achieved by (multi-layer) neural networks (NN) – have brought a wave of optimism that these technologies will speed up scientific discovery [Hey et al., 2009, Agrawal et al., 2018, Cockburn et al., 2018]. NN-based models have been found to be particularly good for discovering representations, invariances, and laws, that is, unusual and interesting patterns that are hidden in high-dimensional data [LeCun et al., 2015, Schmidhuber, 2015, Goodfellow et al., 2016]. In other words, NNs have shown themselves to be particularly suited to addressing scientific problems.

The first question we raise in this article is whether NNs are, in fact, diffusing into the sciences and, if so, what the mechanics of this diffusion process might be. In so doing, we consider five key attributes that allow a technology to be defined as 'emerging' – namely: (i) radical novelty, (ii) fast growth, (iii) coherence, (iv) prominent impact, and (v) uncertainty and ambiguity [Rotolo et al., 2015] – and show that NNs conform to these properties.[1]

The second question we address is how NNs influence scientific discovery. Machines are becoming much more than mere scientific instruments, and might even be described as teammates. Today, intelligent machines can engage in various stages of a (complex) problem-solving process. They can, for example, define the problem(s), identify root causes, propose and evaluate solutions, choose between different op-

[1] Rotolo et al. [2015] conceive of an emerging technology as "*[a] radically novel and relatively fast growing technology characterized by a certain degree of coherence persisting over time and with the potential to exert a considerable impact on the socio-economic domain(s) which is observed in terms of the composition of actors, institutions and patterns of interactions among those, along with the associated knowledge production processes. Its most prominent impact, however, lies in the future and so in the emergence phase is still somewhat uncertain and ambiguous*" (p.1828).

tions, make plans, take actions, and learn from interactions [Seeber et al., 2020]. AI and, in particular, multi-layer NNs have been qualified as a *general-purpose invention in the method of invention* [Cockburn et al., 2018], a conceptual framework that blends the concepts of the method of invention (MI) [Griliches, 1957] and general-purpose technology (GPT)[Bresnahan and Trajtenberg, 1995]. Building on this idea, Agrawal et al. [2018] suggest that NN-based prediction machines can alter the knowledge production function in combinatorial-type research problems by affecting two dimensions: those of 'search' and 'discovery'. NN 'search' methods would support knowledge access by making existing relevant knowledge available to the researcher, whereas NN 'discovery' methods would help identify valuable combinations among elements of that available knowledge. Thus, in a needle-in-a-haystack problem, the 'search' dimension would arrange the haystack and the 'discovery' dimension would find the needle.

This distinction between 'search' and 'discovery' is conceptually interesting. Yet, it tells us little about how AI influences the direction of knowledge development, because it only deals with one body (or one haystack to stick with the analogy) of pre-existing elements of knowledge. However, there are two sides to the knowledge explosion: increasing knowledge within each domain (i.e., larger haystacks) and an increasing number of domains (i.e., more haystacks). A priori, AI can either help scientists explore familiar conceptual spaces – structured styles of thought – in depth or transform the space by making unfamiliar combinations of distant knowledge elements [Boden, 2004, 2009]. The fundamental question, then, is whether AI is currently being used to cope with the knowledge explosion *within* a domain or to facilitate knowledge creation *across* domains – that is, in-depth exploration of a known domain *vis-à-vis* the transformation of the domain through knowledge recombination across other domains.

Hence, we are interested in investigating empirically how NN methods contribute to science in terms of *recombinatorial novelty* and *impact*, an analysis confined here to the *health sciences*. In this study, the concept of recombinatorial novelty refers to novel recombinations across domains, as proxied by scientific journals, whereas the concept of impact refers to the relative importance of a study in the scientific community, as proxied by citation indices. We find that NN adoption is negatively associated with recombinatorial novelty, suggesting that researchers are using NNs as a research tool primarily to cope with the knowledge explosion within domains rather than across domains. Interestingly, our results also reveal a considerable degree of uncertainty as regards impact, reflected by a high variation in citation performance. We suggest that this outcome is consistent with the intrinsic nature of emerging technologies, but also with a sort of 'mode effect' whereby '*everyone wants to be AI and data savvy, but few are ready*'.

The rest of this paper is structured as follows. Section 2 discusses the emergence of the new data-intensive scientific paradigm; Section 3 presents the method for identifying NN-related research and our sample construction; Section 4 documents aspects of the NN diffusion process in the sciences; Section 5 presents our analysis of the contribution of NN methods to the health sciences; and, the final section concludes by identifying a number of areas that might benefit from policy considerations.

1.2 Data-intensive scientific discovery

"*Few fields are untouched by the machine-learning revolution, from materials science to drug exploration; quantum physics to medicine.*" [Nature-Editorial, 2019]

Historically, the process of scientific inquiry has evolved through paradigms, seen as symbolic generalizations, metaphysical commitments, values and exemplars that are shared by a community of scientists and that guide the research of that community [Kuhn, 1962].

For most of human history, scientists have been observing phenomena, postulating laws or principles to generalize the complexity of their observations into simpler concepts – i.e., compressed, elegant mathematical representations that offer insights into the functioning of the universe. Originally there were just two sciences, the *experimental* and the *theoretical*. Indeed, Hey et al. [2009] identify empirical observation and logical (theory) formulation as the first and second scientific paradigms, respectively. Towards the middle of the last century, however, many problems proved too complicated to be solved analytically and researchers had to start simulating. Science entered a third paradigm, one characterized by the development of *computational models* and *simulations* to understand complex phenomena. As the knowledge frontier expands and the landscape gets more complex, it is becoming harder and harder for researchers to know enough to find (useful) combinations of knowledge that produce new (valuable) ideas.

Ongoing developments in AI, especially the impressive achievements made using

NN techniques, have led to mounting pressure to shift from hypothesis-driven to *data-driven scientific discovery*. The emerging scientific paradigm is being built on data-intensive computing with the massive deployment of intelligent machines capable of finding representations, rules, and patterns in an ever-increasing volume of structured and unstructured data [King et al., 2009, Hey et al., 2009, Nature-Editorial, 2019]. Even today, Francis Bacon's basic insight continues to hold: the scientists' job is to search for regularities in the empirical data. Bacon probably could not have foreseen that this search is best achieved today with the support of AI.

What makes NNs particularly powerful is the learning process, that is, they learn from past experience and understand the world in terms of a hierarchy of concepts, where each concept is defined by the way it relates to simpler concepts [Schmidhuber, 2015, Goodfellow et al., 2016]. It is clear that the term 'artificial neural networks' has been coined by analogy with biological neural networks, complete with their neurons, connections and firings. In a general NN model, the variables observed in the data are presented to an input or visible layer composed of several nodes; then a series of hidden layers (also composed of nodes) extracts increasingly abstract features from the data. The term 'hidden' stresses the idea that there is no predetermined structure; rather, it is the model itself that learns which concepts are useful to explain the relationships observed in the data. The nodes in the input, the hidden and output layers are all vaguely similar to biological neurons, and the connections between these nodes can be thought of as reflecting the connections between neurons [Hassabis et al., 2017].

NN-based methods can be applied in scientific settings in a variety of ways (see, e.g., Raghu and Schmidt [2020]). The most common application is to use NNs to tackle complex *prediction problems* – i.e., mapping inputs to predicted outputs. By way of example, the input might be an MRI image and the machine has to output a prediction of whether there are any signs of cancer. A second common application is to obtain interpretable insights into which property of the data led to the observed prediction – that is, *from prediction to understanding*. For example, some tools can be used to analyse the hidden representations of a neural network and detect which features of the input are most critical. A third application is to perform complex *transformations of input data*, such as image super-resolution and data compression, which in turn make data analysis easier and save space. Other recent tools, although in their infancy, would help scientists write better papers and co-write codes.

It is clear that intelligent machines can help shoulder the 'knowledge burden' within a scientific domain, act as a fertilizer of knowledge recombination across domains, and thus enrich and transform the knowledge space. In short, intelligent machines can influence both 'search' and 'discovery' processes.

In the case of the 'search' process, NNs can support access to knowledge by predicting which elements of knowledge and information are most relevant to the researcher. Three examples will serve to illustrate this function. First, NN-based recommender systems can offer high quality cross-domain recommendations by exploiting numeric measurements, images, text and interactions in a unified joint framework [Zhang et al., 2019]. Second, transformational learning can improve learning tasks in one domain by using knowledge transferred from other (related) domains, and in turn capture generalizations and differences across domains [Olier et al., 2021]. And, third, AI can be used for fact-checking, that is, assessing the veracity of scientific claims in sensitive areas such as climate change or Covid-19 pandemic [Wadden et al., 2020].

In the case of the 'discovery' process, NNs provide a better prediction of which elements of knowledge can be combined to produce new knowledge and of the value of that knowledge. Literature-based discovery, for example, is a way to understand implicit (hidden) associations from existing studies, which can result in interesting, surprising, non-trivial hypotheses that are worth studying. Other NN-based tools, such as machine reading comprehension systems, can propose variations on an experiment after having identified gaps in the literature [Baradaran et al., 2022]. Highly efficient forms of deep active learning have also been developed that can reduce the uncertainty associated with those regions of the experiment space that are sparsely populated with results [Daugherty and Wilson, 2018].

A major consequence of considering AI as a research tool – indeed, as a teammate – is that its impact is not limited to its ability to reduce the costs of specific scientific activities, but that it can facilitate a new approach to science itself, by modifying the scientific paradigm in the domains where the new research tool is deployed. Exploring the emergence of NN-based technology in science and its impact on scientific discovery is at the core of our study.

1.3 Identifying neural network research

Our empirical analysis of scientific publications exploits two databases: arXiv.org and Web of Science (WoS). First, we use arXiv.org to draw up an appropriate list of *search terms* referring to NNs based on the natural language processing of scientific abstracts from publications in the subject areas of ‘Computer Science’, ‘Mathematics’, and ‘Statistics’. Second, these search terms are used to query the WoS database and to extract a sample of NN papers across all scientific fields.

Reliance on a list of search terms for document retrieval is a common practice in research on emerging technologies and science in general. Unfortunately, extant studies do not provide us with an authoritative ‘ready-to-use’ list of search terms. Here, we train the word embedding model Word2Vec [Mikolov et al., 2013] with scientific abstracts from arXiv.org’s documents in order to *learn* NN-related terms.

Our training sample consists of scientific abstracts from arXiv.org. AI research tends to be a blend of statistics and informatics, but is developed in the main within the computer sciences. Informatics is a fast-developing field in which conference proceedings traditionally play an important role. More recently, however, the rapid dissemination of research is (best) achieved via open access journals and platforms. Of these, arXiv.org is the most prominent and provides us with a rich corpus for the identification of NN-related terms. We downloaded a total of 197,439 abstracts of papers from the subject areas of ‘Computer Science’, ‘Mathematics’ and ‘Statistics’, for the period 1990–2018. The three areas represented roughly 50% of all arXiv.org documents in 2018, and just 10% in the early 2000s.

Once pre-processed (details in Supplementary Material), the corpus was used to train the Word2Vec model in its skip-gram with negative sampling version. The main outcome of this model is one vector representation for each term in the vocabulary. Hence, we were able to identify the terms that appear in the same cluster as the term ‘neural network’. The resulting list of potential search terms included individual words (uni-grams) as well as technical terms consisting of multiple words (n-grams). We opted to retain only those terms consisting of multiple words – i.e., we removed all uni-grams – in order to err on the side of conservativism and to ensure only the inclusion of terms that relate unambiguously to NNs. Moreover, we retained only the 30 most frequent n-grams after eliminating terms considered as being too generic (e.g., ‘short term’ or ‘supervised learning’). The final list of search terms used in our study is shown in Table 1.

A more complete list of terms for all clusters identified by word embedding can

Table 1.1: NN-related search terms from word embedding

n-gram	Count	n-gram	Count
neural network	402,996	long short term memory	3,122
neural networks	173,470	hidden layers	2,080
artificial neural	100,749	restricted boltzmann	1,635
artificial neural network	99,794	auto encoder	1,444
deep learning	24,104	generative adversarial	1,242
convolutional neural	20,742	encoder decoder	1,198
convolutional neural network	20,595	adversarial network	1,192
recurrent neural	14,355	generative adversarial network	1,085
recurrent neural network	13,965	fully convolutional network	688
deep neural	9,418	convolutional layers	568
multilayer perceptron	9,352	variational autoencoder	216
deep neural network	9,181	adversarial attacks	197
hidden layer	7,810	adversarial examples	92
deep convolutional	4,263	variational autoencoders	75
deep convolutional neural network	3,384	adversarial perturbations	24

Notes: The count refers to how many times a given term occurs in the Web of Science corpus. A document may (and very often does) include several terms. Adding more terms would only slightly change the number of documents retrieved from WoS, as can be seen from the counts of the last few terms.

be found in Appendix.

Our sample for subsequent analysis included all publications in the WoS Core Collection published between 1990 and 2018, and having at least one of the search terms (Table 1) in their title, keywords, or abstract. In total, we identified 260,459 documents (144,095 articles; 39,925 conference proceedings; 76,439 others).

1.4 Technology diffusion in the sciences

This Section documents the diffusion of NN-based methods in the sciences. We show that the diffusion dynamics and the characteristics of the technology largely conform to properties of emerging technologies.

(Relative) fast growth. One of the defining attributes of an emerging technology is the speed of its growth, which is evident in such dimensions as the number of actors involved, the funding made available, and the knowledge output produced.

Our data confirm a ‘burst of research activity’ in all scientific areas (Figure 1), although the volume (blue line) varied markedly. ‘Technology’ (Panel A) is the dominant field, which can be explained in part by the fact that it includes ‘Computer Science’, the main field of origin. It is followed, with about five times fewer papers, by ‘Physical Sciences’ (Panel B), which in turn is closely followed by ‘Life Sciences

Figure 1.1: Trends in NN publication activity by scientific area

Notes: The blue lines show the number of publications and the orange lines plot the growth rates in each scientific area. Growth rates are calculated as three-year moving averages and omitted publications before 2001. Scientific areas correspond to WoS research areas. Panel H refers to research published on arXiv.org, based on the sample discussed in Section 3.

& Biomedicine' (Panel C). Publications in 'Health Sciences' (Panel D) – defined as a subset of 'Life Sciences & Biomedicine' and the focus of the analysis conducted in the next Section – largely parallel those of 'Life Sciences'. Publication counts in 'Social Sciences' (Panel E) are relatively low, becoming negligible for 'Arts & Humanities' (Panel F). Panel G, which combines all WoS documents into one, shows that the (three-year average) growth rate (orange line) in NN publication activity around 2005 was high (at about 10%), it then suffered something of a decline in the years around 2010, before recovering and experiencing steady growth to the end of the observation period (reaching 20%). Indeed, the individual areas exhibited very similar growth patterns.[2]

[2]The overall number of NN-related documents varies according to the sub-disciplines within each scientific area (not shown here). The general trend in 'Technology' is driven mainly by 'Computer

Publication activity on arXiv.org (Panel H) follows essentially the same dynamics. Growth rates mimic the shape described above but are about five times higher than those in the WoS panels. The comparatively higher rates are attributable it would seem to the fact that open platforms are increasingly popular, given their efficiency and speed, as a channel of communication between researchers, particularly within the machine learning and computer science communities [Sutton and Gong, 2017].

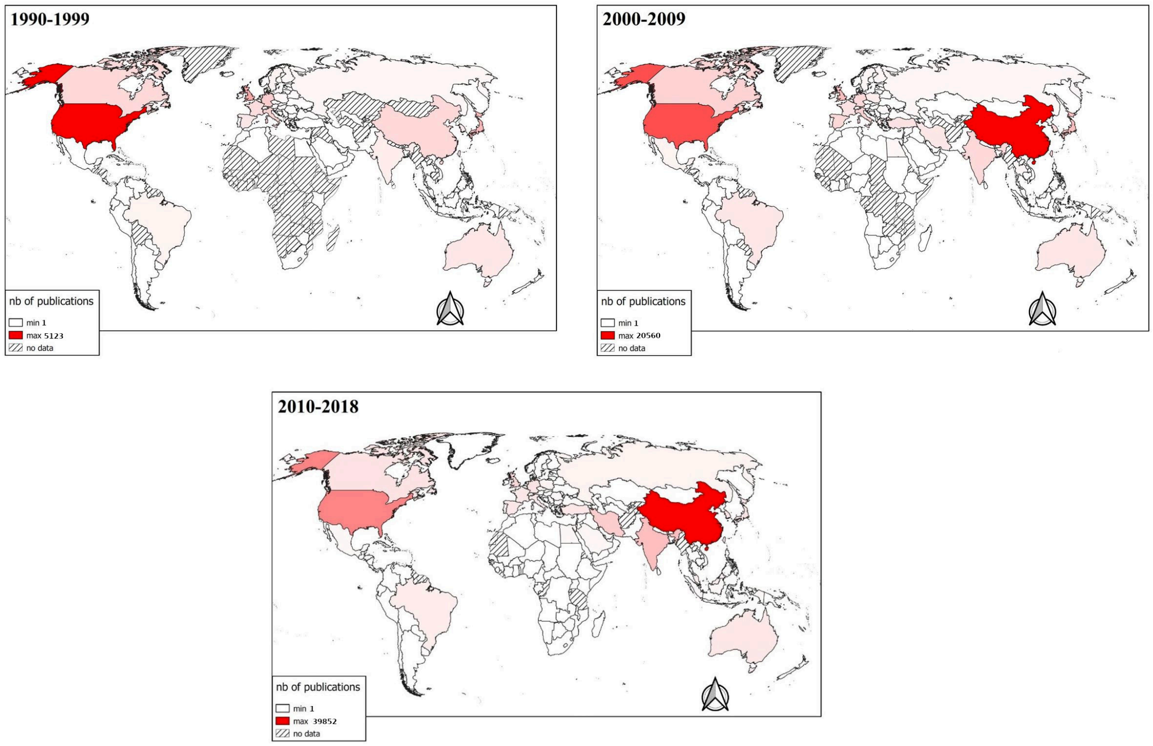
Research output increased not only in absolute numbers but also relative to the overall number of papers in a given scientific area, albeit at a lower level. In 2018, NN documents represented 2.6% of all papers in 'Technology', 1.02% in 'Physical Sciences', and 0.3% in 'Life Sciences & Biomedicine'. This means NN publications still account for only a tiny fraction of the whole research volume, in particular in application domains. However, recent growth rates in these shares are remarkable. NN-related research presents the highest growth rates in the 'Life Sciences & Biomedicine' (47.3% in the period 2017–2018), ranks second in 'Physical Sciences' (42%), and in 'Technology' presents a growth rate of roughly 18%.

Spatial diffusion and actor re-configuration. Another of the defining attributes of an emerging technology is the speed of change in the configuration of actors – e.g., countries, users, and scientists.

Figure 1.2 shows the dynamics of science at the country level. Each document is attributed to a given country when the affiliation of at least one of its authors is in that country. During the first period, 1990–1999, most of the documents (about 5,000) were published by scientists in the United States. Publishing activity was relatively low in absolute numbers in the European countries, Australia and China, and negligible or non-existent in most other countries. In the following decade, 2000–2009, China became the most prolific country with about 20,000 documents. The US ranked second with around 14,000 articles, whereas European countries, Australia, Canada, and India grew sufficiently to maintain their relative strength in the field. Interestingly, in this decade, NN research activity took off in an increasing number of countries. These trends were further reinforced in the last period, 2010–2018. Compared to the previous decade, China doubled its research output,

Science' (103,729 documents), 'Engineering' (95,638) and 'Automation & Control Systems' (24,721). In the case of 'Physical Sciences', it is driven by 'Physics' (7,239), 'Mathematics' (5,123) and 'Chemistry' (3,702), while in 'Life Sciences & Biomedicine', it is driven by 'Environmental Sciences & Ecology' (2,632), 'Neurosciences & Neurology' (2,032), and 'Biochemistry & Molecular Biology' (1,728).

Figure 1.2: Global diffusion of NN in science across countries



Notes: The intensity of colour reflects a country's relative number of NN publications in a given period, with no observed NN publication activity in hatched countries [WoS sample].

widening the gap with the US and, to a lesser extent, with the EU.

In summary, our data seem to suggest that NN research has diffused rapidly at the global scale, and that since the early stages of development there has been a re-configuration of global actors. We consistently observed high volatility in the rankings, with some countries climbing the ladder and others lagging behind.

Radical novelty and 'double-boom' cycle. NNs have experienced a discontinuous wave of major innovations, which points to the radical nature of this technology. (Artificial intelligence has a long, rich history dating back to the 1950s, when researchers from different domains began to explore various paths toward mechanizing intelligence – interested readers may consult Nilsson [2009] and Russell [2010]).

Novelty can also arise from putting the technology to a new use – that is, applying it from one domain to another [Adner and Levinthal, 2002]. The originating domain of NN research is predominately computer science; thus, it seems appropriate to follow Cockburn et al. [2018] and assume that NN publications in all areas other than computer science represent applications of NN methods to address field-specific research problems.

The diffusion of emerging technologies from the originating domain to the application domains typically follows a 'double-boom' cycle [Schmoch, 2007]. Initially, the new technology seems to be of high potential, and high expectations trigger considerable development efforts, especially theoretical – the first boom. However, during these early development activities, several actors discover the difficulties of translating theory into practice. Most fail and cease their innovation activities, putting an end to the first boom. But some continue and, as time passes, they overcome some of the more important practical hurdles and are able to demonstrate genuine advances – starting the second boom. Interestingly, this pattern is largely consistent with the growth patterns recorded in Figure 1 (orange lines), where the first boom, subsequent decline, and second boom are clearly evident.

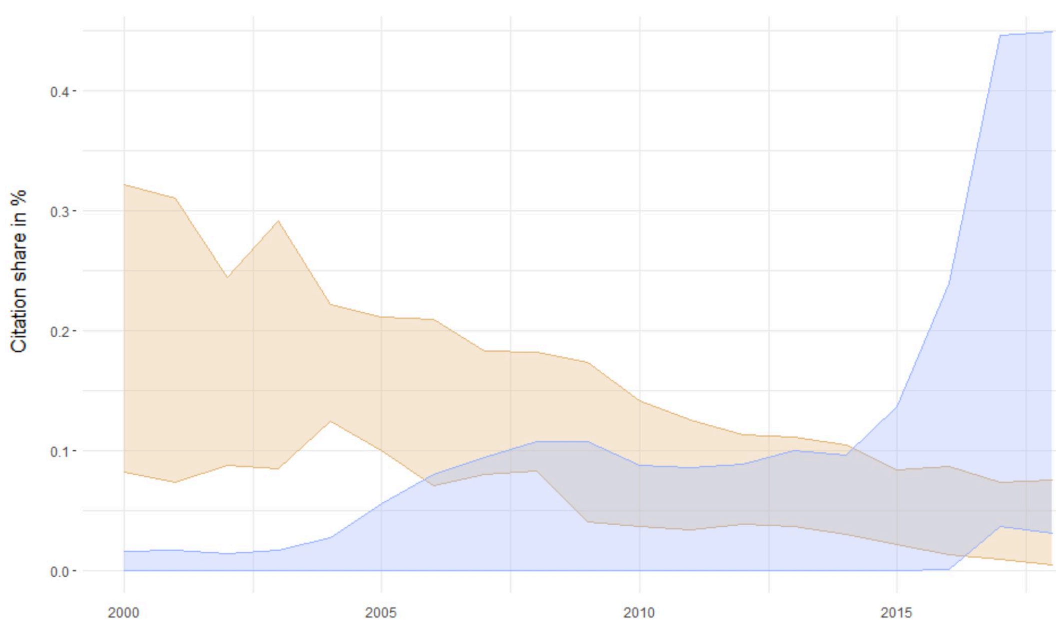
Table 1.2: Influential NN publications

Title | Journal	Cluster	# Citations	Share [%]
Multilayer feedforward networks are universal approximators | NN	1	5,904	0.14
Neural networks and physical systems with emergent ... | PNAS	1	4,658	0.11
Learning representations by back-propagating errors | Nature	1	4,645	0.11
Learning internal representations by error propagation | MIT Press	1	3,921	0.09
Approximation by superpositions of a sigmoidal function | MCSS	1	3,657	0.09
Training feedforward networks with the Marquardt algorithm | IEEE TNNLS	1	3,128	0.07
ANFIS: adaptive-network-based fuzzy inference system | IEEE SMC	1	2,909	0.07
Identification and control of dynamical systems using ... | IEEE TNNLS	1	2,551	0.06
Cellular neural networks: theory | IEEE CAS	1	2,267	0.05
ImageNet classification with deep convolutional neural networks | NeurIPS	2	7,177	0.17
Gradient-based learning applied to document recognition | IEEE Proceedings	2	3,590	0.09
Deep learning | Nature	2	3,542	0.08
Long short-term memory | NC	2	3,074	0.07
A fast learning algorithm for deep belief nets | NC	2	2,710	0.06
Reducing the dimensionality of data with neural networks | Science	2	2,621	0.06
Very deep convolutional networks for large-scale image recognition | arXiv	2	2,582	0.06
Particle swarm optimization | IEEE Proceedings ICNN	2	2,568	0.06
Deep residual learning for image recognition | IEEE Proceedings CVPR	2	2,160	0.05

Notes: This table reports the references (title and journal) of the most cited articles from the WoS publication sample over the period 2000–2018. From a total of 4,190,306 references (1,618,836 unique) cited by the documents in our sample, we selected the five most used references for each year. This gives us 18 time series that were clustered. Clustering is obtained via k-medoid and dynamic time warping. References within clusters ranked by total number of citations.

We also find that the second boom is marked by a shift in emphasis from theoretical principles to practical applications. In support of this evidence, we considered the top five cited references in each year of the observation period (i.e., those documents with the highest annual shares of all cited references in our publications), which gave us a list of 18 unique articles and their corresponding citation counts, as shown in Table 1.2. Using dynamic time warping (DTW) to measure dissimilarity between time series, we then clustered these temporal sequences by means of k-medoids [Berndt and Clifford, 1994]. As shown in Figure 1.3, we obtained two clusters. In the first period, the most cited articles in our sample were theoretical contributions, including a discussion of the possibility of using multilayer feedforward networks as universal function approximators, training algorithms (backprop), and parallel computing theories (cellular NN). In the second period, the most influential articles were no longer theoretical contributions, but rather articles that show how to put theoretical principles into practice. These contributions included inventions that have brought enormous performance gains to real-world tasks, especially for image and text analyses (e.g., deep convolutional neural networks and long short-term memory (LSTM) architectures).

Figure 1.3: Trends in annual citations of influential NN publications



Notes: This figure shows the annual share of all citations in the Web of Science sample for the two clusters of most cited NN articles. The shaded areas are time series intervals defined by minimum and maximum citation shares. In the main, the orange profile represents ‘theoretical’ contributions and the blue profile represents ‘applications’. Due to the limited number of articles that could be cited in the initial period, we clustered the time series from 2000 onwards.

<u>*Coherence*</u>. Another defining attribute of an emerging technology is its coherence, understood as the shared interpretation and acceptance of the technology within a

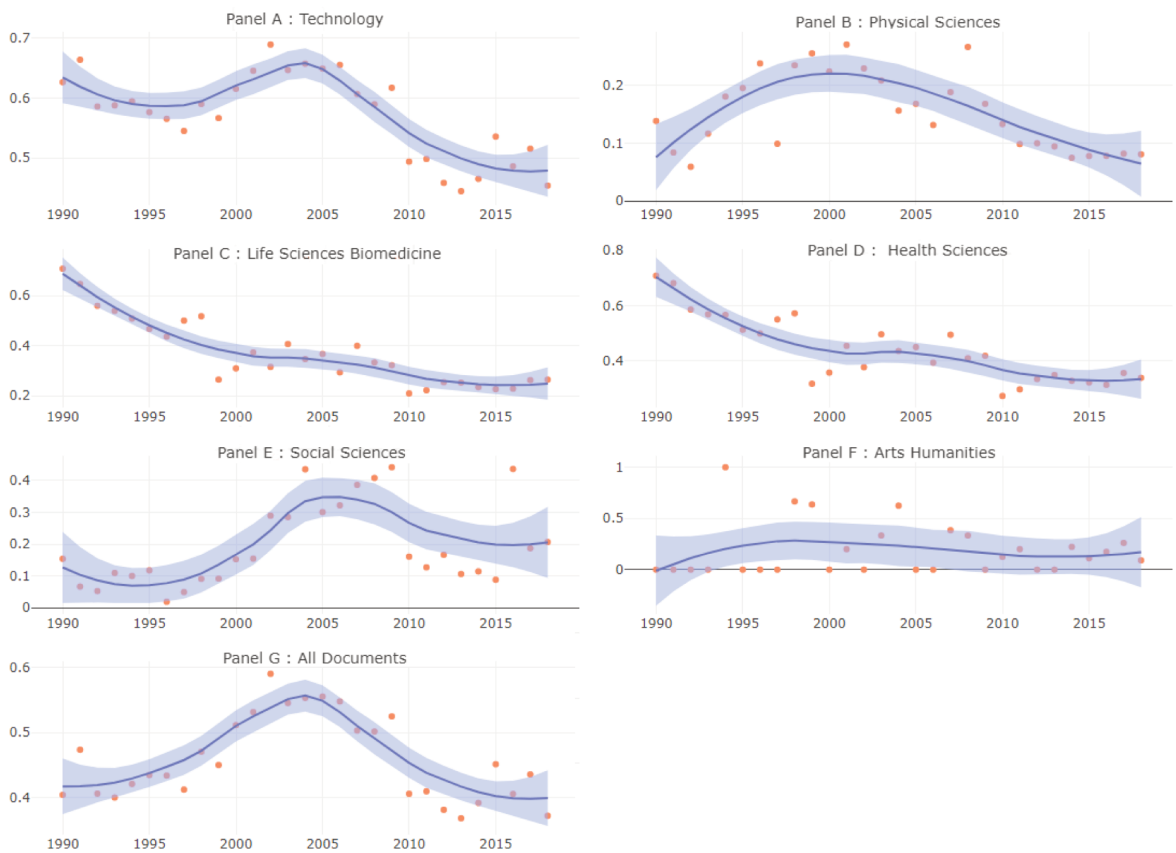
community. Signals of coherence can include the creation of dedicated conference sessions, new specialist journals and new categories in established classification systems. Here, we consider the transition from cross-disciplinary to disciplinary research effort as a sign of coherence, as this would mean that the technology has moved beyond its conceptual stage requiring close interaction between users and developers, and has become 'common practice' in application domains.

Each document is labelled by WoS as belonging to at least one subject category on the basis of the journal in which it was published. In most instances, a document falls into more than one category. The extent to which publications in a given scientific area are cross-classified as computer science contributions can therefore proxy cross-disciplinarity with respect to computer science. Thus, for each broad scientific area and year, we calculated the fraction of NN documents that are (also) labelled as 'Computer Science'.

Figure 1.4 shows the corresponding time trends. Each point of the plot for 'Technology' (Panel A) represents the average number of 'Technology' NN documents cross-classified as 'Computer Science' in a given year. For example, in 1990 about 60% of 'Technology' publications also fell into the 'Computer Science' category (first dot). The overall trend (blue line) follows a flat U-shape that reaches around 70% in 2005, before falling to less than 50% by the end of the observation period. Indeed, in 2018, a large proportion of papers in 'Technology' are no longer labelled as computer science contributions. 'Physical Sciences' (Panel B) also presents an inverse U-shape, with an increase in cross-classified computer science documents that reached 20% in 2000, before falling to 10% by the end of the period. No increase in computer science cross-classification was observed in 'Life Sciences & Biomedicine' (Panel C). From the very high share of 70% at the beginning of the period, a continuous decline was subsequently recorded (with significant drops around 2000 and again in 2010), finishing the period at around 20%. 'Health Sciences' (Panel D) presents the same evolution. 'Social Sciences' (Panel E) increased their share of computer science documents to 40% around 2010, but this was followed by a sharp downturn, while in 'Arts & Humanities' (Panel F), the share of computer science documents is very noisy, and no particular trend can be deciphered.

Taken together, these dynamics suggest that NNs diffuse *from* computer science, the originating discipline, *into* other application-oriented scientific disciplines. Thus, over time, we see a greater propensity of different communities to integrate the technology into their discipline, which is a good signal of coherence.

Figure 1.4: NN publications cross-classified as ‘Computer Science’



Notes: The figures show the fraction of NN documents cross-classified as ‘Computer Science’. Orange dots represent the share of cross-classified papers in each year. The blue curve corresponds to a simple local regression, with the surrounding shaded area representing the 95% confidence interval around the mean.

In short, it is more than apparent that NN technology fulfils many of the conditions to be classified as an emerging technology. It exhibits rapid growth in all domains; it has experienced a turbulent shift and reconfiguration of the actors involved in its development and adoption; and it presents a degree of coherence that persists over time. However, the picture arrived at in the first part of this analysis is incomplete. How does the technology influence scientific discovery in its domains of application? What can be learned about its undoubted impact yet, at the same time, the uncertainty that is often associated with its adoption? We address these questions in the next Section.

1.5 Neural networks in the health sciences

Here, we specifically address the impact of NN-based methods in the 'Health Sciences', one of the application domains with the highest short-term societal impacts [Raghupathi and Raghupathi, 2014, Miotto et al., 2018]. AI, in general, and deep learning, in particular, have already contributed to a variety of data-driven innovations in the health domain – improving healthcare systems, supporting clinicians, and monitoring patient diseases, among others. A review of the literature enabled us to identify applications in virtually all sub-disciplines: health informatics and biomedical research [Marx, 2013, Ravì et al., 2016], computational biology [Angermueller et al., 2016], genomic medicine [Leung et al., 2015], medical imaging [Litjens et al., 2017, Shen et al., 2017, Savadjiev et al., 2019], drug discovery and pharmacogenomics [Ma et al., 2015], real-time patient monitoring [Rajkomar et al., 2018], public health[Miotto et al., 2018, Zhang et al., 2018], and neuroscience and the cognitive sciences [Marblestone et al., 2016, Hassabis et al., 2017, Lake et al., 2017].[3]

Novelty and impact in science. A 'scientific contribution' is typically considered as comprising two elements: *novelty* and *impact*. Different terms for essentially this same idea were used in earlier studies of science, so that debates centred on discussions of the notions of originality, discovery and breakthrough and contributions to scientific progress [de Solla Price, 1963, Merton, 1957, Bourdieu, 1975]. It was Kuhn [1962] who coined the term 'novelty' to describe a more radical contribution that does not simply make an incremental advance in the 'normal science' in place, but rather breaks the current paradigm. More recently, the term novelty has partly lost this radical connotation, but it still carries the idea of a high degree of originality, while the concept of 'recombinatorial novelty' has emerged to highlight the idea that new knowledge arises out of the recombination of previously generated bits of knowledge [Fleming, 2001, Arthur, 2009, Uzzi et al., 2013, Wang et al., 2017].

Only a very small percentage of the potential for useful recombinations in the knowledge space is currently exploited. NNs can change the way science develops by helping to overcome our human limitations [Agrawal et al., 2018, Cockburn et al., 2018, Furman and Teodoridis, 2020]. Yet, how exactly does NN adoption correlate with novelty? The answer to this question depends very much on how the technology

[3] We define the 'Health Sciences' as comprising 83 Web of Science subject categories within the 'Life Sciences & Biomedicine' research area. The complete list of categories included can be consulted in the Supplementary Material.

is used in the scientific complex. Indeed, scientists can adopt new methods either to advance well-established research trajectories within a conceptual space or to explore new avenues by altering the conceptual space with knowledge from other domains, leading to low and high recombinatorial novelty, respectively.

The second element of a 'scientific contribution' concerns its impact, a key attribute of emerging technologies. Impact is related to, but different from, novelty; if research provides novelty, that novelty must be adopted by the scientific community in order for its impact to be felt. And, moreover, research can have an impact on subsequent research for reasons other than (recombinatorial) novelty, especially when providing new insights within established knowledge structures.

Yet, nor should impact be considered fully independent of novelty. Evidence suggests that a high degree of novelty is likely to increase the risk of delays and failures [Azoulay et al., 2011]. Moreover, novel research often requires more complex and risky collaborative social structures [Fleming et al., 2007, Foster et al., 2015]. Thus, highly novel research can be subject to considerable variations in 'quality' Fleming [2001], Wang et al. [2017] and, hence, to greater variations in impact. Uncertainty and ambiguity are common features of the research process, especially because the potential applications of the technology have yet to be explored and understood. Social inertia can further reinforce the uncertainty associated with impact. Emerging technologies typically encounter resistance in society precisely because they cause structural changes in roles and norms [Merton, 1957, Bourdieu, 1975]. This is particularly true of AI which operates at the intersection of ethical and legal considerations and, as such, is shaping the future of both individuals and society as a whole [Lanier, 2011, O'neil, 2017, Zuboff, 2019].

1.5.1 Empirical analysis

We measure scientific knowledge creation in scientific papers published in peer-reviewed journals and conference proceedings in the 'Health Sciences'. Henceforth, the term 'journal' is used to refer interchangeably to both peer-reviewed scientific journals and conference proceedings. We restrict our focus to journals that are not cross-classified as 'Computer Science' journals, ensuring that publications include NN methods as a research tool.

Our approach is to compare publications that involve NNs with those that do not involve NNs, while controlling for a set of confounding factors. Comparisons are made in terms of their recombinatorial novelty and scientific impact. For the main analysis, we operationalize the concept of 'recombinatorial novelty' as the first appearance of a knowledge combination, very much in line with Wang et al. [2017], the details of which we describe below. Novelty *à la* Wang et al. complies with the idea of NNs as a method of invention – i.e., a method for creating something new and valuable. In the case of 'scientific impact', we operationalize this concept as the subsequent use made of a paper, measured by the number of citations received.

Sample. We include all the articles for the whole observation period (2000–2018) published in those health journals where research involving NNs has been most prominent. This provides us with a relatively coherent knowledge base against which we can examine the concepts of novelty and impact. In total, we identified 26,461 NN health papers in about 5,000 health journals and proceedings. Roughly 45% (11,520) of these documents are published in the top 100 health journals in the sample. Hence, we downloaded the entirety of these journals for the period 1990–2018. Our final sample, combining NN and non-NN publications, contains 1,081,223 articles.

Variables. Our main explanatory variable is a binary indicator of a paper's NN content: 1 if the paper involves the use of NN methods, 0 otherwise. Our main dependent variables are (various measures of) recombinatorial novelty and scientific impact based on citation counts.

Recombinatorial novelty is measured in relation to the journals referenced by a paper. Thus, each paper is examined to determine whether it makes 'first-time-ever' combinations of referenced journals – i.e., its list of references contains journal pairs that have never previously appeared jointly in any list of references. In order to exclude journal pairs that simply formed once by happenstance, we further impose the condition that journal pairs be observed again within the next three years. A paper with at least one journal pair in the reference list that is both novel and that has been re-used, is considered as providing some novelty. Thus, we construct a binary indicator of novelty, henceforth referred to as *Novelty Dummy*. A further consideration is that a novel journal pair may span domains that vary in their distance one from another (i.e., more or less distant). This subtlety is captured through the co-citation profiles of the two journals forming a novel pair. The idea is that if both journals are often (rarely) cited with the same third journal(s), they are likely to span less (more) distant domains. In this way, we are able to construct a distance-weighted (continuous) measure of novelty, henceforth referred to as *Novelty*.

Calculations of the binary and weighted novelty measures follow Wang et al. [2017]. However, our procedure differs in two major respects. First, we judge novelty and co-citation distance only on journal pairs that are observed in the reference lists of our sampled papers. Thus, we do not measure novelty *per se* but rather with respect to a knowledge base covered by the sampled health journals.

Second, we calculate different measures of novelty by considering different sets of journals in the references. In this way, we are able to capture the *source* of novelty – i.e., where does this novelty come from? ICT, health, or other domains? While it is true that all the articles in our sample are published in outlets of the 'Health Sciences', they can reference journals in various domains. For instance, a health science paper involving NNs is likely to cite computer science journals where the NN methods were first published. This translates into a recombinatorial novelty 'simply' because of the adoption of the method, but it does not necessarily reflect the recombinatorial potential of NNs to connect and recombine knowledge in complex knowledge landscapes. In other words, we seek to measure whether NN adoption fosters novel recombinations *within* the health sciences and/or *between* the health sciences and disciplines other than the computer sciences. Thus, we calculate novelty not only in journal pairs, as indicated by 'All Sciences', but also limited to journal pairs where (i) no referenced journal is classified as a computer science journal, indicated by 'No CS'; and (ii) both referenced journals are uniquely classified as health sciences, indicated by 'Only HS'. By way of example, the combination of 'Biology & Biochemistry' and 'Computer Science' journals can be regarded as an 'All Sciences' combination; 'Engineering' and 'Molecular Biology & Genetics' as a 'No CS' combination; and 'Neuroscience & Behaviour' and 'Psychiatry/Psychology' as an intra-domain 'Only HS' combination.

Combining these three recombinatorial options with the possibility of calculating novelty as either a binary indicator or a continuous score, we obtain six different novelty measures, namely: *Novelty Dummy (All Sciences)*, *Novelty Dummy (No CS)*, *Novelty Dummy (Only HS)*, *Novelty (All Sciences)*, *Novelty (No CS)*, and *Novelty (Only HS)*.

Impact is measured by the number of citations (*# Citations*) received by a paper from its year of publication up to 2019, the time of data extraction. Furthermore, we code dummy indicators for so-called 'big hit' contributions – i.e., highly cited papers. Whether a paper is among the top 5 or 10% cited papers (*Top 5% Cited* and *Top 10% Cited*) is calculated with reference to other papers published in the same year and falling in the same WoS subject category.

We consider a set of control variables to capture various characteristics of a focal paper. We control for the number of references made by a paper (*# References*) as this might automatically increase the likelihood of its having new combinations. In prior research, the number of authors has been shown to be positively associated with both novelty and impact, hence we control for that (*# Authors*). The adoption of AI in scientific settings can indeed have an ambiguous effect on team size. Size may increase as new members are needed to manage the technology (at least in the early stages), but the technology may also automatize some tasks, thereby generating a replacement effect in the scientific workforce. International collaborations may also be a source of novelty and impact, and may be instrumental in the adoption of the technology. We proxy international collaboration by a dummy (*International Collab.*) taking a value of 1 if there are at least two different countries in the authors' affiliations, 0 otherwise. For the same reason, we construct a dummy for private sector participation (*Private Partic.*) taking a value of 1 if the paper has at least one non-university affiliation in the list. We consider the journal impact factor (*JIF*), since, on the one hand, high impact journals may be biased against novelty, but, on the other, increase visibility and hence citations. We additionally control for the journal age (*Journal Age*). Finally, we include a dummy indicating whether the paper provides a review or survey of extant literature (*Survey*). A survey may in fact cover separate streams of research without really connecting them.[4] Descriptive statistics of the variables are reported in Appendix.

Estimation methods. We model three different types of outcome: (i) binary indicators of novelty and impact, (ii) positive continuous measures of novelty, and (iii) positive discrete measures of impact (number of received citations). Each type of outcome requires a specific econometric setting.

All binary indicators are modelled with a Probit. Our continuous novelty measure is censored at zero, hence we use a Tobit model. Citations are count data for which the Poisson and negative binomial models are natural candidates. Over-dispersion and the conditional mean of the outcome variable being much lower than its vari-

[4] *Private Partic.* takes a value of 1 if we detect in the authors' affiliation at least one of the acronyms present in the Wikipedia page: 'List of legal entity types by country'. We use the SCImago Journal Rank to obtain the impact factor (*JIF*) for each journal in each year. *Journal Age* is calculated as the time elapsed from the date of the journal's creation to the year of publication. *Survey* takes a value of 1 if we detect in the title of the paper the terms 'Survey', 'Overview' or 'Review'.

ance are the most common arguments for favouring the negative binomial over the Poisson model. In our case, both empirical arguments hold; therefore, we opted for the negative binomial to model mean and dispersion separately, each with a linear predictor incorporating our main left-hand side variables and controls.

In all estimations, we include the control variables discussed above and a set of dummies to control for scientific field and cohort effects. We proxy scientific field using WoS categories (field WC). As a paper may fall into several categories, we code dummy variables taking a value of 1 for each category. Throughout the analysis, robust standard errors clustered at the journal-level are obtained via bootstrapping all journals.

Results. Table 1.3, Columns 1–3, shows the Tobit regressions of the continuous measures of novelty, *Novelty*. Columns 4–6 report the Probit estimates of the binary novelty indicators, *Novelty Dummy*.

When considering recombinatorial novelty across all sciences (Column 1), the estimated coefficient is positive but non-significant, but when we exclude computer science references (Column 2) the coefficient becomes negative yet remains non-significant. Restricting references to health sciences only (Column 3) increases the negative coefficient, which is now significant below the 1% significance level. The same pattern is observed when we consider the results of the Probit regression of the novelty dummy.

To what extent does the adoption of NN methods change our expectations of recombinatorial novelty in the health sciences? To a considerable degree, given that adopting NN decreases the degree of novelty by 18.6%. In addition, the marginal effects of Probit (Column 6) tell us that, for the median observation, NN decreases by 0.031 the probability of an article being novel (0.037 for the average observation).

In sum, NN adoption is not significantly correlated with novel recombinations across the entire knowledge landscape, nor with novel recombinations involving anything other than computer sciences. Yet, it is significantly and negatively correlated with novel recombinations within the health sciences. These findings suggest that NN methods tend to be adopted as part of a 'balancing strategy' in which the risk associated with the (emerging) technology is counterbalanced by keeping the knowledge landscape stable. Another way of interpreting this outcome is that NNs are employed mainly as a research tool to support already formalized and well-defined research trajectories in the health sciences community. This evidence is consistent

Table 1.3: Novelty profile of NN publications

	Tobit: Novelty			*Probit: Novelty Dummy*		
	All Sciences	No CS	Only HS	All Sciences	No CS	Only HS
	(1)	(2)	(3)	(4)	(5)	(6)
NN	0.044	-0.031	-0.186***	0.053	-0.008	-0.150***
	(0.038)	(0.034)	(0.040)	(0.037)	(0.033)	(0.037)
# References (log)	1.046***	1.050***	1.029***	0.878***	0.879***	0.843***
	(0.033)	(0.033)	(0.033)	(0.026)	(0.026)	(0.023)
# Authors (log)	0.177***	0.184***	0.227***	0.184***	0.189***	0.223***
	(0.021)	(0.022)	(0.024)	(0.020)	(0.020)	(0.022)
International Collab.	-0.053***	-0.058***	-0.084***	-0.050***	-0.054***	-0.076***
	(0.010)	(0.010)	(0.010)	(0.009)	(0.010)	(0.009)
Private Partic.	-0.004	-0.004	-0.027*	-0.007	-0.008	-0.026**
	(0.012)	(0.012)	(0.014)	(0.012)	(0.013)	(0.013)
JIF	-0.026	-0.024	-0.017	-0.025	-0.024	-0.017
	(0.019)	(0.019)	(0.021)	(0.017)	(0.017)	(0.018)
Journal Age (log)	-0.098	-0.082	-0.044	-0.074	-0.061	-0.030
	(0.099)	(0.100)	(0.108)	(0.090)	(0.090)	(0.095)
Survey	0.225***	0.216***	0.181***	0.206***	0.199***	0.163***
	(0.049)	(0.047)	(0.050)	(0.049)	(0.047)	(0.046)
Log Likelihood	-263,098	-258,255	-221,241	-180,701	-178,639	-161,710
χ^2 [Null Model]	96,074***	94,950***	77,374.6***	75,936***	75,187***	64,730***
χ^2 [w/o NN Model]	4.90*	2.20	60.90***	6.70**	0.10	44.60***
# Obs	356,037	356,037	356,037	356,037	356,037	356,037

Notes: This table reports coefficients of the effect of NN methods (NN, dummy) on recombinatorial novelty built by considering different knowledge landscapes. Bootstrapped (500 replications) standard errors clustered at the journal-level in parentheses: ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively. The effect of NN on the positive continuous novelty measure is estimated using a Tobit regression (Columns 1–3). The effect on the novelty dummy is estimated using a Probit (Columns 4–6). Each novelty measure is calculated on three different sets of journal references: 'All Sciences' – All cited journals, 'No CS' – All cited journals except for computer science journals, and 'Only HS' – Only citations to health science journals. Constant term, scientific field (WoS subject category) and time fixed effects are incorporated in all model specifications. Likelihood-ratio tests are used to compare the goodness-of-fit of two statistical models: (i) null model against complete model; (ii) model without the NN variable against the complete model.

with the idea of extending science while maintaining the advantages of conventional domain-level thinking [Boden, 2004, Uzzi et al., 2013].

Our estimates of the control variables echo previous research. Larger teams are associated with more novelty [Fleming et al., 2007, Lee et al., 2015]; international col-

Table 1.4: Impact profile of NN publications

		NegBin: # Cit. (1)	Probit: Top 5% Cit. (2)	Probit: Top 10% Cit. (3)
Panel A: Mean	NN	0.101**	0.147***	0.155***
		(0.040)	(0.041)	(0.043)
	Novelty (All Sciences)	0.153***	0.200***	0.191***
		(0.023)	(0.016)	(0.015)
	# References (log)	0.491***	0.429***	0.477***
		(0.064)	(0.075)	(0.062)
	# Authors (log)	0.237***	0.166***	0.194***
		(0.026)	(0.039)	(0.036)
	International Collab.	0.064***	0.083***	0.085***
		(0.013)	(0.014)	(0.013)
	Private Partic.	-0.029*	-0.027	-0.034**
		(0.015)	(0.018)	(0.015)
	JIF	0.205***	0.167***	0.179***
		(0.022)	(0.017)	(0.018)
	Journal Age (log)	0.050	-0.066	-0.048
		(0.036)	(0.086)	(0.079)
	Survey	0.541***	0.667***	0.627***
		(0.060)	(0.054)	(0.049)
Panel B: Dispersion	NN	0.136***		
		(0.051)		
	Novelty (All Sciences)	0.093***		
		(0.017)		
	# References (log)	-0.496***		
		(0.038)		
	# Authors (log)	-0.213***		
		(0.044)		
	JIF	0.040		
		(0.031)		
	Journal Age (log)	-0.118***		
		(0.029)		
Log Likelihood		-1,519,720	-69,222	-110,788
χ^2 [Null Model]		318,463***	19,317***	31,564***
χ^2 [w/o NN Model]		8.70***	24.80***	40.00***
# Obs		356,037	356,037	356,037

Notes: This table reports coefficients of the effect of NN methods (NN, dummy) on scientific impact proxied by the number of citations received (Column 1) and 'big hits' (Columns 2 and 3). Bootstrapped (500 replications) standard errors clustered at the journal-level in parentheses: ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The effect of NN on the citation count is estimated using a negative binomial regression. Estimates for the expectation and variance are reported in Panels A and B, respectively. Effects on the binary indicators are estimated using a Probit. Constant term, scientific field (WoS subject category) and time fixed effects are incorporated in all model specifications. Likelihood-ratio tests are used to compare the goodness-of-fit of two statistical models: (i) null model against complete model; (ii) model without the NN variable against the complete model.

laborations are negatively associated with novelty [Wagner et al., 2019]; the chances of providing a new combination of journal references increase with the number of references [Wang et al., 2017]; and, literature reviews also tend to draw from a wider range of sources leading to novel combinations of references. We find a negative effect of private involvement and, finally, a journal's age and impact factor seem to play no role.

How does NN adoption correlate with impact? Table 1.4, Column 1, shows the results of the negative binomial regression of citation counts. Here, the mean and dispersion parameters may vary with various right-hand side factors.[5] We find that NN adoption positively and significantly affects the number of citations received, both in terms of expectation and variance. Compared to non-NN papers, *ceteris paribus*, NN papers receive on average 10.32% more citations. The expectation of citation count increases by a median of 6.01 for NN research. The dispersion of the citation distribution is 19.57% higher for NN papers than for non-NN papers.

The Probit regressions used to model the probability of a paper falling in the right tail (top 5% or 10%) of the year–field citation distribution corroborate the results. The marginal effects suggest that research involving NN has a 0.019 (median value) higher probability of being in the top 10% of the most influential contributions (0.027 mean value), and a 0.009 higher probability of being in the top 5% (0.014 mean value).

As for the controls, the number of authors is positively related to impact [Lee et al., 2015] and reduces impact variation; international collaborations increase citation expectations [Glänzel and Schubert, 2001]; publishing in a high impact factor journal further increases the average number of citations; surveys and other papers with many references tend to attract more citations; and, finally, a negative effect is found between private participation and scientific impact, albeit not particularly significant.

In sum, the econometric analysis shows that research using NN has a high potential for greater impact, on the one hand, but that it is also associated with greater uncertainty of having an impact, on the other. There are several (complementary) explanations for this uncertainty: the 'high-risk/high-gain' that characterizes the adoption of emerging technologies and breakthrough research [Rotolo et al., 2015, Wang et al., 2017]; the challenge of integrating the scientific instrument into existing

[5]We excluded dummy variables other than NN to model the dispersion of citations because these variables caused problems with the convergence of the maximum likelihood estimator. In modelling the dispersion, we also tried simpler specifications by progressively incorporating a few variables at a time.

scientific practices [Rosenberg, 1992]; the ability to extract the true potential from the instrument and not to adopt it simply because 'everybody does'; and the possible social resistance, especially in sensitive domains, as some areas of the health sciences are known to be.

Based on these results, we propose that AI – here, specifically, NN methods – be regarded as an *emerging general method of invention*: 'emerging' because it shares the key attributes of emerging technologies; 'general' because it is increasingly integrated as a research tool in many scientific domains; and, a 'method of invention' because it has great potential for impact in application domains. We consider it more appropriate to consider AI an emerging general method of invention as opposed to a general-purpose method of invention (as in Cockburn et al. [2018]) for two reasons. First, as we have seen in Section 4, although growing, the proportion of scientific contributions related to NNs remains marginal compared to the whole body of scientific activity. Second, whether or not AI can be classified as a general-purpose technology remains open to debate and we find more arguments to support the contention that AI is better considered, for example, as a large technical system with infrastructural properties [Vannuccini and Prytkova, 2021].

1.5.2 Robustness analysis

Our results are robust across a wide range of additional tests. Tables and further material can be found in Appendix and Supplementary Material.

First, we excluded all articles that fall into the WoS 'Neurosciences' category. This domain can be potentially problematic in that some terms (neural network, first and foremost) may not necessarily refer to artificial intelligence but rather to human intelligence and the biological brain. The sample falls by about 30% and the number of NN articles almost halves. However, our results are consistent when replicating the analysis on the sub-sample.

Second, we excluded all articles that contain the terms 'neural network' and 'neural networks' exclusively in their title, keywords, or abstract. Bear in mind that an article may still contain a term such as 'artificial neural network' or 'convolutional neural network' which should now refer to artificial intelligence *stricto sensu*. In this case, neuroscience papers may form part of the sample. This restriction is severe insofar as the number of NN articles falls by more than 70%. Yet our results are robust to this constraint.

The third exercise consists of a different econometric approach. Instead of regression analysis, we compared each NN paper with a 'twin' non-NN paper. More precisely, the empirical strategy considers the adoption of NN as a 'treatment'; hence, we employ exact matching and 1:1 nearest neighbour matching on propensity scores (PSM) to select an appropriate control group of untreated papers. Exact matching is performed considering Web of Science categories, publication year, and journal – that is, we compare a NN article in terms of novelty and impact with an article belonging to the same domain(s), published in the same year and in the same journal. We obtain the propensity scores associated with the binary treatment via the estimation of the Probit model containing the original set of variables. The average treatment effects (ATT) for the selected variables lend further support to our results.

A final test concerns the way novelty is measured. Indeed, some research shows that different novelty indicators are often inconsistent with each other and may return different sets of novel contributions [Fontana et al., 2020]. Thus, we implemented the indicator developed in Uzzi et al. [2013] to define an 'atypical' (novelty/conventionality) quadrant: high-conventionality/high-novelty (HC–HN); high-conventionality/low-novelty (HC–LN); low-conventionality/high-novelty (LC–HN); and low-conventionality/low-novelty (LC–LN). The four categories are employed in a multinomial logistic regression. We find that, within the knowledge landscape of the health sciences, NN articles are more likely to draw on highly conventional combinations of knowledge. *Ceteris paribus*, our estimates suggest that when NN methods inject some highly (field-specific) unusual combinations, they do so primarily in an exceptionally conventional knowledge space.

1.6 Concluding remarks

Most socio-economic analyses of AI have looked at the effects of technology on economic growth [Brynjolfsson and McAfee, 2014, Aghion et al., 2018], labour market and productivity dynamics [Furman and Seamans, 2019, Acemoglu and Restrepo, 2020, Van Roy et al., 2020], changes in skills [Graetz and Michaels, 2018, Brynjolfsson and Mitchell, 2017], and inequality and discrimination [O'neil, 2017, Zuboff, 2019]. Our contribution, here, provides insights into the diffusion and impact of AI methods in the scientific system.

In this paper, we first examined the diffusion of NN research in the sciences in an effort to verify whether NNs conform to certain characteristics of emerging technologies. We found that NN research activity has grown exponentially in almost all sciences and all over the world, and the diffusion process has followed a double-boom cycle with a strong re-configuration of global actors. The diffusion of NN methods into application domains began in a cross-disciplinary fashion involving the computer sciences, breaking their way into 'pure' field-specific research within the various application domains. We then examined the impact of technology adoption on scientific discovery, with a particular focus on the health sciences. We found the adoption of NN methods to be negatively correlated with recombinatorial novelty; however, a positive correlation was found with the expectation and dispersion of citations received, increasing a contribution's likelihood of becoming a 'big hit'.

Conceptually, we considered scientific discovery to be a recombinatorial process in which existing knowledge is recombined to create new knowledge, a process that continues perpetually in a dynamic knowledge landscape. A traditional image of science is one in which the knowledge landscape is made up of islands – i.e., (sub)-disciplines or scientific fields – where most of this recombination takes place. The islands reflect the structure of nature but also the need for a scientific mind to organize the complexity of the world. Seen this way, scientists are sailors whose goal it is to navigate from island to island, figure out their structure, and explore the surrounding landscape. Sailors can opt to stay in the 'comfort zone' and further their knowledge of one (or neighbouring) island(s), or they can sail to more distant islands and connect new areas of the landscape. Both actions enrich the knowledge space, one exploring well-formalized knowledge structures, the other reshaping and rearranging the landscape. Our findings suggest that, at least as it is used today, AI – the boat or the compass, to stick with the analogy – seems to be more in line with the first action. However, the possibilities of discovering new and valuable things about the known islands are far from obvious, as confirmed by our results on scientific impact.

A general-purpose invention in the method of invention? Or a passing fad in science? We think not. Our findings lead us to take up a more moderate stance in the recent debate on how AI affects the development of knowledge. NN methods do not (yet) serve as an autopilot for navigating the sea of knowledge and connecting ideas, but they are, nevertheless, an extremely powerful and versatile research tool that impacts knowledge creation in measurable ways. Thus, we propose that AI should be considered an *emerging general method of invention.* But do not be fooled, we are not simply seeking to win the race to coin the most attractive designation; rather, as we discuss below, thinking of this technology as 'general' and 'emerging' has policy implications that differ substantially from those that might result from thinking of it as a general-purpose technology (for more on the latter, see, e.g., Trajtenberg [2018], Klinger et al. [2021].

First, the diffusion of intelligent machines as input in the research production process calls into question the organization and management of science. AI may trigger a short-term substitution towards capital and away from highly skilled labour in the knowledge production process. Whether such a substitution effect is occurring is doubtful and clearly requires further empirical investigation. In parallel, the arrival of automation technologies in science puts a wide range of research tasks under threat, either by reducing the cost of performing those tasks or by outperforming human scientists in the performance of them. Some tasks within the occupation may be suitable for automation while others may not, and the overall effects on employment in science are very complex. Therefore, research-oriented organizations need a better understanding of the set of tasks performed by their scientists, the coordination of these tasks, and the respective strengths and weaknesses of humans (H) and machines (M), before they can hope to unleash the benefits of H + M cooperation.

Machines are set to become more than tools; they have the potential to become another teammate. As such, H–M interactions will require the coordination of complex activities, including communication, joint actions and human-aware execution. As these machine teammates will operate in different collaborative environments, they need to be designed with different collaborative capabilities. This design area will require considering such aspects as appearance (what machines should look like); learning and knowledge processing (how they should learn); conversation (how they should interact and socialize with their peers); architecture (what their main components should be); reliability, responsibility and liability. (For a more in-depth discussion on design areas for human-machine collaboration, see Seeber et al. [2020]).

It seems that NN methods are being adopted in different scientific fields but that existing knowledge structures are remaining relatively stable. This suggests the full potential of the technology (and its future development) might be better achieved by further spanning the boundaries between scientific areas. The bringing together of expertise and knowledge from various domains could help in the identification of blind spots and opportunities in the knowledge landscape. The concepts of 'knowledge communities' and 'communities of practice' seem particularly apt in this context. Although communities often self-organize and self-sustain themselves, they can also benefit from policy endorsement. It seems crucial to us that institutions and a policy environment be developed that are conducive to enhancing dialogue and cross-fertilization between communities. This could be achieved, for instance, by reinforcing both horizontal (intra-field) and vertical (inter-field) knowledge management. Digital platforms and knowledge hubs could be complemented by physical 'collaborative spaces' where the tacit knowledge of different communities might be transferred face-to-face, documented and made accessible for later use. Another standard instrument is obviously research funding, which should not target individual areas but rather research 'priorities' (e.g., fighting a given disease) involving different communities that can frame their research questions together.

However, promoting collaboration between communities can pose certain challenges in terms of governance and data ownership. Data is a polymorphous category, which means standards, principles and rules governing the various types of data are not homogeneous across communities, let alone across countries. This opens up the question of how data should be generated/used in compliance with different regulations, and also how the value of data should be distributed [Savona, 2019].

The diffusion of AI, as a research instrument, can be self-sustaining only if there is social acceptance – i.e., if the crew trusts the captain and the equipment. Several AI applications represent innovations that can bring about far-reaching changes in all aspects of our daily lives. These social innovations can have unintended yet negative consequences in terms of security, privacy and social equity [O'neil, 2017]. The public will no longer tolerate being excluded from the debate and it is here that the scientific and policy community have a key role to play. Both parties can improve the channelling of scientific evidence into the public arena and fight the risks posed by fake news. Policy can promote communication by setting the right, often intrinsic, incentives to encourage as many scientists as possible to engage with different segments of the public. However, communicating science to non-scientific audiences can be difficult since it requires a different approach from that of communicating science to scientific audiences. This means scientists need to be able to detach the layers of scientific complexity that characterize their research so as to deliver a clear message to the public, a message, moreover, that should include both potential impacts and ethical issues. 'Listening mechanisms' can also be used to inform citizens' knowledge, expectations, and imaginaries about intelligent machines and, why not, about their role in science. There are a variety of means available for achieving these goals, ranging from in-depth interviews and material deliberations to citizen science. We believe that citizen science has the potential to bring the greatest benefits to both the public and the scientific system. The nonprofessional involvement of volunteers in the scientific process, whether in more mundane tasks such as data collection or in other phases of the research, offers great opportunities for the public to become familiar with the technology but also provides researchers with great opportunities to improve their results [Bonney et al., 2014, Sullivan et al., 2018]. However, fully accountable institutional mechanisms are a precondition for guaranteeing trust between scientists and the public and for ensuring continuity in their relationship. For instance, all results and the process used in reaching these results should be open to scrutiny. Policy should promote feedback activities so as to maintain citizen involvement and explain how their inputs were used in meeting research aims; reconcile conflicting values and objectives; and, put in place collective intelligence mechanisms that can help them develop a systemic understanding of the future implications of technological progress and make better consensus decision-making – all very much in line with the notion of 'Decisions 2.0' [Bonabeau, 2009]. Finally, we fully embrace the concept of 'boundary organisations' specifically designed to deal with socio-economic transformations in the digital age. These organisations would sit at the intersection of scientific and political spheres and allow scientists and policy-makers to maintain a constant dialogue with each other.

Although the AI revolution has been the subject under scrutiny here, ironically this revolution offers the tools with the greatest potential for bringing about a radical transformation in the interactions between the public, the scientific community and the policy environment. These interactions, if exploited carefully, should serve to give a boost to human efforts to better understand the greatest mystery of all: the origin and function of the world and our place in it, that is, the tasks of science itself.

1.7 Appendix

Descriptive statistics and results

Table 1.5: Descriptive statistics of the variables

	NN Papers	Non-NN Papers	Total
Re-combinatorial Novelty			
Novelty Dummy (All Sciences)	36.43	30.32	30.40
Novelty Dummy (No CS)	32.39	29.55	29.59
Novelty Dummy (Only HS)	20.96	23.52	23.49
Novelty (All Sciences)	0/0.81 (2.39)	0/0.74 (3.10)	0/0.74 (3.09)
Novelty (No CS)	0/0.65 (2.12)	0/0.71 (3.07)	0/0.71 (3.06)
Novelty (Only HS)	0/0.37 (1.62)	0/0.50 (2.40)	0/0.5 (2.39)
Scientific Impact			
Top 5% Cited	8.33	5.77	5.80
Top 10% Cited	15.68	11.38	11.43
# Citations (Raw Count)	17/38.34 (114.43)	18/35.48 (82.67)	18/35.51 (83.15)
Citations (Yearly Normalized)	2.06/4.06 (8.16)	2.08/3.75 (8.02)	2.08/3.75 (8.02)
Controls			
# References	40/45.92 (29.59)	33/37.12 (25.66)	33/37.23 (25.73)
# Authors	4/4.07 (2.37)	4/4.90 (3.50)	4/4.89 (3.49)
International Collab.	26.21	23.02	23.06
Private Partic.	6.80	7.09	7.09
JIF	1.39/2.12 (2.06)	1.73/2.42 (2.13)	1.73/2.41 (2.13)
Journal Age	22/28.57 (26.07)	33/38.47 (29.08)	32/38.35 (29.06)
Survey	0.72	0.78	0.77
Time Period	[2001 – 2015]	[2001 – 2015]	[2001 – 2015]
# Scientific Fields	46	48	48
# Journals	92	92	92
# Papers	4,560(1.28%)	351,477(98.72%)	356,037(100%)

Notes: Binary indicators in [%], for continuous measures [median/mean (s.d.)]. The statistics refer to the period used for the econometric analysis.

From word embeddings to search terms

This Appendix complements Section 1.3 by adding technical details on the learning of search terms for data retrieval. Source data and codes are fully accessible upon request.

Table 1.6: Atypical profile of NN publications

Category	All Sciences (1)	No CS (2)	Only HS (3)
HC–HN	0.008	0.208	0.308**
	(0.130)	(0.133)	(0.136)
HC–LN	-0.041	0.090	-0.049
	(0.157)	(0.152)	(0.154)
LC–LN	-0.043	-0.086	0.021
	(0.162)	(0.163)	(0.155)
Other variables	Yes	Yes	Yes
Log Likelihood	-374,002	-374,000	-363,855
χ^2 [Null Model]	95,913***	95,488***	115,891***
χ^2 [w/o NN Model]	259***	158.20***	144***
# Obs	320,587	320,587	320,587

Notes: This table reports coefficients of the effect of NN methods (NN, dummy) on atypical profiles. Category LC-HN is the reference category for all models. Bootstrapped (500 replications) standard errors clustered at the journal-level in parentheses: ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively. All variables are incorporated in all model specifications, details in Supplementary Material. Likelihood-ratio tests are used to compare the goodness-of-fit of two statistical models: (i) null model against complete model; (ii) model without the NN variable against the complete model.

Preparation of the training data

The training data is a bulk download of article abstracts from arXiv.org via its API provided by the R package 'aRxiv' [Ram and Broman, 2019]. We obtained in total 197,439 documents submitted between 1990 (2 documents) and 2019 (16,533 documents at the time of downloading – July; 35,807 in 2018) in the 'Computer Science', 'Mathematics' and 'Statistics' sections of arXiv.org. Preprocessing entails removing all abstracts with less than 15 words; a pre-defined set of stop words; all words occurring less than 5 times in the corpus. In addition, we paste unigrams into bi-grams depending on the frequency of co-occurence. Following Mikolov et al. [2013b] and Mikolov et al. [2013] a bi-gram is created when the score of the two words, w_i and w_j, pass a given threshold. The score is calculated as follows: $score(w_i, w_j) = \frac{count(w_i,w_j)-\delta}{count(w_i)\cdot count(w_j)}$, where δ is used as a discounting coefficient and prevents too many bi-grams consisting of very infrequent words to be formed. We choose a threshold of 50 to increase the number of bi-grams generated (default is 100).

After preprocessing, the training data includes 14,458,777 words from a vocabulary of size 87,990. This leads to a weight matrix of dimension 45,050,880 (87,990 ×

512 dimensions).

Estimation of word representations

We estimate word representations with the continuous Skip-Gram model introduced in Mikolov et al. [2013b] and Mikolov et al. [2013]. The Skip-Gram model is one specific variant of a set of word embedding algorithms that have become popular under the label of Word2Vec.

We use negative sampling. In 'old school' parlance, this is essentially a Logit model. The binary dependent variable indicates whether or not two terms are close in the text corpus, at distance c. For each observed neighboring term pair (success), one adds k 'negative samples' (failures). The scalar product of word representations enters the model as the linear predictor. Sequential processing is achieved through stochastic gradient descent.

The results presented along the main text have been obtained with the following parameter settings. The main free parameter is the dimensionality of the dense word representation, which we set to 512 dimensions.[6] We define a context window (distance c) of 7 words from both sides around the target. For each observed neighboring term pair, we draw $k = 15$ negative examples. A negative example is obtained by replacing one word of the observed neighboring terms by another word from the vocabulary that is drawn randomly with probability proportional to its frequency – i.e., $P(w_i) = \frac{f(w_i)^{3/4}}{\Sigma_{j=0}^{n}(f(w_j)^{3/4})}$, which is close to draws uniformly at random. Further, we make use of sub-sampling by specifying 30 epochs, so that the whole dataset is passed 30 times through the network.

Word clustering

Estimation results – i.e., estimated word embeddings – serve as input to a cluster analysis. Term clusters are identified with the k-means clustering procedure. We used the gap statistics to determine the optimal number of clusters. The most frequent n-grams for the 22 identified clusters are reported in Tables 1.7a and 1.7b.

[6]We tried several dimensions to represent dense representation: 256, 300, 512 and 1,024. Our choice was guided by the results of the k-mean clustering; we opted for the dimension for which the DL cluster was best defined.

Table 1.7a: Word embedding obtained via Word2Vec [arXiv.org sample]

Cluster 1 (Deep Learning)	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Cluster 9
n-grams								
adversarial_examples	alternating_direction	audio_visual	adjacency_matrix	action_spaces	asymptotic_properties	alzheimer's_disease	alice_bob	achieves_art
artificial_neural	batch_size	automatic_speech	analysis_pca	actor_critic	bayesian_inference	association_studies	base_station	benchmark_datasets
~~attention_mechanism~~	belief_propagation	bag_words	community_detection	agent_learns	central_limit	brain_activity	brute_force	compared_art
convolutional_neural	boundary_conditions	character_level	component_analysis	atari_games	computationally_efficient	brain_regions	coding_scheme	compared_traditional
convolutional_neural_network	combinatorial_optimization	collaborative_filtering	compressed_sensing	contextual_bandit	confidence_intervals	brain_tumor	cognitive_radio	computational_efficien
~~data_augmentation~~	computational_cost	context_aware	compressive_sensing	contextual_bandits	covariance_matrix	breast_cancer	communication_channels	conduct_experiments
deep_convolutional	constrained_optimization	contextual_information	covariance_matrices	control_policies	cross_validation	causal_relationships	communication_protocols	current_art
deep_learning	convergence_rate	cross_lingual	dimension_reduction	decision_processes	density_estimation	clinical_practice	elliptic_curve	demonstrate_effectiver
deep_neural	convergence_rates	domain_specific	dimensionality_reduction	deep_q_learning	expectation_maximization	clinical_trial	encoding_decoding	demonstrate_efficacy
deep_neural_network	convex_optimization	emotion_recognition	distance_metric	deep_reinforcement	gaussian_process	clinical_trials	encryption_scheme	demonstrate_superiori
~~domain_adaptation~~	differential_equations	fake_news	euclidean_distance	deep_reinforcement_learning	gaussian_processes	computed_tomography	error_correcting	effectiveness_proposed
encoder_decoder	divide_conquer	feature_engineering	euclidean_space	expected_reward	hidden_markov	computer_aided	error_correction	empirical_evaluation
~~feature_extraction~~	dynamic_programming	image_captioning	feature_selection	experience_replay	hypothesis_testing	computer_assisted	fusion_center	error_rate
~~feature_maps~~	dynamical_systems	information_retrieval	fourier_transform	exploration_exploitation	importance_sampling	cross_sectional	hash_function	error_rates
~~feed_forward~~	easy_implement	knowledge_base	gaussian_mixture	game_playing	joint_distribution	ct_images	hash_functions	evaluation_metrics
~~fine_tuning~~	empirical_risk	latent_dirichlet	gaussian_noise	heuristic_search	latent_variable	disease_progression	hoc_networks	experimental_evaluatio
generative_adversarial	evolutionary_algorithms	link_prediction	graph_laplacian	imitation_learning	latent_variables	dna_sequences	homomorphic_encryption	experimental_results
generative_adversarial_network	faster_convergence	low_resource	hilbert_space	imperfect_information	linear_regression	electronic_health	information_leakage	experiments_conducte
~~hand_crafted~~	finite_element	machine_translation	ill_posed	infinite_horizon	log_likelihood	false_negative	information_theoretic	experiments_synthetic
~~latent_space~~	forward_backward	manually_annotated	linear_combination	intelligent_agents	logistic_regression	functional_magnetic	key_exchange	extensive_experimenta
long_short_term_memory	generalization_error	named_entity	linear_combinations	inverse_reinforcement	markov_chain	gene_expression	leader_election	extensive_experiments
~~loss_function~~	gradient_descent	natural_language	low_dimensional	markov_decision	markov_chains	genome_wide	min_entropy	f1_score
~~loss_functions~~	linear_programming	news_articles	low_rank	markov_decision_process	maximum_likelihood	gold_standard	multi_hop	https_github.com
~~multi_label~~	local_minima	nlp_tasks	matrix_completion	motion_planning	monte_carlo	heart_rate	multi_party	magnitude_faster
~~multi_layer~~	min_max	processing_nlp	matrix_factorization	multi_agent	numerical_examples	human_brain	noise_ratio	outperform_art
networks_cnns	mini_batch	question_answering	means_clustering	multi_armed	parameter_estimation	imaging_mri	physical_layer	outperforms_art
neural_network	mixed_integer	relation_extraction	nearest_neighbor	partially_observable	posterior_distribution	low_dose	pseudo_random	outperforms_existing
neural_networks	numerical_experiments	semantic_similarity	nearest_neighbors	policy_gradient	probability_density	lung_cancer	public_key	pascal_voc
~~pre_trained~~	objective_function	sentence_level	nuclear_norm	policy_iteration	probability_distributions	magnetic_resonance	quantum_computers	random_forest
recurrent_neural	optimal_transport	sentiment_analysis	positive_definite	processes_mdps	random_fields	medical_diagnosis	rate_distortion	significant_improveme
recurrent_neural_network	partial_differential	sentiment_classification	principal_component	regret_bound	random_variable	medical_image	ratio_snr	significant_improveme
~~semi_supervised~~	particle_swarm	sequence_sequence	principal_components	regret_bounds	random_variables	medical_images	secret_key	significantly_improve
~~short_term~~	primal_dual	similarity_measures	reproducing_kernel	reinforcement_learning	rate_convergence	medical_imaging	secret_sharing	significantly_improved
~~shot_learning~~	sample_complexity	speech_recognition	signal_noise	reinforcement_learning	sample_size	presence_absence	secure_communication	significantly_improves
~~source_target~~	statistical_physics	text_mining	signal_processing	reward_function	simulation_studies	resonance_imaging	security_protocols	significantly_outperfor
~~spiking_neural~~	stochastic_gradient	topic_modeling	sparse_coding	sequential_decision	smoothing	risk_factors	sensor_networks	simulated_real
~~supervised_learning~~	strongly_convex	web_pages	spectral_clustering	temporal_difference	squared_error	rna_seq	sensor_nodes	source_code
~~target_domain~~	theoretical_guarantees	word_embedding	subspace_clustering	thompson_sampling	time_series	sensitivity_specificity	source_destination	superior_performance
variational_autoencoder	variance_reduction	word_embeddings	support_vector	time_horizon	time_varying	skin_lesion	wireless_communication	synthetic_real
~~weakly_supervised~~	variational_inference	word_vectors	total_variation	upper_confidence	variable_selection	white_matter	wireless_sensor	times_faster
Unigrams								
architecture	approximation	annotation	clustering	action	asymptotic	biological	adversary	accuracy
architectures	approximations	annotations	coefficients	actions	bayesian	brain	bit	achieve
classification	convergence	corpus	decomposition	actor	conditional	cancer	bits	achieved
classifier	convex	document	dictionary	agent	density	clinical	channel	achieves
deep	equations	documents	inverse	agent's	dependence	diagnosis	channels	art
feature	formulation	embeddings	kernel	arm	distributions	disease	codes	benchmark
features	gradient	entities	kernels	arms	estimated	gene	coding	compare
layer	greedy	extraction	manifold	bandit	estimates	genes	communication	compared
learn	heuristic	linguistic	matrices	demonstrations	estimating	genetic	decoding	comparison
learned	iteration	retrieval	matrix	exploration	estimation	imaging	exchange	dataset
learning	iterative	semantic	nonlinear	horizon	estimator	longitudinal	message	datasets
loss	minimization	sentence	projection	learner	estimators	medical	messages	demonstrate
network	numerical	sentences	rank	mdps	gaussian	molecular	protocol	evaluate
networks	optimization	speech	regularization	planner	likelihood	patient	protocols	evaluated
neural	quadratic	style	signal	planning	multivariate	patients	quantum	experiments
representations	smooth	text	sparse	policies	parametric	protein	routing	performance
tasks	solve	topic	sparsity	policy	regression	screening	scheme	proposed
train	solving	translation	spectral	regret	statistics	subjects	schemes	results
trained	stochastic	word	vector	reward	tests	survival	secure	robustness
training	variational	words	vectors	rewards	variance	treatment	transmission	test

Notes: This table reports the most frequent n-grams per cluster, sorted by alphabetical order. Cluster 1 concerns deep learning. The terms excluded from the list used to retrieve data from WoS are crossed out.

Table 1.7b: Word embedding obtained via Word2Vec [arXiv.org sample]

Cluster 12	Cluster 13	Cluster 14	Cluster 15	Cluster 16	Cluster 17	Cluster 18	Cluster 19	Cluster 20
n-grams								
access_control	answer_programming	algebraic_geometry	computational_resources	cooperative_game	computer_science	machine_learning	autonomous_vehicles	
cloud_computing	association_rules	binary_trees	computationally_intensive	decision_maker	decision_makers	real_world	numerical_simulations	
cloud_services	background_knowledge	cayley_graphs	computing_resources	differential_privacy	decision_support			
cyber_physical	boolean_function	combinatorial_interpretation	energy_consumption	expected_utility	health_care			
cyber_security	boolean_functions	combinatorial_objects	energy_efficiency	game_players	public_health			
internet_iot	cellular_automata	dyck_paths	external_memory	game_theoretic	search_engine			
intrusion_detection	constraint_satisfaction	equivalence_classes	fault_tolerance	incentive_compatible	search_engines			
mobile_devices	dempster_shafer	explicit_formula	fault_tolerant	nash_equilibria	social_media			
mobile_phone	expressive_power	explicit_formulas	graphics_processing	nash_equilibrium	social_sciences			
operating_system	finite_automata	expressed_terms	hardware_software	pareto_optimal	software_development			
peer_peer	fuzzy_logic	generating_function	linear_algebra	price_anarchy	software_engineering			
privacy_concerns	knowledge_bases	generating_functions	load_balancing	pure_nash	statistically_significant			
privacy_preserving	kolmogorov_complexity	hopf_algebra	low_latency	resource_allocation				
quality_service	logic_programming	infinite_family	massively_parallel	social_choice				
resource_constrained	object_oriented	lie_algebra	matrix_multiplication	social_welfare				
resource_management	pattern_matching	partition_function	memory_footprint	stable_matching				
safety_critical	programming_languages	schur_functions	memory_requirements					
security_privacy	pspace_complete	simplicial_complex	message_passing					
service_providers	quantum_mechanics	simplicial_complexes	multi_core					
smart_contracts	relational_database	strongly_regular	power_consumption					
smart_grid	relational_databases	symmetric_functions	programming_language					
supply_chain	semantic_web	tensor_product	shared_memory					
user_interface	temporal_logic	tutte_polynomial						
virtual_machines	turing_machine	vector_spaces						
web_service	turing_machines							
web_services								
Unigrams								
access	answer	algebra	code	agents	activities	analysis	activity	algorithm
attack	causal	algebraic	computation	allocation	ai	applications	control	algorithms
attacks	definition	combinatorial	computations	costs	authors	approaches	dynamics	based
cloud	formal	enumeration	distributed	decisions	collected	complex	energy	class
devices	language	families	execution	demand	communities	data	environment	distributio
internet	languages	formula	hardware	equilibrium	community	design	event	efficient
management	logic	formulas	implement	game	individuals	existing	events	function
mobile	notion	integers	implementation	games	papers	framework	flow	functions
platform	operators	lattice	implementations	items	people	information	location	introduce
privacy	probabilistic	odd	implemented	market	project	knowledge	locations	linear
resources	program	partitions	machines	mechanism	public	level	measurements	multiple
security	programs	permutation	memory	mechanisms	research	methods	monitoring	optimal
service	proof	permutations	operations	outcome	researchers	process	physical	parameter
services	proofs	polynomials	parallel	outcomes	science	provide	sensor	properties
sharing	queries	polytope	processing	player	scientific	real	sensors	random
software	query	prime	run	players	social	scale	signals	simple
technology	reasoning	rational	scalable	private	students	system	simulated	size
user	relations	symmetry	scheduling	resource	survey	systems	simulation	space
users	rules	theorem	speed	strategies	topics	task	simulations	structure
web	semantics	theorems	storage	utility	world	techniques	traffic	time

Notes: This table reports the most frequent n-grams per cluster, sorted by alphabetical order. Cluster 1 concerns deep learning. The terms excluded from the list used to retrieve data from WoS

Acronyms and full names

Table 1.8: List of acronyms replaced by full name

Acronym	Full name
ann	artificial_neural_network
anns	artificial_neural_networks
blstm	bidirectional_long_short_term_memory
bns	bayesian_networks
bpn	bidirectional_pyramid_networks
cav	computer_aided_verification
cnn	convolutional_neural_network
cnns	convolutional_neural_networks
crf	conditional_random_fields
ctc	connectionist_temporal_classification
dan	deep_alignment_network
dbm	deep_boltzmann_machine
dbms	database_management_systems
dbn	deep_belief_network
dcn	dynamic_coattention_network
dcnn	deep_convolutional_neural_network
dcnns	deep_convolutional_neural_networks
dl	deep_learning
dek	deep_embedding_kernel
dnn	deep_neural_network
dnns	deep_neural_networks
dqn	deep_q_network

Continued on next page

Table 1.8: List of acronyms replaced by full name – continued

dqns	deep_q_networks
drcn	deeply_recursive_convolutional_network
drl	deep_reinforcement_learning
elm	extreme_learning_machine
fcn	fully_convolutional_network
fhmms	factorial_hidden_markov_model
ga	genetic_algorithm
gan	generative_adversarial_network
gans	generative_adversarial_networks
gcns	graph_convolutional_networks
grnn	general_regression_neural_network
grus	gated_recurrent_units
gsn	generative_stochastic_network
gssl	graph_based_semi_supervised_learning
knn	k_nearest_neighbors
lmnn	large_margin_nearest_neighbor
lstm	long_short_term_memory
lstms	long_short_term_memory
mdp	markov_decision_process
ml	machine_learning
mlp	multilayer_perceptron
mtl	multi_task_learning
nn	neural_network
nns	neural_networks

Continued on next page

Table 1.8: List of acronyms replaced by full name – continued

pmvge	probabilistic_multi_view_graph_embedding
pnn	probabilistic_neural_network
pso	particle_swarm_optimization
psrnns	predictive_state_recurrent_neural_networks
rbf	radial_basis_function
rbfn	radial_basis_function_network
rbms	restricted_boltzmann_machines
rgp	recurrent_gaussian_process
rl	reinforcement_learning
rlns	regularization_learning_networks
rmbs	restricted_boltzmann_networks
rnn	recurrent_neural_network
rnns	recurrent_neural_networks
smffnn	supervised_multilayers_feed_forward_neural_network
snn	spiking_neural_network
snns	spiking_neural_networks
ssrbm	spike_slab_restricted_boltzmann_machine
svm	support_vector_machine
vae	variational_autoencoder
vaes	variational_autoencoders
wae	wasserstein_autoencoder
zsl	zero_shot_learning

End of table.

Diffusion of deep learning in science: the sample

This Appendix complements Section 1.4 with details on the sample used for the analysis on the diffusion of deep learning in science.

Table 1.9: Deep learning documents broken down by period and WoS research areas

Year	All documents (1)	Technology (2)	Physical Sciences (3)	Life Sciences & Biomedicine (4)	Health Sciences (5)	Social Sciences (6)	Art & Humanities (7)
1990	381 [0.15]	193.33 [50.74]	117.17 [30.75]	61.83 [16.23]	106 [27.82]	8.67 [2.27]	0 [0]
1991	836 [0.32]	507.52 [60.71]	193.73 [23.17]	123.58 [14.78]	188 [22.49]	11.17 [1.34]	0 [0]
1992	1,256 [0.48]	759.17 [60.44]	306.92 [24.44]	159.25 [12.68]	234 [18.63]	29.67 [2.36]	1 [0.08]
1993	1,477 [0.57]	841.05 [56.94]	365.37 [24.74]	221.42 [14.99]	315 [21.33]	49.17 [3.33]	0 [0]
1994	1,798 [0.69]	1,074.10 [59.74]	373.60 [20.78]	293.53 [16.33]	385 [21.41]	55.77 [3.10]	1 [0.06]
1995	2,220 [0.85]	1,415.85 [63.78]	436.92 [19.68]	306.23 [13.79]	412 [18.56]	60.50 [2.73]	0.50 [0.02]
1996	2,791 [1.07]	1,843.03 [66.03]	478.30 [17.14]	393.23 [14.09]	475 [17.02]	75.43 [2.70]	1 [0.04]
1997	3,090 [1.19]	2,002.12 [64.79]	530.50 [17.17]	481.38 [15.58]	613 [19.84]	74 [2.39]	2 [0.06]
1998	4,330 [1.66]	2,865.13 [66.17]	566.70 [13.09]	779.25 [18.00]	1,083 [25.01]	113.92 [2.63]	5 [0.12]
1999	4,725 [1.81]	3,302.34 [69.89]	725.92 [15.36]	598.61 [12.67]	627 [13.27]	93.58 [1.98]	4.55 [0.10]
2000	6,259 [2.40]	4,661.05 [74.47]	835.07 [13.34]	621.55 [9.93]	691 [11.04]	138.33 [2.21]	3 [0.05]
2001	6,062 [2.33]	4,376.40 [72.19]	859.52 [14.18]	726.90 [11.99]	806 [13.3]	94.68 [1.56]	4.50 [0.07]
2002	6,676 [2.56]	5,191.35 [77.76]	762.67 [11.42]	614.42 [9.20]	700 [10.49]	104.37 [1.56]	3.20 [0.05]
2003	7,230 [2.78]	5,430.80 [75.11]	923.27 [12.77]	768.00 [10.62]	897 [12.41]	100.43 [1.39]	7.50 [0.10]
2004	7,765 [2.98]	5,907.90 [76.08]	921.27 [11.86]	811.58 [10.45]	879 [11.32]	119.75 [1.54]	4.50 [0.06]
2005	9,023 [3.46]	7,026.45 [77.87]	1,072.60 [11.89]	790.40 [8.76]	896 [9.93]	129.80 [1.44]	3.75 [0.04]
2006	10,654 [4.09]	8,206.27 [77.03]	1,424.57 [13.37]	885.2 [8.31]	859 [8.06]	136.60 [1.28]	1.36 [0.01]
2007	11,086 [4.26]	8,234.85 [74.28]	1,551.97 [14.00]	1,072.22 [9.67]	1,246 [11.24]	217.38 [1.96]	9.58 [0.09]
2008	11,891 [4.57]	9,053.63 [76.14]	1,562.03 [13.14]	1,015.39 [8.54]	1,067 [8.97]	256 [2.15]	3.95 [0.03]
2009	13,049 [5.01]	10,066.02 [77.14]	1,601.43 [12.27]	1,113.80 [8.54]	1,102 [8.45]	258.05 [1.98]	9.70 [0.07]
2010	10,467 [4.02]	7,702.98 [73.59]	1,399.65 [13.37]	1,117.50 [10.68]	992 [9.48]	242.87 [2.32]	4 [0.04]
2011	10,872 [4.17]	8,033.20 [73.89]	1,462.38 [13.45]	1,110.13 [10.21]	943 [8.67]	261.78 [2.41]	4.50 [0.04]
2012	12,227 [4.69]	9,238.63 [75.56]	1,571.02 [12.85]	1,189.90 [9.73]	1,047 [8.56]	220.95 [1.81]	6.50 [0.05]
2013	12,691 [4.87]	9,439.40 [74.38]	1,779.75 [14.02]	1,248.40 [9.84]	1,106 [8.71]	217.78 [1.72]	5.67 [0.04]
2014	14,355 [5.51]	11,044.90 [76.94]	1,747.07 [12.17]	1,263.52 [8.80]	1,067 [7.43]	293.02 [2.04]	6.50 [0.05]
2015	16,764 [6.44]	12,934.47 [77.16]	1,978.12 [11.80]	1,476.93 [8.81]	1,267 [7.56]	367.65 [2.19]	6.83 [0.04]
2016	18,425 [7.07]	13,927.08 [75.59]	2,265.67 [12.30]	1,700.87 [9.23]	1,449 [7.86]	515.55 [2.80]	15.83 [0.09]
2017	24,046 [9.23]	18,488.38 [76.89]	2,993.48 [12.45]	2,099.37 [8.73]	2,008 [8.35]	449.93 [1.87]	14.83 [0.06]
2018	28,013 [10.76]	20,223.48 [72.19]	4,192.73 [14.97]	3,078.15 [10.99]	3,001 [10.71]	491.90 [1.76]	26.73 [0.10]

Notes: Number of deep learning documents (Column 1). Weighted count for all other columns. For 'All Documents' the shares [%] are calculated on the basis of the entire DL sample. For all other columns the share refers to the period. For example, the number of documents published in 2018 represents 10.76% of all DL documents; of the 28,013 documents, 72.19% belong to 'Technology', 14.97% to 'Physical Sciences', and so on.

Deep learning in health sciences: data construction and sample details

This Appendix provides additional statistics on the empirical analysis of Section 1.5. The perimeter of the domain 'health sciences' has been delineated using the WoS subject categories reported in Table 1.11. Health sciences can be viewed as a subset of the broader WoS research area 'Life Science & Biomedicine'.

Table 1.10: Deep learning publication activity broken down by country and period

1990-1999		*2000-2009*		*2010-2019*	
Country	# Documents	Country	# Documents	Country	# Documents
[UE]	6,205	[UE]	24,047	China	39,852
USA	5,123	China	20,560	[UE]	28,358
United Kingdom	1,695	USA	13,665	USA	17,320
Japan	1,368	United Kingdom	5,151	India	10,349
Germany	991	Japan	5,076	Iran (Islamic Republic of)	7,008
Italy	851	Taiwan	3,611	United Kingdom	4,917
China	764	Italy	3,269	Japan	4,471
Canada	721	India	3,225	Taiwan	4,027
France	704	Canada	3,204	Korea (Republic of)	3,902
Spain	474	Spain	2,898	Turkey	3,895
Taiwan	456	Korea (Republic of)	2,872	Spain	3,438
Australia	436	Germany	2,802	Canada	3,377
Korea (Republic of)	427	Turkey	2,228	Germany	3,271
India	354	France	2,183	Italy	3,063
Netherlands	286	Iran (Islamic Republic of)	1,964	Australia	2,925
Brazil	224	Brazil	1,874	Malaysia	2,631

Notes: Top 15 countries for each period. [EU] represents EU28 as in 2018.

Meta data on the estimation sample

This Appendix provides details on the sample constructed to carry out the empirical analysis on health sciences (Section 1.5.1). To benchmark deep learning publications, we download all the articles for the whole observation period published in the top 100 journals where research involving deep learning has been the most prominent.

Table 1.12: Sampled papers by journal and period

Journal | Foundation date	1990–1999	2000–2009	2010–2019
ACADEMIC RADIOLOGY | 1994	1,250	2,179	2,265
ANALYTICAL AND BIOANALYTICAL CHEMISTRY | 1862	0	5,235	8,300
ANNALS OF BIOMEDICAL ENGINEERING | 1972	684	1,624	2,389
BASIC & CLINICAL PHARMACOLOGY & TOXICOLOGY | 1945	0	1,944	8,411
BEHAVIORAL AND BRAIN SCIENCES | 1978	4,162	3,573	2,273
BEHAVIOURAL BRAIN RESEARCH | 1980	1,719	2,861	5,861
BIOLOGICAL PSYCHIATRY | 1959	6,477	10,589	13,323

Continued on next page

Table 1.12: Sampled papers per journal and period – continued

Journal | Foundation date	1990–1999	2000–2009	2010-2019
BIOMED RESEARCH INTERNATIONAL | 2001	0	0	16,302
BIOMEDICAL ENGINEERING ONLINE | 2002	0	160	1,309
BIOMEDICAL SIGNAL PROCESSING AND CONTROL | 2006	0	150	1,398
BIORESOURCE TECHNOLOGY | 1991	1,434	4,464	15,677
BIOSYSTEMS | 1967	651	1,062	927
BMC BIOINFORMATICS | 2000	0	3,455	6,060
BMC MEDICAL INFORMATICS AND DECISION MAKING | 2001	0	171	1,350
BRAIN | 1878	1,457	2,820	3,395
BRAIN AND LANGUAGE | 1974	1,317	1,899	914
BRAIN RESEARCH | 1966	15,725	11,563	6,503
CEREBRAL CORTEX | 1991	517	1,877	3,006
CLINICAL NEUROPHYSIOLOGY | 1949	286	3,047	3,493
COGNITIVE NEURODYNAMICS | 2007	0	94	418
COGNITIVE SCIENCE | 1977	178	411	851
COMBINATORIAL CHEMISTRY & HIGH THROUGHPUT SCREENING | 1998	47	766	849
COMPUTATIONAL AND MATHEMATICAL METHODS IN MEDICINE | 1997	0	44	1,632
COMPUTATIONAL INTELLIGENCE AND NEUROSCIENCE | 2007	0	0	855
COMPUTERIZED MEDICAL IMAGING AND GRAPHICS | 1988	565	615	664
CORTEX | 1964	562	1,021	2,213
CURRENT BIOLOGY | 1991	2,744	7,273	7,714
CURRENT OPINION IN NEUROBIOLOGY | 1991	527	1,035	1,400

Continued on next page

Table 1.12: Sampled papers per journal and period – continued

Journal | Foundation date	1990–1999	2000–2009	2010-2019
EPILEPSIA | 1909	6,736	17,762	10,780
EUROPEAN JOURNAL OF MEDICINAL CHEMISTRY | 1966	1,164	1,990	7,632
EUROPEAN JOURNAL OF NEUROSCIENCE | 1989	6,282	8,968	3,260
EXPERIMENTAL BRAIN RESEARCH | 1966	3,062	4,110	3,517
FOOD CHEMISTRY | 1976	2,077	6,114	16,416
FRONTIERS IN COMPUTATIONAL NEUROSCIENCE | 2007	0	38	1,136
FRONTIERS IN HUMAN NEUROSCIENCE | 2008	0	91	5,451
FRONTIERS IN NEUROINFORMATICS | 2007	0	1	482
FRONTIERS IN NEUROSCIENCE | 2009	0	114	4,778
FRONTIERS IN PSYCHOLOGY | 2010	0	0	14,466
HIPPOCAMPUS | 1991	490	1,007	1,231
HUMAN BRAIN MAPPING | 1993	182	1,110	3,052
IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING | 1964	1,594	2,528	3,218
IEEE TRANSACTIONS ON NEURAL SYSTEMS AND REHABILITATION ENGINEERING | 2001	0	553	1,366
INTERNATIONAL JOURNAL OF COMPUTER ASSISTED RADIOLOGY AND SURGERY | 2006	0	664	1,337
INTERNATIONAL JOURNAL OF ENVIRONMENTAL RESEARCH AND PUBLIC HEALTH | 2004	0	214	11,117
INTERNATIONAL JOURNAL OF MOLECULAR SCIENCES | 2000	0	804	18,697
INVESTIGATIVE OPHTHALMOLOGY & VISUAL SCIENCE | 1962	17,439	2,973	2,640
JOURNAL OF CHROMATOGRAPHY A | 1958	7,664	11,861	9,715
JOURNAL OF COGNITIVE NEUROSCIENCE | 1989	1,302	3,950	2,978

Continued on next page

Table 1.12: Sampled papers per journal and period – continued

Journal | Foundation date	1990–1999	2000–2009	2010-2019
JOURNAL OF COMPUTATIONAL NEUROSCIENCE | 1994	113	415	545
JOURNAL OF DIGITAL IMAGING | 1988	401	625	974
JOURNAL OF MECHANICS IN MEDICINE AND BIOLOGY | 2001	0	308	1,080
JOURNAL OF MEDICAL IMAGING AND HEALTH INFORMATICS | 2011	0	0	1,706
JOURNAL OF MEDICAL SYSTEMS | 1977	56	243	2,125
JOURNAL OF MEDICINAL CHEMISTRY | 1959	5,619	6,783	7,387
JOURNAL OF MOLECULAR BIOLOGY | 1959	7097	9691	4,251
JOURNAL OF NEURAL ENGINEERING | 2004	0	331	1456
JOURNAL OF NEUROPHYSIOLOGY | 1938	4,750	6,040	5,195
JOURNAL OF NEUROSCIENCE | 1981	6,705	13,443	13,766
JOURNAL OF NEUROSCIENCE METHODS | 1979	1,638	2,690	2,874
JOURNAL OF NUCLEAR MEDICINE | 1964	12,218	11,603	20,672
JOURNAL OF PHARMACEUTICAL AND BIOMEDICAL ANALYSIS | 1983	2,362	4,376	5,347
JOURNAL OF PHYSIOLOGY-PARIS | 1992	383	476	227
JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA | 1929	7,323	6,801	7,806
JOURNAL OF THEORETICAL BIOLOGY | 1961	2,371	3,270	4,203
JOURNAL OF UROLOGY | 1917	12,499	29,396	39,207
JOURNAL OF VIBROENGINEERING | 2007	0	268	2,562
MEDICAL ENGINEERING & PHYSICS | 1994	537	1,153	1,699
MEDICAL IMAGING 2018: COMPUTER-AIDED DIAGNOSIS | 2018	0	0	136
MEDICAL PHYSICS | 1997	2,268	14,402	28,629

Continued on next page

Table 1.12: Sampled papers per journal and period – continued

Journal | Foundation date	1990–1999	2000–2009	2010-2019
MOLECULES | 1996	210	1,875	15,389
NATURE NEUROSCIENCE | 1998	414	2,899	2,727
NEUROIMAGE | 1993	373	7,286	9,626
NEURON | 1988	2,393	3,977	4,693
NEUROPSYCHOLOGIA | 1963	1,338	2,465	3,584
NEUROREPORT | 1990	5,112	5,152	2,125
NEUROSCIENCE | 1976	5,846	7,472	7,491
NEUROSCIENCE AND BIOBEHAVIORAL REVIEWS | 1977	617	772	2,190
NEUROSCIENCE LETTERS | 1975	10,062	9,976	7,188
NEUROSCIENCE RESEARCH | 1984	1,081	7,801	4,976
NUCLEIC ACIDS RESEARCH | 1974	11,010	10,326	12,648
PERCEPTION | 1972	4,762	7,581	7,639
PHYSICS IN MEDICINE AND BIOLOGY | 1956	1,853	4,283	5,381
PHYSIOLOGICAL MEASUREMENT | 1980	367	1,167	1,620
PLOS COMPUTATIONAL BIOLOGY | 2005	0	1,149	5,187
PROTEINS-STRUCTURE FUNCTION AND BIOINFORMATICS | 1986	437	2,991	2,190
PSYCHOLOGICAL REVIEW | 1894	379	493	390
RADIOLOGY | 1923	19,517	12,402	5,188
RADIOTHERAPY AND ONCOLOGY | 1983	1,623	10,706	16,163
SCHIZOPHRENIA RESEARCH | 1988	5,257	8,757	7,323
TRENDS IN COGNITIVE SCIENCES | 1997	332	1,263	1,092
VISION RESEARCH | 1961	4,295	3,164	1,890
2007 ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY, VOLS 1-16 | 2007	0	1,703	0

Continued on next page

Table 1.12: Sampled papers per journal and period – continued

Journal | Foundation date	1990–1999	2000–2009	2010-2019
2011 ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY (EMBC) | 2011	0	0	2,083
2015 37TH ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY (EMBC) | 2015	0	0	2,008
2017 39TH ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY (EMBC) | 2017	0	0	1,123
2017 IEEE 14TH INTERNATIONAL SYMPOSIUM ON BIOMEDICAL IMAGING (ISBI 2017) | 2017	0	0	285
2018 11TH INTERNATIONAL CONGRESS ON IMAGE AND SIGNAL PROCESSING, BIOMEDICAL ENGINEERING AND INFORMATICS (CISP-BMEI 2018) | 2018	0	0	249
2018 IEEE 15TH INTERNATIONAL SYMPOSIUM ON BIOMEDICAL IMAGING (ISBI 2018) | 2018	0	0	364

End of table.

Table 1.11: WoS subject categories defining 'health sciences'

Category	Count [Share]	Category	Count [Share]
Neurosciences	6,683 [2.56]	Geriatrics & Gerontology	75 [0.03]
Biology	6,084 [2.33]	Anatomy & Morphology	73 [0.03]
Mathematical & Computational Biology	3,386 [1.30]	Orthopedics	73 [0.03]
Radiology, Nuclear Medicine & Medical Imaging	2,678 [1.03]	Transplantation	63 [0.02]
Medical Informatics	2,218 [0.85]	Dentistry, Oral Surgery & Medicine	62 [0.02]
Psychology	1,932 [0.74]	Virology	61 [0.02]
Microbiology	1,908 [0.73]	Hematology	59 [0.02]
Biochemistry & Molecular Biology	1,852 [0.71]	Nursing	50 [0.02]
Biotechnology & Applied Microbiology	1,727 [0.66]	Reproductive Biology	41 [0.02]
Pharmacology & Pharmacy	1,221 [0.47]	Integrative & Complementary Medicine	36 [0.01]
Biophysics	863 [0.33]	Emergency Medicine	35 [0.01]
Psychiatry	733 [0.28]	Rheumatology	24 [0.01]
Cell Biology	582 [0.22]	Tropical Medicine	21 [0.01]
Health Care Sciences & Services	549 [0.21]	Mycology	18 [0.01]
Oncology	548 [0.21]	Allergy	7 [0]
Surgery	465 [0.18]	Medical Ethics	4 [0]
Genetics & Heredity	443 [0.17]	Psychology, Experimental	0 [0]
Physiology	431 [0.17]	Psychology, Applied	0 [0]
Behavioral Sciences	419 [0.16]	Psychology, Multidisciplinary	0 [0]
Toxicology	396 [0.15]	Psychology, Biological	0 [0]
Public, Environmental & Occupational Health	384 [0.15]	Neuroimaging	0 [0]
Endocrinology & Metabolism	283 [0.11]	Engineering, Biomedical	0 [0]
Pathology	249 [0.10]	Biochemical Research Methods	0 [0]
Medical Laboratory Technology	241 [0.09]	Clinical Neurology	0 [0]
Ophthalmology	237 [0.09]	Psychology, Developmental	0 [0]
Urology & Nephrology	235 [0.09]	Cardiac & Cardiovascular Systems	0 [0]
Rehabilitation	230 [0.09]	Psychology, Social	0 [0]
Gastroenterology & Hepatology	184 [0.07]	Critical Care Medicine	0 [0]
Immunology	174 [0.07]	Medicine, Research & Experimental	0 [0]
Obstetrics & Gynecology	161 [0.06]	Psychology, Mathematical	0 [0]
Respiratory System	129 [0.05]	Chemistry, Medicinal	0 [0]
Evolutionary Biology	116 [0.04]	Medicine, Legal	0 [0]
Developmental Biology	112 [0.04]	Medicine, General & Internal	0 [0]
Anesthesiology	111 [0.04]	Peripheral Vascular Disease	0 [0]
Pediatrics	108 [0.04]	Psychology, Clinical	0 [0]
Nutrition & Dietetics	99 [0.04]	Health Policy & Services	0 [0]
Otorhinolaryngology	96 [0.04]	Psychology, Educational	0 [0]
Infectious Diseases	82 [0.03]	Social Sciences, Biomedical	0 [0]
Audiology & Speech-Language Pathology	81 [0.03]	Primary Health Care	0 [0]
Gerontology	81 [0.03]	Andrology	0 [0]
Dermatology	78 [0.03]	Psychology, Psychoanalysis	0 [0]
Substance Abuse	76 [0.03]		

Notes: Number of deep learning papers by WoS subject category. A document can belong to several categories. Shares in [%].

Table 1.13: Health sciences sample and deep learning articles

		Full sample health sciences		*Sample for econometrics*	
Year	# Journals	# Articles	# DL Articles	# Articles	# DL Articles
1990	44	14,317	25		
1991	48	17,809	37		
1992	52	21,029	87		
1993	55	21,295	97		
1994	57	24,458	119		
1995	60	24,632	171		
1996	60	25,072	155		
1997	62	24,155	186		
1998	65	29,891	203		
1999	65	29,254	226		
2000	66	30,239	222		
2001	68	27,272	217	14,427	139
2002	70	31,120	235	14,580	132
2003	70	31,225	256	15,463	162
2004	72	34,686	300	16,924	182
2005	72	35,177	327	17,586	198
2006	77	41,966	412	20,762	250
2007	83	42,947	520	23,510	366
2008	84	41,931	431	23,044	292
2009	86	46,195	420	23,480	293
2010	85	47,384	485	25,103	328
2011	89	52,550	554	30,082	417
2012	89	48,763	559	29,497	426
2013	89	49,814	500	32,112	381
2014	89	57,045	586	34,341	462
2015	90	55,277	701	35,126	532
2016	89	56,232	729		
2017	90	57,146	1,114		
2018	91	62,342	1,646		
Total		1,081,223	11,520	356,037	4,560

Notes: The articles published in the period 1990–2000 are used to build the novelty measures for the first focal year 2001. The articles published in the period 2016–2018 are used to check whether the new combinations of referenced journals are reused in the following three years after the last focal year 2015. The discrepancy between the number of articles in the whole sample and the number in the sample used for econometric analysis is due to the presence of missing information in the variables considered.

Re-combinatorial novelty: indicators

This Appendix complements Section 1.5 with details on the procedure for the construction of novelty measures (Section 1.5.1). It also reports some statistics on the most frequent combinations of Web of Science subject categories, Tables 1.14–1.16. Codes for the variable construction and analysis are fully accessible upon request.

Algorithm for the construction of novelty indicators

The novelty indicators are calculated at the year-level. Let y be the focal year, we compute combinations of referenced journals in scientific documents belonging to three groups:

- All papers published in the focal year y.
- All papers published before the focal year y, B_y
- All papers published 3 years after the focal year y, A_y

In our study the focal year, y, is moving from 2001 to 2015, while the first year for which B_y is calculated remains fixed. We choose the year 2001 as the first focal year to guarantee a sufficiently long time window (1990–2000) over which all previous combinations of referenced journals are assessed.

Suppose a paper P published in year y cites three different journals J_1,J_2 and J_3. This gives rise to three unique combinations: (J_1, J_2), (J_1, J_3), and (J_2, J_3).

- For each of these combinations, we check whether $(J_i, J_j) \in B_y$, and if not, the pair is removed from the analysis – i.e., the combination is simply not new.
- If $(J_i, J_j) \notin B_y$, we examine whether $\sum_{P_{A_y} \in A_y} \{(J_i, J_j) \in A_y\} \geq 5$. If the last statement is FALSE, we remove this pair from the analysis – i.e., the new combination is not reused in the future.[7]
- If $(J_i, J_j) \notin B_y$ & $\sum_{P_{A_y} \in A_y} \{(J_i, J_j) \in A_y\} \geq 5$, then the journal pair combination is considered new and non trivial, hence we add that pair to the set of novel combinations N_y.

[7] As robustness checks, we also considered different thresholds for the re-use, i.e. 3 and 10. By construction, the number of combinations considered as novel increases (decreases) significantly when the threshold is lower (higher). However, as shown in Wang et al. [2017], the dynamics of novelty are not much affected by these alternative specifications.

The difficulty of making new journal combinations are not equally distributed. Journals can share 'common friends' making it possible to create more or less difficult new combinations. For example, P_{iy} is making for the first time the combination (J_1, J_2), but J_1 is usually cited with J_3 and J_2 is also sometimes cited with J_3. Creating this new combination is therefore less difficult compared to two journals that do not share any 'common friends'. To investigate the difficulty of citing J_1 and J_2 for the first time, we construct a co-occurrence matrix of pairs of cited journals on the 3 years preceding the focal year y, and compute a cosine similarity:

$$COS_{(J_1,J_2)} = \frac{J_1.J_2}{\|J_1\| \, \|J_2\|}$$

The difficulty of making the (J_1, J_2) combination is then $1 - COS_{(J_1,J_2)}$. To construct the novelty indicator for the article P_{iy}, we sum up all the difficulties for pairs $\in N_y$ and apply the $log(x+1)$ transformation:

$$Novelty(P_{iy}) = log\Big[\sum_{(J_i,J_j)\in N_y}(1 - COS_{(J_i,J_j)}) + 1\Big]$$

WoS subject categories combinations

Table 1.14: Subject categories combinations (All Sciences)

Combinations [Category A | Category B]	# Combinations	Share [%]
DL articles / 2001–2005	450	
NEUROSCIENCE & BEHAVIOR | NEUROSCIENCE & BEHAVIOR	51	11
NEUROSCIENCE & BEHAVIOR | PSYCHIATRY/PSYCHOLOGY	49	11
CLINICAL MEDICINE | NEUROSCIENCE & BEHAVIOR	22	5
BIOLOGY & BIOCHEMISTRY | COMPUTER SCIENCE	17	4
COMPUTER SCIENCE | NEUROSCIENCE & BEHAVIOR	14	3
Non-DL articles / 2001–2005	39,018	
NEUROSCIENCE & BEHAVIOR | NEUROSCIENCE & BEHAVIOR	2,618	7
CLINICAL MEDICINE | NEUROSCIENCE & BEHAVIOR	2,378	6
BIOLOGY & BIOCHEMISTRY | MOLECULAR BIOLOGY & GENETICS	2,369	6
CLINICAL MEDICINE | CLINICAL MEDICINE	2,036	5
MOLECULAR BIOLOGY & GENETICS | MOLECULAR BIOLOGY & GENETICS	1,927	5
DL articles / 2006–2010	2,266	
NEUROSCIENCE & BEHAVIOR | NEUROSCIENCE & BEHAVIOR	167	7
NEUROSCIENCE & BEHAVIOR | PSYCHIATRY/PSYCHOLOGY	150	7
BIOLOGY & BIOCHEMISTRY | NEUROSCIENCE & BEHAVIOR	108	5
NEUROSCIENCE & BEHAVIOR | PHYSICS	86	4
BIOLOGY & BIOCHEMISTRY | CHEMISTRY	81	4
Non-DL articles / 2006–2010	118,363	
CLINICAL MEDICINE | CLINICAL MEDICINE	6,164	5
NEUROSCIENCE & BEHAVIOR | NEUROSCIENCE & BEHAVIOR	5,444	5
BIOLOGY & BIOCHEMISTRY | MOLECULAR BIOLOGY & GENETICS	4,644	4
NEUROSCIENCE & BEHAVIOR | PSYCHIATRY/PSYCHOLOGY	4,547	4
CLINICAL MEDICINE | NEUROSCIENCE & BEHAVIOR	4,389	4
DL articles / 2011–2015	3,986	
NEUROSCIENCE & BEHAVIOR | PSYCHIATRY/PSYCHOLOGY	302	8
COMPUTER SCIENCE | ENGINEERING	249	6
CLINICAL MEDICINE | NEUROSCIENCE & BEHAVIOR	200	5
PSYCHIATRY/PSYCHOLOGY | PSYCHIATRY/PSYCHOLOGY	188	5
ENGINEERING | ENGINEERING	181	5
Non-DL articles / 2011–2015	328,197	
CLINICAL MEDICINE | CLINICAL MEDICINE	29,295	9
BIOLOGY & BIOCHEMISTRY | CLINICAL MEDICINE	17,817	5
CLINICAL MEDICINE | MOLECULAR BIOLOGY & GENETICS	15,581	5
PSYCHIATRY/PSYCHOLOGY | PSYCHIATRY/PSYCHOLOGY	13,583	4
CLINICAL MEDICINE | NEUROSCIENCE & BEHAVIOR	13,027	4

Notes: This table reports the number and share of the most frequent combinations of WoS subject categories broken down by period and DL status.

Table 1.15: Subject categories combinations (No CS)

Combinations [Category A | Category B]	# Combinations	Share [%]
DL articles / 2001–2005	375	
NEUROSCIENCE & BEHAVIOR | NEUROSCIENCE & BEHAVIOR	51	14
NEUROSCIENCE & BEHAVIOR | PSYCHIATRY/PSYCHOLOGY	49	13
CLINICAL MEDICINE | NEUROSCIENCE & BEHAVIOR	22	6
BIOLOGY & BIOCHEMISTRY | NEUROSCIENCE & BEHAVIOR	13	3
BIOLOGY & BIOCHEMISTRY | BIOLOGY & BIOCHEMISTRY	12	3
Non-DL articles / 2001–2005	37,666	
NEUROSCIENCE & BEHAVIOR | NEUROSCIENCE & BEHAVIOR	2,618	7
CLINICAL MEDICINE | NEUROSCIENCE & BEHAVIOR	2,378	6
BIOLOGY & BIOCHEMISTRY | MOLECULAR BIOLOGY & GENETICS	2,369	6
CLINICAL MEDICINE | CLINICAL MEDICINE	2,036	5
MOLECULAR BIOLOGY & GENETICS | MOLECULAR BIOLOGY & GENETICS	1,927	5
DL articles / 2006–2010	1,989	
NEUROSCIENCE & BEHAVIOR | NEUROSCIENCE & BEHAVIOR	167	8
NEUROSCIENCE & BEHAVIOR | PSYCHIATRY/PSYCHOLOGY	150	8
BIOLOGY & BIOCHEMISTRY | NEUROSCIENCE & BEHAVIOR	108	5
NEUROSCIENCE & BEHAVIOR | PHYSICS	86	4
BIOLOGY & BIOCHEMISTRY | CHEMISTRY	81	4
Non-DL articles / 2006–2010	114,806	
CLINICAL MEDICINE | CLINICAL MEDICINE	6,164	5
NEUROSCIENCE & BEHAVIOR | NEUROSCIENCE & BEHAVIOR	5,444	5
BIOLOGY & BIOCHEMISTRY | MOLECULAR BIOLOGY & GENETICS	4,644	4
NEUROSCIENCE & BEHAVIOR | PSYCHIATRY/PSYCHOLOGY	4,547	4
CLINICAL MEDICINE | NEUROSCIENCE & BEHAVIOR	4,389	4
DL articles / 2011–2015	3,188	
NEUROSCIENCE & BEHAVIOR | PSYCHIATRY/PSYCHOLOGY	302	9
CLINICAL MEDICINE | NEUROSCIENCE & BEHAVIOR	200	6
PSYCHIATRY/PSYCHOLOGY | PSYCHIATRY/PSYCHOLOGY	188	6
ENGINEERING | ENGINEERING	181	6
NEUROSCIENCE & BEHAVIOR | NEUROSCIENCE & BEHAVIOR	154	5
Non-DL articles / 2011–2015	319,990	
CLINICAL MEDICINE | CLINICAL MEDICINE	29,295	9
BIOLOGY & BIOCHEMISTRY | CLINICAL MEDICINE	17,817	6
CLINICAL MEDICINE | MOLECULAR BIOLOGY & GENETICS	15,581	5
PSYCHIATRY/PSYCHOLOGY | PSYCHIATRY/PSYCHOLOGY	13,583	4
CLINICAL MEDICINE | NEUROSCIENCE & BEHAVIOR	13,027	4

Notes: This table reports the number and share of the most frequent combinations of WoS subject categories broken down by period and DL status.

Table 1.16: Subject categories combinations (Only HS)

Combinations [Category A | Category B]	# Combinations	Share [%]
DL articles / 2001–2005	251	
NEUROSCIENCE & BEHAVIOR / NEUROSCIENCE & BEHAVIOR	51	20
NEUROSCIENCE & BEHAVIOR / PSYCHIATRY/PSYCHOLOGY	49	20
CLINICAL MEDICINE / NEUROSCIENCE & BEHAVIOR	22	9
BIOLOGY & BIOCHEMISTRY / NEUROSCIENCE & BEHAVIOR	13	5
BIOLOGY & BIOCHEMISTRY / BIOLOGY & BIOCHEMISTRY	12	5
Non-DL articles / 2001–2005	31,917	
NEUROSCIENCE & BEHAVIOR / NEUROSCIENCE & BEHAVIOR	2,618	8
CLINICAL MEDICINE / NEUROSCIENCE & BEHAVIOR	2,378	7
BIOLOGY & BIOCHEMISTRY / MOLECULAR BIOLOGY & GENETICS	2,369	7
CLINICAL MEDICINE / CLINICAL MEDICINE	2,036	6
MOLECULAR BIOLOGY & GENETICS / MOLECULAR BIOLOGY & GENETICS	1,927	6
DL articles / 2006–2010	1,293	
NEUROSCIENCE & BEHAVIOR / NEUROSCIENCE & BEHAVIOR	167	13
NEUROSCIENCE & BEHAVIOR / PSYCHIATRY/PSYCHOLOGY	150	12
BIOLOGY & BIOCHEMISTRY / NEUROSCIENCE & BEHAVIOR	108	8
BIOLOGY & BIOCHEMISTRY / CHEMISTRY	81	6
CLINICAL MEDICINE / NEUROSCIENCE & BEHAVIOR	68	5
Non-DL articles / 2006–2010	85,342	
CLINICAL MEDICINE / CLINICAL MEDICINE	6,164	7
NEUROSCIENCE & BEHAVIOR / NEUROSCIENCE & BEHAVIOR	5,444	6
BIOLOGY & BIOCHEMISTRY / MOLECULAR BIOLOGY & GENETICS	4,644	5
NEUROSCIENCE & BEHAVIOR / PSYCHIATRY/PSYCHOLOGY	4,547	5
CLINICAL MEDICINE / NEUROSCIENCE & BEHAVIOR	4,389	5
DL articles / 2011–2015	1,921	
NEUROSCIENCE & BEHAVIOR / PSYCHIATRY/PSYCHOLOGY	302	16
CLINICAL MEDICINE / NEUROSCIENCE & BEHAVIOR	200	10
PSYCHIATRY/PSYCHOLOGY / PSYCHIATRY/PSYCHOLOGY	188	10
NEUROSCIENCE & BEHAVIOR / NEUROSCIENCE & BEHAVIOR	154	8
BIOLOGY & BIOCHEMISTRY / NEUROSCIENCE & BEHAVIOR	109	6
Non-DL articles / 2011–2015	238,226	
CLINICAL MEDICINE / CLINICAL MEDICINE	29,293	12
BIOLOGY & BIOCHEMISTRY / CLINICAL MEDICINE	17,817	7
CLINICAL MEDICINE / MOLECULAR BIOLOGY & GENETICS	15,581	7
PSYCHIATRY/PSYCHOLOGY / PSYCHIATRY/PSYCHOLOGY	13,583	6
CLINICAL MEDICINE / NEUROSCIENCE & BEHAVIOR	13,026	5

Notes: This table reports the number and share of the most frequent combinations of WoS subject categories broken down by period and DL status.

Robustness analysis: descriptive statistics and results

This Appendix complements our analysis with descriptive statistics and estimation results for regressions and matching. Tables 1.17–1.19 refer to the sample of articles that are not classified as 'Neurosciences'. Tables 1.20–1.22 refer to the sample of articles that do not contain the terms 'neural_network' and 'neural_networks' in their title, keywords or abstract. Table 1.23 reports the results of the matching exercises. Table 1.24 reports the estimates for the Multinomial Logistic regression to model the novelty/conventionality quadrant [Uzzi et al., 2013, Wagner et al., 2019]. Codes for the variable construction and analysis are fully accessible upon request.

Neuroscience articles excluded

Table 1.17: Descriptive statistics of the variables – Neuroscience articles excluded

	DL Papers	Non-DL Papers	Total
Re-combinatorial Novelty			
Novelty Dummy (All Sciences)	38.17	32.48	32.54
Novelty Dummy (No CS)	32.65	31.56	31.57
Novelty Dummy (Only HS)	18.77	23.69	23.64
Novelty (All Sciences)	0/0.82 (2.16)	0/0.84 (3.41)	0/0.84 (3.4)
Novelty (No CS)	0/0.62 (1.82)	0/0.82 (3.38)	0/0.81 (3.37)
Novelty (Only HS)	0/0.29 (1.17)	0/0.53 (2.58)	0/0.53 (2.57)
Scientific Impact			
Top 5% Cited	7.73	5.59	5.62
Top 10% Cited	14.61	11.02	11.06
# Citations (Raw Count)	14/31.27 (140.01)	17/31.73 (83.93)	17/31.72 (84.69)
Citations (Yearly Normalized)	1.75/3.23 (8.52)	2/3.48 (8.46)	2/3.48 (8.46)
Controls			
# References	32/37.84 (25.46)	30/33.24 (23.26)	30/33.28 (23.29)
# Authors	4/4.03 (2.26)	4/5.03 (3.61)	4/5.01 (3.6)
International Collab.	23.65	21.95	21.97
Private Collab.	6.37	7.56	7.55
JIF	0.86/1.33 (1.26)	1.63/1.98 (1.51)	1.62/1.98 (1.51)
Journal Age	22/29.16 (28.88)	35/41.08 (32.33)	35/40.95 (32.32)
Survey	0.89	0.98	0.98
Time Period	[2001 – 2015]	[2001 – 2015]	[2001 – 2015]
# Scientific Fields	41	43	43
# Journals	54	54	54
# Papers	2,355(1.03%)	225,748(98.97%)	228,103(100%)

Notes: Binary indicators in [%], for continuous measures [median/mean (s.d.)]. The statistics refer to the period used for the econometric analysis.

Table 1.18: Novelty profile of deep learning publications – Neuroscience articles excluded

	Tobit: Novelty			*Probit: Novelty Dummy*		
	All Sciences	No CS	Only HS	All Sciences	No CS	Only HS
	(1)	(2)	(3)	(4)	(5)	(6)
DL	0.030	-0.065	-0.225***	0.046	-0.035	-0.181***
	(0.052)	(0.049)	(0.066)	(0.052)	(0.049)	(0.059)
# References (log)	1.100***	1.104***	1.076***	0.948***	0.949***	0.894***
	(0.037)	(0.036)	(0.034)	(0.033)	(0.032)	(0.025)
# Authors (log)	0.124***	0.129***	0.167***	0.131***	0.135***	0.166***
	(0.021)	(0.021)	(0.026)	(0.020)	(0.020)	(0.022)
International Collab.	-0.036***	-0.041***	-0.074***	-0.035***	-0.039***	-0.068***
	(0.012)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Private Collab.	0.017	0.017	-0.008	0.016	0.016	-0.006
	(0.013)	(0.014)	(0.018)	(0.014)	(0.014)	(0.017)
JIF	0.022	0.025	0.039	0.023	0.026	0.036
	(0.07)	(0.068)	(0.073)	(0.067)	(0.065)	(0.067)
Journal Age (log)	0.007	0.031	0.059	0.015	0.033	0.057
	(0.145)	(0.143)	(0.160)	(0.135)	(0.134)	(0.144)
Survey	0.126***	0.119***	0.077*	0.123***	0.115***	0.074*
	(0.041)	(0.038)	(0.042)	(0.045)	(0.041)	(0.041)
Log Likelihood	-172,590	-168,967	-139,119	-115,102	-113,562	-100,187
χ^2 [Null Model]	74,312***	73,797***	59,245***	57,591***	57360***	49,127***
χ^2 [w/o DL Model]	1.30	5.30**	44.60***	2.60	1.40	31.1***
# Obs	228,103	228,103	228,103	228,103	228,103	228,103

Notes: This table reports coefficients of the effect of deep learning (*DL*, dummy) on recombinatorial novelty built by considering different knowledge landscapes. Bootstrapped (500 replications) standard errors clustered at the journal-level in parentheses: ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The effect of DL on the positive continuous novelty measure is estimated using a Tobit regression (Columns 1–3). The effect on the novelty dummy is estimated using a Probit (Columns 4–6). Each novelty measure is calculated on three different sets of journal references: 'All Sciences' – All cited journals, 'No CS' – All cited journals except for computer science journals, and 'Only HS' – Only citations to health science journals. Constant term, scientific field (WoS subject category) and time fixed effects are incorporated in all model specifications. Likelihood-ratio test are used to compare the goodness of fit of two statistical models: (i) null model against complete model; (ii) model without the *DL* variable against the complete model.

Table 1.19: Impact profile of deep learning publications – Neuroscience articles excluded

		NegBin: # Citations	Probit: Top 5% Cited	Probit: Top 10% Cited
		(1)	(2)	(3)
Panel A: Mean	DL	0.090	0.107*	0.120**
		(0.060)	(0.059)	(0.059)
	Novelty (All Sciences)	0.165***	0.210***	0.194***
		(0.028)	(0.022)	(0.022)
	# References (log)	0.470***	0.367***	0.416***
		(0.062)	(0.103)	(0.086)
	# Authors (log)	0.211 ***	0.154***	0.172***
		(0.032)	(0.054)	(0.050)
	International Collab.	0.068***	0.093***	0.089***
		(0.014)	(0.016)	(0.017)
	Private Collab.	-0.011	-0.009	-0.007
		(0.016)	(0.021)	(0.016)
	JIF	0.222***	0.202***	0.192***
		(0.035)	(0.062)	(0.065)
	Journal Age (log)	0.078*	0.025	0.045
		(0.044)	(0.103)	(0.111)
	Survey	0.551***	0.693***	0.630***
		(0.050)	(0.070)	(0.060)
Panel B: Dispersion	DL	0.164**		
		(0.075)		
	Novelty (All Sciences)	0.097***		
		(0.017)		
	# References (log)	-0.473***		
		(0.040)		
	# Authors (log)	-0.199***		
		(0.036)		
	JIF	0.107**		
		(0.054)		
	Journal Age (log)	-0.123***		
		(0.033)		
Log Likelihood		-955,206	-45,382	-72,968
χ^2 [Null Model]		193,546***	7,890***	12,715***
χ^2 [w/o DL Model]		2.10	6.80**	12.60***
# Obs		228,103	228,103	228,103

Notes: This table reports coefficients of the effect of deep learning (*DL*, dummy) on scientific impact proxied by the number of received citations (Column 1) and 'big hits' (Columns 2 and 3). Bootstrapped (500 replications) standard errors clustered at the journal-level in parentheses: ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The effect of DL on the citation count is estimated using a Negative Binomial regression. Estimates for the expectation and variance are reported in Panel A and B, respectively. The effects on the binary indicators is estimated using a Probit. Constant term, scientific field (WoS subject category) and time fixed effects are incorporated in all model specifications. Likelihood-ratio test are used to compare the goodness of fit of two statistical models: (i) null model against complete model; (ii) model without the *DL* variable against the complete model.

Neural network(s) articles excluded

Table 1.20: Descriptive statistics of the variables – Neural network(s) articles excluded

	DL Papers	Non-DL Papers	Total
Re-combinatorial Novelty			
Novelty Dummy (All Sciences)	37.97	30.05	30.08
Novelty Dummy (No CS)	32.64	29.22	29.23
Novelty Dummy (Only HS)	18.57	22.72	22.71
Novelty (All Sciences)	0/0.78 (1.92)	0/0.74 (3.17)	0/0.74 (3.16)
Novelty (No CS)	0/0.61 (1.67)	0/0.72 (3.14)	0/0.72 (3.13)
Novelty (Only HS)	0/0.26 (0.87)	0/0.49 (2.43)	0/0.49 (2.42)
Scientific Impact			
Top 5% Cited	7.33	6.00	6.00
Top 10% Cited	14.15	11.7	11.71
# Citations (Raw Count)	15/27.88 (41.87)	17/34.95 (84.48)	17/34.93 (84.36)
Citations (Yearly Normalized)	1.78/3.21 (4.95)	2/3.73 (8.23)	2/3.73 (8.22)
Controls			
# References	32/36.56 (22.63)	31/36.13 (25.60)	31/36.14 (25.59)
# Authors	4/4.15 (2.16)	4/4.75 (3.28)	4/4.75 (3.27)
International Collab.	23.06	22.45	22.45
Private Collab.	7.58	6.91	6.92
JIF	0.96/1.3 (1.27)	1.57/2.37 (2.20)	1.57/2.37 (2.20)
Journal Age	23/29.58 (28.77)	31/37.38 (29.04)	31/37.35 (29.04)
Survey	1.17	0.83	0.83
Time Period	[2001 – 2015]	[2001 – 2015]	[2001 – 2015]
# Scientific Fields	45	48	48
# Journals	84	84	84
# Papers	1,201(0.37%)	319,755(99.63%)	320,956(100%)

Notes: Binary indicators in [%], for continuous measures [median/mean (s.d.)]. The statistics refer to the period used for the econometric analysis.

Table 1.21: Novelty profile of deep learning publications – Neural network(s) articles excluded

	Tobit: Novelty			*Probit: Novelty Dummy*		
	All Sciences	No CS	Only HS	All Sciences	No CS	Only HS
	(1)	(2)	(3)	(4)	(5)	(6)
DL	0.083	0.003	-0.171***	0.091*	0.014	-0.137**
	(0.051)	(0.052)	(0.061)	(0.053)	(0.057)	(0.058)
# References (log)	1.046***	1.050***	1.025***	0.880***	0.880***	0.838***
	(0.032)	(0.032)	(0.033)	(0.026)	(0.026)	(0.023)
# Authors (log)	0.186***	0.194***	0.241***	0.191***	0.197***	0.233***
	(0.023)	(0.024)	(0.027)	(0.022)	(0.022)	(0.024)
International Collab.	-0.058***	-0.064***	-0.095***	-0.055***	-0.061***	-0.086***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.009)
Private Collab.	0.001	0.001	-0.023	0.001	-0.001	-0.021
	(0.012)	(0.013)	(0.015)	(0.012)	(0.012)	(0.014)
JIF	-0.040**	-0.037*	-0.029	-0.037**	-0.034*	-0.026
	(0.020)	(0.021)	(0.021)	(0.018)	(0.018)	(0.018)
Journal Age (log)	-0.092	-0.077	-0.040	-0.069	-0.056	-0.026
	(0.103)	(0.106)	(0.115)	(0.094)	(0.096)	(0.101)
Survey	0.204***	0.195***	0.160***	0.192***	0.184***	0.146***
	(0.042)	(0.040)	(0.043)	(0.045)	(0.042)	(0.042)
Log Likelihood	-234,600	-230,021	-194,470	-160,685	-158,739	-142,454
χ^2 [Null Model]	90,036***	88,839***	70,498***	71,192***	70,357***	58,980***
χ^2 [w/o DL Model]	4.70*	0.02	12.80***	5.30**	0.10	9.4***
# Obs	320,956	320,956	320,956	320,956	320,956	320,956

Notes: This table reports coefficients of the effect of deep learning (*DL*, dummy) on recombinatorial novelty built by considering different knowledge landscapes. Bootstrapped (500 replications) standard errors clustered at the journal-level in parentheses: ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The effect of DL on the positive continuous novelty measure is estimated using a Tobit regression (Columns 1–3). The effect on the novelty dummy is estimated using a Probit (Columns 4–6). Each novelty measure is calculated on three different sets of journal references: 'All Sciences' – All cited journals, 'No CS' – All cited journals except for computer science journals, and 'Only HS' – Only citations to health science journals. Constant term, scientific field (WoS subject category) and time fixed effects are incorporated in all model specifications. Likelihood-ratio test are used to compare the goodness of fit of two statistical models: (i) null model against complete model; (ii) model without the *DL* variable against the complete model.

Table 1.22: Impact profile of deep learning publications – Neural network(s) articles excluded

		NegBin: # Citations (1)	Probit: Top 5% Cited (2)	Probit: Top 10% Cited (3)
Panel A: Mean	DL	0.110*	0.136*	0.153**
		(0.067)	(0.070)	(0.064)
	Novelty (All Sciences)	0.138***	0.190***	0.181***
		(0.022)	(0.017)	(0.016)
	# References (log)	0.517***	0.436***	0.485***
		(0.061)	(0.075)	(0.063)
	# Authors (log)	0.248***	0.179***	0.206***
		(0.031)	(0.040)	(0.038)
	International Collab.	0.070***	0.088***	0.090***
		(0.014)	(0.015)	(0.014)
	Private Collab.	-0.034**	-0.031	-0.04**
		(0.017)	(0.019)	(0.016)
	JIF	0.202***	0.155***	0.168***
		(0.022)	(0.018)	(0.019)
	Journal Age (log)	0.063*	-0.043	-0.032
		(0.038)	(0.093)	(0.089)
	Survey	0.522***	0.646***	0.607***
		(0.055)	(0.056)	(0.051)
Panel B: Dispersion	DL	0.075		
		(0.053)		
	Novelty (All Sciences)	0.086***		
		(0.017)		
	# References (log)	-0.488***		
		(0.039)		
	# Authors (log)	-0.202***		
		(0.043)		
	JIF	0.037		
		(0.03)		
	Journal Age (log)	-0.116***		
		(0.032)		
Log Likelihood		-1,360,967	-63,884	-101,311
χ^2 [Null Model]		282,883***	17,961***	29,217***
χ^2 [w/o DL Model]		1.60	5.50**	10.40***
# Obs		320,956	320,956	320,956

Notes: This table reports coefficients of the effect of deep learning (*DL*, dummy) on scientific impact proxied by the number of received citations (Column 1) and 'big hits' (Columns 2 and 3). Bootstrapped (500 replications) standard errors clustered at the journal-level in parentheses: ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The effect of DL on the citation count is estimated using a Negative Binomial regression. Estimates for the expectation and variance are reported in Panel A and B, respectively. The effects on the binary indicators is estimated using a Probit. Constant term, scientific field (WoS subject category) and time fixed effects are incorporated in all model specifications. Likelihood-ratio test are used to compare the goodness of fit of two statistical models: (i) null model against complete model; (ii) model without the *DL* variable against the complete model.

Table 1.23: Novelty and impact profile – Matching

	Exact Matching		*Propensity Score Matching*	
	(1)	(2)	(3)	(4)
Novelty (All Sciences)	0.054***	0.053***	0.035***	0.023
Novelty (No CS)	0.026**	0.026**	0.008	-0.001
Novelty (Only HS)	-0.005	-0.005	-0.025**	-0.033***
# Citations	0.192***	0.195***	0.102***	0.063**

Notes: This table reports Average Treatment Effect on the Treated (ATT) for novelty and impact variables. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The set of variables used for each matching is composed as follows: (1) Journal / WoS Categories / Publication Year; (2) All dummy variables in our set of control variables / Journal / WoS Categories / Publication Year; (3) Number of authors (log) / Number of References (log) / Journal / WoS Categories / Publication Year; (4) All Variables.

Atypical combinations in deep learning publications

Table 1.24: Atypical profile of deep learning publications

	Category	All Sciences	No CS	Only HS
		(1)	(2)	(3)
DL	HC–HN	0.008	0.208	0.308**
		(0.130)	(0.133)	(0.136)
	HC–LN	-0.041	0.090	-0.049
		(0.157)	(0.152)	(0.154)
	LC–LN	-0.043	-0.086	0.021
		(0.162)	(0.163)	(0.155)
# References (log)	HC–HN	-0.198***	-0.216***	-0.168***
		(0.066)	(0.065)	(0.061)
	HC–LN	-0.687***	-0.668***	-0.711***
		(0.066)	(0.064)	(0.063)
	LC–LN	-0.460***	-0.463***	-0.550***
		(0.063)	(0.060)	(0.062)
# Authors (log)	HC–HN	-0.392***	-0.393***	-0.433***
		(0.060)	(0.060)	(0.066)
	HC–LN	-0.557***	-0.597***	-0.603***
		(0.078)	(0.077)	(0.086)
	LC–LN	-0.260***	-0.254***	-0.299***
		(0.048)	(0.047)	(0.050)

Continued on next page

Table 1.24: Atypical profile of deep learning publications – continued.

	Category	All Sciences	No CS	Only HS
		(1)	(2)	(3)
International Collab.	HC–HN	0.103**	0.160***	0.128***
		(0.042)	(0.043)	(0.044)
	HC–LN	0.096**	0.155***	0.141***
		(0.041)	(0.040)	(0.043)
	LC–LN	-0.013	0.052	0.119***
		(0.047)	(0.044)	(0.044)
Private Collab.	HC–HN	-0.050	-0.067	0.045
		(0.069)	(0.071)	(0.072)
	HC–LN	0.010	-0.108*	-0.093
		(0.063)	(0.060)	(0.062)
	LC–LN	0.052	-0.016	0.025
		(0.068)	(0.069)	(0.071)
JIF	HC–HN	0.134***	0.145***	0.146***
		(0.035)	(0.035)	(0.038)
	HC–LN	0.117***	0.105***	0.092***
		(0.032)	(0.033)	(0.035)
	LC–LN	-0.087	-0.114*	-0.116
		(0.062)	(0.062)	(0.075)
Journal Age (log)	HC–HN	-0.068	-0.064	-0.050

Continued on next page

Table 1.24: Atypical profile of deep learning publications – continued.

	Category	All Sciences	No CS	Only HS
		(1)	(2)	(3)
		(0.196)	(0.189)	(0.194)
	HC–LN	-0.207	-0.158	-0.178
		(0.173)	(0.168)	(0.176)
	LC–LN	-0.055	-0.089	-0.224
		(0.241)	(0.24)	(0.258)
Survey	HC–HN	-0.399	-0.294	-0.492
		(0.339)	(0.348)	(0.328)
	HC–LN	0.458**	0.096	0.472**
		(0.225)	(0.209)	(0.204)
	LC–LN	0.892***	0.592***	0.779***
		(0.224)	(0.211)	(0.211)
Log Likelihood		-374,002	-374,000	-363,855
χ^2 [Null Model]		95,913***	95,488***	115,891***
χ^2 [w/o DL Model]		259***	158.20***	144***
# Obs		320,587	320,587	320,587

End of table.

Chapter 2

Barriers and Drivers of AI Adoption in Science

This chapter was co-authored with

Stefano Bianchini and Moritz Müller

Summary of the chapter

This article explores the factors influencing the adoption and reuse of AI in scientific research. We focus on the role of scientific and technical human capital (STHC) of domain scientists, assessed through the institutional and social environment in which they are embedded along with their individual characteristics. Using data from OpenAlex over the period 2012-2020, we show that collaborations with early-career researchers and past interactions with scientists with backgrounds in computer science and AI are strongly correlated with AI adoption. The institutional environment also plays a significant role in the first part of the process (trying out AI), but is less influential in determining AI reuse. Also, access to computational resources does not generally correlate with AI adoption. At the individual level, we show that scholars with a taste for exploration are more likely to adopt new computational technologies, but at the same time, the likelihood decreases when scientists have acquired a dominant position in their research domain.

2.1 Introduction

Artificial Intelligence (AI) is making its way into science. And let's be honest, this comes as no surprise. The number of publications on AI has witnessed an almost five-fold increase compared to a decade ago, with more than 200,000 papers by 2022, accounting for about 5 % of the total volume of scientific publications. Most of this research has gradually shifted from core AI to its application, which currently represents some 70 % of scientific activity[1]. This trend indicates a growing inclination among scientists to integrate AI tools into their research methodologies. Taken together, these figures raise two important questions: What motivates scientists to adopt AI? And what underlying factors influence researchers to incorporate AI in their work?

Answering the first question is relatively straightforward. Scientists adopt AI because of its high perceived benefits. Although some may argue that the rapid adoption of AI is merely a "fad effect", it is undeniable that the technology has shown tremendous potential for enhancing research in various fields, especially at a time when new ideas are ostensibly getting harder to find [Bloom et al., 2020]. In recent years, indeed, AI have been successfully used in such diverse areas as predicting the 3D structure of proteins [Jumper et al., 2021], regulating nuclear fusion plasma in the tokamak configuration [Degrave et al., 2022], predicting the formation of the structure of the Universe [He et al., 2019], and creating a map of the brains of small insects [Winding et al., 2023]. Take note that examples could be multiplied *ad nauseam*, in virtually all scientific fields, largely exhausting the available space for discussion. Motivated by this pervasiveness, in some recent research we have shown that the impact of AI on research outcomes can be significant, though highly uncertain, which has led us to conclude that AI – deep learning in particular – qualifies as a "emerging general method of invention" [Bianchini et al., 2022].

However, the picture we have provided in our previous works, although we think interesting, is incomplete in that we have focused exclusively on the impact of technology but not on what happens upstream, that is, what are the factors that motivate researchers to integrate AI into their work. So, as the significance of AI in science grows, it becomes critical to understand what factors support the adoption of AI and can facilitate the democratization of the technology throughout the scientific system, ensuring that no one is left behind. This study aims to provide novel insights into

[1]These numbers are part of a project named 'Trends on the diffusion of AI in science' for the European Commission (Arranz D., Bianchini S., Ravet J. and De Girolamo L.) – Forthcoming

this issue[2].

The theory of scientific and technical (S&T) human capital provides us with a solid conceptual framework for studying the mechanisms of AI diffusion in the sciences. The framework describes the various resources that individuals continuously draw upon to create knowledge, which we believe are also viable explanations for why some researchers may adopt AI for the first time, and possibly continue to use the technology thereafter. Resources can be divided into two broad categories: those that reside within the individual and those that are anchored in the relationships between the individual and their working environment [Bozeman et al., 2001, Bozeman and Corley, 2004].[3]

Among the resources that are "internal" to the researcher – *human capital* endowments – we find any individual's scientific capabilities, often classified into three (presumably) overlapping categories, namely cognitive skills, scientific and technical knowledge, and contextual skills. Cognitive skills can be thought of as those innate abilities, such as problem-solving and memory, that are largely independent of context, although can interact and change with context. Scientific and technical knowledge, on the other hand, is obtained through formal training and education on specific theories and explanations. And, finally, contextual knowledge is the type of craft knowledge acquired by doing research and provides heuristics for problem-solving in particular circumstances, although these heuristics can also be transferred to other contexts. Different individuals have different internal endowments, some of which may be more conducive to pushing an individual to adopt an emerging technology for scientific aims – i.e., AI in the case at hand.

Yet scientists do not exist in a social vacuum. The production of scientific knowledge is inherently a social enterprise, just like, as we argue in this paper, the integration of new tools into scientific practices. Scientists employ therefore a wide variety of network-mediated resources – *social capital* endowments – to do their work. Two types of "external" resources are of particular importance here. First, social network

[2]Two observations are worth noting. First, our study focuses exclusively on scientists who use AI in their work, rather than those who develop it. Second, a researcher may not be directly using AI technology, but still qualifies as an adopter because they are involved in a project that uses it in some capacity. Therefore, we can reasonably assume that the researcher has some understanding of the advantages and limitations of AI in their field of research, even if they are not the primary user. Hereafter, we will use the terms adoption and integration of AI technology into scientific practice interchangeably.

[3]Formally, S&T human capital is defined as the sum of scientific, technical and social knowledge, skills and resources embodied in a particular individual – that is, "*an expanded notion of human capital when paired with a productive social capital network*" [Bozeman et al., 2001, p. 6].

ties with peers. Scientists do not have the same education and training, they belong to disciplines with unique cultures and routines, and thus the tacit components of their understanding of science are not the same. This is crucial because through formal collaborations and informal communications, scientists can acquire and employ complementary skills and technical resources to create and transform knowledge and ideas in ways that would not be possible in an isolated context [Bozeman and Corley, 2004, Taylor and Greve, 2006, Lee et al., 2015]. There is broad agreement that some degree of team diversity can actually facilitate knowledge creation [Phillips and Malone, 2014, Leahey, 2016, Ayoubi et al., 2017]. Second, we must not forget the institutional setting in which the research process takes place, the second type of external resource(s). The institutional setting encompasses several factors that shape the practice of science within institutions and organizations, from physical infrastructure and funding to ethical norms of scientific conduct; and some environments are more conducive to the production of impactful science than others [Fox, 1991, Heinze et al., 2009, Fortunato et al., 2018].

A key implication of the S&T human capital framework is that the adoption and benefits of AI in research are contingent upon a conjunction of multiple factors, including equipment, material resources, organizational and institutional frameworks, and the human capital embodied in individuals. A second implication is that while personal knowledge and know-how are certainly important, they may not always be necessary, as social capital can serve as a suitable substitute, and vice versa. And a final implication is that the mere pooling of resource elements is not sufficient to ensure success. It takes a final "ingredient", namely, the quality of fit – or amalgamation – between all available resources.

In this paper, we consider three dimensions that can motivate domain scientists – defined as those individuals who have never published in computer science outlets in their lifetime – to adopt AI in their work: their pre-existing knowledge, skills, and taste for experimentation; the knowledge and expertise of their peers; and the institutional setting in which the researcher is embedded. To measure the scientific knowledge and expertise of individuals, we rely on their past publication activity in terms of thematic diversity, impact, and other dimensions. In assessing the institutional setting, we consider the quality of the researcher's home institution and whether it has a computer science department. Our analysis also takes into account the accessibility of computational resources such as high-performance computing. While earlier AI research relied on a synergy of algorithms, hardware, and specialized software, modern AI heavily depends more and more on computational power. Some studies have highlighted a disparity in computational resources across AI research, with non-elite universities and small non-technological firms struggling to produce impactful research due to limited access to these resources [Ahmed and Wahed, 2020].

We rely on OpenAlex to follow domain scientists who started to use AI in their research between 2012 and 2020. Focusing on three dimensions: institutional, social, and individual factors, we found that AI specialization of institutions substantially fosters AI adoption across various scientific fields, while the influence of access to high-performance computing and institutional ranking may be less decisive. Social connections, particularly with AI experts, computer scientists and early-career researchers, significantly promote AI adoption. Furthermore, researchers with diverse cognitive profiles are more likely to embrace AI. When considering the reuse of AI in subsequent articles, our findings showed that it is considerably influenced by the composition of the research team in the initial AI article. Working with individuals with prior AI experience or early-career researchers seems crucial to facilitate future usage of the technology. Importantly, these patterns of AI adoption exhibit field-specific variations, reflecting the unique contexts and demands of each scientific discipline.

2.2 Conceptual framework and hypotheses

In this section, we apply a revised version of the S&T framework to the context of AI adoption in science and establish a set of testable hypotheses. We begin by exploring the influence of social relations and network ties among scientists, as well as the role of the institutional environment in which they operate, including access to computational resources. We then discuss the role of internal resources, such as formal education and past experience, along with other individual traits of the researcher. In articulating our hypotheses, we provide some information on the measurement of variables.

2.2.1 External resources

2.2.1.1 Human capital and about-knowledge from social ties

Existing social connections may play a significant role in AI adoption through at least two channels: becoming part of a productive team dealing with AI and the ability to judge potential relevance of AI for one's own research.

Social ties offer a form of social capital that can be advantageous for researchers when adopting AI, as they might directly contribute to the initial AI project. As Bozeman (2001) emphasized, *At the project ST human capital level, the focus is on the aggregate of all project participants' endowments and social connections, as well as the physical and economic resources available to a project* [Bozeman et al., 2001, p. 20]. For instance, an applied chemist may collaborate with computer scientists to use AI methodologies in a joint research endeavor. In this case, collaborating computer scientists could be either previous collaborators or discovered through an existing social network. Either way, having prior collaborations with computer scientists may be helpful in establishing collaborations with them in the (future) AI project.

The second pathway entails how past interactions may influence a scientist's perceptions and interpretations of novel technological advancements such as AI. To embark on a new field, what knowledge should a scientist possess? The prevailing belief is that for a scientist who wants to venture into a new field or incorporate methodological tools from that field into their own research, it is desirable to 'know more' about *what is going on* in the field. While we do not dismiss that knowing more is desirable (albeit one may wonder in what amount) we believe it is not the only prescription for a scientist to build bridges between domains. Nor is it perhaps the most efficient prescription. Science is simply too big and, as a result, there are cognitive limits that prevent individuals from fully understanding disciplinary tools, knowledge architectures, and associated 'best practices' across its various domains. This is especially true for emerging technologies, whose potential and applications in science are not yet entirely clear [Rotolo et al., 2015]. What specifically needs to be known then? And for what purposes?

The theory of 'about-knowledge' offers a compelling answer. It suggests the existence of a specific type of knowledge, namely *about-knowledge* or *connective knowledge*, which helps scientists recognize the potential that could be realized by merging their own expertise with knowledge from other fields – in our case, realize the potential of AI as a scientific tool in a given application domain.

About-knowledge can be thoughts as the kind of 'know-how' that is necessary to achieve an early point of connection with another field – "*[A] range of fairly simple facts and information about the sort of problem domains and approaches that populate different fields and specialisms*" [Priaulx and Weinel, 2018, p. 8]. It is not an intimate understanding of how the scientist's core domain expertise interacts with another field, but rather a set of cognitive foundations into the kind of contributions that the field can make to their own expertise, as well as a broad understanding of the practical settings, languages, sub-cultures, expectations and reward models that regulate that field.[4]

It is clear that the lack of insights into what AI scientists do and/or where AI research is headed significantly decreases the likelihood that a scientist will recognize the relevance of the field's contribution to their own work. In this sense, about-knowledge should fill the deficit by providing a wide-angled lens of the potential of the technology, hence prompting the scientist to think "*Maybe it's worth giving it a try!*".

One may argue that about-knowledge is not knowledge at all. Instead, it is simply a series of decontextualized facts or accounts of popular understanding. But this is precisely what makes the concept of about-knowledge so important to our research, as our hypothesis is that the 'big picture' about AI acts as an initiating force that motivates scientists to pursue the idea of incorporating it into their research. This brings us to our first research hypothesis, that is:

H1: *A broader about-knowledge of AI increases the likelihood that a domain scientist incorporates AI into their research.*

But H1 opens the way for a couple of other important questions. First, how do researchers acquire about-knowledge related to AI? Second, how does this knowledge cross disciplinary boundaries? And, third, how can we eventually operationalize the concept of about-knowledge in practice?

In recent years numerous initiatives have encouraged the next generation of scientists to pursue interdisciplinary programs with a focus on AI and data science more generally. These efforts are rooted in the conviction that formal education and training are necessary to close the knowledge gap in this field – which may be

[4] It is noteworthy that this type of knowledge and resulting connections, which we believe can be immensely valuable, particularly in the initial stages of the collaborative life cycle, are often overlooked in the literature on interdisciplinarity and transdisciplinarity.

partially true, although discussing in depth the true benefits of interdisciplinarity training is beyond the scope of our research. The mechanisms of knowledge transmission, especially about-knowledge, however, may be much simpler. We will simply contend that even a little insight into other fields can help the scientist understand the role that these fields can play in their work and, as in the context of AI, help see the potential of the technology and address misconceptions and non-conceptions that would otherwise remain mistakenly overlooked.

And here is where the S&T framework finds a link with the theory of about-knowledge: it is about people embedded within the collaboration networks and populating the same institutional environment who can facilitate knowledge exchange and mediate interactions. Social ties are simply the most critical vehicle for enhancing about-knowledge connectivity across diverse domains.

So let us start from collaboration networks. Here, we use the network of past collaborations of domain scientists as a valuable historical record of their interactions and collaborations with peers, particularly those with some experience in AI technology within the same application domain or in computer science more in general. It is reasonable to assume, in fact, that scientists who are embedded in a network where their peers have already proven experience with AI (e.g., as evidenced by at least one publication) have more incentives to adopt the technology in their own research, especially if their colleagues have achieved successful outcomes. This is because, even if a collaboration does not occur with a past collaborator, being part of a network reinforces a scientist's rudimentary knowledge about the potential of the technology. More formally, we argue that:

H1a: *Prior ties to scientists with AI relevant human capital increase the likelihood that a domain scientist incorporates AI into her research.*

The second channel through which knowledge and about-knowledge can reach a domain scientist is through social interactions with their peers who work in the same institution. Organization science has long established that the location of an actor's contacts in the social structure can offer advantages to the actor when it comes to acquiring information and resources, as do attributes that are rooted in their interactions, such as trust and trustworthiness [Tsai and Ghoshal, 1998]. This is because communication is a complex and often an arduous process that requires individuals to converge on a common sense and is thus facilitated by both spatial and cognitive proximity[5].

It is plausible to assume that individuals who work in institutions that specialize in AI research have easier access to knowledge and about-knowledge pertaining to AI. On the one hand, members of the same institution have more opportunities to spend time together on social occasions and hence more opportunities to exchange ideas and resources freely.[6] On the other hand, members of an institution working in close (geographical and cognitive) proximity may exhibit mimetic isomorphic behaviors, whereby they tend to adopt similar structures, practices, and strategies to their peers, a concept first described to explain what makes organizations so similar [DiMaggio and Powell, 1983, Mizruchi and Fein, 1999]. We believe that this phenomenon can also be observed in scientific research, where researchers may adopt similar research designs, methods, and theoretical frameworks to those used by their colleagues within the same institution, particularly when facing uncertainty or ambiguity, as we can assume in the context of AI adoption. Taken together the above arguments lead us to the following research hypothesis:

H1b: *A prevalence of AI research within an institution increases the likelihood that a domain scientist incorporates AI into their research.*

2.2.1.2 Mentorship and newbies

What is more important than the mentor-newbie relationship when it comes to social ties and knowledge transfer? The term ‘mentor’ typically denotes an experienced individual who imparts their skills and knowledge to a younger person, often someone identified as promising and part of the next generation (e.g., post-doctoral researchers, PhD students, or junior untenured researcher) [Archibugi, 2021]. Under the right circumstances, a mentorship collaboration can facilitate the transfer of various S&T human capital assets, such as craft-skills, know-how, contacts with other peers, industry and funding agents, and more. However, in the context of new methods for scientific discovery and fresh ways to approach scientific problems through

[5]It should be noted that in this context, ‘cognitive’ refers more to the understanding of collective objectives that are shared by a group of individuals or an organization (see, e.g., Coleman [1988]), rather than the similarities of knowledge bases between individuals. The literature suggests that reciprocity (i.e., a favor for a favor; an action for an action) and a sense of contribution to the organization are two key factors that encourage knowledge and information sharing between individuals within an organization [Cummings, 2004, Wang and Noe, 2010]

[6]Admittedly, brilliant suggestions for our own research come more often from casual conversations and informal communications than from formalized meetings and events, don’t they?

AI, it is reasonable to assume that the flow of assets could also occur in the reverse direction, that is *from the junior to the mentor*. And we have strong evidence to support this conjecture.

The academic job market is rich in human resources specializing in AI. According to some recent statistics, the number of AI/ML-related curricula has increased more than any other curriculum in recent years and is unlikely to slow down in the years to come. For instance, in 2020 alone, over 30,000 undergraduate students in the US completed a computer science degree, and one in every five students who earned a PhD degree in computer science specialized in AI/ML [7]. We can expect similar figures in many other countries. AI-related courses are no longer limited to computer science departments at the undergraduate level; rather, a growing number of universities offer interdisciplinary programs that combine AI/ML with other fields. The new generation of scientists also has at their disposal a plethora of online resources offered by universities and private companies that focus specifically on AI/ML. One example is Massive Open Online Courses (MOOCs), which are emerging as an affordable and popular option for those who want to deepen their knowledge of AI/ML, from introductory courses to others on cutting-edge algorithms and advanced applications.[8]

In summary, we are confident that young researchers who are well-versed with AI techniques and tools can bring new perspectives and insights to more experienced colleagues who are often stuck in doing science "as usual". Empirically, we will identify newbies as authors who have published for the first time in a given year. We posit that:

H2: *Collaboration with early-career researchers increases the likelihood that a domain scientist incorporates AI into their research.*

2.2.1.3 Computational resources

While AI is commonly perceived as an intangible technical system, it is *de facto* rooted in physical infrastructure and hardware. Yet, the role of physical assets and

[7] See Stanford AI Index Report from 2022 here: https://aiindex.stanford.edu/report/

[8] The importance of AI literacy from the early stage of education has also been recognized globally. A recent report by UNESCO (2022) highlights the commitment of several countries to developing AI literacy and competencies in K-12 schools. Generally, these initiatives aim to prepare new generations for a world in which AI will be ubiquitous, and thus understand the power and versatility of this technology along with its ethical dilemmas.

their associated computing capabilities – also known as AI compute – have been largely overlooked in policy circles and scholarly literature.[9]

AI compute can be understood as "*one or more stacks of hardware and software used to support AI workloads and applications in an efficient manner*" [OECD, 2023, p. 20]. For machine learning systems, it is clear that compute can facilitate three key steps in scientific pipeline: (i) processing and cleaning large data, (ii) training models and calibrating them (e.g., determining the value of weights of a neural network from the data presented to the model), and (iii) inferencing, which is using the trained model for a specific application to determine an output. Of course, the computing requirements can vary considerably depending on the application, ranging from large high-performance computing (HPC) clusters to smaller laptops and workstations.

Cutting-edge research in ML has become synonymous with access to large computing infrastructures and expertise to exploit them. Sevilla et al. [2022] carried out a detailed investigation of the computational requirement of 123 milestone ML models over time and showed that since the 2010s, the amount of computation required to accommodate modern machine learning systems has soared, with an impressive 5.7-month doubling time (see also Amodei and Hernandez [2018] for estimates with different assumptions) – just for comparison, Moore's law has a 2-year doubling period. While not all researchers use state-of-the-art and computationally intensive ML systems, having access to computing resources can still make a significant difference and, reasonably, be a major driver of AI adoption. How then can scientists access computing resources?

Researchers have various options for accessing AI compute, including data centers or supercomputers located in physical facilities, public or private cloud computing services, and decentralized access at the edge of devices, such as mobile IoT devices. It can be difficult to empirically determine which resource(s) a researcher relies on for their work, yet we contend that the local availability of computing resources, whether within their institution or through collaborators, may serve as a motivating factor for researchers to adopt AI for the first time and potentially use it again.This is not just because researchers can handle larger and more complex datasets and get results faster than they could with limited computing resources, but also – and particularly – because of the institutional culture that embraces AI, as we discussed in Section 2.2.1. Scientists are well-aware that computing resources are readily available and

[9]One reason for that is the lack of standardized and validated data on computing resources. National and institutional and data on the supply and demand of AI compute is not easily accessible and, in some cases, considered sensitive proprietary information.

can potentially support their work; they also know that they can rely on support services to help them use these resources more efficiently and effectively.

Here, we measure the availability of AI compute by the presence of an HPC cluster within the focal researcher's organization (although it should be noted that such compute infrastructure can also be used for non-AI workloads such as mathematical modeling and simulations). In very general terms, HPC is a technology that uses clusters of powerful processors, working in parallel, to process data and solve complex problems at high speeds [OECD, 2023]. Unlike standard computing systems, HPC systems can handle multiple tasks simultaneously across multiple computer servers or processors with a centralized scheduler that manages the computing workload. The high cost of HPC can put this technology out of reach for most organizations, resulting in a significant "compute divide" within and between countries and institutions, as well as between the private sector and academia Ahmed and Wahed [2020]. This is especially true for AI applications in some data-intensive scientific fields such as bioinformatics or particle physics where ML training and inferencing can be highly demanding in terms of memory and computational resources. The existence of a computational divide can therefore impede the adoption of AI and generate disparities in the productivity gains that AI can offer to science. In summary, our hypothesis is that:

H3: *The presence of HPC cluster within a researcher's organization increases the likelihood of integrating AI into research.*

2.2.2 Internal resources

We now turn to the internal resources of the domain scientist, which can be broadly classified into three, somewhat overlapping, categories: cognitive skills, scientific and technical knowledge, and contextual skills [Bozeman et al., 2001].

2.2.2.1 Scientific background and experience

Let us start with the most straightforward, scientific and technical knowledge. This is the type of knowledge acquired through formal scientific education. It involves a thorough understanding of particular theories, experimental and research findings, and the ability to anticipate where research in a particular area is heading. From a Kuhnian perspective, scientific and technical knowledge enables the scientist to feel part of a specific epistemic community, to be accepted by their peers as a member of that community, and ultimately to adopt the shared scientific paradigm [Kuhn, 1962, Ch.2 and 3]. Contextual skills can be viewed instead a subset of scientific and technical knowledge and relate more closely to the type of knowledge gained from practical research experience. Unlike scientific and technical knowledge, contextual skills often involve a tacit component that can only be obtained "on-the-job", that is, in the process of doing research.

In our study, we use a scientist's first field of activity (i.e., domain of the first publication) as a proxy for formal scientific education and context skills. While we do not advance any specific research hypotheses, we believe that this variable is essential to account for idiosyncratic differences in the propensity to adopt AI that may arise due to an individual's scientific background.

2.2.2.2 Taste for exploration

The third dimension is about cognitive skills, that can be viewed as those mental abilities and processes that allow individuals to perceive, process, and use information in a given environment. As such, they necessarily relate to science, but not exclusively to it; they include skills such as reasoning, learning, and others. Here, we are particularly interested in the dispositions or traits that underlie many cognitive skills and processes. One of these is a *taste for exploration*.

Exploration is intimately linked to curiosity, a personal trait that prompts individuals to explore uncharted territories. We think that curiosity is a useful construct for understanding scientists' behavior in terms of technology adoption. Although psychologists have not reached a consensus on its definition, it is generally accepted that curiosity involves an intrinsic motivated desire for new information – an "appetite for knowledge", or more formally "*a form of cognitively induced deprivation that arises from the perception of a gap in knowledge or understanding*" [Loewenstein, 1994, p. 75] [10].

But the curiosity to explore uncharted territories creates some tensions. In the sociology of science, this strategic tension is commonly referred to as 'succession' versus 'subversion' [Bourdieu, 1975]; in organization science and innovation as 'ex-

[10] It could be argued that an individual may also be curious about the topics she knows best. However, it should be noted that our definition of curiosity extends beyond the inclination of a scientist to expand her understanding within the area she is most knowledgeable about – a characteristic that should be common to every scientist! – but rather encompasses the search for knowledge and information far out in the knowledge space.

ploitation' versus 'exploration' [March, 1991, Gupta et al., 2006]. Where do the tensions come from? Science can be viewed as a competitive territory in which scientists have to strategically choose what to study and what to cite. Compared to the returns from staying within the boundaries of the discipline, the returns from exploring other fields are systematically less certain, more distant in time, and often negative. Hence, once a scientist occupies a dominant position in a specific field, it is clear that deviating from the *habitus* can be perceived as a "risky gamble". Increasing returns from experience can trap individuals in exploiting old certainties, refining and extending existing skills, whose returns are proximate and predictable [March, 1991]. A conservative strategy allows scientists to secure publication more likely and benefit from the S&T human capital they have accumulated. On the other hand, transcending local search space and accessing more distant knowledge opens up opportunities for originality, a prime requisite of academic reward and long-term reputation [Foster et al., 2015]

How does all of this relate to the integration of AI into a scientist's research practice? We argue that *epistemic-specific* curiosity, or the desire for new knowledge or a particular piece of information [Wagstaff et al., 2021], is a relevant driver of AI adoption. In the realm of AI, curiosity can arise spontaneously when some situational factors alert an individual to the existence of potential in that domain. Situational factors can be of various kind, from exposure to a sequence of events (e.g., seminars, online information) to the possession of information by someone else – in line with our discussion in Section 2.1.1. Regardless of the specific factor, scientists with a general inclination to explore 'new stuff' will be more likely to envision potential applications of the technology and recognize its relevance to their field of expertise. However, as mentioned earlier, strong tensions may arise when scientists hold dominant positions within their fields, which may lead them to resist solutions that diverge from established practices and avoid venturing beyond their disciplinary boundaries.

Past scientific activity is a visible consequence of research choices, including a taste for exploration; and citations provide evidence of others' judgment of the relevance of a scientist's work. We will use both measures to test the following hypotheses:

H4a: *A higher 'taste for exploration' increases the likelihood of integrating AI into research.*

H4b: *A higher scientific reputation and recognition decrease the likelihood of integrating AI into research.*

2.3 Data and Methods

2.3.1 Data

Our main interest is the adoption of AI as a research method in science. More specifically, we investigate whether and how the STHC endowment of a non-computer scientist is related to her decision to adopt AI methods in her research. We measure AI adoption on scientific publications, namely i) publishing one first paper applying AI, and ii) reusing AI in at least a second paper.

Data set This question is investigated using scientific articles included in the database OpenAlex [Priem et al., 2022]. OpenAlex is an open-source database with more than 230 million scientific papers[11]. We use this dataset because of its large coverage, and because it provides relevant information to our study: Titles and abstracts of papers are used to identify papers dealing with machine learning. Authors are disambiguated such that we are able to trace the paper trail of our focal scientists as well as of their co-authors. This allows for measuring the evolving co-author network and the (publication) experience of scientists in that network. Authors' affiliations are also cleaned and geographically localized, and openAlex provides bibliometric measures at the organizational level that we use. We add further information on the organisations in our sample; notably the Shanghai ranking of the university and the availability of high performance computing in town (see below). Finally, we rely on a system of scientific categorization of journals provided by openAlex termed concept[12]. In particular, we indicate the scientific field of a paper by OpenAlex' 0-level concept assigned to the journal of the paper.

Sample We focus on the trajectory of scientists that eventually applied AI in their scientific domain (other than computer sciences), and we restrict the analysis

[11]We work with the entire database as of August 2022.

[12]Concepts are automatically ascribed to a journal by a classifier, trained on the MAG corpus, that takes as input the title and the abstract of papers published in that journal as well as the title of the journal (see https://docs.openalex.org/api-entities/concepts).

to scientists with at least two publication records before the year of their first AI-related paper.[13] This allows for measuring the STHC endowment of a scientist before AI adoption. In order to judge on the persistence of AI use, we also require at least one publication record after the year of the first AI-related paper. Thus, a focal scientist is observed over a period of at least three years.

The development of AI as a research tool is relatively recent, with diffusion of serious AI applications taking off in the early 2010s [Bianchini et al., 2022]. More precisely, the year 2012 may be considered the beginning of the AI revolution, with significant advancements in deep learning leading to its widespread diffusion [Krizhevsky et al., 2017]. This prompts us to restrict attention to the first AI use in the period from 2012 (the year when AI took off) to 2020 (end of sample period). We further limit the analysis to researchers who started research after 1980; excluding older scientists at the end of their careers and, hence, in general less susceptible to adopt AI in research.[14] Our sampling definition of focal scientists is concisely described in Figure 2.3.1.

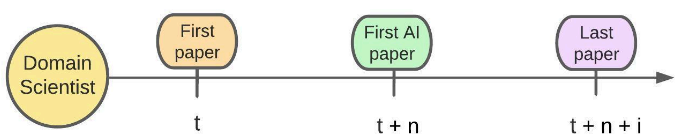


Figure 2.1: Focal scientists [15]

The sampling proceeds as follows. In the first step, we scanned the abstracts and titles of all papers to identify AI papers. To qualify a scientific article as an 'AI paper', we build on a list of keywords provided by the Baruffaldi et al. [2020].

[13]Note that this restriction excludes scientists with a first non-computer science paper building on AI methods. This is a deliberate decision. Given the current development of AI, it is probably not far fetched to assume that future generations of scientists, in one way or another, will use AI in their research as naturally as we use our computers today. Thus, the diffusion process will depend to a large extent on i) the development of AI for various applications and, given a certain state of development, ii) the adoption of AI methods by the generation of currently active scientists. The latter is what we focus on.

[14]Researchers who began before 1980 are approaching the end of their careers, and thus their AI adoption dynamics may be less influenced by their scientific and technical human capital, and more affected by impending retirement.

[15]A focal scientist is active in a domain other than computer science ('domain scientist'), has a first paper and at least a second paper after year $t = 1980$, a first AI-related paper in the period $t + n = (2012, 2020)$, and a subsequent last paper in the period $t + n + i = (2013, 2022)$, with $n, i \geq 1$.

After manually cleaning the keyword list to preserve only terms related to machine learning models we obtain a list of 47 terms (provided in the appendix). If any of the selected terms are mentioned in an abstract or a title of an article, then that article is considered as using AI. This approach results in a total of 1.62 million papers written by 2.83 million authors. Based on the authors' publication history, we define a scientist to be a non-computer scientist if she has no publication in a computer science journal (as indicated by openAlex' main concept assigned to each journal). In total, we identify 1,280,857 non-computer scientists with AI-related publications in the period from 2012 to 2022.

We subsequently limit our analysis to individuals who were active both before and after their exposure to AI, as illustrated in Figure 2.3.1. This ensures that a transition toward AI was made and allows us to investigate whether AI became an integral part of a researcher's toolbox (by reusing AI subsequently). To approximate STHC during the year of AI exposure, it is necessary for the researchers to have published at least two papers prior to the exposure. After computing the variables (see below) and excluding authors with missing information, our sample reduces to 76,344 authors. These focal scientists collectively authored 2,695,096 articles, among which 56,733 were AI-based. This is the sample we work with.

A short reflection on how we categorize scientists as computer scientists may be appropriate at this point. To accurately label computer scientists, we consider a researcher's entire career rather than focusing on their early publications. This approach accounts for the fact that computer scientists may initially publish in domain-specific journals outside of their primary field. We do not have access to individual degrees, and it is common for computer scientists to have publications in non-computer science journals. To determine if a researcher is a computer scientist, we examine whether they have at least one publication in a journal with "Computer Science" as the main concept in OpenAlex. This method ensures that we capture the computer science-related skills of a researcher even if they published in non-computer science journals during the initial years of their career. While this criterion may be restrictive, we believe that the ability to publish in a pure computer science journal indicates a researcher possesses specific computer science expertise.

Our econometric strategy entails a matching approach to compare focal scientists (AI adopters) and non-focal but similar scientists that did not adopt AI. However, the implemented matching procedure makes use of various measures (similar in what?)

and needs to be understood in light of the econometric strategy (why matching?). This is where we turn to now. Basic descriptives of the expanded, matched sample follow at the end of this section.

2.3.2 Measures

This paper attempts to provide empirical insights on how AI adoption of a focal scientist relates to her STHC endowment. The following details how we measure AI adoption (the response variables), STHC endowment (explanatory variables), and further measures (control variables) for the empirical analysis.

AI adoption is measured on scientific papers written by a focal scientist. Whether or not a paper uses AI is determined through AI keywords found in the title or abstract (see above, in the Data section, the definition of a focal scientist). Conceptually, we think of AI adoption as a process that consists of (at least) two steps. The first step is to use AI methods in research for the first time (henceforth 'first use of AI'). Then, given that first AI experience, a scientist may or may not employ AI methods subsequently ('re-use of AI').

STHC endowment of a focal scientist is measured in three dimensions, i.e. institutional capital, social capital from peers, and individual human capital. Figure 2.3.2 summarizes our modelization of Scientific & Technical Human Capital, the construction of the different metrics is detailed in the following.

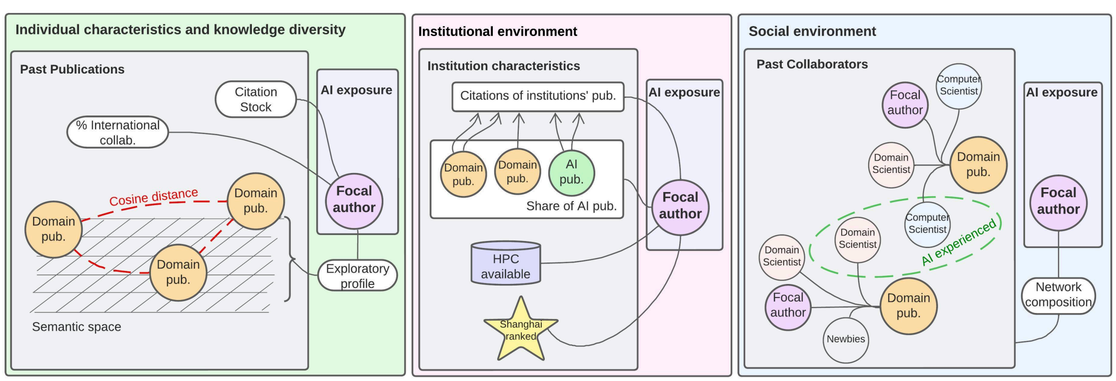


Figure 2.2: S&T Human capital [16]

[16]**Left figure:** The institutional environment potentially provides information, directs attention, and offers resources (computation facilities, human capital) related to AI; and institutions enjoy a

Institutional capital is assessed by measuring a university's prestige and scientific excellence through Shanghai Ranking and citations. Institutional focus on AI is captured through AI-related papers. Availability of relevant physical infrastructure through the presence of high-performance computing in town. In detail, the following measures are calculated:

Scientific excellence: Scientific excellence of an institution is captured by the average number of citations per paper and year. In detail, we consider all papers in OpenAlex with at least one affiliation to the institution for which we observe the number of citations each paper received up to 2022. For each paper, we calculate the average number of citations per year, and then average over all papers of the institution to obtain our measure of 'scientific excellence'.

Prestige: The university's prestige is based on the Shanghai ranking[17]. The variable 'prestige' indicates whether a particular institution is listed in the Shanghai ranking in a given year[18].

Specialization: We proxy the resource in terms of competence present in a given institution with the degree of specialization in the technology, i.e. if many members contribute to research using AI, it is expected that the environment is relatively supportive for conducting this type of research. The degree of specialization in AI is assessed by the proportion of papers that are related to AI in a given institution and a given year. From this, we construct the binary variable 'specialized' that is one if the institution is among the top 10% institutions in terms of degree of specialization and is zero else.

High-Performance computing (HPC) infrastructure: Access to massive computing resources is sometimes necessary when using AI for research; suggesting that HPC infrastructure at the university is potentially a relevant physical infrastructure. In order to determine whether a university hosts HPC infrastructure we searched the web through the platform perplexity.ai. Roughly, perplexity.ai is a sophisticated question-answering system that utilizes large language models to provide answers to

certain level of reputation and scientific excellence. **Middle figure:** The prior co-author network provides Human capital that is relevant to the focal scientists' domain, computational analysis, and/or AI. **Right figure:** The focal scientist's human capital is described by her past research output in terms of scientific content, quality, and internationality.

[17] see https://www.shanghairanking.com/rankings

[18] Note that our measures of 'scientific excellence' and 'prestige' are not imperative; there are alternative measures. Neither do they exclusively proxy these concepts, as they are also correlated with further relevant features of institutions. A university's prestige for example helps to (and allows for) the accumulation of various resources — as more recognition brings more resources and vice versa.

complex queries that include the internet links relevant to the answer given.[19] We ask the following question: *'Is there a High-Performance computing infrastructure in the university of ...'* As the answers given to our question are formulated in a very recurrent manner, it is relatively easy to code the answer text into 'Yes' and 'No' simply with regular expressions (we used 16 regular expressions). This way, we submit queries to perplexity.ai asking for HPC availability in 12,500 cities and obtain an explicit indication of HPC availability ('Yes'/'No') for 12,050 cities. We checked some of the 450 cities for which perplexity.ai gave no clear indication, found that these are mostly small cities with probably no HPC, and coded all of these as having no HPC.

Social capital in relation to peers is measured through the co-author network of the focal author, classifying collaborators into domain scientists and computer scientists. In addition, we note how many collaborators have experience with AI (AI-experienced collaborators), and how many collaborators wrote their very first paper with the focal scientist (collaboration with newbies).

More precisely, the overall co-author network consists of 25,348,325 authors with joint papers published between 1990 and 2020. The prior co-author network in year t builds on all papers (i.e. their authors and revealed co-authorship ties) from 1990 up to year $t-1$. Each scientist in the prior network is classified as a domain scientist or computer scientist based on her individual publication history over the whole observation period. Additional co-author features taken into account are 'AI experienced' and 'newbie' (of the past). The social (network) capital of a focal scientist is then simply the number of prior co-authors of different types.

Domain collaborators: The number of prior co-authors without any paper in a computer science journal and without any AI-related paper. More formally: The variable 'domain collaborators' of a focal scientist in year t corresponds to the number of the focal scientist's co-authors in the prior co-author network that have no computer science publication and no AI experience up to year $t-1$.

Computer science collaborators: The number of prior co-authors with at least one paper in a computer science journal up to year $t-1$.

AI experienced collaborators: The number of prior co-authors with at least one AI-related paper up to year $t-1$.

Collaboration with newbies: The number of collaborators who had never published

[19] While it employs OpenAI's GPT-3 technology, it remains unclear how it determines the relevance and ranking of web pages.

before the year of the collaboration, these individuals are new to the database in year t.

Individual human capital is assessed on the focal scientist's prior publications. Measures include the type of research conducted ('scientific domain'), the propensity to gravitate toward diverse scientific topics ('scientific diversity'), recognition from past publications ('citations received'), and the propensity to engage in international research collaboration ('international collaboration').
Scientific domain: The scientific domain of a domain scientist, i.e. a non-computer scientist, is proxied by the highest level concept of her first paper.
Scientific diversity: To capture an individual's ability to work on diverse topics, we compute the diversity of articles preceding year t for a given researcher. For each item, we represent its abstract in a vector space through word embedding methods (i.e. word2vec – Mikolov et al. [2013]). Once these articles are represented in a vector space, we can compute their cosine distance with all possible articles' combinations. The average distance gives us then an indication on the explorative profile of the researcher.
Citations received: The variable 'citations received' proxies the number of citations a focal scientist's papers received up to a year $t-1$. In detail, we use the citation count from 2022 to estimate the citation count of a given article published in year $t-x$ up to a certain year $t-1$ simply by assuming that the paper received in each year the same number of citations.
International collaboration: We measure 'international collaboration' by the share of prior articles that have multiple country affiliations.

2.3.3 Econometric strategy

We model the adoption of AI in the production of research papers essentially as a combination of AI technology and STHC. An important aspect of AI research technology is that it is not a monolithic, single technology. Rather, it should be considered a bundle of various technologies undergoing specific developments, and pertaining to science specialties with differing degrees (this has been pointed out in the Introduction and Background sections). Therefore, we allow the state of AI technology (A) to vary not only over time t but also over the science specialty of the focal individual, denoted $s(i)$, and write $A_{s(i),t}$.

In order to usefully apply AI in research, a focal scientist i may build on certain aspects of its STHC endowment accumulated up to time $t-1$, $\mathbf{H}_{i,t-1}$. Note that $\mathbf{H}$ is a vector incorporating organizational capital, social (network) capital, and individual human capital of the focal scientist.

A scientist does not necessarily employ all STHC he or she is endowed with, i.e. $\mathbf{H}_{i,t-1}$, to publish an individual paper. Therefore, we also consider the *realized STHC* employed in a given paper p, denoted $\mathbf{H}_{p(i),t}$. Realized STHC ($\mathbf{H}_{p(i),t}$) will be to a large extent a part of the past STHC endowment ($\mathbf{H}_{i,t-1}$), but some capital may be acquired during research in year t, and some may be lost over time.

In order to fix ideas, we postulate a simple AI paper production function $F(\cdot)$ that emphasizes the complementarity (or interaction) between AI technology and the various aspects of capital:

$$F(A_{s(i),t}, \mathbf{H}_{p(i),t}) = A^{\gamma}_{s(i),t}\ \mathbf{H}^{\beta}_{p(i),t}$$

Our main argument is that in the presence of AI certain aspects of STHC will be more valuable compared to research not dealing with AI. One immediate implication is that the realized STHC employed for an AI paper is likely to differ from that realized for a non-AI paper. Descriptive statistics at the end of this section suggest that this is indeed the case.

Our focus is a second implication, namely, that the value of AI adoption increases with the endowment (or presence) of AI-relevant STHC. The probability that an AI paper is produced, rather than a non-AI paper, will thus depend on the availability of those different factors of production in combination with the specific AI technology applied. For the estimation we rely on the log-transformation and a logit regression:

$$\begin{aligned}
p(y_{i,t}=1) &= f(A_{s(i),t}, \mathbf{H}_{i,t-1}) \\
p(y_{i,t}=1) &= \phi\left(\gamma \log(A)_{s(i),t} + \beta \log(\mathbf{H}_{i,t-1}) + \nu_{i,t}\right) \\
&= \phi\left(\gamma_{s,t} + \beta \mathbf{h}_{i,t-1} + \nu_{i,t}\right)
\end{aligned}$$

where $p(y_{i,t}=1)$ is the probability that a paper produced applies AI (rather than not) conditional on a paper produced at all, ϕ denotes the logit-function. $\mathbf{h}_{i,t-1}$ is the log-transform of our measurements $\mathbf{H}_{i,t-1}$. In principle, some relevant capital and/or individual specific tendencies may be also unobserved $\nu_{i,t}$.[20] With this estimation equation, the dynamics of AI technologies is effectively dealt with by introducing intercepts $\gamma_{s,t}$ that vary over time and science specialty.

Our estimation strategy is based on a matching approach where we match scientists belonging to the same science field and same cohort, but with different adoption behavior. Conditional logit regressions on the matched pairs removes all their common factors from the estimation regression. One common factor is the AI technology, another are cohort specific (unobserved) human capital and preferences. As outlined above, it seems likely that scientists in the same field and belonging to the same cohort face similar (exogenous) dynamic AI technology $A_{s(i),t}$. Furthermore, scientists of same field and cohort may share some similarities in unobserved preferences and skills, $\nu_{s(i),t}$. This leads us to write individual unobserved components as the sum of average cohort effects $\bar{\nu}_{s(i),t}$ and individual deviations from that average $\tilde{\nu}_{i,t}$, i.e. $\nu_{i,t} = \bar{\nu}_{s(i),t} + \tilde{\nu}_{i,t}$. By matching same cohort scientists i and j, we obtain $A_{s(i),t} = A_{s(j),t}$ and $\bar{\nu}_{s(i),t} = \bar{\nu}_{s(j),t}$.

These common factors can be removed in a conditional logits approach. First note that matching is on the outcome such that in all matches one individual i adopts AI and the other scientist j does not adopt AI ($y_{i,t} = 1, y_{j,t} = 0$), or vice versa. The conditional logit model, takes into account that only two possible outcomes are possible, and we estimate the probability of one of them:

$$\begin{aligned} p(y_{i,t} = 1, y_{j,t} = 0) &= \frac{exp(\gamma_{s,t} + \beta h_{i,t-1} + \bar{\nu}_{s(i),t} + \tilde{\nu}_{i,t})}{exp(\gamma_{s,t} + \beta h_{i,t-1} + \bar{\nu}_{s(i),t} + \tilde{\nu}_{i,t}) + exp(\gamma_{s,t} + \beta h_{j,t-1} + \bar{\nu}_{s(j),t} + \tilde{\nu}_{j,t})} \\ &= \frac{exp(\beta h_{i,t-1} + \tilde{\nu}_{i,t})}{exp(\beta h_{i,t-1} + \tilde{\nu}_{i,t}) + exp(\beta h_{j,t-1} + \tilde{\nu}_{j,t})} \end{aligned}$$

where common factors have been pulled out of the second equation.

In case individual unobserved components $\tilde{\nu}_{i,t}$ are correlated with observed factors in **h**, coefficient estimates are biased. Imagine for example that an individual has an unobserved 'taste' (or unobserved capacity) for data-intensive research, which led him in the past to collaborate with computer scientists that we measure as a specific

[20]For convenience, the coefficients associated with AI research technology and STHC have the same names in the paper production function and the adoption function, but they are of course not the same.

kind of social capital. Not including the unobserved component in a simple regression would in that case over-estimate the effect of such social capital on the probability of applying AI in research.[21] That must be taken into account in the interpretation of the estimation results, where we speak of correlations rather than of causal effects.

Note that in the estimation equation above the variation of AI technology is dealt with as a simple scaling factor of the valuation of STHC. A more flexible and probably more realistic formulation would be to allow for variation of the exponents (β's) of STHC with the dynamics of AI technologies. For example computer scientists may be indeed needed in the early stages of AI development, whereas in, say, two generations domain scientists may well be capable of autonomously using AI (because of both different human capital formed during training and different AI technology). We investigate the variation of STHC coefficients across scientific fields and time by estimating additional regressions.

2.3.4 Matching strategy

Technology available at a time t, age of the researcher, initial training and research trajectory influence the opportunity to use AI. When matching individuals with similar potential to use AI, we consider only the technology's advancement and applicability in a given field. To answer our two questions on adoption and reuse we have performed two different matchings on the basis of the same criteria.

As shown in Figure 2.3, we construct two matched samples. The first matching is used to investigate first-time AI use. Each focal scientist is matched with a non-focal scientist who never published an AI-related paper, but published a first paper in the same year and same scientific field as did the focal scientist, published a paper in the same scientific field and same year as the focal scientists' first AI paper, and published at least one paper subsequently (as the focal scientist). Note that individuals with computer science papers are excluded (focal and matched scientist).

The second matching is used to investigate re-use of AI. Here we compare two focal scientists where both had a first-time AI paper in the same scientific field and

[21] The econometric literature discusses various ways to deal with the potential bias due to unobserved effects. One possible avenue would be conditional logit, where one conditions on all outcomes of an individual in order to extract an unobserved fixed effect. This however would be reasonable only if one neglects the rapid technological development that is currently undergoing as time varying coefficients ($\gamma_{s,t}$) can not be introduced and other factors (e.g. AI relevant social capital) are strongly trending (and detrending is never 'perfect'). Another possibility would be to instrument $\mathbf{h}$, for example by effects of (unexpected) job switches. A proper causal identification is left for future research.

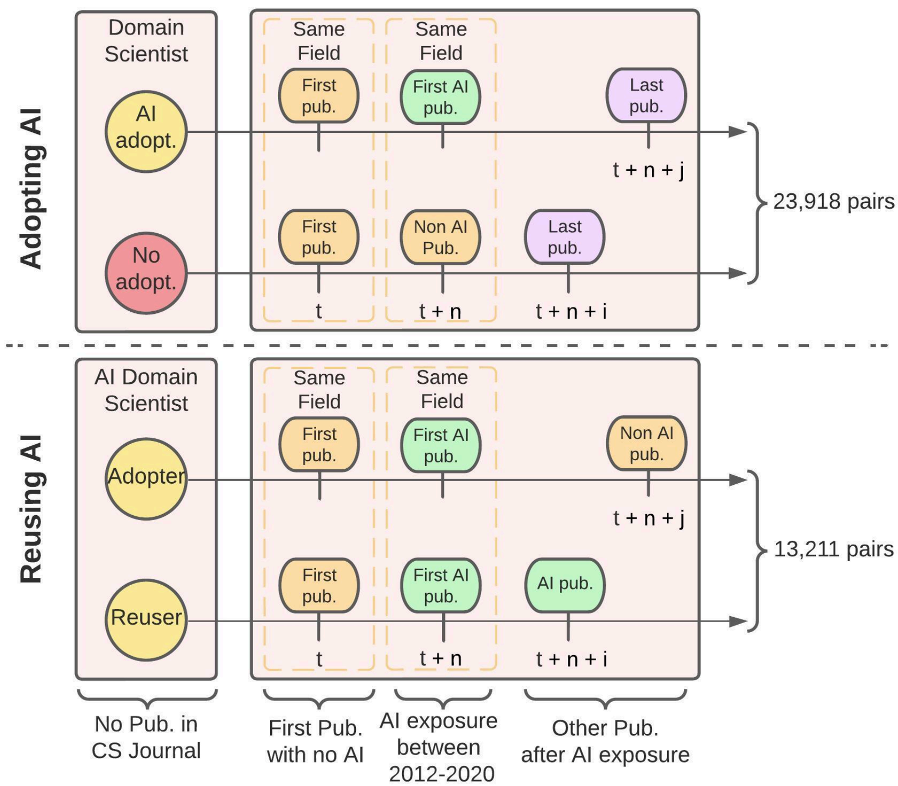


Figure 2.3: Matched samples to investigate AI adoption, i.e. first-time AI research and re-using AI.

year, but one scientist re-uses AI in a subsequent year and the other does not. As in the first matching, we also require that both published their first paper in the same field and year.

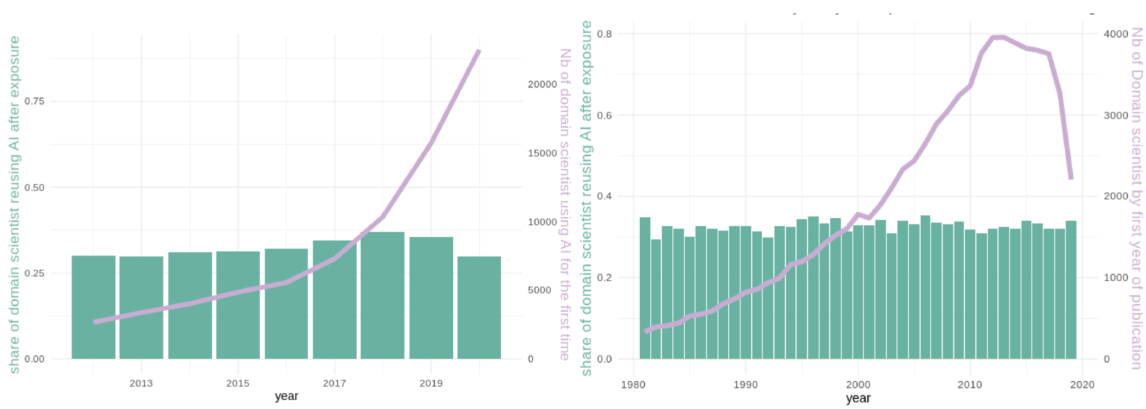
In both cases we realized an exact matching, for the adoption part, we used all Openalex to get a correct tween. We manage to match 23,918 AI adopters with non-adopters after cleaning our data. Concerning the second matching exercise we were able to connect 13,211 adopting pairs, in each pair only one researcher produced another AI-based document.

2.3.5 Descriptive statistics

This section provides in a first step aggregate statistics on the adoption of AI in research by our focal scientists (i.e. scientists of scientific domains other than computer sciences). In a nutshell, Figure 2.4a shows that the number of scientists with first-time AI use is strongly increasing in particular since 2017, with a stable share of around one third re-using AI subsequently. There is also a strong cohort effect such that in particular early career scientists adopt AI (Fig. 2.4b).

Turning to the factors affecting AI adoption, we see that i) focal scientists co-author more frequently with computer scientists and AI experienced scientists in AI-related papers than in non-AI papers (Table 2.1), and that ii) AI adopters, be it first-time AI use or re-use of AI, tend to have a larger endowment of AI relevant STHC than those that do not adopt AI (Table 2.2).

The last two tables, Tables 2.3 and 2.4, provide correlations among all the variables on the matched samples used to estimate the effect of STHC endowment on first-time AI use and re-use of AI respectively.



(a) Domain Scientist by year of AI exposure (b) Domain Scientist by first year of pub.

Figure 2.4: Number of author per year and share that reuse it

<u>*A growing body of scientists adopting AI*</u>: The number of scientists integrating AI for the first time is increasing massively, with over 20,000 people are integrating AI for the first time in 2020 (see solid line in Figure 2.4a with scale on the right vertical axis). Interestingly, and this was one of the reasons that encouraged us to further invest in the reuse dimension, the share of those first-time users who will one day re-use AI technology again is rather stable at around 35% (see bars in Figure 2.4a with scale on the left vertical axis). Despite the advancement of technology and the increasing availability of resources to facilitate AI utilization, the probability to re-use AI conditional on having a first-time AI experience remains rather stable over time.

Due to our sample construction, we cope with a population with a fairly mature academic career. The first year of publication is on average 2006, and focal scientists have on average an academic age of 11.05 years during their first exposure to AI (the median is 9 years). The mode of the first year of publication for our focal scientists in the sample is 2012 with 4,000 scientists (see solid line in Figure 2.4b with scale on the right vertical axis). Thus, the sample is comprised mostly by early- and mid-career scientists, suggesting that older generations are less engaged in AI-related research.[22] As shown by the bars in Figure 2.4b with scale on the left vertical axis, the share of individuals reusing AI remains stable around 35% regardless of the seniority of the individual. This suggests that persistent adoption of AI is indeed feasible across generations.

Requirement of a different team: A basic tenet underlying this study is that AI research is realized with specific STHC that differs from STHC used in non-AI research. Ideally, one would like to observe the extent to which a focal scientist with given STHC endowment leverages each dimension of her STHC for her different research projects. This would allow to see which aspects of STHC are particularly relevant when it comes to AI. Our publication data at hand does in general not provide such information. The exception is the social capital dimension of STHC in form of co-authors with their individual scientific background.

Table 2.1 compares therefore co-authors of non-AI papers with co-authors of AI papers. The underlying sample includes for each focal scientist one (randomly chosen) non-AI paper and one (randomly chosen) AI paper that both appeared in the year of the focal scientist's first-time AI use. On average, focal scientists work with one more co-author in AI papers (11.65 authors in total) compared to non-AI papers (10.73 authors in total). This 'additional co-author' tends to have a computer science background. Besides, the research team is more likely to include AI experienced scientists. On the other hand, the number of domain (non-computer) scientists is on average the same in AI as in non-AI papers. Finally, newbies without any prior publication contribute somewhat more often to AI papers than to non-AI papers. The t-test strongly rejects the null hypothesis of no difference between AI papers and non-AI papers for all co-author types except for the number of domain scientists.

STHC endowment and AI adoption: Table 2.2 presents descriptive statistics for both samples employed in the regression analysis of first-time AI use (left part) and re-use of AI (right part). The upper part of the table highlights the disparities in STHC endowment between those who engage with AI and those who do not. The lower part shows differences in co-authors and citations received of the focal scientist's

[22] A thorough comparison of academic age of focal scientists and non-focal, other scientists in the overall population is not given here due to our sample restrictions.

Table 2.1: Co-authors of first-time AI users in non-AI papers and AI papers

	non-AI papers	AI papers	t-test
# authors	10.73 (10.31)	11.65 (12.51)	14.32***
# CS aut.	2.19 (4.09)	3.12 (4.82)	36.66***
# AI exp. aut.	1.71 (3.48)	2.84 (4.54)	49.34***
# Domain aut.	8.53 (7.65)	8.53 (9.62)	0.04
# Newbies aut.	1.19 (1.7)	1.41 (2.82)	16.13***
Observations	62712	62712	

Notes: This table presents the descriptive statistics for Realized STHC for AI adopters. It compares an AI paper with a non-AI paper published in the first year of AI exposure. The table provides averages and standard deviations (in parentheses) of the number of different types of co-authors. T-tests determine whether the mean differences between the groups are statistically significant. Significance levels are denoted by ***, **, and * for the 1%, 5%, and 10% levels, respectively. Note that for this table focal scientists without a non-AI paper in the year of their first-time AI paper are excluded.

first AI paper and the matched scientist's paper (on the left side the matched paper is non-AI of a non-adopter, on the right side the matched paper is the first AI paper of a scientist trying out once AI but not re-using AI).

Consider first the upper part, STHC endowment. The t-tests indicate that nearly all variables representing STHC endowment show significant mean differences between individuals who explore AI and those who abstain. At the institutional level, the percentage of individuals in specialized institutions is elevated (AI inst. spec.), and the citation impact of these institutions appears to be positively correlated with AI adoption (Inst. cit.). This is true for first-time AI use and, to a lesser extent, for re-using AI. Whether or not the university is listed in the Shanghai ranking seems to make no difference.[23] External resources associated with social networks seem more crucial for individuals who embrace AI technology (be it first-time AI use or re-use of AI). On average, the scope of these individuals' networks is more extensive across all categories of collaborators. In the case of first-time AI use, the strongest difference (as indicated by the t-statistic) is in the number of computer science collaborators and the number of collaborators with AI experience. In the case of re-using AI, the t-statistic is similarly strong across all co-author types. One may note however that focal scientists re-using AI are particularly inclined to collaborate with early-career researchers, as their average experience working with such researchers is higher. Finally, at the individual level, the ability to navigate the knowledge space (exploratory profile) and prior accomplishments (citation stock) also appear to be linked to AI

[23] One potential explanation is that our measure indicating whether the university is among the top 1000 universities is too rough to capture existing 'elite-effects'.

Table 2.2: Descriptives statistics for both matching strategies

	First-time AI			Re-use AI		
Variable	Matched scientists (without AI)	Focal scientists (with AI)	T-test	Matched scientists (not re-using AI)	Focal scientists (re-using AI)	T-test
STHC endowment						
AI inst. spec.	0.05 (0.22)	0.1 (0.29)	19.07***	0.08 (0.27)	0.09 (0.29)	2.73***
Inst. cit.	2.86 (3.88)	3.04 (1.98)	6.57***	3.09 (2.09)	3.14 (2.03)	2.06**
Shanghai ranked	0.03 (0.17)	0.03 (0.17)	-0.09	0.03 (0.17)	0.03 (0.17)	-0.47
HPC	0.71 (0.45)	0.74 (0.44)	6.16***	0.73 (0.44)	0.74 (0.43)	2.58***
# Domain col.	146.84 (181.62)	155.58 (220.97)	4.73***	134.27 (190.6)	167.48 (245.87)	12.27***
# CS col.	21.08 (33.88)	27.83 (44.68)	18.6***	23.86 (40.68)	31.46 (49.54)	13.62***
# AI col.	9.3 (18.43)	14.56 (25.89)	25.62***	12.36 (24.12)	17.18 (29.74)	14.46***
# Newbies col.	51.95 (68.31)	55 (81.34)	4.45***	46.83 (69.11)	59.05 (89.77)	12.41***
Exploratory profile	0.18 (0.06)	0.18 (0.06)	2.29**	0.18 (0.06)	0.18 (0.06)	3.01***
Citation stock	918.2 (2040.94)	1031.02 (2441.89)	5.48***	826.93 (1959.27)	1088.13 (2625.3)	9.16***
% International pub.	0.3 (0.25)	0.3 (0.25)	-1.73*	0.3 (0.26)	0.29 (0.25)	-0.95
Co-authors and citations of (matched) paper						
# Domain aut.	10.88 (9.99)	8.4 (8.9)	-28.68***	8.38 (8.96)	7.67 (7.26)	-7.04***
# AI exp. aut.	1.31 (2.68)	2.77 (4.25)	44.9***	2.67 (4.38)	2.87 (4.39)	3.7***
# CS aut.	1.78 (3.46)	3 (4.45)	33.42***	2.99 (4.73)	3.07 (4.55)	1.36
# Newbies aut.	1.28 (2.86)	1.24 (2.12)	-1.63	1.23 (2.21)	1.21 (2.17)	-0.66
# Citations	4.15 (13.49)	4.75 (10.82)	5.37***	4.54 (10.48)	5.43 (12.2)	6.37***
Total	23918	23918		13211	13211	

Notes: This table presents the descriptive statistics for various variables, including their mean values and standard deviations (in parentheses) for different matching strategies: those who didn't use AI and those who did, as well as those who didn't reuse AI and those who did. The table also provides results from t-tests to determine if the mean differences between the groups are statistically significant. Significance levels are denoted by ***, **, and * for the 1%, 5%, and 10% levels, respectively. In the AI re-use section, the matched individuals who do not reuse AI are also part of the focal scientists using AI for the first time in the first analysis.

adoption.

Now turn to the lower part of Table 2.2. AI papers authored by first-time AI users (see left part of table) exhibit a composition of co-authors that is different to papers of matched individuals who do not use AI. Specifically, the number of domain scientists tends to be smaller for first-time AI users, while the presence of individuals with computer science specialization and AI experience nearly doubles. Furthermore, AI articles tend to have a higher average impact.

A similar pattern is observed when examining the co-authors in papers of re-users and non-re-users (see right part of table). The citation count of the initial AI article is higher among those who persist in using AI technology. Moreover, individuals who incorporate AI into their subsequent research already possess a more specialized STHC endowment in computer science and AI compared to other researchers who will not continue using the technology. They are also more likely to be affiliated with highly specialized AI institutions and have increased access to computing centers. Ultimately, individuals who maintain the use of AI in their research exhibit a more exploratory profile and demonstrate stronger past success than their counterparts who do not continue.

Tables 2.3 and 2.4 display basic descriptive statistics for our two estimation samples. We can immediately observe that the measures derived from collaboration networks are highly correlated among each other. In particular, the number of newbie collaborators and the number of domain scientist collaborators are strongly correlated with a Pearson correlation coefficient of above 0.9 in both estimation samples. This is however expected because many newbies are likely to count as domain scientists. Similarly, the number of collaborators with AI experience is strongly related to the number of collaborators in computer science; in both samples the correlation coefficient is above 0.8. Additionally, a clear positive correlation is visible between the number of collaborations and the stock of citations. For completeness, the last variable included in each table is the outcome variable, i.e. first-time AI use and re-use of AI respectively. Note however that Table 2.2 is more appropriate to shed light on the relation between our (binary) left-hand-side variable and right-hand-side variables.

Table 2.3: Descriptive Statistics - first-time AI regression

Variables	Mean	Std	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) AI inst. spe.	0.07	0.26	1											
(2) Inst. cit.	1.37	0.34	-0.01	1										
(3) Shanghai ranked	0.05	0.21	0	0	1									
(4) HPC	0.73	0.44	-0.03	0.22	0	1								
(5) # Domain col.	4.29	1.14	-0.11	0.06	-0.03	0.04	1							
(6) # CS col.	2.47	1.16	-0.05	0.22	-0.04	0.11	0.77	1						
(7) # AI col.	1.61	1.11	-0.04	0.24	-0.05	0.12	0.64	0.82	1					
(8) # Newbies col.	3.24	1.21	-0.07	-0.01	0	0	0.93	0.67	0.53	1				
(9) Exploratory profile	0.18	0.06	-0.08	-0.1	0	-0.03	0.25	0.12	0.09	0.26	1			
(10) Citation stock	5.45	1.77	-0.04	0.21	-0.02	0.09	0.73	0.68	0.54	0.69	0.06	1		
(11) % International pub.	0.29	0.25	0	0.1	-0.03	0	0.18	0.27	0.25	0.09	-0.15	0.12	1	
(12) First-time AI (yes/no)	0.50	0.50	0.11	0.07	0	0.03	-0.08	0.09	0.18	-0.06	0	-0.03	0	1

Notes: This table presents the descriptive statistics for the whole sample on AI adoption.

<u>*Massive adoption by researchers in life science*</u>: The majority of researchers in our sample published in journals related to life sciences. Publications in medicine, biology and chemistry account for 80% of first-time AI papers. The six fields presented in our regression (namely, Medicine, Biology, Chemistry, Physics, Psychology and Materials science) represent 95% of our sample of AI papers. Researchers adopting AI tend to apply this technology in a field they are familiar with, i.e. the one where their first research was published. In detail, 62% of researchers publish their first AI article in the journal of the same scientific field as their first publication. We provide estimation results on individual scientific fields after the main results on the pooled sample in the next section.

Table 2.4: Descriptive Statistics - re-using AI regression

Variables	Mean	Std	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
(1) AI inst. spe.	0.10	0.29	1																
(2) Inst. cit.	1.40	0.34	-0.04	1															
(3) Shanghai ranked	0.05	0.21	-0.03	-0.01	1														
(4) HPC	0.74	0.44	-0.06	0.22	-0.01	1													
(5) # Domain col.	4.14	1.31	-0.14	0.09	-0.03	0.04	1												
(6) # CS col.	2.56	1.23	-0.1	0.23	-0.04	0.11	0.8	1											
(7) # AI col.	1.84	1.16	-0.1	0.24	-0.06	0.11	0.68	0.84	1										
(8) # Newbies col.	3.10	1.34	-0.1	0.01	-0.01	0	0.94	0.71	0.58	1									
(9) Exploratory profile	0.17	0.06	-0.11	-0.05	0	-0.01	0.34	0.2	0.15	0.33	1								
(10) Citation stock	5.29	1.89	-0.09	0.2	-0.03	0.08	0.77	0.72	0.59	0.73	0.16	1							
(11) % International pub.	0.29	0.26	0.01	0.09	-0.03	-0.01	0.17	0.26	0.24	0.1	-0.11	0.12	1						
(12) # Domain aut.	1.91	0.61	-0.07	0.03	-0.02	0	0.34	0.19	0.22	0.29	0.13	0.13	0.15	1					
(13) # AI exp. aut.	0.90	0.73	-0.03	0.14	-0.03	0.07	0.1	0.24	0.38	0.04	-0.01	0.06	0.14	0.26	1				
(14) # CS aut.	1.07	0.73	-0.02	0.17	-0.03	0.08	0.14	0.35	0.31	0.08	0.02	0.09	0.19	0.2	0.65	1			
(15) # Newbies aut.	0.49	0.57	0.02	-0.01	0.02	-0.03	0.07	-0.02	-0.02	0.1	0.05	-0.02	0.03	0.45	0.01	0.05	1		
(16) # Citations	1.37	0.87	0.01	0.18	-0.01	0.02	0.01	0.08	0.08	-0.02	-0.04	0.11	0.1	0.09	0.19	0.23	0.02	1	
(17) Re-use AI (yes/no)	0.50	0.50	0.02	0.04	0	0.01	0.08	0.13	0.16	0.08	0.02	0.07	0.01	-0.02	0.04	0.04	0.01	0.06	1

Notes: This table presents the descriptive statistics for the whole sample on re-using AI.

2.4 Results

2.4.1 Main results

2.4.1.1 Conditional Logit with matching: Adopting AI

Table 2.5 presents the results of the conditional logit regression of first-time AI use on three dimensions of STHC, i.e. institutional, social, and individual factors. Recall that in the estimation sample the outcome, first-time AI use, is one for focal scientists and zero for matched scientists; and that STHC is measured up to the year before first-time AI use (see also previous section).

Table 2.5 provides coefficient estimates of four models — one column for each dimension of STHC separately and the fourth column for estimating all coefficients jointly. The Loglikelihood ratio tests (LR Test) at the bottom of the table confirm that all models improve significantly over the intercept-only-model. Considering the log likelihood in increasing order, we see that the social dimension (Column 2) is the most informative dimension for first-time AI use, followed by institutional factors (Column 1), and individual factors (Column 3). This is consistent with the idea that AI relevant STHC does not solely reside in the individual scientist, but is also a result of the granularity with which we measure the various factors.

<u>Institutional Factors</u>: The positive and significant coefficients of AI institution specialization (AI inst. spe.) and institutional citation impact (Inst. cit.) in models (1) and (4) suggest that researchers affiliated with institutions specialized in AI and with higher citation impact are more likely to adopt AI in their work, validating Hypothesis H1b. One explanation could be that researchers are heavily influenced

Table 2.5: Conditional Logit with matching (first-time AI use)

	Dependent variable: First-time AI use			
	Institutional	Social	Individual	Full Model
	(1)	(2)	(3)	(4)
AI inst. spe.	0.719***			0.585***
	(0.039)			(0.043)
Inst. cit.	0.418***			0.241***
	(0.029)			(0.033)
Shanghai ranked	-0.019			-0.016
	(0.058)			(0.065)
HPC	0.053**			-0.006
	(0.021)			(0.024)
# Domain col.		-1.503***		-1.422***
		(0.034)		(0.036)
# CS col.		0.139***		0.173***
		(0.020)		(0.021)
# AI col.		0.793***		0.783***
		(0.019)		(0.019)
# Newbies col.		0.575***		0.549***
		(0.025)		(0.026)
Exploratory profile			0.331**	1.699***
			(0.165)	(0.193)
Citation stock			-0.044***	-0.083***
			(0.007)	(0.011)
% International pub.			-0.015	-0.191***
			(0.038)	(0.046)
Observations	47,836	47,836	47,836	47,836
Log Likelihood	-16,260.790	-14,171.280	-16,555.990	-13,969.870
LR Test	635.801*** (df = 4)	4,814.822*** (df = 4)	45.419*** (df = 3)	5,217.655*** (df = 11)

Notes: This table reports coefficients of the effect STHC on AI adoption on all fields. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The effect of STCH on AI adoption is estimated using a conditional logit with matching.

by their institutional environment, including shared norms, directives, and funding focused on specific subjects. Thus, an institution with a strong AI specialization and recognized research output can more easily leverage the financial means to facilitate AI adoption. However, being affiliated with a Shanghai-ranked institution does not significantly affect first-time AI use. Additionally, we partially reject Hypothesis H3 since access to high-performance computing (HPC) resources is significantly associated with AI adoption in model (1) but not in the full model (4). This suggests that HPC may not be crucial. One reason could be that many AI models are available pretrained and hence using them does not necessarily require elevated computational resources. On the other hand, steep learning curves for using an HPC may create a barrier — that may be overcome with the help of skilled co-authors as the second model suggests.

Social Factors: Model (2) and the full model (4) show that number of domain collaborators (#Domain col.) has a negative and significant effect on AI adoption, suggesting that a more extensive network of domain collaborators may actually reduce the likelihood of AI adoption. On the other hand, the number of computer science collaborators (#CS col.) and AI collaborators (#AI col.) both have positive and significant effects on AI adoption. This implies that having collaborators experienced in computer science and AI increases the likelihood of AI adoption, thereby confirming Hypothesis H1a.

The relevance of having prior contacts to computer scientists and/or AI experienced scientists may stem from two different effects. For one, such prior contacts can facilitate subsequent collaboration in AI research through repeating ties, referrals and other social processes. Moreover, individuals within a researcher's network help acquire *about-knowledge*, enabling the researcher to integrate more easily into teams using AI. These individuals may have a greater capacity to interact with experienced individuals, increasing their likelihood of becoming users themselves. Researchers with too many domain-specific collaborators may be less exposed to 'AI-related thinking' or be part of niches where AI is still not widely used. Additionally, The positive influence of early-career (#Newbies col.) research collaborators on AI adoption confirms Hypothesis H2. Researchers can expand their research subjects through the fresh perspectives of young researchers, and have an advantage in trying out AI, as newer generations are better trained in modern statistical and computational approaches. Many universities, even in the social sciences, now offer data science courses; thus, working with younger researchers may increase the probability of using AI due to the growing proportion of individuals with the necessary skills among the next generation.

Individual Factors: As shown in Models (3) and (4), we verify Hypothesis H4a as having an exploratory profile positively impacts AI adoption. This suggests that individuals with a more diverse cognitive profile are more used to exploring subjects distant from their prior knowledge. These individuals could have an advantage in integrating into teams using methods they have not previously used or re-orienting their research focus to seize opportunities opened by AI methods.

Citation stock has a negative and significant effect on AI adoption, suggesting that a higher number of past citations may reduce the likelihood of adopting AI.

This may indicate that individuals who have already demonstrated success in their past publications have fewer incentives to transition to AI. The negative impact of citation stock on AI adoption validates Hypothesis H4b, indicating that scientists with a higher scientific reputation and recognition are actually less likely to integrate AI into their research.

Lastly, the percentage of international publications (% International pub.) does not have a significant effect in model (3) but demonstrates a negative and significant effect in the full model (4). Taken at face value, this could imply that researchers who tend to work more with local colleagues have stronger connections that facilitate better knowledge transfer, thereby contributing more effectively to developing a reasonably understanding of AI. Any interpretation must be cautiously however, because the coefficient becomes significant with the introduction of (correlated) variables which hints to identification issues.

2.4.1.2 Conditional logit with matching: Re-using AI

Table 2.6 analyzes the effect of realized scientific and technical human capital (STHC) employed in the first AI paper and STHC endowment on the re-use of AI in subsequent studies. The models are similar to those in the regression of first-time AI, but now the sample includes only focal scientists with AI experience; with scientists re-using AI matched to those not re-using AI.

Again, log-likelihood tests indicate that all models improve significantly over the intercept-only model. The order of model fit indicated by the log-likelihood slightly changes: Prior co-authors (Column 2) still provide the best fit, but now individual factors follow (Column 3), with institutional factors ranked third (Column 1). Note that each model takes into account the experience of the first-use of AI (variables '# Domain author' to '# Citations').

First-use of AI experience: The team composition of the first AI article seems to be crucial for understanding how this technology will be integrated into the future research of scientists. The stronger presence of individuals with prior AI experience in a team indicates the team's specialization in AI. It is this specialized social environment that allows the researcher to increase his chances of reusing AI later on.

Consistently across all models (1-4), we observe that the number of domain experts (# Domain aut.) negatively effects the re-use of AI. The number of computer science authors (# CS aut.) also has a negative effect throughout. In contrast, the number of AI experts (# AI exp. aut.) shows positive and significant effects. It

Table 2.6: Conditional logit with matching (re-using AI)

	Dependent variable: re-using AI			
	Institutional	Social	Individual	Full Model
	(1)	(2)	(3)	(4)
# Domain aut.	-0.303***	-0.330***	-0.324***	-0.315***
	(0.026)	(0.028)	(0.026)	(0.028)
# AI exp. aut.	0.173***	0.063**	0.176***	0.061**
	(0.024)	(0.027)	(0.024)	(0.027)
# CS aut.	-0.055**	-0.113***	-0.061**	-0.108***
	(0.024)	(0.026)	(0.024)	(0.026)
# Newbies aut.	0.155***	0.184***	0.177***	0.183***
	(0.024)	(0.025)	(0.024)	(0.025)
# Citations	0.151***	0.165***	0.133***	0.166***
	(0.016)	(0.016)	(0.016)	(0.016)
AI inst. spe.	0.102**			0.114**
	(0.046)			(0.047)
Inst. cit.	0.116***			-0.021
	(0.039)			(0.041)
Shanghai ranked	-0.020			-0.052
	(0.078)			(0.080)
HPC	0.042			0.033
	(0.029)			(0.030)
# Domain col.		-0.258***		-0.251***
		(0.038)		(0.040)
# CS col.		0.079***		0.101***
		(0.027)		(0.028)
# AI col.		0.274***		0.280***
		(0.025)		(0.025)
# Newbies col.		0.245***		0.230***
		(0.030)		(0.030)
Exploratory profile			0.957***	0.779***
			(0.232)	(0.242)
Citation stock			0.132***	0.004
			(0.010)	(0.014)
% International pub.			-0.137***	-0.337***
			(0.052)	(0.055)
Observations	26,422	26,422	26,422	26,422
Log Likelihood	-9,008.229	-8,717.901	-8,923.676	-8,687.683
LR Test	297.877*** (df = 9)	878.533*** (df = 9)	466.984*** (df = 8)	938.969*** (df = 16)

Notes: This table reports coefficients of the effect of STHC endowment and realized STHC on re-using AI in all fields. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The effect of STCH on re-using AI is estimated using a conditional logit with matching.

seems that the team shouldn't be composed of too many computer scientists and domain scientists, i.e., the team should not be too large (since the sum of the two gives the number of authors). Also, we see that the presence of early-career researchers (#

Newbies aut.) is crucial, as they are more likely to bring new skills and may remain connected to the team. For example, the re-use of AI could be induced by doctoral students who publish several AI papers with their supervisors.

Interpretation of these findings is possible in light of an additional regression of re-using AI based on a sample with focal scientists having a first AI paper earlier, up to 2018 instead of 2020 (see appendix, Table 2.16). In that regression, the number of collaborators in the first AI paper with prior experience in AI do not play a significant role. Instead, it is the presence of newcomers in the first AI paper that is crucial. This highlights the extent to which the new generation is initiating this paradigm shift and transmitting it to researchers that are more advanced in their careers. Indeed, in the larger sample that extends up to 2020, which is the basis for our main regression results provided here, many of these newcomers have become individuals with AI experience. This is why their contribution has become significant in re-using AI.

Finally, the reward associated with the first AI publication in terms of citations positively influences re-using AI. This shows the incentives that exist when peers validate the article with which a researcher has transitioned to AI.

STHC Endowment: The results of this section are very similar to those of first-use of AI for all dimensions of the measured STHC. However, there is one exception: the citation stock of researchers and their institutions no longer seem to play a role in encouraging continuing AI-based research.

Institutional Factors: In Model 1, AI institutional specialization (AI inst. spe.) and institutional citation impact (Inst. cit.) have positive and significant effects on re-using AI. However, the significance of the institutional citation impact disappears in the full model (4). The Shanghai ranking and HPC variables are not significant in any model. This suggests that, what sets apart scientists persistently applying AI from those trying out once AI is not primarily their institutional resources.

Social Factors: Model 2 reveals findings similar to the regressions on the first use of AI. Our analysis shows a negative effect of the number of domain collaborators (# Domain col.) on re-using AI, while the number of computer science collaborators (# CS col.) and the number of AI collaborators (# AI col.) exhibit positive effects. Furthermore, the number of new collaborators (# Newbies col.) positively influences re-using AI. All this is consistent with the full model (4).

Individual Factors: In Model 3, we see a strong positive and significant effect of an exploratory profile on re-using AI. The citation stock exhibits a positive influence, but this effect disappears in the full model (4). Lastly, the percentage of international publications (% International pub.) has a negative and significant impact on re-using AI in both Model 3 and Model 4. This finding is consistent with the idea that collaborating with geographically proximate researchers is essential for building long-term relationships to work on AI-related topics. Such proximity facilitates communication and learning from colleagues [van der Wouden and Youn, 2023].

2.4.2 Extension: AI adoption across scientific fields and time

Table 2.7 delves into a field-level analysis, confirming that our results are generally stable across fields but that there are also some relevant differences.

Table 2.7: Conditional Logit with matching across fields (first-use of AI)

	Dependent variable: first-use of AI					
	Medicine	Biology	Chemistry	Physics	Psychology	Materials science
	(1)	(2)	(3)	(4)	(5)	(6)
AI inst. spe.	0.650***	0.653***	0.271**	0.607***	0.700**	0.716***
	(0.082)	(0.069)	(0.123)	(0.144)	(0.295)	(0.182)
Inst. cit.	0.303***	0.260***	-0.198	0.253*	0.212	0.796***
	(0.051)	(0.053)	(0.131)	(0.142)	(0.243)	(0.293)
Shanghai ranked	0.012	0.029	-0.094	-0.544	-0.393	-0.239
	(0.110)	(0.099)	(0.223)	(0.360)	(0.392)	(0.430)
HPC	-0.032	-0.001	0.237***	0.114	0.026	-0.184
	(0.037)	(0.038)	(0.087)	(0.104)	(0.175)	(0.184)
# Domain col.	-1.290***	-1.450***	-1.684***	-1.620***	-1.417***	-1.285***
	(0.058)	(0.057)	(0.141)	(0.145)	(0.215)	(0.265)
# CS col.	0.154***	0.243***	0.153**	0.054	-0.176	0.528***
	(0.033)	(0.033)	(0.074)	(0.093)	(0.155)	(0.149)
# AI col.	0.799***	0.668***	0.701***	1.100***	1.290***	0.579***
	(0.030)	(0.030)	(0.070)	(0.083)	(0.132)	(0.130)
# Newbies col.	0.605***	0.560***	0.537***	0.332***	0.311*	0.136
	(0.041)	(0.041)	(0.095)	(0.103)	(0.163)	(0.175)
Exploratory profile	2.271***	1.197***	1.866**	2.585**	2.303**	-1.361
	(0.288)	(0.308)	(0.867)	(1.131)	(1.108)	(1.925)
Citation stock	-0.088***	-0.095***	-0.023	-0.050	-0.105	-0.201**
	(0.016)	(0.018)	(0.043)	(0.041)	(0.074)	(0.084)
% International pub.	-0.338***	-0.060	-0.124	-0.452**	0.594**	0.055
	(0.073)	(0.072)	(0.174)	(0.206)	(0.295)	(0.338)
Observations	19,684	18,484	3,400	2,758	1,374	1,022
Log Likelihood	-5,818.122	-5,437.961	-966.666	-687.362	-333.878	-277.148
LR Test (df = 11)	2,007.665***	1,936.211***	423.369***	536.977***	284.627***	154.101***

Notes: This table reports coefficients of the effect STHC on AI adoption across fields. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The effect of STCH on AI adoption is estimated using a conditional logit with matching.

Institutional factors: Our analysis shows that AI institution specificity (AI inst.

spe.) impacts is consistent across all fields. Institutional citations (Inst. cit.) are also significantly positively related to first-use of AI in most fields, except in Chemistry, where the effect is insignificant. The Shanghai ranking is insignificant in any of the fields. Our results indicate that HPC may be particularly relevant in Chemistry, where it is significantly positively related with first-use of AI. One possible explanation for this finding is that Chemistry often involves computationally intensive tasks, such as molecular simulations, which require substantial computational power provided by HPC resources. This may lead to a stronger dependence on HPC in Chemistry compared to other fields, thus making it a more influential factor in AI adoption.

Social factors: The number of domain collaborations (# Domain col.) is negatively related to AI adoption across all fields. In contrast, the number of computer science collaborators (# CS col.) positively relates to AI adoption in Medicine, Biology, Chemistry, and Materials Science. At the same time, it is not significant in Physics and Psychology. This may reflect different AI integration levels but also differences in prevalence of AI relevant skill sets across domains. Moreover, the number of AI collaborations (# AI col.) relates consistently and significantly to AI adoption in all fields.

Individual factors: The exploratory profile has a significant positive relation to first-time AI use in Medicine, Biology, Chemistry, Physics, and Psychology, but not in Materials Science. The citation stock has a significant negative effect on first-time AI use in Medicine, Biology, and Materials Science, while it is not significant in the other fields. The percentage of international publications (% International pub.) has a significant negative effect on first-time AI use in Medicine and Physics, a significant positive impact in Psychology, and is not significant in the other fields. Why that is the case is not clear at the current state of research.

Table 2.8 shows results on AI re-use at the field level. Considering the realized STHC in the first AI paper, we find that the number of domain authors (# Domain aut.) is negatively and significantly associated with AI re-use across all fields, except for Psychology, where the association is insignificant. This observation is consistent with the pooled regression analysis. In contrast, the number of AI expert authors (# AI exp. aut.) exhibits mixed results. It is positively and significantly related to AI re-use in Medicine, yet negatively related in Chemistry and Psychology, and not significant in other fields. This may indicate that the effectiveness of AI expertise in promoting AI re-use varies between fields, depending on the particularities of each

Table 2.8: Conditional Logit with matching across fields (Reusing AI)

	Dependent variable: Reusing AI					
	Medicine	Biology	Chemistry	Physics	Psychology	Materials science
	(1)	(2)	(3)	(4)	(5)	(6)
# Domain aut.	-0.278***	-0.297***	-0.646***	-0.393***	0.221	-0.469*
	(0.041)	(0.050)	(0.127)	(0.102)	(0.180)	(0.268)
# AI exp. aut.	0.137***	0.065	-0.246**	-0.136	-0.579***	0.001
	(0.039)	(0.047)	(0.103)	(0.112)	(0.194)	(0.232)
# CS aut.	-0.095**	-0.138***	-0.037	-0.028	0.307*	-0.155
	(0.040)	(0.045)	(0.102)	(0.104)	(0.180)	(0.231)
# Newbies aut.	0.229***	0.100**	0.270***	0.333***	-0.410**	0.629**
	(0.036)	(0.044)	(0.102)	(0.105)	(0.179)	(0.244)
# Citations	0.118***	0.199***	0.388***	0.139**	0.269**	0.491***
	(0.023)	(0.029)	(0.073)	(0.068)	(0.110)	(0.134)
AI inst. spe.	0.040	0.093	0.220*	0.013	0.537*	0.175
	(0.089)	(0.082)	(0.132)	(0.129)	(0.293)	(0.222)
Inst. cit.	0.058	-0.141**	-0.076	0.160	0.237	-0.311
	(0.063)	(0.069)	(0.166)	(0.153)	(0.289)	(0.372)
Shanghai ranked	0.029	-0.067	-0.408	0.191	-0.055	0.080
	(0.132)	(0.126)	(0.262)	(0.329)	(0.458)	(0.618)
HPC	0.139***	-0.078	0.178*	-0.099	0.007	0.022
	(0.047)	(0.051)	(0.104)	(0.116)	(0.195)	(0.230)
# Domain col.	-0.246***	-0.329***	-0.229*	-0.107	-0.140	-0.313
	(0.066)	(0.068)	(0.134)	(0.128)	(0.220)	(0.274)
# CS col.	0.104**	0.199***	0.123	-0.103	-0.110	-0.229
	(0.045)	(0.049)	(0.092)	(0.105)	(0.179)	(0.198)
# AI col.	0.293***	0.194***	0.311***	0.375***	0.649***	0.425**
	(0.039)	(0.042)	(0.090)	(0.092)	(0.169)	(0.178)
# Newbies col.	0.265***	0.307***	0.131	-0.002	0.144	0.172
	(0.049)	(0.052)	(0.102)	(0.097)	(0.175)	(0.213)
Exploratory profile	1.386***	0.425	-1.002	0.790	2.039	-2.230
	(0.355)	(0.413)	(0.956)	(1.163)	(1.289)	(2.513)
Citation stock	0.015	-0.031	-0.076	0.048	0.066	0.091
	(0.020)	(0.025)	(0.053)	(0.048)	(0.084)	(0.117)
% International pub.	-0.331***	-0.334***	-0.228	-0.170	-0.460	-0.689*
	(0.087)	(0.095)	(0.211)	(0.195)	(0.317)	(0.418)
Observations	11,982	8,752	2,042	1,808	886	520
Log Likelihood	-3,862.273	-2,881.568	-656.582	-596.359	-270.364	-159.776
LR Test (df = 16)	580.743***	303.288***	102.243***	60.493***	73.400***	40.884***

Notes: This table reports coefficients of the effect STHC on re-using AI across fields. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The effect of STHC on re-using AI is estimated using a conditional logit with matching.

discipline. The number of computer science authors (# CS aut.) is negatively and significantly related to AI re-use in Medicine and Biology. Conversely, it is positively related in Psychology and not significant in other fields. These contrasting effects in different fields may be due to the varying roles of computer science expertise in driving AI adoption, contingent on the subject matter. Further research however is needed to pinpoint the strong variation of coefficient estimates across scientific fields.

The number of new authors (# Newbies aut.) and the number of citations (# Citations) are consistent with our findings at the aggregated level. An exception is found in Psychology, where new authors do not appear to foster the re-use of the technology. This might be because training in data science for psychologists is not well-developed yet, or applications of AI in this field require skills beyond the scope of individuals trained in this area.

Regarding STHC endowments, AI institutional specialization (AI inst. spe.) is positively related to AI re-use in Chemistry and Psychology, while not significant in other fields. This suggests that institutional support for AI may be more critical in some fields than others and that these two fields primarily drive the positive effect seen in the pooled regression. Although not significant in the pooled regression, institutional citation impact (Inst. cit.) is negatively associated with AI re-use in Biology. This could mean that higher institutional research quality reduces AI re-use in Biology, possibly due to a reliance on expertise in more traditional methods. High-performance computing (HPC) is positively related to AI re-use in Medicine and Chemistry and not significant in other fields, emphasizing the importance of computational resources in these domains for persistent integration of AI into research.

Medicine and Biology drive the majority of the social factors observed in the pooled sample, such as the number of domain collaborators (# Domain col.), computer science collaborators (# CS col.), and new collaborators (# Newbies col.), which are only significant in these fields. Also, the number of AI collaborators (# AI col.) is consistent across all fields, highlighting that the availability of AI knowledge through colleagues' expertise is crucial in every field.

Regarding individual factors, the exploratory profile and citation stock variables do not display a consistent pattern across fields. The effect of exploratory profiles is only visible in Medicine. The percentage of international publications (% International pub.) is negatively related to AI re-use in Medicine, Biology, and Materials Science.

2.5 Conclusion

In this study, we examined factors influencing the adoption and reuse of AI in scientific research, focusing on institutional, social, and individual factors.

First, we highlighted the importance of AI specialization within institutions: when AI 'is in the air', the likelihood of adoption increases, which is not all that surprising. Second, and contrary to our expectations, access to high-performance computing resources and affiliation with top-tier universities may not be decisive factors, at least in most fields of application (except, for example, medicine). Third, regarding social factors, we showed that network position matters, particularly when scientists are closely connected to peers who have already used AI in their research. Also, collaborations with early-career researchers – i.e., newbies – seem to contribute significantly to the adoption process, presumably because of their up-to-date training and expertise. Finally, we found that some individual characteristics are important when it comes to integrating AI into scientific practices. A taste for exploration, for instance, seems to enhance the ability of individuals to recognize the potential of AI in their application domains and prompt them to 'give it a try'; conversely, having a dominant position and high reputation within a field tends to hinder this propensity. Taken together, our findings offer some insights for policymakers and science administrators aiming to enhance the diffusion of AI tools in the sciences, providing them with a broader understanding of the complex interplay between these factors. Some critical reflections are therefore in order.

Let us begin with how one might rethink the institutional context in which research takes place. An organizational climate that emphasizes individual competition over cooperation may pose a barrier to knowledge sharing and circulation. Indeed, given the relative ease with which funding for AI research can be obtained (at least at present), scientists may be reticent to share their AI-related knowledge with their colleagues to avoid intensifying competition. Incentives can be put in place to create a supportive culture and foster knowledge circulation within institutions as well as among epistemic communities. Such incentives can be intrinsic, such as recognition and praise, and further supplemented with extrinsic rewards, such as bonuses and higher salaries.

While incentives are valuable, facilitating knowledge sharing can also be achieved by establishing a work environment that promotes interactions among scientists and communication across departments. This can be accomplished, for instance, through informal AI-focused events rather than (often futile) interdisciplinary ambitions. In this regard, organizations should be able to identify 'boundary-spanning' individuals who are eager to share their (tacit) AI knowledge and expertise with their peers, while also possessing effective communication skills to engage non-AI experts and pique their interest. Lastly, we think that organizations could set some research priorities around AI, thus creating a sense of group identity and personal responsibility [Cabrera and Cabrera, 2002]. A relevant theory for this purpose is that of 'organization-based self-esteem' (OBSE), which refers to the degree to which an individual considers him/herself capable, significant, and worthy as a member of the organization (Wang and Noe, 2010). As such, scientists may be more likely to share their knowledge with others if they feel their competencies align with the organization's goals.

Our results seem to suggest that computational resources are not a major determinant of AI adoption, except for some areas. Thus, if the policy ambition is to democratize AI in as many application domains as possible, one may wonder whether large investments in computational facilities such as HPC – which come with substantial overheads and the need for specialized human resources – are the most effective strategy. Alternatively, a more modest but widespread investment in data science/ML laptops and workstations can be a powerful vehicle for AI adoption in the sciences and, why not, a mechanism to broaden access to technology and close computing divides.

We do not rule out, however, that computational resources are a significant asset when it comes to cutting-edge AI research, as evidence suggests (Sevilla et al., 2022). According to a recent study by the OECD [2023], when asked about the main barriers or challenges to accessing AI computation, about 50 % of respondents cited the cost of AI compute. Thus, the lack of financial resources for most public and private organizations can give a group of big players an unfair advantage that results in a concentration of power. More in general, we believe that further research is needed to understand better the demand for AI compute, particularly in domain-specific applications, and not solely for core AI research.

Finally, we refrain from making recommendations on how to nudge individual choices toward exploration while maintaining some degree of exploitation, which is nonetheless essential for scientific progress. Yet we can state with some confidence that current trends in science policy and scientific communities, from impact assessments to targeted research funding (see, e.g., Franzoni et al., 2022), are unlikely to favor exploratory research paths. Perhaps it is time to rethink these habits as well.

2.6 Appendix

Data

This section details some aspects of the data set and how it was treated.

Table 2.9: Number of authors per concept for first AI publication and first publication in the sample

Concept	AI pub. first concept	First pub. first concept
Medicine	29762	31823
Biology	29469	23200
Physics	5727	5995
Chemistry	5691	7525
Psychology	2549	2489
Materials science	1763	1856
Geology	1213	1085
Economics	1080	414
Engineering	856	905
Geography	357	297
Mathematics	180	243
Political science	103	187
Philosophy	54	114
Environmental science	41	53
Business	32	49
Sociology	17	34
Art	10	57
History	3	18
Total	78907	76344

Notes: This table reports the number of authors per concept for their first AI publication and their first publication in the sample

Table 2.10: Number of authors per first publication concept on all OpenAlex

	concept	share (all)	ai_concept	share (AI)
NaN	133883095	0.55	1621837	0.57
Art	2992083	0.01	2843	0.00
Biology	22657217	0.09	288427	0.10
Business	1123248	0.00	9979	0.00
Chemistry	12103929	0.05	98775	0.03
Computer science	8907260	0.04	837505	0.30
Economics	2590735	0.01	44780	0.02
Engineering	4904995	0.02	149083	0.05
Environmental science	150569	0.00	1206	0.00
Geography	789458	0.00	22084	0.01
Geology	1919726	0.01	34395	0.01
History	1351900	0.01	782	0.00
Materials science	4353568	0.02	48982	0.02
Mathematics	2000396	0.01	51949	0.02
Medicine	47610267	0.20	395575	0.14
Philosophy	2531771	0.01	7312	0.00
Physics	10253095	0.04	221476	0.08
Political science	4234858	0.02	13403	0.00
Psychology	3947498	0.02	38739	0.01
Sociology	681340	0.00	1722	0.00
Nb unique author	242118813	1.00	2837138	1.00

Notes: This table reports the number of authors per concept of first publication on all OpenAlex, and for AI articles.

Table 2.11: AI terms used to label articles

Terms
adversarial network
generative adversarial network
artificial intelligence
autoencoder
backpropagation
Bayesian learning
bayesian network
deep belief network
deep learning
ensemble learning
hebbian learning
instance-based learning
Kernel learning
K-means
latent dirichlet allocation
latent semantic analysis
long short term memory
machine learning
extreme machine learning
Markovian
hidden Markov model
multi-layer perceptron
naïve Bayes classifier
natural language generation
natural language processing
natural language understanding
nearest neighbour algorithm
neural network
artificial neural network
convolutional neural network
deep convolutional neural network
deep neural network
recurrent neural network
neural turing
neural turing machine
Q-learning
random forest
regression tree
reinforcement learning
semi-supervised learning
stochastic gradient
supervised learning
support vector regression
transfer learning
unsupervised learning
variational inference
vector machine
support vector machine

Notes: This table reports the list of AI terms used to identify AI articles.

Table 2.12: Regular expression used to label HPC availability

Label	Regular Expressions
Yes	
	'ŷes, '
	't̂he .{0,50} has a high(—-)performance computing'
	't̂he .{0,50} does have a high(—-)performance computing'
	't̂he university of .{0,50} has an infrastructure for high(—-)performance computing'
	't̂here is a high(—-)performance computing'
	't̂he university of .{0,50} has a computational infrastructure'
	'it also has a high(—-)performance computing'
No	
	'n̂o, '
	't̂here is no information available'
	't̂he university of .{0,50} is not mentioned'
	'there is no evidence'
	'there is no mention of'
	'there is no evidence that the university .{0,50} has a high(-—)performance computing infrastructure'
	'it is unclear (whether—if) the'
	'ît is not clear if the university of .{0,50} has a high(—-)performance computing'
	'does not appear'
	't̂he university of .{0,50} does not have a high'
	't̂here is no clear information'
	'i could not find .{0,4}information'
	'does not have its high performance'
	'it appears that the university of .{0,50} does not have'
	't̂here is no information in the provided search'

Notes: This table reports the regular expressions used to label HPC availability in a given institution based on the answer received from Perplexity.ai after asking '...'

Robustness check

As a robustness check, the same analysis has been reported on focal scientists with a first AI paper in the years 2012 to 2018 (instead of 2020), and re-using AI up to 2022. This is of interest because it provides insights on earlier first-time AI papers, and it increases the observation period after the first AI paper, which provides more time to observe re-using AI.

Table 2.13: Descriptives statistics for both matching strategies (2012-2018)

	First-time AI			Re-use AI		
Variable	Matched scientists (without AI)	Focal scientists (with AI)	T-test	Matched scientists (not re-using AI)	Focal scientists (re-using AI)	T-test
STHC endowment						
AI inst. spe.	0.04 (0.21)	0.1 (0.3)	17.03***	0.09 (0.29)	0.1 (0.3)	1.99**
Inst. cit.	3.11 (5.34)	3.27 (1.86)	2.9***	3.26 (2)	3.34 (2)	2.41**
Shanghai ranked	0.05 (0.21)	0.05 (0.21)	0.26	0.05 (0.21)	0.05 (0.21)	-0.09
HPC	0.72 (0.45)	0.75 (0.43)	4.77***	0.74 (0.44)	0.75 (0.43)	1.65*
# Domain col.	131.26 (161.92)	137.2 (202.37)	2.45**	117.15 (170.04)	156.82 (231.27)	11.57***
# CS col.	18.72 (31.15)	25.28 (42.07)	13.39***	21.17 (36.41)	30.11 (46.68)	12.64***
# AI col.	6.58 (14.07)	11.14 (20.46)	19.63***	9.14 (17.65)	14.35 (24.3)	14.51***
# Newbies col.	47.65 (62.73)	49.82 (75.89)	2.36**	41.94 (64.41)	56.24 (85.36)	11.2***
Exploratory profile	0.18 (0.06)	0.18 (0.06)	-0.42	0.17 (0.06)	0.18 (0.06)	2.38**
Citation stock	814.59 (1800.97)	883.4 (1893.5)	2.81***	723.18 (1642.2)	994.24 (2111.29)	8.49***
% International pub.	0.29 (0.25)	0.29 (0.25)	0.74	0.29 (0.26)	0.29 (0.25)	0.79
Realized STHC						
# Domain aut.	10.26 (9.67)	7.5 (7.78)	-23.75***	7.49 (7.61)	7.22 (6.88)	-2.2**
# AI exp. aut.	0.94 (2.28)	2.25 (3.29)	34.97***	2.16 (3.22)	2.41 (3.62)	4.35***
# CS aut.	1.75 (3.47)	2.9 (4.4)	21.98***	2.86 (4.7)	2.95 (4.21)	1.21
# Newbies aut.	1.17 (2.55)	0.98 (1.58)	-6.88***	0.97 (1.71)	0.99 (1.62)	0.59
# Citations	4.55 (13.59)	5.28 (13.14)	4.13***	4.96 (12.49)	5.91 (14.38)	4.16***
Total	11415	11415		7013	7013	

Notes: This table presents the descriptive statistics for various variables for the period 2012-2018, including their mean values and standard deviations (in parentheses) for different matching strategies: those who didn't use AI and those who did, as well as those who didn't reuse AI and those who did. The table also provides results from t-tests to determine if the mean differences between the groups are statistically significant. Significance levels are denoted by ***, **, and * for the 1%, 5%, and 10% levels, respectively. In the AI re-use section, the matched individuals who do not reuse AI are also part of the focal scientists using AI for the first time in the first analysis.

Table 2.14: Conditional logit with matching (first-use of AI) (2012-2018)

	Dependent variable: first-use of AI			
	Institutional	Social	Individual	Full Model
	(1)	(2)	(3)	(4)
AI inst. spe.	0.969***			0.826***
	(0.058)			(0.065)
Inst. cit.	0.420***			0.179***
	(0.042)			(0.050)
Shanghai ranked	0.001			0.025
	(0.069)			(0.079)
HPC	0.078**			-0.008
	(0.032)			(0.036)
# Domain col.		-1.613***		-1.520***
		(0.051)		(0.053)
# CS col.		0.208***		0.233***
		(0.029)		(0.030)
# AI col.		0.846***		0.838***
		(0.027)		(0.027)
# Newbies col.		0.602***		0.570***
		(0.038)		(0.039)
Exploratory profile			-0.176	1.307***
			(0.230)	(0.276)
Citation stock			-0.059***	-0.084***
			(0.010)	(0.016)
% International pub.			0.096*	-0.046
			(0.056)	(0.068)
Observations	22,832	22,832	22,832	22,832
R^2	0.019	0.118	0.001	0.127
Max. Possible R^2	0.500	0.500	0.500	0.500
Log Likelihood	-7,695.999	-6,483.322	-7,896.683	-6,366.616
Wald Test	388.390*** (df = 4)	1,999.590*** (df = 4)	32.420*** (df = 3)	2,105.850*** (df = 11)
LR Test	433.939*** (df = 4)	2,859.293*** (df = 4)	32.570*** (df = 3)	3,092.704*** (df = 11)
Score (Logrank) Test	418.019*** (df = 4)	2,502.995*** (df = 4)	32.518*** (df = 3)	2,683.228*** (df = 11)

Notes: This table reports coefficients of the effect STHC on AI adoption on all fields for the period 2012-2018. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The effect of STCH on AI adoption is estimated using a conditional logit with matching.

Table 2.15: Conditional logit with matching (first-use of AI; 2012 – 2018)

	Dependent variable: first-use of AI					
	Medicine	Biology	Chemistry	Physics	Psychology	Materials science
	(1)	(2)	(3)	(4)	(5)	(6)
AI inst. spe.	1.080***	0.812***	0.557***	0.556**	0.999**	0.744**
	(0.135)	(0.098)	(0.194)	(0.229)	(0.408)	(0.331)
Inst. cit.	0.103	0.310***	-0.314	0.195	0.332	1.395***
	(0.085)	(0.076)	(0.204)	(0.202)	(0.327)	(0.540)
Shanghai ranked	0.208	-0.003	-0.127	-0.312	-0.621	0.628
	(0.137)	(0.118)	(0.279)	(0.424)	(0.401)	(0.695)
HPC	0.017	-0.012	0.157	-0.057	0.005	-0.344
	(0.062)	(0.054)	(0.136)	(0.156)	(0.224)	(0.344)
# Domain col.	-1.409***	-1.536***	-1.781***	-1.872***	-1.270***	-1.942***
	(0.094)	(0.079)	(0.214)	(0.224)	(0.253)	(0.491)
# CS col.	0.257***	0.297***	0.292***	-0.027	-0.031	0.392
	(0.053)	(0.046)	(0.109)	(0.129)	(0.204)	(0.266)
# AI col.	0.850***	0.706***	0.796***	1.274***	1.345***	0.792***
	(0.047)	(0.041)	(0.105)	(0.124)	(0.176)	(0.231)
# Newbies col.	0.684***	0.532***	0.636***	0.570***	0.137	0.101
	(0.068)	(0.058)	(0.149)	(0.163)	(0.201)	(0.350)
Exploratory profile	1.434***	1.241***	1.718	2.511	2.917**	-2.420
	(0.448)	(0.417)	(1.254)	(1.718)	(1.331)	(3.703)
Citation stock	-0.083***	-0.093***	-0.096	-0.038	-0.142	-0.078
	(0.026)	(0.025)	(0.062)	(0.061)	(0.097)	(0.155)
% International pub.	-0.209*	0.114	0.206	-0.372	0.049	0.532
	(0.116)	(0.104)	(0.262)	(0.298)	(0.390)	(0.674)
Observations	8,182	9,814	1,632	1,350	902	378
R^2	0.121	0.122	0.142	0.198	0.204	0.222
Max. Possible R^2	0.500	0.500	0.500	0.500	0.500	0.500
Log Likelihood	-2,309.781	-2,761.502	-440.586	-318.644	-209.607	-83.588
Wald Test (df = 11)	734.490***	885.070***	160.130***	160.190***	109.460***	48.710***
LR Test (df = 11)	1,051.769***	1,279.543***	250.045***	298.460***	206.005***	94.834***
Score (Logrank) Test (df = 11)	924.707***	1,115.180***	210.264***	237.112***	165.562***	75.269***

Notes: This table reports coefficients of the effect STHC on AI adoption across fields for the period 2012-2018. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The effect of STCH on AI adoption is estimated using a conditional logit with matching.

Table 2.16: Conditional logit with matching (re-using AI; 2012 – 2018)

	Dependent variable: re-using AI			
	Institutional	Social	Individual	Full Model
	(1)	(2)	(3)	(4)
# Domain aut.	-0.175***	-0.211***	-0.198***	-0.189***
	(0.034)	(0.038)	(0.035)	(0.039)
# AI exp. aut.	0.125***	-0.025	0.132***	-0.027
	(0.032)	(0.037)	(0.033)	(0.037)
# CS aut.	-0.003	-0.082**	-0.015	-0.073**
	(0.032)	(0.036)	(0.032)	(0.036)
# Newbies aut.	0.122***	0.149***	0.141***	0.143***
	(0.035)	(0.036)	(0.035)	(0.037)
# Citations	0.125***	0.142***	0.098***	0.145***
	(0.021)	(0.022)	(0.021)	(0.022)
AI inst. spe.	0.116**			0.133**
	(0.059)			(0.062)
Inst. cit.	0.147***			-0.062
	(0.053)			(0.058)
Shanghai ranked	0.009			-0.023
	(0.088)			(0.092)
HPC	0.031			0.026
	(0.041)			(0.042)
# Domain col.		-0.335***		-0.330***
		(0.052)		(0.055)
# CS col.		0.115***		0.140***
		(0.037)		(0.039)
# AI col.		0.358***		0.364***
		(0.033)		(0.033)
# Newbies col.		0.321***		0.299***
		(0.042)		(0.043)
Exploratory profile			1.058***	0.952***
			(0.314)	(0.331)
Citation stock			0.166***	0.014
			(0.014)	(0.019)
% International pub.			-0.101	-0.356***
			(0.072)	(0.077)
Observations	14,026	14,026	14,026	14,026
R^2	0.008	0.044	0.018	0.046
Max. Possible R^2	0.500	0.500	0.500	0.500
Log Likelihood	-4,807.279	-4,547.324	-4,736.299	-4,528.260
Wald Test	104.830*** (df = 9)	551.500*** (df = 9)	235.850*** (df = 8)	581.250*** (df = 16)
LR Test	107.525*** (df = 9)	627.435*** (df = 9)	249.483*** (df = 8)	665.563*** (df = 16)
Score (Logrank) Test	106.605*** (df = 9)	599.997*** (df = 9)	244.735*** (df = 8)	634.974*** (df = 16)

Notes: This table reports coefficients of the effect STHC on re-using AI in all fields for the period 2012-2018. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The effect of STCH on re-using AI is estimated using a conditional logit with matching.

Table 2.17: Conditional logit with matching (re-using AI; 2012 – 2018)

	Dependent variable: re-using AI					
	Medicine	Biology	Chemistry	Physics	Psychology	Materials science
	(1)	(2)	(3)	(4)	(5)	(6)
# Domain aut.	-0.058	-0.200***	-0.559***	-0.567***	0.155	-0.321
	(0.062)	(0.065)	(0.191)	(0.145)	(0.208)	(0.564)
# AI exp. aut.	0.024	0.001	-0.385***	0.003	-0.673***	0.379
	(0.057)	(0.061)	(0.146)	(0.160)	(0.227)	(0.514)
# CS aut.	-0.014	-0.104*	-0.073	-0.015	0.305	-0.086
	(0.059)	(0.058)	(0.149)	(0.143)	(0.209)	(0.526)
# Newbies aut.	0.152***	0.122**	-0.029	0.663***	-0.261	1.073*
	(0.058)	(0.059)	(0.150)	(0.154)	(0.215)	(0.573)
# Citations	0.038	0.236***	0.352***	0.096	0.331***	1.102***
	(0.034)	(0.037)	(0.107)	(0.095)	(0.128)	(0.326)
AI inst. spe.	-0.003	0.129	0.292*	0.103	0.580*	0.183
	(0.122)	(0.104)	(0.177)	(0.181)	(0.337)	(0.521)
Inst. cit.	0.039	-0.206**	0.093	0.092	-0.171	-1.010
	(0.100)	(0.090)	(0.230)	(0.203)	(0.383)	(0.858)
Shanghai ranked	0.098	-0.119	-0.448	0.383	0.121	-0.472
	(0.159)	(0.140)	(0.307)	(0.409)	(0.484)	(1.025)
HPC	0.157**	-0.058	0.041	-0.104	0.256	-0.423
	(0.071)	(0.068)	(0.149)	(0.163)	(0.240)	(0.547)
# Domain col.	-0.320***	-0.474***	-0.522***	-0.098	-0.176	0.289
	(0.099)	(0.089)	(0.188)	(0.186)	(0.267)	(0.561)
# CS col.	0.244***	0.209***	0.103	-0.289*	-0.230	0.181
	(0.067)	(0.061)	(0.122)	(0.153)	(0.211)	(0.412)
# AI col.	0.340***	0.287***	0.409***	0.465***	0.809***	0.261
	(0.056)	(0.053)	(0.127)	(0.130)	(0.206)	(0.432)
# Newbies col.	0.314***	0.422***	0.488***	-0.073	0.261	0.272
	(0.075)	(0.071)	(0.150)	(0.148)	(0.202)	(0.455)
Exploratory profile	1.548***	0.456	-0.268	3.018*	3.045*	-3.208
	(0.529)	(0.533)	(1.231)	(1.680)	(1.590)	(6.220)
Citation stock	0.042	-0.040	-0.122*	0.233***	0.037	-0.722***
	(0.029)	(0.032)	(0.073)	(0.070)	(0.097)	(0.275)
% International pub.	-0.464***	-0.148	-0.277	-0.015	-0.721*	-1.094
	(0.130)	(0.126)	(0.292)	(0.283)	(0.399)	(1.058)
Observations	5,724	5,272	1,046	984	644	160
R^2	0.073	0.045	0.058	0.062	0.086	0.167
Max. Possible R^2	0.500	0.500	0.500	0.500	0.500	0.500
Log Likelihood	-1,766.548	-1,705.258	-331.409	-309.618	-194.274	-40.847
Wald Test (df = 16)	352.880***	213.340***	51.710***	52.600***	44.570***	17.220
LR Test (df = 16)	434.478***	243.757***	62.215***	62.821***	57.838***	29.210**
Score (Logrank) Test (df = 16)	404.309***	232.733***	58.259***	59.133***	52.944***	24.419*

Notes: This table reports coefficients of the effect STHC on re-using AI across fields for the period 2012-2018. ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The effect of STCH on re-using AI is estimated using a conditional logit with matching.

Chapter 3

Novelpy: A *Python* Package to Measure Novelty and Disruptiveness of Bibliometric and Patent Data

This chapter was co-authored with

Kevin Wirtz

Summary of the chapter

Novelpy (v1.2) is an open-source *Python* package designed to compute bibliometric indicators. The package aims to provide a tool for the scientometrics community that centralizes various measures of novelty and disruptiveness, enables their comparison, and fosters reproducibility. This paper offers a comprehensive review of the different indicators available in *Novelpy* by formally describing these measures (both mathematically and graphically) and presenting their advantages and limitations. We then compare the different measures on a random sample of 1.5M articles drawn from the Pubmed Knowledge Graph to demonstrate the module's capabilities. We encourage anyone interested to participate in the development of future versions.

3.1 Introduction

Identifying and tracking relevant pieces of knowledge remains a core issue in the Science of Science research. A better understanding of knowledge flow dynamics, mechanisms behind the emergence of new ideas, and identification of novel or impactful documents are crucial for fostering effective science, which will, in turn, help address future societal challenges [Fortunato et al., 2018, Foster et al., 2021, OECD, 2021]. This article proposes integrating various bibliometric indicators within a *Python* package. It assembles within a single module novelty or, more broadly, creativity measurements through combinatorial novelty indicators [Uzzi et al., 2013, Foster et al., 2015, Lee et al., 2015, Wang et al., 2017, Shibayama et al., 2021], as well as several impact measures, including disruptiveness metrics [Wu et al., 2019, Wu and Yan, 2019, Wu and Wu, 2019, Bu et al., 2019, Bornmann et al., 2019a].

This module is intended for researchers in the emerging and multidisciplinary field of Science of Science. There is an increasing tendency to create new scientometric indicators, but there are fewer initiatives to design reproducible experiments. For novelty indicators, there is minimal reference to prior approaches when creating a new indicator; thus, the flexibility in the choice of measures raises the temptation to choose the measure that produces the intended outcome [Foster et al., 2021]. Only a few studies attempt to establish a conceptual background of creativity and the formalization of the indicators [Foster et al., 2021]. This article provides a mathematical and graphical description of these indicators. To the best of our knowledge, it is the first tool that enables the computation of these metrics.

Two macro types of analysis can describe Scientometrics: performance analysis and Science Mapping Analysis (SMA) [Moral Muñoz et al., 2020]. Performance analysis aims to assess the activities of scientific actors and their impact. Its purpose is to assign a value to the productivity and pervasiveness of research conducted by a unit (article, author, institution). SMA "is mostly directed at monitoring a scientific field to determine its (cognitive) structure, its evolution, and main actors within" [Noyons et al., 1999]. It captures a snapshot of a part of the scientific system at a given moment to analyze its structure. The present package allows performing analysis through disruptiveness measures; it also assesses the creative potential of papers using novelty indicators. Both metrics require science mapping analysis to be measured since they are generated through maps of the structure of science. Inputs, outputs, and impacts of these scientific activities are the three perspectives used in bibliometric analysis [Sugimoto and Larivière, 2018][1]. Entities involved in most combinatorial novelty indicators use only the output part of documents to compute their measures [Uzzi et al., 2013, Foster et al., 2015, Lee et al., 2015, Shibayama et al., 2021], except for Wang et al. [2017], which uses references from future articles to control for re-utilization. Disruptiveness indicators [Wu et al., 2019, Bu et al., 2019, Bornmann et al., 2020] take the outputs and impacts of a given document to construct their metrics. They are based on both the references and citations of a given document. This module focuses on metrics using outputs (references/keywords) and impact features (citations/references and keywords from future articles).

While citation is an invaluable source of information, several limitations exist when using the sheer number of citations to evaluate impact. Inter-field (and even intra-field) comparisons can be challenging, as the sheer number of scientists and the way science is performed vary significantly depending on the research domain (methodology, solo author vs. team publication, citation habits). The gap in the number of citations is mainly due to the field's structure and does not necessarily represent the documents' quality. This phenomenon becomes an issue when raw numbers are used to measure the importance of research [Purkayastha et al., 2019]. The same problem arises with self-citation, comparing national and international journals, or document languages [Van Leeuwen et al., 2001].

Network effects have been observed in citation dynamics. Wallace et al. [2012] showed that scholars tend to cite researchers with whom they have a deeper social connection. They also found that researchers are more likely to cite collaborators of collaborators, thereby creating a citation continuum. Articles with international collaborations are more cited due to network effects [Wagner et al., 2019]. Other negative citation behaviors arise in Bornmann and Daniel [2008]; scholars tend to cite papers to satisfy editors and reviewers, showing an apparent disconnection between citation and actual importance during the creation process. Field-specific issues can be addressed using normalization methods or different counting methods of citations (see Waltman [2016] for a comprehensive review). One family of normalized indicators is disruptiveness [Wu et al., 2019, Wu and Yan, 2019, Wu and Wu, 2019, Bu et al., 2019, Bornmann et al., 2019a]. These measures analyze how a focal article

[1] Input refers to human and financial resources and captures the different interactions of agents in the system at various levels (authors/institutional/country levels). Output results from the research process, the different entities that characterize a document. Finally, impact measures knowledge dissemination generated by an article through citations, attention by the general public, or re-utilization of a document's components.

acts as a bottleneck between future papers and the references of the focal papers. They capture whether a document consolidates a domain (i.e., future papers rely on the same pieces of knowledge as the focal paper) or constitutes a starting point for documents from various areas (i.e., future papers only use information from the document).

Scientific advancement is the result of individuals' creativity, where *creativity* is defined as "*held to involve the production of high-quality, original, and elegant solutions to complex, novel, ill-defined, or poorly structured problems*" [Hemlin et al., 2013]. Scholars have proposed measurements to complement these impact indicators with creativity indicators, usually called "atypicality", originality", or "novelty" indicators. The need for quantifying novelty comes from its position as an essential component of the structure of the scientific and economic system. Novelty is at the origin of peer recognition, which acts as a "reward system" for individuals. The "priority rule" grants recognition to the first person making the discovery [Merton, 1957, Carayol et al., 2019]. Novelty is also at the core of the theory developed in evolutionary economics, in which technological progress and creativity influence the cyclical nature of the economy [Schumpeter et al., 1939, Nelson, 1985, Amendola et al., 2014]. Scientific progress remains elusive, and novelty indicators are intended to approach creativity, as making relevant novel combinations is perceived as innovative [Burt, 2004, Rodríguez-Navarro, 2016, Bornmann et al., 2019b]. The earliest novelty indicators focused mainly on past information (i.e., using an entity created the same year) or the distance between articles from a given year, based on their references' overlapping [Dahlin and Behrens, 2005].

More recently, scholars have integrated the conceptual framework of knowledge recombination (a combination of pre-existing ideas that leads to invention) into novelty indicators. This concept was already developed by Poincaré [1910]. Although he refers to the specific case of science, it can be extended to any type of non-scientific creative process where combinations can be both material and conceptual [Winter and Nelson, 1982]. Weitzman [1998] discussed how knowledge could be generated through a combinatorial process of past ideas and how this can generate economic growth as long as potential new ideas are exploitable. At the same time, an invention does not necessarily arise from combining two components together for the first time. Indeed, it can also arise from creating a new relationship between two already linked components [Schumpeter et al., 1939, Henderson and Clark, 1990]. This deepens the idea brought by Jacob [1977] that scientific advancement emerges from looking at something from a new angle rather than incorporating a new instrument. Scientists have proposed a more probability-based approach to capture this combinatorial process. Instead of focusing solely on the degree of novelty of a combination, they look at how unlikely this combination is to happen. The more distant the items in the combination, the more complex and unlikely it is to make this combination. Therefore, the combination is more novel. To solve mathematical problems, Poincaré used the knowledge he found in another field [Poincaré, 1910]. The more distant the fields were, the more insight he gained. However, novel documents exhibit higher variance in citation performance. Academics adopting an exploration strategy face a higher risk of failure [Fleming, 2001, Foster et al., 2015, Wang et al., 2017, OECD, 2021]. Indeed, scientific documents that have a fair mix of novel and conventional ideas are more likely to be "sleeping beauties" than other documents (see Ke et al. [2015] and Wang et al. [2017]). The idea of March [1991] that organizations which explore and consolidate existing processes/technologies are more likely to survive can also be applied in the scientific realm[2]. Novelty indicators can be applied to different entities (patents, papers, webpages, etc.) using various units of knowledge (references, keywords, MeSH terms, text, and others).

Most of the packages available in *R* and *Python* deal with performance or SMA. Moral Muñoz et al. [2020] carried out a detailed and up-to-date review of the different tools and libraries that help researchers in their daily work. Although much work has been done to study citation, co-authorship, or any coupling, novelty and disruptiveness indicators are still unavailable, and researchers have to code these metrics themselves. Concerning the reproducibility of novelty studies, only Shibayama et al. [2021] shared their code on Github to calculate their new novelty indicator, but this is still an isolated event. This tool, therefore, ensures that indicators of novelty and disruption used in future studies will be replicable.

The rationale for incorporating novelty and disruptiveness indicators in a single package comes from the fact that they both capture different aspects of the documents: the former aims at quantifying the risky profile of research, looking at the balance between exploitation and exploration [March, 1991] of the knowledge space. At the same time, the latter analyzes how impactful an article is for science. The link between novelty and citation count has been of interest in previous research [Uzzi et al., 2013, Wang et al., 2017], and more recently, Lin [2021] studied the relationship between novelty and disruptiveness indicators. The different studies only look

[2]Here, survival can be expressed as a high citation count.

at specific novelty indicators, and a complete benchmark is still missing. This paper contributes to an ongoing effort to systematically benchmark and compare multiple indicators of impact and novelty by proposing an open-source tool to the community.

This article contributes to the Science of Science literature by providing an open-source *Python* package, *Novelpy*, to compute Novelty and Disruptiveness measurements. It unifies the existing indicators in a common framework using a formalization based on graph theory and provides some hands-on experience. We hope that *Novelpy* will contribute to homogenizing our practice in the science of science and support researchers in their work. The package will be available in *Python*, one of the most popular open-source programming languages (hence with the most prominent community support), and will be maintained long-term. The package currently works with a specific and documented data structure, but tools to easily use well-known data sources are under development. The package will be hosted on *PyPI* and also on *Github*, which allows the creation of bug reporting and/or proposition of development[3]. The rest of the paper is structured in the following way. In Section 3.2 contains the formalization of the indicators that are implemented in *Novelpy*. Section 3.3.2 demonstrates the package's capabilities on a random sample drawn from PubMed. We close the paper with a discussion on the remaining limitations of novelty indicators' usages and the purpose of the package.

3.2 Supported indicators

This section details the content of *Novelpy*, describes the computation for each indicator, and the data required. The *Novelpy* Python package provides a set of functions to perform quantitative analysis in scientometrics. The structure of the module is divided between novelty and disruptiveness indicators. Novelty indicators are also separated between indicators based on co-occurrence matrices and ones based on text embedding techniques, as represented in figure 3.1.

Practically, disruptiveness indicators are all calculated through the same function, while novelty indicators have a function for each measure. All functions are explained in the module's documentation (`https://novelpy.readthedocs.io/`).

Different data types can be employed depending on the indicator, as shown in 3.1. All indicators working with a co-occurrence matrix can use references, journals, or keywords, and disruption indices rely on the citation network. Shibayama et al.

[3]Documentation is available here `https://novelpy.readthedocs.io/en/latest/usage.html`

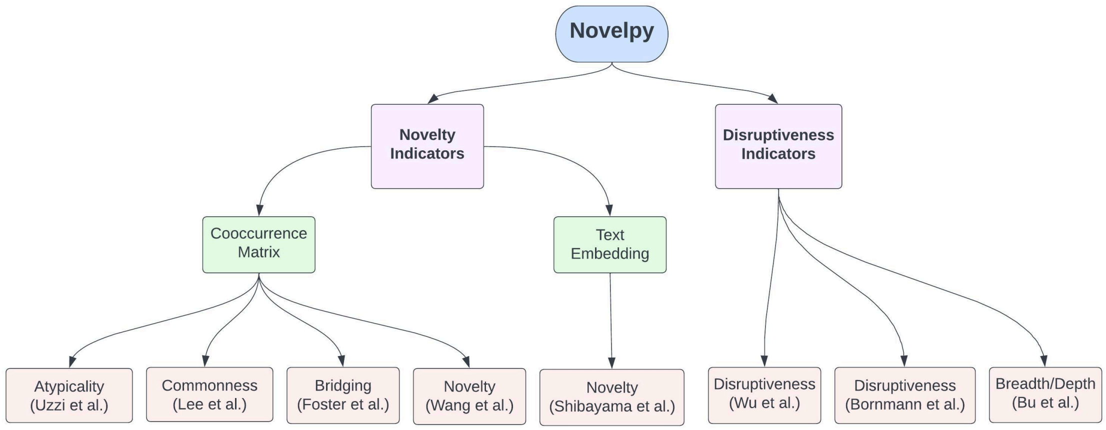


Figure 3.1: *Novelpy*'s module structure

[2021]'s indicators use the citation network and title or abstract to represent the article's semantics in a vector space. Various tools to preprocess bibliometric data are also included within the package to simplify the computation of proposed measures[4] (e.g., co-occurrence matrix construction, text embedding, citation and co-authorship network creation). Table 3.1 summarizes the indicators available in the module, their strengths and weaknesses, and the possible variables to compute them.

The module supports a wide range of data sources as long as they are in the proper format; note that transforming data to the expected structure is relatively simple. Helper functions are available to directly transform PubMed Knowledge Graph data into the desired structure[5]. For other databases, further backend to OpenAlex, Web of Science, Scopus, and PATSTAT are under construction. The package currently works with documents in JSON or MongoDB format. Mongo will be preferred for large databases to avoid overflowing the RAM.

3.2.1 Novelty Indicators

We focus on novelty indicators in the package based on the combinatorial idea. As discussed in section 3.1, novelty indicators can be differentiated into two groups regarding how they compute the distance between items. The first group uses a combination of items, such as keywords and journals, to create a co-occurrence matrix. Algorithms make use of this matrix to compute the distance. The more distant,

[4] see `https://novelpy.readthedocs.io/en/latest/utils.html`

[5] Expected structure is presented here: `https://novelpy.readthedocs.io/en/latest/usage.html#format-supported`

Type	Indicator	Pros	Cons	Ref. Journals	Keywords	Citation net.	Title/Abs.
Novelty	Uzzi et al. [2013]	Conserve dynamical citation structure	Computationally intensive	X	X		
	Lee et al. [2015]	Computationally lightweight Data-saving	Conceptually less advanced	X	X		
	Foster et al. [2015]	Consider undirect link Computationally lightweight	Discret distances	X	X		
	Wang et al. [2017]	Computationally lightweight	Data-Intensive	X	X		
	Shibayama et al. [2021]	High granularity	Computationally and data-intensive			O	O
Disruptiveness	Wu et al. [2019]	Normalized	Data-intensive Issue with term K_{FP}			X	
	Bornmann et al. [2019a]	Normalized	Data-intensive			X	
	Bu et al. [2019]	Normalized	Data-intensive			X	

Table 3.1: *Novelpy*'s available indicators. X means that you can run the indicator on either variable. O Means you need both variables to run it

the more unexpected and, therefore, novel the combination. The second type of indicator maps items in a Euclidean space with text embedding techniques like word2vec [Mikolov et al., 2013b]. The distance is then computed in this semantic space. As shown in Figure 3.1, novelty indicators are split between those using co-occurrence of entities such as journals or keywords and those using word embedding techniques. For the first group of indicators, we first need to create a co-occurrence matrix for each year of the given dataset. While some indicators only use the focal year to compute the score for each combination [Uzzi et al., 2013, Lee et al., 2015, Carayol et al., 2019], others take into account past combinations in the score calculation [Foster et al., 2015] and future re-utilization [Wang et al., 2017].

Atypicality [Uzzi et al., 2013], Commonness [Lee et al., 2015], and Novelty [Wang et al., 2017] are all indicators that use references of an article at a journal level. Previous studies usually focused on one type of knowledge unit, but as long as one can create a co-occurrence matrix between items, it becomes trivial to generalize. Carayol et al. [2019] reformulate Lee et al. [2015] and apply it to keywords and construct the indicator accounting for inter-field heterogeneity by splitting the analysis.

Fleming [2001] computes a combination of patent subclasses, a prevalent practice in patentometrics. Dahlin and Behrens [2005] propose a novelty measure based on the overlapping between documents' references that was reused by Trapido [2015]. Based on this work Matsumoto et al. [2021] propose an extension that computes the average share of references that are shared between a focal paper and all other documents in the same field. These indicators are not present in *Novelpy* (v1.2) but will be added in future versions.

Although the co-occurrence matrix can be considered an adjacency matrix, only a handful of indicators use graph theory to compute the distance between items. Indeed, indicators *à la* Uzzi et al. [2013], or Lee et al. [2015] take into account only the direct neighborhood during distance calculation. If items A and B are close, items B and C are close, and D is unrelated to any of them, then the combination of A and C is more likely to happen than A and D. This logic is completely ignored if one considers the direct neighbors. Wang et al. [2017] integrated this into their indicator by considering the cosine similarity between nodes' neighbors, which considers common friends (A and C in the example above). Using community detection as in Foster et al. [2015], one can better represent the distance between two units by using the global structure of the network. However, the discrete nature of the novelty score can be argued. Using text embedding, one can have a continuous representation of the distance between items. This distance is related to the text's structure since word similarity depends on their neighborhood. Some initiatives used these techniques with different purposes but could be used to create a novelty score. Hain et al. [2020] create a similarity measure between patents using word2vec [Mikolov et al., 2013b]. Shibayama et al. [2021] was the first to apply word embedding techniques in a novelty context. They embed references in a Euclidean space using spaCy and then compute a distribution of cosine distances between documents present in the references for a given document.

We propose a mathematical formalization of these indicators. Setting up this framework offers a basis for defining future new indicators. These indicators are formulated based on graph theory, where the network's nodes are units of knowledge (journals, keywords, or references), and edges represent the co-occurrence of these units in entities (documents or patents).

- Co-occurrence matrix can be written as a graph $G = (V, E, w)$.
- Set of nodes V of dimension v represent here the entities (e.g. keywords,

journals), a given entity is defined as V_i.

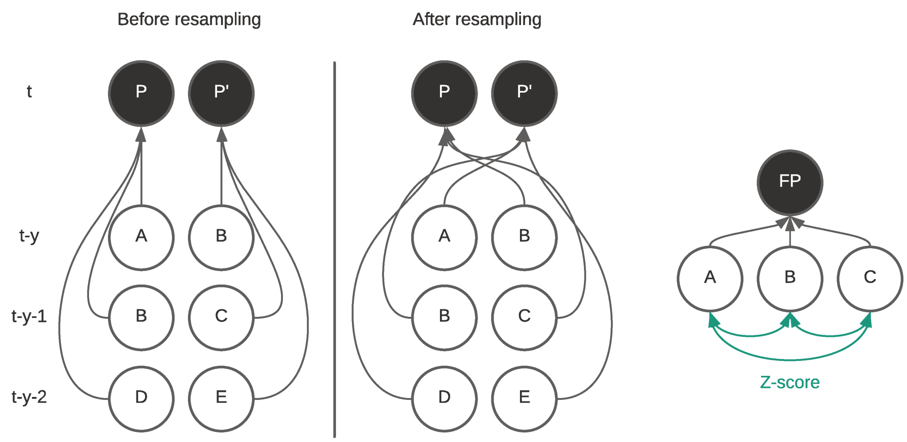
- Set of edges is noted E.
- Number of combinations between V_i and V_j is the weight for the edge (V_i, V_j) and is written $w(V_i, V_j)$.
- Degree of a node V_i is written k_i. N is the sum of the weighted edges in G without self-loops, $N = \Sigma_{i=1}^{v-1}\Sigma_{j=i+1}^{v} w(V_i, V_j)$.

D define our set of documents of dimension n. Each focal paper, FP, has its network, which can be defined as G_{FP}, E_{FP} is the subset of edges present in document FP. G_{FP} uses the same set of nodes V as G and can be express as $G_{FP} = (V, E_{FP}, w_{FP})$. In some cases, G_{FP} is an unweighted network and will be written then $G_{FP} = (V, E_{FP})$. The number of links, $w(V_i, V_j)$, is then defined as the sum of all combinations of two given entities overall document in D, $w(V_i, V_j) = \Sigma_{d=1}^{n} w_d(V_i, V_j)$ where $w_d(V_i, V_j)$ is binary if the graph is unweighted at the document level. $G(V, E, w)$ can be split at a year level. For example, in year t, and the associated network will be noted $G_t(V, E_t, w_t)$. Uzzi et al. [2013], Lee et al. [2015], use only the subgraph G_t for calculation. Foster et al. [2015] use the accumulation of past networks. For Wang et al. [2017], several subgraphs are involved in computing the indicator. The novelty indicators *à la* Wang et al. [2017] deal with four subgraphs of G. One needs to consider two different past sets of documents (noted P and B) and a set of future documents (noted F).

3.2.1.1 Uzzi et al. [2013]: Atypicality

The goal of the measure proposed by Uzzi et al. [2013], called "Atypicality", is to compare an observed network with a random network. The network is shuffled, preserving the temporal distribution of references at the paper level. As shown in Figure 3.2, a document citing two articles from, for example, 1985 and one from 1987 will still cite articles published the same year, but the journal can change. The frequency of the combination (V_i, V_j) at time t is defined as $w_t(V_i, V_j)$, and we extract the adjacency matrix of observed frequencies. The idea is basically to compute the frequency Z-score for each journal combination. The Z-score is defined as $z = (obs - exp)/\sigma$; an observed frequency is compared with a theoretical one.

The theoretical frequency is generated through Markov chain Monte Carlo simulation, preserving the dynamical structure of citations. In the case of Atypicality,



(a) Resampling strategy (b) Score calculation

Figure 3.2: Uzzi et al. [2013] [6]

we are dealing with $s+1$ different networks for the year t, the existing network and s resampled ones. The existing network is G_t, as defined above. The others are generated by preserving an article's temporal distribution of references. For each document FP, we want to keep the number of references published in year $t-y$ stable for all y to ensure that the global age distribution of the pieces of knowledge used at time t remains stable.

One needs to generate s random networks G_t. After re-sampling, the publishing year of references is no longer taken into account. Edges' weights are then aggregated to fit with G_t edge structure E_t by summing over all combinations. The observed frequency for each sample is computed for each edge (V_i, V_j). We write the set of frequencies for the combination of V_i and V_j in the s samples $w_t^s(V_i, V_j)$. One can then compute the mean and standard deviation for each edge's frequency and compute a z-score.

$$Z - score_{ijt} = \frac{w_t(V_i, V_j) - mean(w_t^s(V_i, V_j)}{std(w_t^s(V_i, V_j))}$$

For each paper, taking all combinations made (E_{FP}), a distribution of z-score written Z_{FP} is computed, and the 10th percentile (P_{10}) of this distribution (the novelty) and the median (P_{50}) (the conventionality). The novelty and conventionality for document FP are then written:

[6](a): P and P' are two distinct papers, P cites journals A, B, and D. P' cites journals B, C, and E. The goal is to shuffle the network by conserving the dynamic structure of citations at the paper level. P is no longer citing A from $t-y$ but cites B from year $t-y$. (b): Comparing the observed and resampled networks, we can compute a z-score for each journal combination.

$$Novelty_{FP} = P_{10}(Z_{FP})$$

$$Conventionality_{FP} = P_{50}(Z_{FP})$$

While this indicator only requires data from a specific year, it is still computationally greedy. Indeed, generating the s samples and the computation of the average and the standard deviation for each possible combination is expensive. On the contrary, this indicator allows for keeping the temporal structure stable, which is more in line with the reality of the availability of the knowledge pieces.

3.2.1.2 Lee et al. [2015]: Commonness

Lee et al. [2015] compares an observed network with a theoretical network (Observed vs Expected frequency of edges) at a year level. The observed number of combinations (V_i, V_j) at time y_t is the number of edges $w_t(V_i, V_j)$, the theoretical number of combinations is $\frac{k_i * k_j}{N_t}$, the degree for entity i and j multiplied together and divided by the total number of combinations made in year t.

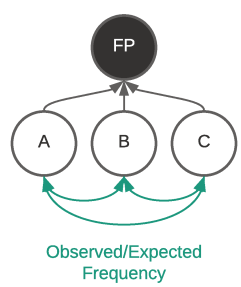


Figure 3.3: Lee et al. [2015]

$$Commonness_{ijt} = \frac{w_t(V_i, V_j) * N_t}{k_i * k_j}$$

For each paper, taking all combinations made in document FP (E_{FP}), a distribution of commonness-score written C_{FP} is computed. The commonness for document FP is the 10th percentile (P_{10}) of this distribution and is written as:

$$Commonness_{FP} = -log(P_{10}(C_{FP}))$$

The main advantage of the commonness indicator is its speed of calculation; it is the least demanding indicator in terms of the execution time of the package. The indicator only requires data from a specific year. Note that this indicator is very close to Uzzi et al. [2013]'s one. Both would be equal if Uzzi et al. [2013] resampling method would not consider the references' publishing year.

3.2.1.3 Foster et al. [2015]: Bridging

Foster et al. [2015] propose a novelty indicator based on community detection algorithms. It captures the distance between two entities taking into account undirected edges. The goal of the measure is to identify the network's community studied and capture proximity through the community in which the combined entities are clustered.

Any community algorithm can be applied to this indicator. We rely on the Louvain algorithm in *Novelpy* following Foster et al. [2021], but we intend to add further options. After applying the community algorithm on $G(V, E, w)$, we are left with multiple clusters of entities. C_i is the community to which the entity i belongs.

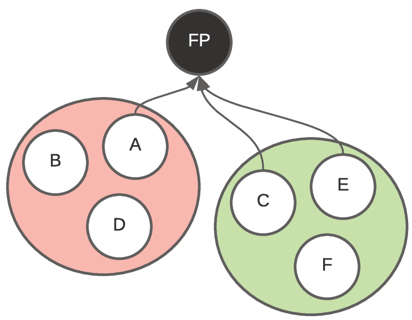


Figure 3.4: Foster et al. [2015] [7]

$$Novelty_{FP} = \frac{\sum_{(i,j)\in E_{FP}} 1 - \delta(C_i, C_j)}{|E_{FP}|}$$

Where $\delta(C_i, C_j) = 1$ if $C_i = C_j$ (i.e., both entities, i and j, are in the same community), $\delta(C_i, C_j) = 0$ otherwise. The novelty score of an entity is the proportion of pairwise combinations that are not in the same community.

This indicator brings into the field algorithms that capture the global network structure and only require data from a specific year. At the same time, this indicator

[7] FP cites different journals which belong to different communities. The novelty is the number of journal combinations from two different communities. Communities of journals are computed through a community detection algorithm.

does not allow measuring distances between communities and proposes only a binary distinction.

3.2.1.4 Wang et al. [2017]: Novelty

Wang et al. [2017] propose a measure of difficulty for pairs of references that have never been made before. These new pairs need to be reused after the given publication's year (scholars do not have to cite directly the paper that creates the combination, but only the combination itself). The idea is to compute the cosine similarity for each journal combination based on their co-citation profile b years before t. The cosine similarity between W_i^B and W_j^B is defined:

$$COS(W_i^B, W_j^B) = \frac{W_i^B . W_j^B}{\|W_i^B\| \|W_j^B\|}$$

where W_i^B represent all links of entity i, B years before year t.

Novelty *à la* Wang et al. [2017] relies on four subgraphs of G constructed using two different past sets of documents, a set of future documents, and the set of documents for the focal year. These different subgraphs are defined as follows (note the first year of the dataset y_0 and the last as y_n):

- $G_t = (V, E_t, w_t)$ is a subgraph of G from year t (documents published year t)
- $G_P = (V, E_P, w_P)$ is a subgraph of G from year t_0 to $t-1$ (documents published before year t)
- $G_B = (V, E_B, w_B)$ is a subgraph of G from year $t-b$ to $t-1$ is used to measure the cosine similarity between nodes. This set is a subgraph of G_P (documents are published in a given window before year t)
- $G_F = (V, E_F, w_F)$ is a subgraph of G from year $t+1$ to $t+f$ (documents published in a given window after year t)

This indicator focuses on new combinations reused afterwards and not achieved before the given year y_t. One needs to keep all elements of $E_t \notin E_P$ and $E_t \in E_F$. More precisely, edges belonging to the following subset (that we call E_N) are the only edges used to compute this indicator $E_N = (E_t \cap E_F) \cap \overline{E_P}$

Cosine similarities are calculated using G_B. For each document, we compute an undirected and unweighted network. The novelty is the sum of all edges from $E_{FP} \in E_N$, that is:

$$Novelty_{FP} = \sum_{(i,j) \in E_N} 1 - COS(W_i^B, W_j^B)$$

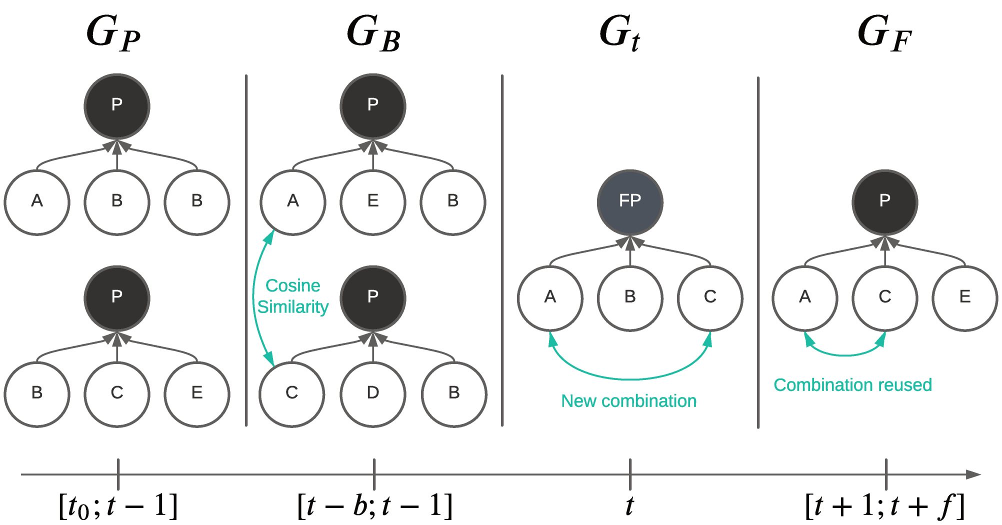


Figure 3.5: Wang et al. [2017] [8]

The main issue with this indicator is the amount of data needed to compute the measure. One needs as much data as possible before the focal year to ensure that the combination has never been made. At the same time, some hyperparameters involved in this measurement can drastically modify the results. For example, the time window to capture the re-utilization of a combination or the number of times reused needed to be novel is very arbitrary.

3.2.1.5 Shibayama et al. [2021]: Novelty

Shibayama et al. [2021] propose to incorporate semantic distances to capture diversity in the set of references from a given article following Hain et al. [2020] and their similarity measure between patents. Document centroids are computed by summing all word representations for each document.

Consider a directed unweighted graph $G(V, E)$ containing the citation network. For a given document FP, a referenced document is denoted by r, and the set of nodes that are cited by FP is then $In_{FP} = r : (FP, r) \in E$. Shibayama et al. [2021] compute all distances between each document's centroids ($C^2_{|In_{FP}|}$ combinations).

[8] For a given article at time t, we check if the journal combined were already combined in the past (G_P). We then check if the combination is reused in the future (G_F). If the combination is new and reused, the difficulty of making such a combination is calculated on the recent past (G_B)

All documents have a vectorial representation in a semantic space of length 200. Distances between two references $i, j \in In_{FP}$ are calculated through cosine similarity: $n_{ij} = 1 - COS(T_i, T_j)$, where T_i is the dense vector text representation for a document i. A distribution of novelty scores $N_{FP} = n_{ij} : i, j \in Out_{FP}$ is then computed, and for each document, the final score is a percentile of N_{FP}.

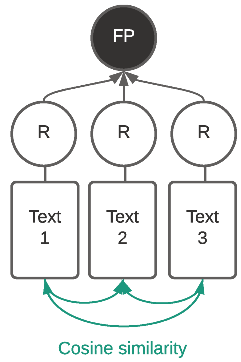


Figure 3.6: Shibayama et al. [2021] [9]

Shibayama et al. [2021]'s indicator is both data-intensive and computationally intensive. One needs to obtain all references' titles/abstracts for a given set of articles. The package currently works with a pre-trained Word2Vec model, *en_core_sci_lg* from *spacy*, to compute the dense representation of a document. Future versions will incorporate a back-end to use any pre-trained model.

3.2.2 Disruptiveness Indicators

Disruptiveness indicators offer alternative measures of impact to the number of citations. They allow understanding if a given article behaves as a bottleneck between the knowledge mobilized in a given article and the articles that will cite it. Disruptiveness was introduced in scientometrics by Wu et al. [2019] and was previously proposed for patents by Funk and Owen-Smith [2017]. Following Azoulay [2019]'s definition, a paper can either consolidate existing knowledge or disrupt it. If future papers that cite a focal paper and its references do not use fundamental new pieces of knowledge created in it (i.e., the focal paper consolidates the existing knowledge space but does not disrupt the playing field). On the other hand, if future papers cite only the focal paper and not its references, then the focal paper is considered

[9] For a given article, each reference's abstract (or title) is represented in a semantic space through text embedding techniques. The distance between two references is then computed through cosine similarity.

disruptive. Quoting Bornmann et al. [2019a], "[...] many citing documents not referring to the FP's cited references indicate disruptiveness. In this case, the FP is the basis for new work which does not depend on the context of the FP, i.e., the FP gives rise to new research." All presented measures normalize citation and give a relative perspective on a publication's impact [Bu et al., 2019]. Disruptiveness indicators consider the importance of pieces of knowledge (references) in a given article for other articles, whereas Depth and Breadth, as proposed in Bu et al. [2019], capture how the knowledge generated by that given item is reused and whether it allows for the consolidation of a domain or is instead used in a disparate manner.

Consider a directed unweighted graph $G(V, E)$ containing the citation network.

- For a given document FP we note a document cited by FP, r. The set of nodes that are cited by FP is then $In_{FP} = \{r \in V|(FP, r) \in E\}$
- For a given document FP we note a document citing FP, c. The set of nodes that are citing FP is then $Out_{FP} = \{c \in V \ |(c, FP) \in E\}$
- The number of citations for FP is then $deg^{-}(FP) = |Out_{FP}|$ and number of references $deg^{+}(FP) = |In_{FP}|$
- The set of references for an article citing FP is then noted In_c

3.2.2.1 Wu et al. [2019]: Disruptiveness

By adapting Wu et al. [2019] notation, we called I_{FP} the set of nodes with FP as a parent that does not have FP's parents as parents. More formally $I_{FP} = \{c \in Out_{FP}|In_c \notin In_{FP}\}$. The set of J^l_{FP} is the set of nodes with FP as a parent that share at least l parents with FP. We note $J^l_{FP} = \{c \in Out_{FP}||\{In_c \in In_{FP}\}| > l\}$. Finally, K_{FP} is the set of nodes that share parents with FP but that do not have FP as a parent: $K_{FP} = \{v \in V|v \in In_{FP}\}$.

The disruptiveness *à la* Wu et al. [2019] is then noted :

$$DI_1 = \frac{|I_{FP}| - |J^1_{FP}|}{|I_{FP}| + |J^1_{FP}| + |K_{FP}|}$$

Some variants that consider only paper sharing at least l references have been proposed:

$$DI_5 = \frac{|I_{FP}| - |J^5_{FP}|}{|I_{FP}| + |J^5_{FP}| + |K_{FP}|}$$

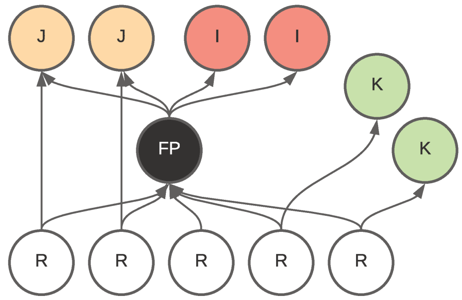


Figure 3.7: Wu et al. [2019], Bornmann et al. [2019a] [10]

3.2.2.2 Bornmann et al. [2019a]: Disruptiveness

A variant that removes the term $|K_{FP}|$ has been proposed by Wu and Yan [2019] because the number of documents that cite references from the focal documents without citing the focal documents is often too large compared to the paper from other sets. Wu and Wu [2019] show how considering the set K_{FP} can lead to a decrease in disruptiveness when the term $|I_{FP}| - |J^1_{FP}|$ is negative. In that configuration, more papers that do not cite FP ($|K_{FP}|$) lead to higher disruptiveness, which is different from how the indicators conceptually work. Defined as DI^{nok}_l by Bornmann et al. [2019a], we note:

$$DI^{nok}_l = \frac{|I_{FP}| - |J^l_{FP}|}{|I_{FP}| + |J^l_{FP}|}$$

3.2.2.3 Bu et al. [2019]: Breadth and Depth

Bu et al. [2019] propose an alternative to the above disruptiveness indicators. It calculates the proportion of articles citing the focal paper that also cites other articles citing it. The indicator allows us to understand whether the document contributes to a restricted research domain; the documents citing the focal paper are interdependent and cite each other. On the contrary, the documents using the focal paper's research may also be unconnected and belong to more extensive research space.

Let FP be the focal paper, the articles citing it the set Out_{FP}. We are interested in the articles cited by the documents of the set Out_{FP}. For each element c of Out_{FP}, we observe a set of associated references named In_c. The proportion of documents citing document FP and also citing documents that are citing FP is then written

[10]For a given article FP, we retrieve: (a): Articles citing FP and references from FP (named J). (b): Articles citing FP but no references from FP (named I). (c): Articles citing references from FP but do not cite FP (named K).

[11]For all articles citing FP, we check if they also cite papers citing FP.

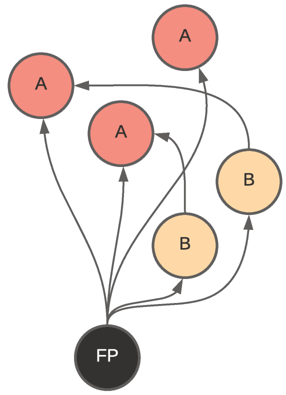


Figure 3.8: Bu et al. [2019] [11]

as:

$$Depth_{FP} = \frac{|\{c \in Out_{FP} : |In_c \in Out_{FP}| > 0\}|}{|Out_{FP}|}$$

On the contrary, the breadth, the proportion of papers citing FP that do not cite other publications also citing FP, is written:

$$Breadth_{FP} = \frac{|\{c \in Out_{FP} : |In_c \in Out_{FP}| = 0\}|}{|Out_{FP}|} = 1 - Depth_{FP}$$

Bu et al. [2019] also propose a measure of dependence. It captures the average number of references shared between the focal paper FP and documents citing it. In_{FP} is the set of references of FP. For all document c that cite FP (Out_{FP}), we want to know the number of references shared:$|\{In_c \in In_{FP} : c \in Out_{FP}\}|$. The average number of references shared between document FP and all documents citing it ($c \in Out_{FP}$) is then:

$$Dependence_{FP} = \frac{\sum\limits_{c \in Out_{FP}} |In_c \in In_{FP}|}{|Out_{FP}|}$$

Two other indicators from Bu et al. [2019] are not computed in our function: Independence and Dependence. However, they represent the proportion of publications citing a focal paper that also cites references from the focal paper. Using notation from 3.2.2.1: $\frac{|I_{FP}|}{|I_{FP}|+|J_{FP}^1|}$ one can easily derive this value from disruptiveness indicators DI_1^{nok}. Indeed from $DI_1^{nok} = \frac{|I_{FP}|-|J_{FP}^1|}{|I_{FP}|+|J_{FP}^1|}$ we can compute the independence, the proportion of articles citing the focal paper that do not cite articles cited by the focal paper

$$\frac{|I_{FP}|}{|I_{FP}|+|J_{FP}^{1}|}=\frac{DI_{1}^{nok}+1}{2} \quad \text{if} \quad |Out_{FP}|>0$$

All these measures are quite demanding in terms of data requirements. Indeed, for each given article, we need to access the references, the articles citing the focal paper, and the articles citing the references of the focal paper.

3.3 Sample analysis

3.3.1 Descriptive statistics

This section provides examples of applications that could be performed with *Novelpy*. We use the Pubmed Knowledge Graph (PKG) sample [Xu et al., 2020], which stores research articles published on Pubmed and offers metadata for all papers. This analysis is proposed as an example to demonstrate our module features after computing the indicators[12]. All figures and tables can be found in the appendix. The sample is restricted from 1995 to 2015; the focal period is 2000-2010. The sample is composed of 1,469,352 papers and 2,959,650 distinct authors. Authors are disambiguated in PKG using advanced heuristics and algorithms. The sample was selected so that every article has the attributes needed to run the indicators. Each paper lists references, mesh terms, authors, titles, and abstracts. Table 3.2 and Figure 3.9 summarize the statistics of the sample. On average, the number of references used in a paper is 23, consistent with typical citation behavior [Abt and Garfield, 2002]. The number of papers almost doubled in 10 years, which is in line with the literature [Fortunato et al., 2018].

3.3.2 Results

As discussed in previous sections, research on novelty indicators still needs to be conducted across multiple dimensions. *Novelpy* will facilitate computing different indicators on various entities. Researchers can then use the novelty scores provided by the package to perform their analyses. Individual-level analysis can be conducted by examining the distribution of novelty scores, as shown in Figure 3.10. Comparing indicators and studying the evolution of novelty over the years are the primary mo-

[12] Interested readers will find code and resources to create tables, plots, and indicators here `https://novelpy.readthedocs.io/en/latest/usage.html#id5`

tivations for this package. Only a few studies examine the dynamics of novelty over time. Nevertheless, understanding the evolution of creativity in papers, patents, or other entities can offer insights into the trade-off between exploration and exploitation of the research space in a given field. Figure 3.11 displays the evolution of the mean novelty score for each indicator, given the variable (references, mesh terms). We cannot draw conclusions since the sample is random and aggregated across all fields within Pubmed. The pattern of trends varies significantly depending on the indicator and variable. This heterogeneity might be evidence that further investigation is required to understand precisely what these indicators capture and in which cases they best predict novelty. This question is even more relevant, considering the lack of correlation between indicators in Figure 3.12.

3.4 Discussion

This paper aims to demonstrate the capabilities of the new *Python* package *Novelpy*. We presented a sample analysis using the functions within this package to showcase how it can assist interested readers in computing and analyzing existing indicators or addressing current challenges related to novelty measurement. Several critiques can be made on current novelty measurements, and addressing these points is crucial for solidifying our understanding and usage of these indicators.

The diversity and convergence in how novelty indicators are created raise questions about what they measure. As observed in our sample analysis, the results are highly dependent on the indicator used, which confirms previous concerns about cherry-picking the indicator [Shibayama et al., 2021, Foster et al., 2021]. Simultaneously, indicators often focus on the same entity (keywords or reference journals). Recent measures like Shibayama et al. [2021] and Arts et al. [2021] broaden this domain by utilizing text information from references. Novelty indicators are rarely conceptualized and often require a qualitative background. Qualitative studies like Tahamtan and Bornmann [2018] question the significance of literature in authors' creative processes. The link between references and creativity is debated, and further investigation is needed to determine if references can be reliably used as a proxy variable for creativity.

Research evaluation was once performed solely by experts in the scientometric field and specialists working for public institutions. Open access data has recently led to entrusting this evaluation to a broader range of researchers and public workers.

These new actors need the necessary tools to compute scientometric indicators and some understanding of their relevance. Using software creates a gap between the user and the actual data, which may lead to issues if the assumptions necessary for the indicators' relevance are overlooked. Data-driven decisions can become inefficient if the algorithm used is a black box and is misused. A solid background in how and why these indicators are created is necessary to limit bias in selecting indicators when used in research. As seen in Section 3, every indicator has its pros and cons, different hyperparameters (time window, re-utilization, number of samples, and others), and is highly dependent on the database used. The coverage varies greatly depending on the database (language, fields, nationality, and others) [Sugimoto and Larivière, 2018]. These aspects and the increasing number of novelty indicators create arbitrary decision-making when using them. Sugimoto and Larivière [2018] suggests that indexing and classification of documents differ between databases, making it challenging to reproduce studies on other databases. Constructing a general indicator applicable to all scientific disciplines is difficult, as citation habits are heterogeneous, making comparisons between fields risky [Carayol et al., 2019] (proposing to compute scores by field, but this is not the norm). Depending on the country, methods and standards may differ within a discipline, and the historical practice of a field may change the representations.

Improving novelty measurement is essential for supporting innovative research. Highly novel documents are less likely to be cited in the short run and are less likely to be published in high-impact factor journals [Wang et al., 2017, Mairesse et al., 2021]. Due to the pressure from citation count evaluation, the exploration of science is less likely to occur. Researchers may tend to conform to conventional references within their field, which is already accentuated during the submission process. Documents that are already highly cited, considered stepping stones in the field, will thus receive even more citations, creating a vicious circle. This vicious circle has the consequence of narrowing research, where only those who agree with the existing paradigm are rewarded with citations.

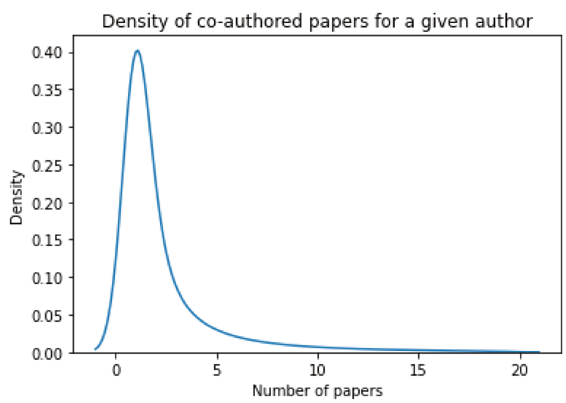
This phenomenon is already observed in AI research, where topics become increasingly less diverse [Klinger et al., 2020]. The goal of science is not to persist with merely satisfactory solutions but to explore a range of possibilities, even those that may prove fruitless. Citation indicators typically do not emphasize researchers who take risks by attempting novel approaches. Various funding methods exist to support high-risk, high-reward (i.e., highly novel) research [OECD, 2021]. Experts are not free from bias when evaluating novelty, funding processes are not uniform, and many decisions remain arbitrary. Currently, none of them uses novelty indicators to evaluate proposals. Novelty measurement might be relevant in providing reliable information when awarding grants to research proposals.

We conclude this discussion with a roadmap and our aspirations for *Novelpy*. The primary feature we aim to develop in future versions is automatic execution using well-known databases (PATSTAT, Microsoft Academic Knowledge Graph, Arxiv, etc.). At present, users must pre-process data to match our format. Although we provide a comprehensive example and make the sample available here `https://novelpy.readthedocs.io/en/latest/usage.html#id5`, we believe that expanding the accepted inputs will aid researchers in working on improving novelty indicators. The second feature we plan to add is a time complexity analysis. To conduct a proper benchmark between indicators, we need to compare their computing speeds. Users can currently perform this manually, but we intend to streamline the process and add plots to address this gap. Finally, we will selectively add new and past indicators. Anyone interested in contributing to the module can visit GitHub `https://github.com/Kwirtz/novelpy` and create a pull request.

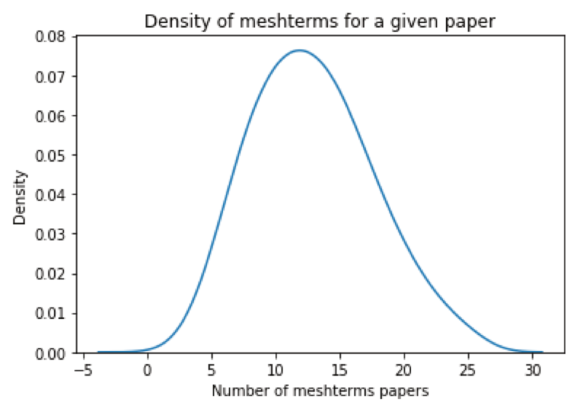
3.5 Appendix

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
n papers	49,872	52,046	54,721	58,439	62,241	67,361	70,501	75,717	81,228	84,496	89,168
mean of cited paper	27.3871	27.3672	27.9704	28.5654	28.8562	29.3572	30.0423	30.3297	31.0576	31.5393	32.3128
var of cited paper	707.008	708.596	742.314	709.619	807.733	758.342	809.695	795.216	845.84	944.337	896.461
mean of meshterms per paper	13.3097	13.4067	13.2431	13.3788	13.2862	13.1364	12.8499	12.8425	12.8575	12.9128	12.8867
var of meshterms per paper	26.5811	27.6517	26.0774	26.9045	27.2265	26.4599	22.9795	23.3933	23.3725	24.6855	24.9734

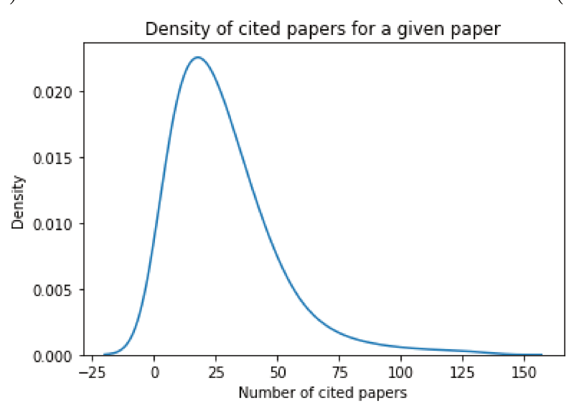
Table 3.2: Sample Statistics



(a)



(b)



(c)

Figure 3.9: (a) Density of contribution of authors. On average an author has 2.6 publications (solo or co-authored) in 10 years.
(b) Density of the number of mesh terms between 2000-2010. On average, a paper is labelled with 13 mesh terms.
(c) Density of references between 2000-2010. On average, a paper has 23 references

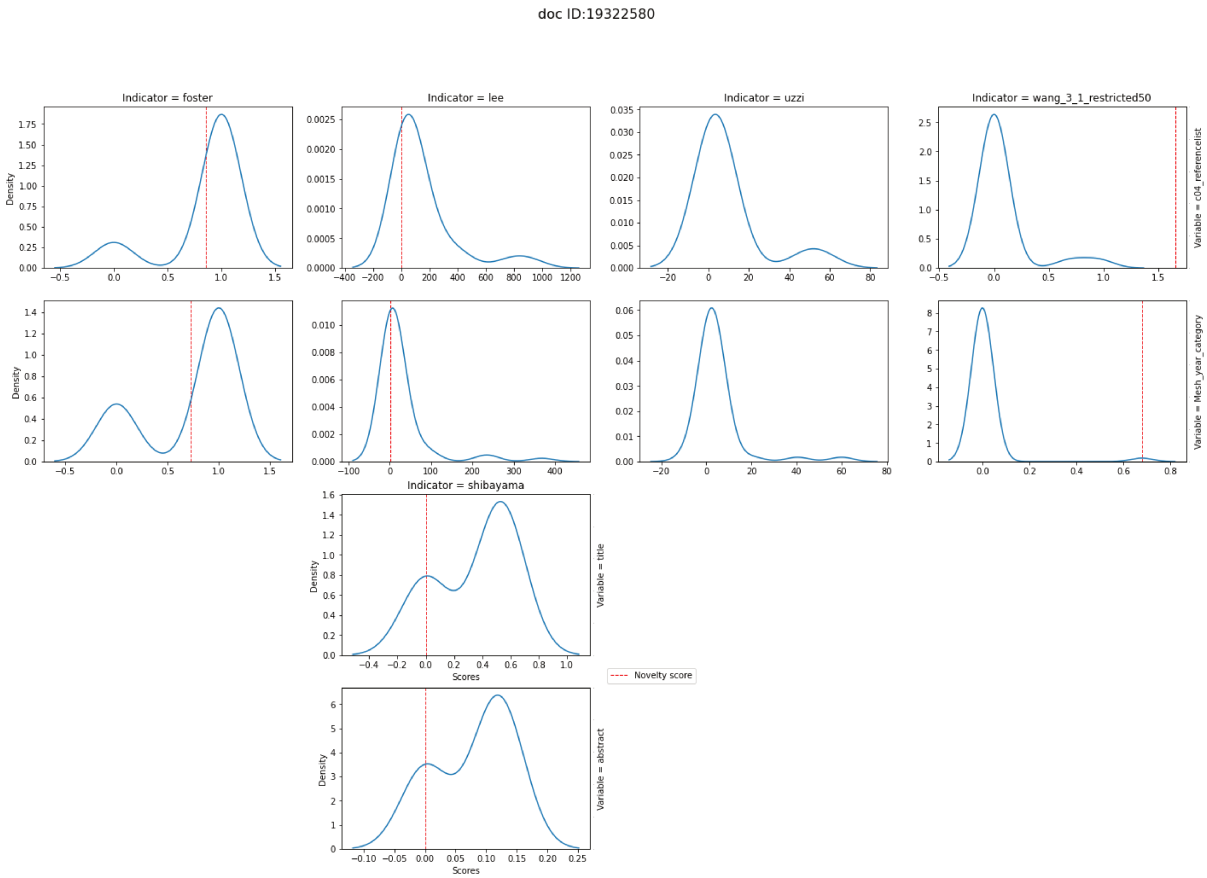


Figure 3.10: Each combination has a novelty score. A single plot represents the density of score combinations for a specific paper (PMID 10698680) for a specific indicator and entity (i.e. mesh terms, journals, title, abstract). The scores for the first row were computed using a combination of cited journals. The scores for the second row were computed on mesh terms combinations. Each column represents an indicator. The last two rows are for text embedding-based indicators (i.e. Shibayama et al. [2021]: Novelty, Author proximity) on the title of the paper or abstract.

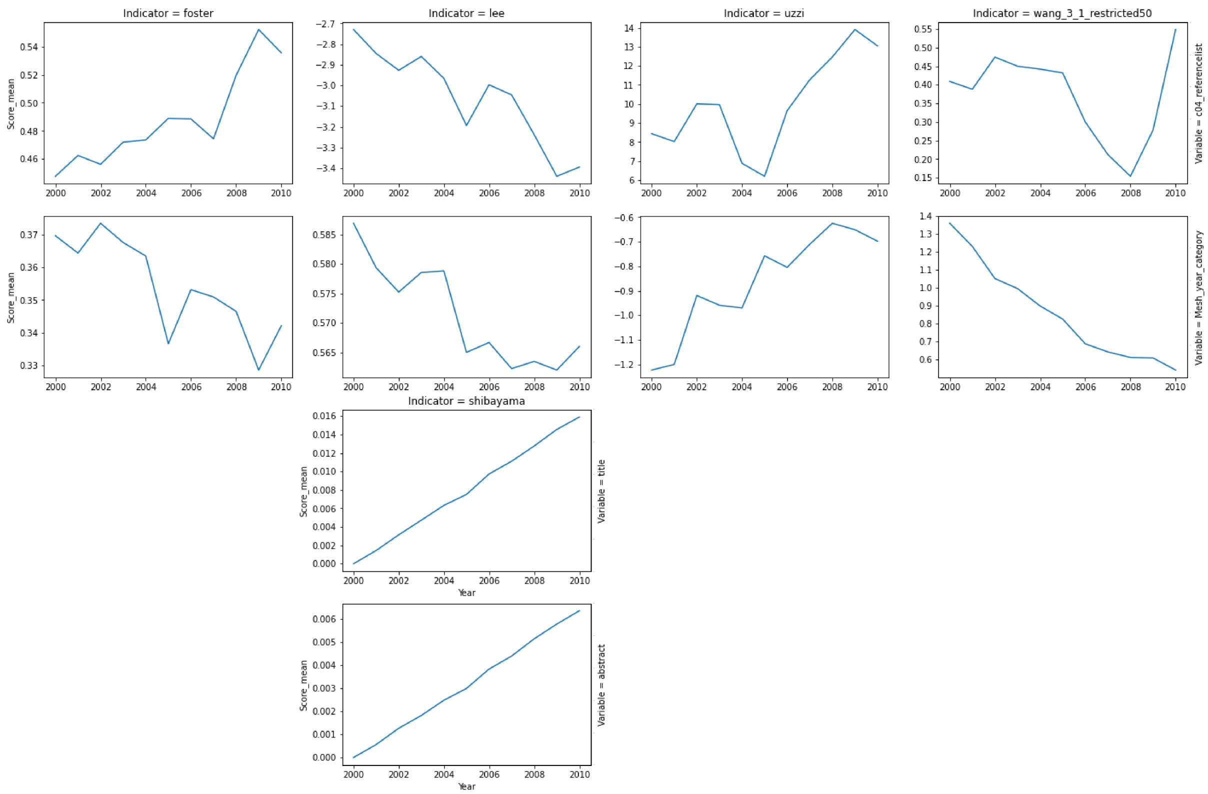


Figure 3.11: The mean novelty score on every document for a given year. Columns and rows represent respectively indicators and variables.

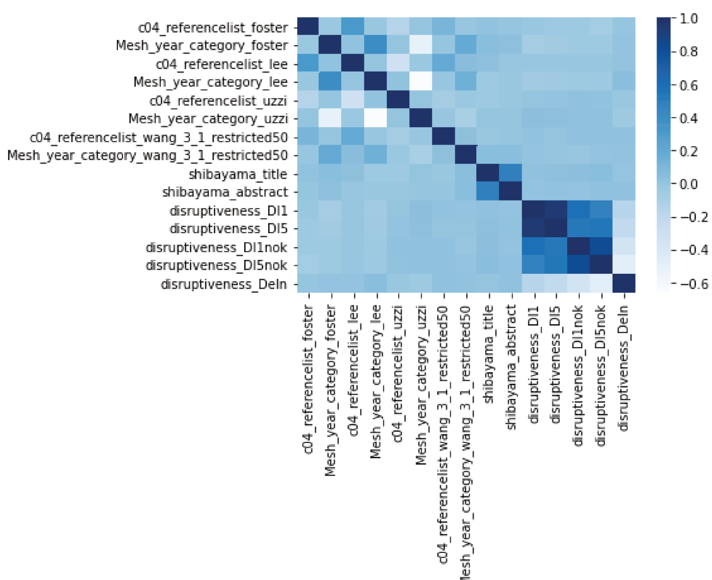


Figure 3.12: The correlation between the novelty score for each indicator, given the entity, for the period 2000-2010

Chapter 4

Unpacking Scientific Creativity: A Team Composition Perspective

This chapter was co-authored with

Kevin Wirtz

Summary of the chapter

This paper investigates the relationship between cognitive diversity within scientific teams and their ability to generate innovative ideas and gain scientific recognition. We propose a novel author-level metric based on the semantic representation of researchers' past publications to measure cognitive diversity at individual and team levels. Using PubMed Knowledge Graph (PKG), we analyze the impact of cognitive diversity on novelty, as measured by combinatorial novelty indicators and peer labels on Faculty Opinion. We assessed scientific impact through citations and disruption indicators. Cognitive diversity between team members appears to be always beneficial to combine more distant knowledge. We show that while the effect is positive, it is marginally decreasing. Our findings reveal also that within-team average exploratory profiles follow an inverse U-shaped relationship with combinatorial novelty and citation impact. We show that the presence of highly exploratory individuals is profitable to generate distant knowledge combinations only when balanced by a significant proportion of highly exploitative individuals. Also, teams with a high share of exploitative profiles consolidate science, while those with a high share of both profiles disrupt it. These results emphasise the implication of team composition in scientific creativity, suggesting that combining these two types of individuals leads to the most disruptive and distant knowledge combinations.

4.1 Introduction

Creativity is a crucial driving force in fostering the production of new knowledge in an ever-growing landscape of scientific research and technological innovation [Geuna, 1999, Amendola et al., 2014, Witt, 2016]. A broadly accepted definition of creativity assumes a bipartite composition involving a combination of novelty and effectiveness [Runco and Jaeger, 2012]. As science moves towards a team-based model [Wuchty et al., 2007], the creativity of scientific publications should be studied from a social perspective. The cognitive dimension (i.e. differences in thinking, problem-solving approaches, and perspectives among individuals) plays a crucial role in enabling exchange of information and creation of new knowledge [Nooteboom, 2000, Nooteboom et al., 2007]. It is induced by individuals' characteristics and the trade-off carried out between *exploration* and *exploitation* [March, 1991] of the knowledge space during their career. In the context of science, exploration involves actively pursuing the expansion of one's understanding and curiosity across various areas of knowledge. On the other hand, exploitation refers to individuals specializing in a specific field and continuously building upon their expertise in that area. The presence of individuals with exploratory profiles appears to facilitate communication among team members who are cognitively distant and foster creativity as the intersection of different perspectives is commonly required to solve complex scientific problems [Page, 2008].

This paper aims to study the extent to which the exploratory nature of scholars and the cognitive diversity of scientific teams shape their ability to generate innovative ideas and obtain scientific recognition. We propose a new author-level measure of cognitive diversity based on the semantic representation of their past papers; this metric allows us to proxy both intra-individual and inter-individual cognitive dimensions and their impacts on creativity in science.

In scientific creativity, originality and success emerge as two essential components [Runco and Jaeger, 2012]. However, the focus has predominantly shifted towards success. The excessive emphasis on success through measures such as citation counts for articles or authors has been found to constrain novelty and originality by providing limited incentives for researchers, ultimately leading to suboptimal research choices. The reliance on an impact metric to reward and evaluate researchers created a harmful behavior whereby scientists maximize the metric to become more appealing to funding agencies or institutions. As Goodhart's law states, "*when a measure becomes a target, it ceases to be a good measure*" [Goodhart, 1984]. The h-index, although heavily criticized, became a central evaluation instrument of researchers [Costas and

Franssen, 2018]. And this negatively impacts novelty as innovative research tends to be less cited in the short run [Wang et al., 2017]. Researchers are discouraged from opting for a more exploratory approach when developing a research question, as work that is too innovative tends to be rejected when it deviates too much from the established paradigm [Carayol and Dalle, 2007, Trapido, 2015].

Career choices are directly affected by this phenomenon. Given the heterogeneity in terms of the impact of novel research, researchers have less incentive to produce too innovative work because of the uncertainty linked with novel research. In the short term, individuals might turn to more conventional research questions to maximise their h-index while minimizing the risk associated with novel research.The bias toward maximizing the h-index already has a tangible impact on limiting novelty in various research fields. The imbalance between growth in the scientific workforce and research funding has led to 'hyper-competition' in the medical sciences; the scientific system favors individuals who can ensure outcomes over those with potentially groundbreaking ideas that might disrupt the field [Alberts et al., 2014]. Such a focus of the researcher on his or her impact is done at the expense of his or her novelty, showing a clear disconnection between the goal of science and its operationalization.

One of the goals of science is to advance the boundary of the knowledge space [Shi et al., 2015, Witt, 2016, Veugelers and Wang, 2019]. The novelty (also referred to as originality or invention) lies at the cornerstone of innovative research, bridging existing knowledge and unexplored scientific territories. Effectiveness, on the other hand, refers to the recognition attributed to this novelty. Novelty is at the foundation of peer recognition and acts as a "reward system" wherein the individual credited with the initial discovery garners recognition. [Merton, 1957, Stephan, 1996, Carayol et al., 2019]. Novelty is crucial for scientists to develop new solutions to the grand challenges of the century (climate change, poverty, global pandemics, and others) [Petersen et al., 2021]. Highly innovative research is frequently referred to as "High-Risk High-Reward" (HRHR) to reflect its high volatility of outcomes (i.e., novelty does not imply effectiveness). In particular, highly novel research receives more citations on average, but the uncertainty is also more considerable [Wang et al., 2017]. Funding opportunities are limited for innovative research due to its risky nature [Ayoubi et al., 2021, OECD, 2021, Franzoni et al., 2022]. Multiple grant initiatives try to support HRHR research, and funding decisions are all based on expert judgment [OECD, 2021]. But there is a direct bias towards novelty when scholars evaluate a peer's work [Wang et al., 2017, Ayoubi et al., 2021] and the effect is accentuated by the intellectual distance with the examiner [Boudreau et al., 2016]. Measures such as novelty indicators attempt to estimate the originality of a document and might guide experts to support innovative research. Yet, these novelty indicators are still relatively recent and understudied as it is mostly intended to explain success. As a result, it is essential to explore and validate new methods to understand better how to detect potential innovative and impactful research based on different criteria than past novelty or previous success.

Not all idea combinations are worth exploring, hence the challenge of distinguishing between novel and impactful ones. [March, 1991] distinguishes two different strategies for invention in organizations: "Exploration and exploitation". Exploitation focuses on a combination of ideas that are closely related to each other, thus representing a low-risk strategy. On the other hand, exploration represents the navigation through the knowledge space to combine more distant ideas, inducing more volatile results. March [1991] supports the idea that a mix of exploitation and exploration is the key to organization's survival. Put differently, producing a valuable invention would require a proper mix of typical and atypical combinations of knowledge, as seen in Uzzi et al. [2013]. This dichotomy has been studied in different domains, as mentioned in Foster et al. [2015] (e.g., "conformity" versus "dissent" in the philosophy of science), and can also be applied to research. As the body of knowledge in science expands, researchers increasingly specialize their competencies [Jones et al., 2008, Jones, 2009] and thus are better able to recombine information locally in the knowledge space, facing incentives to collaborate [Fleming, 2001, Boudreau et al., 2016]. Science is seen as a social phenomenon [Fleck, 2012]. Indeed, agents that recombine knowledge are individuals embedded in a social context, and cognitive and social phenomena strongly influence the invention process [Fleming, 2001]. Team size has been shown to impact creativity [Paulus and Nijstad, 2003, Shin and Zhou, 2007, Wuchty et al., 2007, Falk-Krzesinski et al., 2011, Erren et al., 2017, Mueller, 2019]; however, the authors' characteristics have not been adequately considered in the process as current novelty indicators primarily focus on the information *within* a document[1]. We contend here that the cognitive distance between co-authors and the team composition of a research paper may be among the most critical factors influencing knowledge creation. So, based on the concept of exploration and exploitation, we propose an indicator that serves as a proxy for exploratory *vs.* exploitative trade-

[1]E.g. references, text, keywords. A detailed review of classical re-combinatory novelty indicators can be found in Pelletier and Wirtz [2022].

off at both the individual and team levels through past publications. In a nutshell, our indicator measures the cognitive distance between team members as well as the individual propensity to work on various subjects.

We are unaware of previous studies that have used individuals' past research experiences to investigate how the cognitive dimension influences the novelty and recognition of the resulting articles. Note that we do not consider our indicator as a replacement for current novelty indicators but rather as a tool that could enhance our understanding of the mechanisms behind creativity. In fact, by incorporating the cognitive dimension into novelty studies, we can develop a more comprehensive understanding of the complex relationship between cognitive aspects, interdisciplinary efforts, and the nature of scientific innovation. Furthermore, examining these questions enables us to provide valuable insights and guidance for researchers and institutions striving to enhance scientific progress while avoiding potentially misleading interpretations of research performance measurement.

Using PubMed Knowledge Graph (PKG), we empirically investigate the role of these cognitive diversities in the production of novel research outcomes and the ability to obtain scientific recognition. We performed the analysis on novelty on five combinatorial novelty indicators [Uzzi et al., 2013, Lee et al., 2015, Foster et al., 2015, Wang et al., 2017, Shibayama et al., 2021], both on references and MeSH terms, as well as on perceived novelty, using labels submitted by researchers to qualify the contribution of an article (Faculty Opinion)[2]. For scientific recognition, we rely on the traditional number of citations and six indicators of disruption and consolidation [Wu et al., 2019, Bu et al., 2019, Bornmann et al., 2019a].

Our findings emphasize the crucial role of cognitive dimensions in creativity, significantly impacting originality and success. We show that cognitive diversity always seems beneficial to combine more distant knowledge. In contrast, the within-team average exploratory profile follows an inverse U-shaped relation with combinatorial novelty (i.e. there is a turning point where it is no longer beneficial). The same relation can be found with citation counts, but we show that the cognitive dimension also strongly influences the nature of citations. Teams with more exploitative profiles consolidate science, while those with high exploratory profiles disrupt it only if they are associated with exploitative researchers. The union of those two types of individuals leads to the most disruptive and distant knowledge combinations. To maximize the relevance of these combinations, maintaining a limited number of highly ex-

[2] More information can be found here: https://facultyopinions.com/

ploratory individuals is essential, as highly specialized individuals must question and debate their novel perspectives. These specialized individuals are the most qualified to extract the full potential from novel ideas and situate them within the existing scientific paradigm.

The remainder of the paper is organized as follows. In section 4.2 we review the existing literature. Section 4.3 details the creation of our metrics and the methodology for addressing our research questions. Section 4.4 presents the results of our analysis. Section 4.5 concludes the paper and outlines future directions for developing novelty indicators.

4.2 Background and literature review

This section highlights the team's relevance in fostering creativity in science and emphasises how team size can influence this process. We also underscore the importance of identifying the social dimensions of the team, a crucial factor in generating new knowledge. Finally, we propose a new approach based on the semantic representation of authors' past publications that allows studying the role of the cognitive dimension in a team's ability to produce new and impactful knowledge.

4.2.1 Team science as an engine of creativity

Over the past two decades, there has been a significant increase in interest surrounding the Science of Team Science (SciTS) [Falk-Krzesinski et al., 2011][3]. Since the 1950s, the average number of authors per paper has risen across all scientific disciplines [Wuchty et al., 2007]. Research collaborations have also become more diverse, inter-institutional collaborations in science and engineering and social science grew by 32.8% and 34.4%, respectively, between 1975 and 2005 [Jones et al., 2008]. In addition, international collaboration has also expanded, with one in five research projects now involving multiple countries [Xie and Killewald, 2012].

Teamwork has proven to be a practical approach to producing impactful scientific results. Articles written by teams tend to have a higher impact, receiving more citations on average and are more likely to become influential than articles authored solely [Wuchty et al., 2007, Whitfield, 2008]. Researchers benefit from collaboration in various ways. Collaborative efforts can enhance rigour through co-authors'

[3]For an up-to-date and comprehensive review, see Wang and Barabási [2021].

verification [Leahey, 2016] and facilitate the dissemination of their work beyond their immediate networks [Leahey, 2016]; this effect is further amplified when collaborations are international or inter-institutional [Adams, 2013, Jones et al., 2008]. Additionally, teams have better access to resources, as projects executed by groups are more likely to apply for funding and succeed in obtaining it [Rawlings and McFarland, 2011]. Teams are more likely to produce novel articles than solo-authored publications [Carayol et al., 2019, Uzzi et al., 2013, Wagner et al., 2019]. As highly cited work is often associated with a combination of novel and conventional ideas [Uzzi et al., 2013], teams of researchers may be more adept at generating novel ideas or striking a balance between novel and traditional concepts than individual authors.

Successful team performances put individuals and their interactions at the heart of the creative process. Over recent decades, the perception of teamwork has undergone significant changes. In the early 1990s, the prevailing belief was that groups should not be used for creativity because of inherent process loss in the creative process. This perspective has shifted dramatically, and team collaboration is now considered a critical factor in promoting creativity [Paulus and Nijstad, 2003]. Creativity relies on individual's existing knowledge base: "*Creative thinking cannot happen unless the thinker already possesses knowledge of a rich and/or well-structured kind*" [Boden, 2001]. Knowledge exists on a continuum, ranging from explicit to tacit [Nonaka, 1994]. The generation of new knowledge occurs through interactions between explicit and tacit knowledge via a process known as the socialization, externalization, combination, and internalization (SECI) spiral. Tahamtan and Bornmann [2018] highlighted various approaches reported by researchers for fostering creativity. Engaging in conversations with colleagues seems to remain central to problem-solving and generating new, practical ideas. Hence, new ideas are becoming more challenging to discover as the idea space expands linearly while scientific publications grow exponentially [Bloom et al., 2020, Milojević, 2015]. As scientific knowledge increases, team sizes grow, and agents increasingly specialize their competencies [Jones et al., 2008, Jones, 2009].

The burst of possible combinations in the knowledge space suggests that agents can more effectively recombine information locally [Fleming, 2001]. "Local search" for an inventor involves exploiting existing combinations or using standard technological components. Agents tend to direct their research towards familiar subjects, focusing on topics related to their expertise or that of their co-authors (local search/-exploitation) [Fleming, 2001, Nelson, 1985, March, 1991]. Conversely, exploration (or "distant search") is characterized by using new components or testing novel combinations [Fleming, 2001, March, 1991]. The nature of the new combinations realized depends on agents' trade-offs between exploiting and exploring the knowledge landscape. Exploitation reduces the risk of failure, as researchers draw from experience with combinations and architectures that have previously failed [Vincenti, 1990]. Researchers must then collaborate with others to explore the knowledge space more efficiently, and the team's composition might determine this balance between exploration and exploitation.

4.2.2 Team characteristics in the creative process

We review here some dimensions of the team composition that affect the scientific process.

Size dimension: The importance of co-authors during the process of creativity has been debated in the literature, and the effect of team size and composition on creativity has been the focus of multiple studies [Paulus and Nijstad, 2003, Shin and Zhou, 2007, Wuchty et al., 2007, Falk-Krzesinski et al., 2011, Erren et al., 2017, Mueller, 2019]. Team size shapes and is shaped by the nature of the work carried out. Large teams tend to be more risk-averse and consolidate a field rather than introducing new opportunities [Christensen and Christensen, 2003, Paulus et al., 2013, Lakhani et al., 2013, Wu et al., 2019]. Larger teams use more up-to-date and influential research in their work, consequently fostering greater engagement within their scientific community and further increasing their impact [Wu et al., 2019]. However, large teams are more prone to coordination and communication failures as the entire team must have faith in the project to succeed, as agreement and communication between team members can be challenging and time-consuming [Bikard et al., 2015]. In fact, the number of people involved in a project can have heterogeneous effects on creativity, and no optimal team size fits every project. A small team may be more useful in the conceptualization phase, while a larger team might be beneficial in the implementation and testing phase of the project [Wang and Barabási, 2021]. Shin and Zhou [2007] highlight the organization's importance for creativity. Using evidence from Cambridge and AT&T's Bell Laboratories (home to numerous Nobel Prize winners), they discuss researchers' ideal context for fostering creativity and conclude that the presence of a healthy environment for a small group of people (up to seven) promotes creativity. These results are further confirmed by Lee et al. [2015] and Carayol et al. [2019], indicating that the relationship between team size and novelty appears U-shaped and is highly heterogeneous across disciplines.

Structural and relational social capital: Nahapiet and Ghoshal [1998] conceptualize three dimensions of social capital that impact intellectual capital development: structural, relational and cognitive. Though primarily used to understand intellectual capital development in organizations and firms, the dimensions of social capital presented in Nahapiet and Ghoshal [1998] can be applied to the context of knowledge production in science due to their intrinsic relevance to relationship and network dynamics [Liao, 2011]. Structural capital examines the links between individuals, and structural distances have been widely studied through collaboration networks (see Kumar [2015] for an extensive review on network collaborations). Relational capital represents the nature and intensity of the connections between team members. A critical factor in intellectual development is the ability to communicate with each other, and the actors' experience reinforces the phenomena [Taylor and Greve, 2006, Liao, 2011, Kelchtermans et al., 2020]. For instance, McFadyen and Cannella Jr [2004] emphasize the role of the intensity of past relationships between scientists in fostering new knowledge. Indeed, members with strong relationships, norms, obligations, and mutual trust tend to communicate more easily [Liao, 2011]. Other relational aspects, such as hierarchical or geographical dimensions, also impact the knowledge space exploration. For example, supervising doctoral students is not only associated with entering new areas but also extending towards more distant fields [Kelchtermans et al., 2020] – See also Chapter 2 of this thesis.

Cognitive social capital: The cognitive capital remains challenging to measure as it is linked to the shared background between coauthors and their common language. Cognitive diversity is often encouraged through interdisciplinary projects as the intersection of different perspectives is commonly required to solve complex scientific problems [Page, 2008]. Indeed, people from outside a domain may have some advantage to offer fresh ideas through their distinct knowledge [Jeppesen and Lakhani, 2010, Kuhn, 1962]. The effectiveness of generating new knowledge is impacted by factors such as variations in background, belief and reasoning styles among scientists, all of which contribute to cognitive diversity. The cognitive distance between team members is expected to display an inverted U-shaped correlation with both learning and innovation [Nooteboom et al., 2007], as people being too distant will face difficulty in communicating, and those being cognitively too similar benefit less from distinct perspectives in the knowledge creation process.

Cognitive distances between individuals can be studied through various metrics. Kumar et al. [2017] used, for example, citations networks and citations context in full text. Boudreau et al. [2016] represented the cognitive distance between funding evaluators and the proposal through MeSH terms similarity. Similarly, Ayoubi et al. [2017] represent the distance between the focal scientist and her team by comparing cosine similarities of referenced journals from scientists' past publications. Other measurements, without being explicit, may relate to cognitive dimensions, Wagner et al. [2019] discovered that international collaborations negatively affect novelty and produce more conventional knowledge combinations, highlighting barriers and transaction costs that influence the production of creative work. Finally, measures of cognitive distance strongly relate to interdisciplinarity. Petersen et al. [2021] represent author diversity using the discipline of the institution. Using authors' disciplinary diversity, Abramo et al. [2018] show that more distant coauthors produce articles with more diverse references.

Exploratory profile: Individual characteristics and the ability to interact with individuals from different fields are essential to efficiently managing cognitive diversity in a team. When the distance between disciplines is too high, a "Renaissance" individual [Jones, 2009] can ease their connection [Wu et al., 2022]. The presence of a scientist with a multifaceted profile bridges the gap between the different backgrounds of other team members. This is crucial as a shared knowledge base between researchers streamlines the socialization process and facilitates knowledge recombination, fostering creativity. Shin and Zhou [2007] focused on the relationship between diversity (interdisciplinarity) and creative ideas in groups. Shin and Zhou [2007]'s idea is that the presence of a "transformational leader", whose role is to mediate between individuals, each specialized in a different field, leads to greater team creativity. Xu et al. [2022] provided a first answer to this hypothesis by examining the share of team members engaged in the conceptual work, the L-ratio, which was deduced from the analysis of author contribution reports. The findings suggest that hierarchical teams generate less novelty than egalitarian teams and tend to develop existing ideas more frequently.[4] We argue that the notion of transformational leader

[4]Interestingly, their method was expanded in an article with no contribution reports. Through Louvain algorithms, they identified clusters of co-occurring research activities in their first dataset. They then built a neural network to infer author roles based on their characteristics and predicted it for 16 million articles on Microsoft Academic Graph (MAG).

or renaissance individual is connected to exploratory profile *à la* March [1991], individuals enabled to link others in the knowledge space due to their ability to navigate in different spaces.

4.2.3 Exploring the cognitive dimension

We investigate scientific impact through citation networks and recent indicators of disruption and breath and depth Wu et al. [2019], Wu and Wu [2019], Bu et al. [2019], Bornmann et al. [2019a]. These indicators determine whether a document consolidates a domain or constitutes a founding step. To explore its influence on novelty, we use two approaches, one based on combinatorial novelty indicators [Uzzi et al., 2013, Lee et al., 2015, Foster et al., 2015, Wang et al., 2017, Shibayama et al., 2021] and one based on external validation via Faculty Opinion (previously called F1000) following Bornmann et al. [2019b]. Faculty Opinion is a website hosting reviews of papers tagged as presenting "New Results", "Novel Drug target", "Technical advancement", "Interesting hypothesis", and "Controversial results", among other categorizations labelled by experts in the field. It allows us to empirically assess the capacity of novelty indicators and our indicators to predict the novelty as perceived by other researchers in the community.

Novelty indicators have been compared and evaluated based on citation count [Uzzi et al., 2013, Lee et al., 2015, Foster et al., 2015, Wang et al., 2017]. Fontana et al. [2020] compared Wang et al. [2017] and Uzzi et al. [2013], Lee et al. [2015] using randomized citation networks and demonstrated the ability of the Uzzi et al. [2013], Lee et al. [2015] indicators to better track novelty. Their findings are supported by using some Nobel Prize winners' articles and a list of APS milestone articles. Other studies have evaluated these indicators based on surveys, such as Shibayama et al. [2021] and Matsumoto et al. [2021], whereas Bornmann et al. [2019b] have evaluated them based on labels collected on Faculty Opinion and found similar results as in Fontana et al. [2020]. However, only a few indicators have been compared and tested simultaneously. This study intends to validate the effect of the cognitive dimension on a large variety of metrics.

Our indicator is not a substitute for other novelty indicators. It does not represent the novelty of an article as it is based upon previous information and would be similar even without the focal article. Instead, it provides an understanding of team composition that would benefit creativity in science. We can think of our measure as a measure of *potential novelty*, i.e. opportunities for new knowledge recombination available through the diversity of background in the team and the capacity of individuals to bridge the gap between other team members. In comparison, combinatorial novelty indicators would capture then the *realized novelty*, i.e. the output of the research conducted by this team in terms of pieces of knowledge used. Finally, Faculty Opinion labelling and other external validation methods can describe the *perceived novelty*, i.e. the peers' perception of this study. Hence, in these terms, we ask whether potential novelty contributes to realized and perceived novelty and its scientific recognition. Two research questions can be drawn regarding the effect of the cognitive dimension on creativity. Do teams with higher cognitive diversity are more likely to approach a subject creatively, demonstrating originality (*perceived* and *realized*) and recognition? Does the presence of *explorative* individuals within a team enhance communication among members and facilitate their exploration of the knowledge space to develop new and relevant solutions to research problems? Studying the cognitive dimension of creativity in science is of great interest, especially as it can help identify how to improve collaboration and communication among researchers with diverse cognitive profiles. Through our metric, we also offer a different approach to resource allocation decisions, giving another picture of teams with a high potential for creative output.

4.3 Data and methods

4.3.1 Measuring cognitive diversity and exploratory profile

The proposed metric examines the semantic heterogeneity of researchers' work as a proxy for their cognitive diversity. It thus offers an alternative to using categories, keywords, or citation networks, more complex to be monitored directly by the researchers themselves. Following Hain et al. [2020] and [Shibayama et al., 2021], we can embed this list of documents in a vectorial space to apply a distance measure such as cosine similarity [Mikolov et al., 2013b]. We assume that an author of a paper in a specific position within the semantic space possesses knowledge embedded around that position. Our indicator has two properties: it offers a measure of researchers' profiles at the individual level and a measure of distances between them. Consequently, we can proxy the trade-off between exploitation and exploration that a researcher undergoes throughout their career (intra-individual) and the trade-off materializing during the formation of a team (inter-individual) within the same mathematical space.

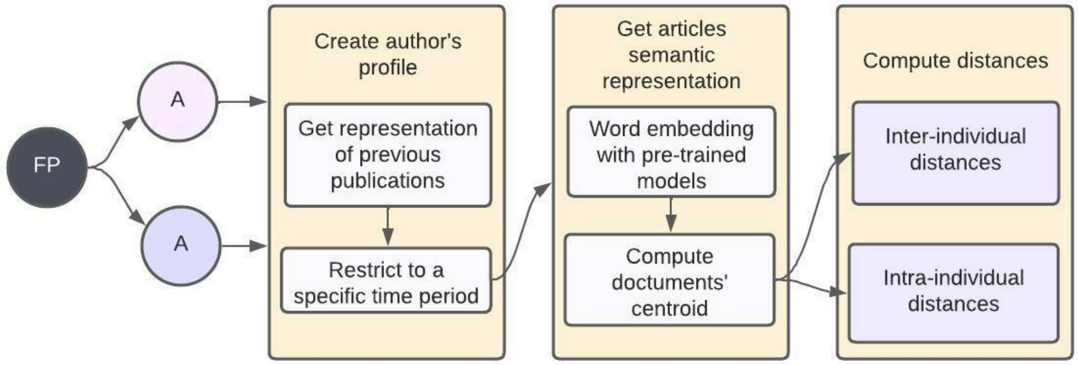


Figure 4.1: Construction of the indicator

As explained in Figure 4.1, we track authors to create a list of authors' past publications. Then, we can create a cognitive profile for each author at a given time t; each publication is embedded in the semantic space and represents the cognitive landscape of the author. We restrict to publications up to b years before t to account for researchers' current topics of interest and difficulty retaining information [Argote et al., 1990]. We can finally define a researcher's exploratory profile at time t by calculating pairs of cosine distances between past papers published. This will create a density of cosine distances which, using the taxonomy of March [1991], can be interpreted the following way: the fatter the right (left) tail is, the more exploratory (exploitative) the researcher. The same holds for the team. A sizeable right tail indicates a cognitively distant researchers team. This provides us with information on how distant their knowledge base is from others. The greater the distance, the less likely their respective knowledge space can be combined, thus affecting the probability of combining novel ideas. An intra-author and inter-author distribution enables a wide exploration of the relationship between novelty, creativity, and teams.

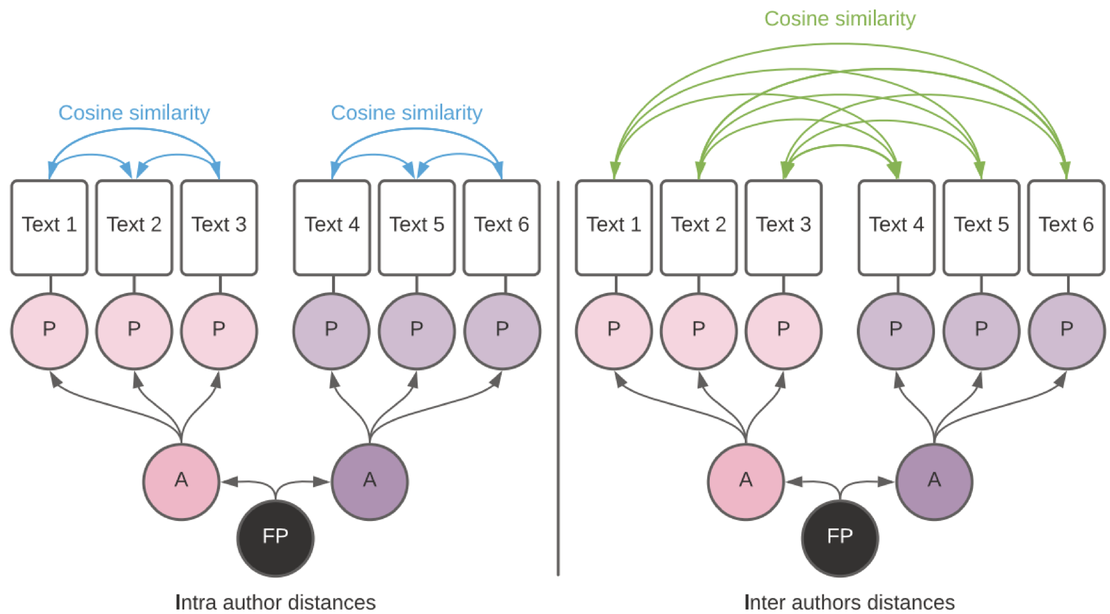


Figure 4.2: Exploratory profile and cognitive diversity

We model our measures on two different perspectives as represented in Figure 4.2: intra-author distances, which asses exploratory profile, and inter-authors distances, to capture cognitive diversity. A given focal paper (FP) is written by two authors named 'A'. We retrieve each author's past production, named 'P'. On the one hand, we can then calculate the distance between all publications from a given author (intra-author distances). On the other hand, we can also compare past publications from two authors (inter-author distances). We can build our framework using directed bipartite networks, defined as $G(U, V, E)$. U represents the set nodes for authors, V is the set for articles, and E is the set of links between authors and articles.

We only consider collaborations between authors when looking at a given article; collaboration is implicit since the set of parents of a given document $FP \in V$ corresponds to the set of authors that collaborate. For a given document FP, an author that has contributed to FP is noted a, and the set of nodes that contributes to FP is then $In_{FP} = \{a \in U : (a, FP) \in E\}$.

We want to retrieve all past publications for all authors in In_{FP}. The global set of publications before FP is noted V_{FP}^{t-b}, the set of articles published b years before the document FP. The set of past publications for author $a \in In_{FP}$ is noted $Out_a^{t-b} = \{v \in V_{FP}^{t-b} : (a, v) \in E\}$. For a document FP and an author $a \in In_{FP}$, we retrieve the set of past publications Out_a^{t-b}. All Out_a^{t-b} elements vectorial representations are compared, from which distribution of cosine distances are calculated. The distance between two documents $i, j \in Out_a^{t-b}$ is $d_{ij} = 1 - COS(T_i, T_j)$ where T_i is the dense vector text representation for document i.

Intra-author semantic distances: A distribution of semantic distance score D_a is computed through cosine similarity using all document $i, j \in Out_a^{t-b}$, the process is repeated for each authors $a \in In_{FP}$. The intra-author distance for a given author a is the q-th percentile (P_q) of this distribution and is written as:

$$Intra_a = P_q(D_a)$$

A general distribution of the intra-authors publication distances is constructed using the set of distances for all authors $A_{FP} = \{D_a : a \in In_{FP}\}$, the individual trade-off between exploitation/exploration is then captured through the average of the exploratory profiles in a given team.

$$Intra_{FP} = \frac{\sum_a (P_q(D_a))}{|In_{FP}|}$$

Inter-authors semantic distances: A distribution of semantic distance score between authors' previous work is constructed by comparing different authors' publications. For two given authors $a, e \in In_{FP}$, $|Out_a^{t-b}| \times |Out_e^{t-b}|$ distances are used to construct the distribution of distances $D_{a,e}$ between a and e. The final distribution then groups together all distances between authors' previous works $B_{FP} = \{D_{a,e} : a, e \in In_{FP}\}$, the trade-off between exploitation/exploration in team composition is captured through the percentile of B_{FP}:

$$Inter_{FP} = P_q(B_{FP})$$

Current techniques for large-scale author disambiguation allow the investigation of individual trajectories in science. However, the use of this information comes with a computational cost. This indicator pushes towards a massive use of data because one needs all authors' past publications for a given set of documents. Structuring the data to compute the measure is time-consuming and data-intensive. One needs indeed all papers' text from all authors in a given database. However, using pre-trained embedding models allows direct computing indicators without the requirement of complete database access. Therefore, measures are not dependent on the study sample as indicators of novelty based on cooccurrence matrices but rather on the sample used to train the model. Also, by processing titles and abstracts through embedding techniques, the authors' background is represented with greater granularity than through the keywords or the journals where the authors have been published.

4.3.2 Data

Our analysis relies on two databases. The first, PubMed Knowledge Graph (PKG), allows us to test the effect of the cognitive dimension on scientific impact and *realized* novelty of articles, while the second, Faculty Opinion verifies whether the cognitive dimension affects the *perceived* novelty by peers.

We use Pubmed Knowledge Graph (PKG), a collection of 35 million scientific papers and books from life science and biomedical journals provided by the National Library of Medicine (NLM) at the National Institutes of Health (NIH). Authors are disambiguated by leveraging Natural Language Processing (NLP) and online data, as outlined by Xu et al. [2020]. We based our analysis on all the 3.5M articles written by 3,276,250 authors and published in 9,348 journals between 2000 and 2005. We selected fairly old data due to the nature of the process studied. Indeed, novel articles are more likely to become "sleeping beauties" and accumulate citations in the long run [Lin et al., 2021]. Also to compute novelty indicators, we require information about references. We rely both on abstracts of references to embed their semantics and calculate the distance as in Shibayama et al. [2021]. Also, we use past publication references' journals to build past cooccurrence matrices used to capture combination existence and difficulty for other novelty indicators. For this purpose, we used the database between 1980 and 2005 to get all information needed, representing 11,261,955 documents.

To test if our indicators affect the novelty perceived by peers, we used Faculty Opinion following Bornmann et al. [2019b]. Faculty Opinion is a database featuring papers tagged as presenting 'New Results', 'Novel Drug target', 'Technical advancement', 'Interesting hypothesis', and 'Controversial results', among other categorizations determined by the platform users. The platform hosts reviews of the most significant research in Biology and Medicine. This makes it easy to match the articles in the database with PKG. Indeed, from the 190k articles in Faculty Opinion, we found 27,122 in our sample (2000-2005).

4.3.3 Empirical strategy

To explore the relationship between the team's cognitive dimension and its ability to recombine pieces of knowledge in novel ways and achieve recognition, we start with a basic exploratory data analysis followed by three econometric analyses to test our hypotheses.

The first two analyses aim to understand how a team's cognitive diversity and the exploratory profiles of its members impact *perceived* novelty (i.e., peer labelling on Faculty Opinion) and *realized* novelty (i.e., indicators of combinatorial novelty). Then our analysis seeks to comprehend the effect of the cognitive dimension on scientific recognition using citation and disruption measures.

Realized novelty and scientific impact connections with cognitive dimension are both investigated through PKG, the normalization performed at the field and year levels of this measure provides a measure ranging between 0 and 1, which we model using linear models with cluster robust standard errors at the journal level. Lastly, we examine how the presence of highly exploratory and exploitative individuals influences the team's creativity. This analysis will help determine if cognitive diversity and the presence of exploratory profiles are explicitly visible in an article's knowledge composition.

For the analysis of *perceived* novelty, we employ the Faculty Opinion database and model, through Logit and Poisson regressions, the likelihood of an article being labelled with "novel" categories ("Technical Advance", "Interesting Hypothesis", "Novel Drug Target"). In our sample, 80% of the observations are labelled as 'New Findings', and 95% of the total sample would be considered new using the top 4 most represented categories (22,216 novel articles versus 1,750 not-novel). The fact that most articles are labelled as new findings makes this category less informative; therefore, we decided to exclude it and remove articles solely labelled with this category. As a result, our prediction is based on a more balanced sample (8,950 novel articles versus 3,605 not-novel). This will enable us to understand whether the cognitive dimension is associated with *perceived* novelty. We do not expect a direct effect but rather hypothesize that cognitive diversity influences a latent variable representing the article's actual contribution. This actual contribution of the paper may or may not be visible in the *realized* novelty measured by novelty indicators but might be then reflected in labelling made by peers.

4.3.4 Variables

Variables used in our empirical analysis can be separated into four categories: novelty indicators, scientific impact, cognitive, and control variables. For control variables, aside from data from PKG, we use journals listed in Scimago to control for scientific domains and measure of the impact associated with the journal. Each of our variables is at the paper level. For the empirical strategy, novelty, impact and cognitive measures will be field weighted by year using the percentile rank procedure – noted (FW). We use the first category of the journal from Scimago to approximate the field.

Novelty indicators

The indicators used in our analysis are Uzzi et al. [2013], Lee et al. [2015], Foster et al. [2015], Wang et al. [2017], Shibayama et al. [2021]. A formal mathematical description of them can be found in Chapter 3 of this thesis. Note that we have inversed the sign of the measures related to Uzzi et al. [2013] for simplicity and comparison with other indicators. The computation is done with *Novelpy*[5].

Scientific impact variables

For impact measures, we use citation counts and disruptiveness indicators, also described in Chapter 3. We used all available indicators in *Novelpy*, namely: Wu et al. [2019], Bu et al. [2019] and Bornmann et al. [2019a].

Cognitive variables

Team cognitive diversity: The mean of the inter-authors semantic distance as defined in Section 4.3.1 with q=90 for a given paper. It measures to what extent a team is composed of highly cognitively distant authors (i.e. Author 1 background is vastly dissimilar to Author 2 background). Furthermore, we suppose the relation between the team's cognitive diversity and other measures is not linear. We take the square of the team's cognitive diversity to test this.

Average exploratory profile: The mean of the intra-authors semantic distance as defined in Section 4.3.1 with q=90 for a given paper. It captures to what extent a team

[5] *Novelpy* is a python package that allows computing novelty and disruptiveness indicators. More details can be found here: https://novelpy.readthedocs.io/

comprises authors with distant past publications (i.e. Author 1 worked on diverse subjects). As for team cognitive diversity, we add a square term in the regressions.

Number of highly exploratory authors: To have more information on the team structure, we decided to define a threshold to identify highly exploratory authors. Looking at the intra- author's semantic distance as defined in Section 4.3.1. An author is considered highly exploratory if its 90^th percentile is in the top 10% of all $Intra_{FP}$ in our sample.

Number of highly exploitative authors: We expect highly exploratory authors to work best with highly exploitative authors (i.e. Novelty is probably most successful with a combination of typical and atypical individuals). We construct this measure following the same procedure as exploratory authors. Looking at the intra- author's semantic distance as defined in Section 4.3.1. An author is considered highly exploitative if its 90^th percentile is below our sample's median of all $Intra_{FP}$.

Interaction term between highly exploratory and highly exploitative authors: We added an interaction term between the two types of profiles as both competencies might complement each other.

Control variables

We included as control variables the number of authors, references and MeSH terms. We also controlled for the year and information related to the journal of publication.

Scimago Journal Ranking (SJR): An indicator of a journal's prestige based on weighted citation and eigenvector centrality derived from Scopus' citation networks by Scimago [González-Pereira et al., 2009].

Scimago Journal Category: Scimago provides a classification of journals based on various fields. We used the first category linked to a journal; our database contains journals from 271 categories.

4.3.5 Descriptive statistics and preliminary evidence

We further clean our database and restrict it to papers with at least 2 references/ MeSHterms/ authors and with a journal ISSN. Our final dataset represents approximately 2.1M articles.

Table 4.1 presents the descriptive statistics for the variables in our sample. Examining their distribution, it is worth noting that some indicators concentrate novelty around a small number of articles, as in Foster et al. [2015] or in Wang et al. [2017], merely 21% of the articles possess non-zero values (measured on references). Also, indicators such as citation count or Uzzi et al. [2013] among others, display relatively extreme values. Specifically for Uzzi et al. [2013], it is highly dependent on the z-score computation, when the variance of the journal combination is minimal, the z-score can rapidly become substantial. These disparities in distribution prompted us to apply a percentile rank procedure by field and year, as explained in the previous subsection.

Table 4.1: Descriptive statistics

Statistic	Min.	Pctl(25)	Median	Mean	Pctl(75)	Max	St. Dev.	N
# References	2	12	22	27.37	36	2690	25.76	2108280
# Meshterms	2	9	13	13.25	16	51	5.19	2108280
# Authors	2	3	4	5	6	282	2.94	2108280
# Citations	0	9	22	46.99	50	81577	129.47	2108280
SJR	0.1	0.627	1.130	1.787	2.035	39.946	2.22	2094669
Disruption_1	-1	-0.007	-0.001	0.003	05179	1	0.06	2108280
Disruption_{1noK}	-1	-0.588	-0.269	-0.192	0.111	1	0.51	2108280
Disruption_5	-1	0	0.001	0.018	0.009	1	0.07	2108280
Disruption_{DeIn}	0	0.79	1.662	2.067	2.875	92.5	1.81	2108280
Breadth	0	0.307	0.5	0.517	0.714	1	0.26	2108280
Depth	0	0.258	0.5	0.458	0.672	1	0.26	2108280
Share Exploratory	0	0	0	0.063	0	1.0	0.14	2108280
Share Exploitative	0	0	0.333	0.365	0.6	1	0.32	2108280
Author intra $_{abs}$	0	0.22	0.29	0.29	0.36	1.02	0.09	1837749
Author inter $_{abs}$	0	0.26	0.33	0.33	0.40	1.02	0.09	1837748
Shibayama $_{abs}$	0	0.222	0.274	0.275	0.327	0.991	0.07	2081854
Uzzi_{Ref}	-62396.32	-7.34	3.66	-18.03	14.02	199.49	206.82	1891079
Lee_{Ref}	-17.581	0.145	0.840	0.567	1.466	6.006	1.45	2092283
Foster_{Ref}	0	0.117	0.4	0.366	0.583	1	0.25	2092283
Wang_{Ref}	0	0	0	0.583	0	2872.106	4.79	2092283
Uzzi_{Mesh}	-287.0	-1.1	0.9	2.7	4.5	189.1	8.19	765751
Lee_{Mesh}	-7.996	0.4562	0.807	0.794	1.174	4.717	0.60	2105186
Foster_{Mesh}	0	0.274	0.476	0.424	0.591	1	0.22	2105186
Wang_{Mesh}	0	0	0	0.299	0.307	28.668	0.76	2105186

The correlogram in Figure 4.3 illustrates the various indicators' interconnection. A hierarchical clustering algorithm is applied to the correlation matrix and several clusters emerge. It includes citation and consolidation indicators, novelty indicators, cognitive dimension indicators, and disruption indicators. Regardless of whether MeSH terms or references are used to derive the indicators, the novelty indicators group remains consistent, suggesting that combinatorial novelty indicators capture a shared underlying dimension of innovation in scientific research. The correlation between Lee et al. [2015] and Uzzi et al. [2013] is particularly robust since both measures are nearly identical except for the incorporation of the reference's publication year in Uzzi et al. [2013]'s resampling process. It should be noted that a negative correlation is expected since low values signify atypicality in Uzzi et al. [2013], while high values represent novelty in Lee et al. [2015], this is why we inverse the sign of Uzzi et al. [2013] to get positive correlation between indicators. A strong correlation is observed between Shibayama et al. [2021] and our indicators, as it employs the same measurement on references, and some elements may overlap. Specifically, self-citation increases the correlations between Shibayama et al. [2021] and our indicator since the same combinations are calculated in the author and reference parts. Moreover, the clustering differentiates between citation count, consolidation indicators (Depth, DeIN), and disruption indicators (DI1, DI5, DI1nok, and Breadth). These distinctions emphasize how consolidation indicators are more closely related to citation count and demonstrate how disruption indicators capture other dimensions of scientific impact.

The development of an author-level indicator necessitates examining its relationship with team size. Figure 4.4 illustrates how intra- and inter-individual cognitive indicators are strongly associated with team size. Although it is unclear whether cognitive diversity generates a specific team size or if team size produces this diversity, it is visible that as the cognitive diversity within a team increases, the average exploratory profile must also rise to maintain a comparable team size. The hump-shape relationship on both sides is easily observable, suggesting that the more diverse the team and/or the more exploratory the individuals, the smaller the team. Conversely, highly homogeneous teams typically imply smaller average team sizes, even if the average exploratory profile is high. This pattern is partially attributable to the construction of our indicator, which averages distance. In larger teams high distance between members might be compensated by other members that are close to each other. This counterbalancing is less pronounced in smaller teams, resulting

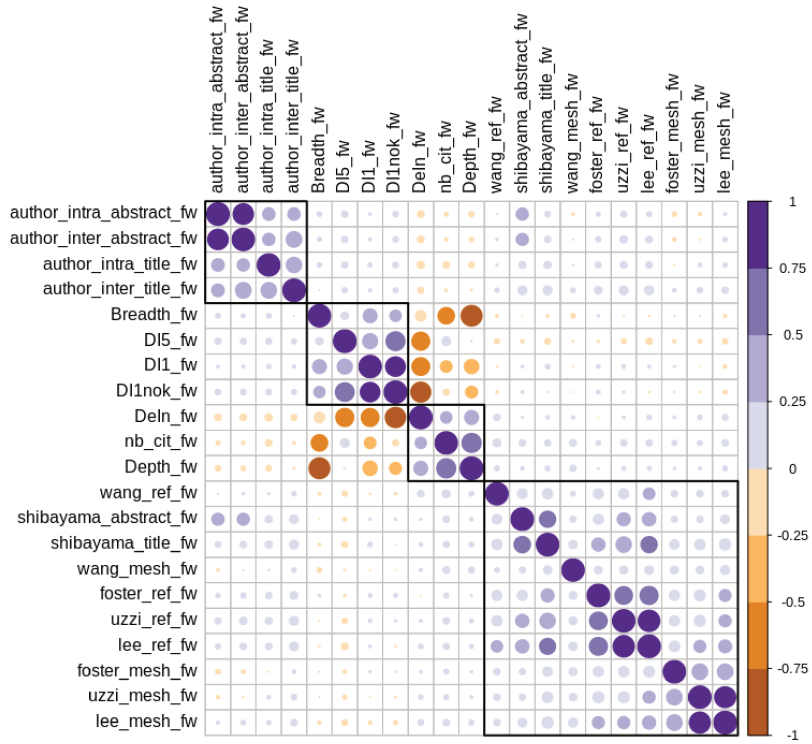


Figure 4.3: Correlogram with hierarchical clustering

in more extreme values. However, several explanations for this phenomenon can be offered. For instance, substantial cognitive diversity might create communication barriers among team members, particularly when individuals are less explorative. Consequently, smaller teams are formed due to potential coordination and knowledge exchange difficulties. In cases with a high average exploratory profile combined with high cognitive diversity, forming smaller teams may be more convenient, as researchers might explore the knowledge space too broadly. Smaller teams could help prevent efforts from dispersing in various directions. As for teams with low cognitive diversity, the absence of cognitive diversity and exploratory profiles could relate to niches where individuals possess similar knowledge and expertise. As a result, many team members might not be necessary, as they can efficiently navigate the local knowledge space. The same argument can be made for individuals with comparable skills and exploratory profiles, as they may represent teams that regularly collaborate on diverse topics. The distinct skill requirements for these teams may be lower, leading to smaller team sizes.

Interestingly, when comparing the analysis of team size with disruption, we confirm the findings of Wu and Wu [2019]. As illustrated in Figure 4.6 in the Appendix, peripheral observations are more disruptive (represented with DI1nok), corresponding to the location of our smaller teams in Figure 4.4 right panel. Teams consolidating science, as indicated by the Depth variable, are also, on average, the most prominent teams. Small teams that disrupt science tend to have exploratory profiles and/or diverse team compositions. Science disruption seems to occur through small teams with either highly distinct skills or very exploratory profiles. What seems essential is the ability to access a broader knowledge space, regardless of whether this space is reached through the team's highly explorative profiles or the team's diversity. Teams composed of individuals who are, on average, highly exploratory but with low team cognitive diversity represent teams with similar skills that cover the knowledge space effectively. In contrast, highly diverse teams with specialized individuals also span the knowledge space to propose disruptive ideas, although they may face communication challenges. The combination of these two factors also appears to contribute to disruption, albeit less prominently, suggesting the detrimental effect of excessive diversity. Another inverted U-shaped relationship exists between a team's average exploratory profile and novelty indicators. When balanced by a relatively explorative average profile, cognitive diversity appears beneficial without showing a saturation point.

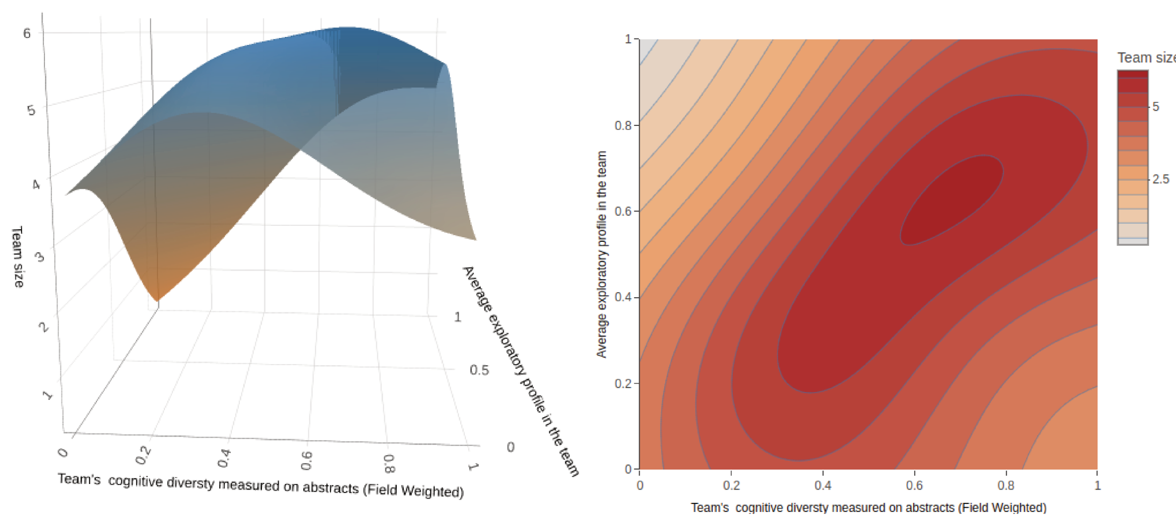


Figure 4.4: Team size, exploratory profiles and cognitive diversity

The relationship differs when we adopt an alternative perspective and consider the proportion of highly exploratory and exploitative individuals within scientific teams. A dome is visible in each indicator, signifying successful trade-offs between exploitation and exploration. Figure 4.5 offers insight into the relationship between these two aspects and scientific recognition and combinatorial novelty. Teams with fewer highly exploratory individuals and a higher proportion of highly exploitative individuals typically contribute to consolidating the field (Depth metric). Conversely, groups with a higher proportion of highly exploratory individuals and a smaller proportion of highly exploitative individuals are more likely to initiate disruptions in their fields (DI1nok metric). These observations complement the findings of Uzzi et al. [2013], which suggest that a balance between conventional and atypical knowledge combinations produces the most impactful research. Moreover, this analysis enables us to examine how the balance between exploratory and exploitative individuals affects knowledge creation itself.

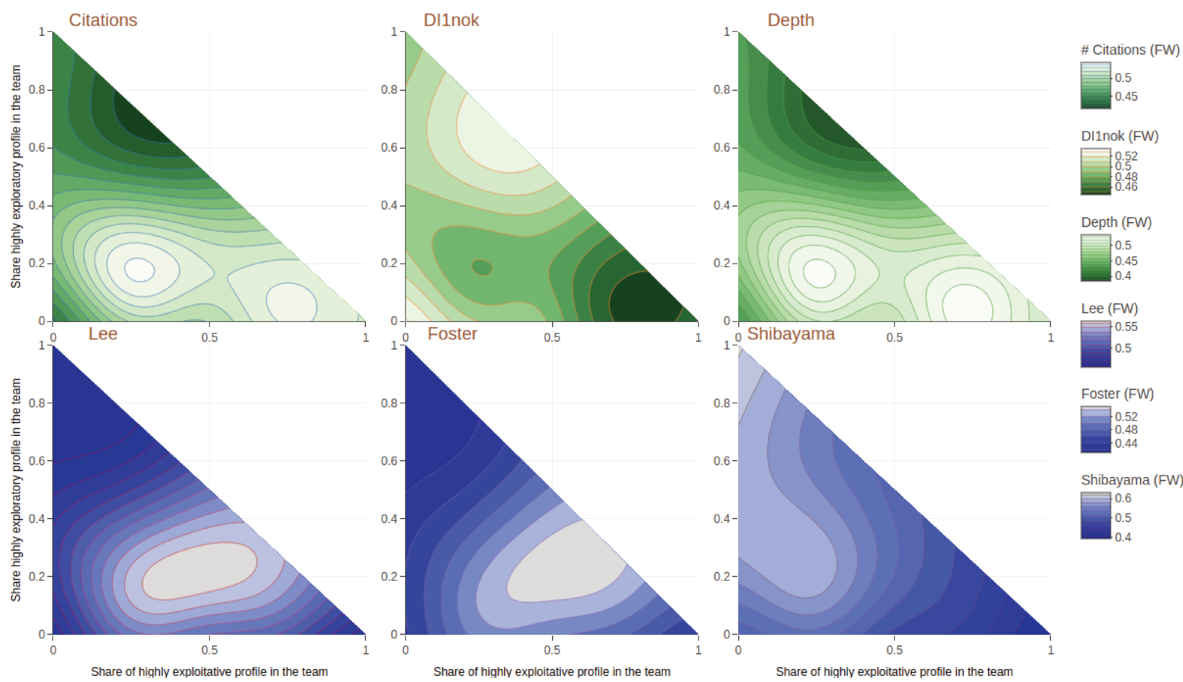


Figure 4.5: Relation between the share of highly exploitative and highly exploratory profile in a team with and Novelty/ Scientific Impact

Teams featuring a fair proportion of exploratory individuals and a more sustained level of exploitative individuals seem to be most likely to generate compelling new combinations of knowledge. Figure 4.5 suggests that an optimal team composition would consist of approximately 50% highly exploitative and 20% highly exploratory individuals to increase the likelihood of combining distant knowledge. The situation is less clear for Shibayama et al. [2021], where a high proportion of highly exploratory individuals appears to be beneficial[6]. Exploratory individuals contribute to the team

[6]This might be connected with the relationship between our measure and the measure of

by introducing fresh and innovative ideas from their extensive knowledge. These individuals can challenge conventional thinking and steer the team in new directions. Simultaneously, they might foster communication among group members with distant knowledge. In contrast, highly exploitative individuals are crucial for refining and optimizing these novel ideas. Their specialized expertise allows the team to identify feasible and effective solutions, ensuring the creative potential of the exploratory individuals is appropriately channelled into tangible outputs. Additionally, their deep understanding of a specific field facilitates effective communication. The highly exploratory profile complements the specialized knowledge and proficiency of the highly exploitative team members. This dynamic enables the team to capitalize on the full potential of their diverse cognitive abilities, optimizing the innovation process and yielding scientific advancements.

4.4 Results

4.4.1 Cognitive dimension and novelty

4.4.1.1 Realized novelty

This subsection examines the relationship between the team's cognitive dimension and novelty indicators. To this end, we report the results of an OLS to identify the joint impact of authors' intra-diversity and inter-diversity on the indicators. The outcomes of these models are presented in Table 4.2.

First, we confirm that cognitive diversity in a scientific team fosters realized novelty. Team cognitive diversity (Row 1-2) reveals a significant positive effect on combinatorial novelty. This suggests distant individuals can ease the combination of distant journals in the references. The squared term has negative coefficients. However, the turning point is higher than 1, meaning the relationship is strictly increasing (See Table 4.13 in Appendix). However, it means that the marginal benefit of cognitive distance is decreasing. When interpreting the coefficients, it is important to remember that the independent and dependent variables are expressed in percentile rank within a given field and year. A one percentage point increase in the independent variable's percentile rank implies a β percentage point increase in the dependent variable. In our case, the marginal effect of a quadratic term depends

Shibayama et al. [2021] as it is measured in a similar manner. Self-citation also directly impacts the relationship between these two metrics as the same combination of articles will be calculated in both metrics.

Table 4.2: Combinatorial Novelty: cognitive diversity and average exploratory profile (Field-Weighted/ References)

	Dependent variable:				
	Uzzi	Lee	Foster	Wang	Shibayama
	(1)	(2)	(3)	(4)	(5)
Author inter $_{abs}$ (FW)	0.169***	0.166***	0.116***	0.098***	0.284***
	(0.008)	(0.007)	(0.010)	(0.006)	(0.007)
Author inter $_{abs}$ˆ2 (FW)	-0.031***	-0.034***	-0.023**	-0.028***	-0.118***
	(0.007)	(0.007)	(0.009)	(0.006)	(0.007)
Author intra $_{abs}$ (FW)	0.056***	0.043***	0.041**	-0.002	0.188***
	(0.014)	(0.013)	(0.019)	(0.008)	(0.009)
Author intra $_{abs}$ˆ2 (FW)	-0.088***	-0.094***	-0.084***	-0.026***	-0.047***
	(0.011)	(0.010)	(0.015)	(0.006)	(0.010)
# References	0.002***	0.002***	0.001***	0.005***	0.002***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
# Meshterms	0.004***	0.006***	0.005***	-0.001***	0.004***
	(0.0004)	(0.0004)	(0.0004)	(0.0002)	(0.0004)
# Authors	0.008***	0.007***	0.007***	0.001***	0.007***
	(0.0004)	(0.0004)	(0.0005)	(0.0003)	(0.0003)
SJR	-0.012***	-0.011***	-0.014***	-0.008***	-0.011***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Year	Yes	Yes	Yes	Yes	Yes
Journal Cat.	Yes	Yes	Yes	Yes	Yes
Observations	1,647,430	1,815,603	1,815,603	1,815,603	1,809,155
R^2	0.055	0.062	0.039	0.122	0.130
Adjusted R^2	0.055	0.062	0.039	0.122	0.130
Residual Std. Error	0.281	0.278	0.310	0.345	0.267
F Statistic	406.544***	512.283***	315.079***	1,065.575***	1,143.840***

Notes: This table reports coefficients of the effect of cognitive diversity and average exploratory profile on combinatorial novelty using PKG. Standard errors are cluster robust at the journal level: ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively. The effects are estimated with an OLS. Variables are field-weighted and constant term, scientific field (Scimago Journal Category), and time-fixed effects are incorporated in all model specifications.

on the value of the independent variable. We can calculate marginal effects at the mean values of the independent variable. For example, in Uzzi et al. [2013] (model 1), the marginal effect of Author inter$_{abs}$ (FW) at the mean value is calculated this way: $\frac{\Delta y}{\Delta(inter)} = 0.169 - 2 * (-0.031) * Mean(Inter)$. Since variables are expressed in percentile rank, the mean and the median are 0.5. The marginal effect can be then calculated easily, $\frac{\Delta y}{\Delta(inter)} = 0.169 - *(-0.031) = 0.2$. This means that by increasing one percentage point on the ranking of team diversity in a given field and year, one can increase by 0.2 percentage points in the ranking of the most novel articles in the field and year.

On the contrary, the average exploratory profile must remain reasonable to maximize novelty. As visible in Table 4.13, the turning points are around 30% for all indicators except Shibayama et al. [2021], for which it is upper than one. This can mean two things, and this is what we will examine in the second part of this results section, either the researchers have a rather moderate explorative profile, or there is a balance between exploratory and exploitative individuals. A set of profiles that are too exploratory seems detrimental, as does a set of too exploitative profiles. As shown in Table 4.2, this holds for all indicators on references, except for Wang et al. [2017], for which the individual effect is negative, one explanation can be the fact that Wang et al. [2017] control for future reutilization of the novel combination. Indeed this gives a 'scientific impact' dimension to the metrics and the presence of more specialized individuals may impact the relevance of the combination for the community, making it more likely to be reused.

On MeSH terms, as visible in Table 4.11 in the Appendix, individual exploratory aspects appear to have a direct negative impact. Indexers assign the MeSH terms and may be subject to bias or misinterpretation. In contrast, the references directly relate to the researchers' choices and reflect their interests and preferences. There are two possibilities, indexers may be unable to capture all the nuances and subtleties of research conducted by individuals with high-average exploratory profiles. Alternatively, the novelty of references could be induced by an author bias in citing previous works irrelevant to the contribution. Researchers' past publications do not directly impact indexers, so she might not need to qualify the article with distant MeSH terms because the novelty is not sufficiently explicit. This suggests that MeSH terms do not reflect the diversity of knowledge and ideas present in individual past work but rather the diversity of competencies between team members.

These relations remain consistent when regressions are not performed using percentage rank information, and indicator behavior with MeSH terms and references seems to be much more corroborated, as visible in Table 4.14 and 4.15 in the Appendix. The fact that the effect is nearly the same on most of the indicators of novelty demonstrates the robustness of this analysis - our measure captures something similar regardless of the construction of the novelty indicator and the information used.

The potential for novelty seems more apparent when looking at the exact composition in terms of exploratory profiles, i.e., the share of explorative individuals and the share of highly exploitative individuals. In Table 4.3, we replace the average exploratory profile variables with the exploitative and exploratory individual shares and the interaction of these two variables.

Table 4.3: Combinatorial Novelty: Cognitive diversity, highly exploratory and exploitative profile (Field-Weighted/ References)

	Dependent variable:				
	Uzzi	Lee	Foster	Wang	Shibayama
	(1)	(2)	(3)	(4)	(5)
Author inter $_{abs}$ (FW)	0.168***	0.163***	0.107***	0.066***	0.400***
	(0.014)	(0.012)	(0.020)	(0.008)	(0.012)
Author inter $_{abs}$ˆ2 (FW)	-0.007	-0.006	0.028	0.001	-0.160***
	(0.012)	(0.011)	(0.018)	(0.007)	(0.012)
Share exploratory	-0.166***	-0.173***	-0.214***	-0.084***	-0.022***
	(0.007)	(0.007)	(0.010)	(0.004)	(0.006)
Share exploitative	0.027***	0.053***	0.057***	0.002	-0.092***
	(0.003)	(0.003)	(0.005)	(0.002)	(0.004)
Share exploratory * Share exploitative	0.298***	0.273***	0.390***	0.080***	-0.112***
	(0.016)	(0.016)	(0.020)	(0.011)	(0.018)
# References	0.002***	0.002***	0.001***	0.005***	0.002***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
# Meshterms	0.004***	0.006***	0.005***	-0.001***	0.004***
	(0.0004)	(0.0003)	(0.0003)	(0.0002)	(0.0004)
# Authors	0.008***	0.007***	0.007***	0.001***	0.006***
	(0.0004)	(0.0004)	(0.0005)	(0.0002)	(0.0003)
SJR	-0.012***	-0.011***	-0.014***	-0.008***	-0.011***
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)
Year	Yes	Yes	Yes	Yes	Yes
Journal Cat.	Yes	Yes	Yes	Yes	Yes
Observations	1,647,430	1,815,603	1,815,603	1,815,603	1,809,155
R^2	0.059	0.068	0.046	0.122	0.129
Adjusted R^2	0.059	0.068	0.046	0.122	0.129
Residual Std. Error	0.280	0.277	0.308	0.345	0.267
F Statistic	436.681***	556.617***	372.829***	1,065.763***	1,132.467***

Notes: This table reports coefficients of the effect of cognitive diversity and highly exploratory and exploitative profiles on combinatorial novelty using PKG. Standard errors are cluster robust at the journal level: ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively. The effects are estimated with an OLS. Variables are field-weighted and constant term, scientific field (Scimago Journal Category), and time-fixed effects are incorporated in all model specifications.

While cognitive diversity appears to be always beneficial to combine new knowledge, the presence of too many explorative individuals is harmful. Indeed, its presence only becomes beneficial when counterbalanced by a higher share of exploitative individuals. We can clearly see how this trade-off is necessary to create novelty in the regressions. In the same way as before, the coefficients can be interpreted directly, a percentage point increase in the share of highly explorative individuals increases by β percentage point in the ranking of the most novel articles in the field and year.

Exploratory individuals will develop new perspectives that specialized individuals will capitalize on to make them succeed. A larger share of specialized individuals facilitates communication among members if they are in the same field; otherwise, scientists with diverse backgrounds appear to facilitate communication among team members who are cognitively distant [Page, 2008]. This mirrors the "Renaissance" individual of [Jones, 2009] or the "transformational leader" of Shin and Zhou [2007] who can ease connections between distant members and foster the team's creativity.

Too many such individuals would make the exploration less efficient, and the emerging ideas would potentially not be successfully implemented because the embedding of the conducted research in a scientific paradigm would not be sufficient. The results are similar across novelty indicators, except for Shibayama et al. [2021], in which the best team composition is made from non-exploitative, non-highly exploratory researchers. Table 4.12 in the Appendix shows that the results also hold for indicators based on MeSH terms.

The two sets of results on the impact of cognitive distance and researcher profile show that combining specialized and exploratory profiles is a good proxy for potential novelty as it enhances the realized novelty in the team[7]. While Uzzi et al. [2013] show that this trade-off between conventional and atypical combinations of knowledge is the most impactful, we demonstrate that this idea holds at the team level as well and that these configurations are most likely to achieve atypical combinations.

4.4.1.2 Perceived novelty

In this subsection, we examine the relationship between the cognitive dimension and novelty as assessed by experts. Specifically, we employ a Logit model to identify the impact of authors' intra-diversity and inter-diversity on the likelihood of being classified in at least one novel category. The results of these models are presented in Table 4.4. The effect of team cognitive diversity plays a positive role in *perceived* novelty, as seen in the first and second specifications. This effect is less clear when considering individual characteristics. The average exploratory profile has a negative impact. In

[7]Table 4.17 provided in the Appendix shows that the results are similar when considering unnormalized indicators

model 3, we can see that our previous results on *realized* novelty (Table 4.2) only holds for the cognitive distance between individuals when tested on *perceived* novelty. In contrast, when examining the specifications with the share of highly exploratory and exploitative individuals, the results corroborate the regressions performed on *realized* novelty. The proportion of highly exploratory individuals has a negative effect. Instead, typical individuals play a positive role, and the intersection of both types of researchers is indeed positive for predicting novelty. Note that in this specification, cognitive diversity between members is no longer significant.

Table 4.4: Faculty Opinions: cognitive diversity and average exploratory profile, highly exploratory and exploitative profile (Field-Weighted)

	Dependent variable:				
	Novelty Perceived				
	(1)	(2)	(3)	(4)	(5)
Author inter $_{abs}$ (FW)	0.306**	0.715*			0.330
	(0.126)	(0.388)			(0.302)
Author intra $_{abs}$ (FW)	-0.532***	-0.196			
	(0.155)	(0.419)			
Author inter $_{abs}$^2 (FW)		-0.438			-0.270
		(0.376)			(0.325)
Author intra $_{abs}$^2 (FW)		-0.364			
		(0.379)			
Share exploratory			-0.675**	-1.233***	-1.238***
			(0.275)	(0.371)	(0.384)
Share exploitative			0.339***	0.317***	0.337***
			(0.117)	(0.118)	(0.115)
Share exploratory * Share exploitative				2.360**	2.289**
				(1.052)	(1.062)
Control variables	YES	YES	YES	YES	YES
Observations	12,555	12,555	12,555	12,555	12,555
Log Likelihood	-7,076.944	-7,073.965	-7,072.608	-7,070.408	-7,069.551
AIC	14,423.890	14,421.930	14,415.220	14,412.820	14,415.100

Notes: This table reports coefficients of the effect of cognitive diversity, average exploratory profile, highly exploratory and exploitative profiles on perceived novelty from Faculty Opinions. Standard errors are cluster robust at the journal level in parentheses: ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The effects is estimated using a Logit model. Variables are field-weighted and constant term, scientific field (Scimago Journal Category) and time fixed effects are incorporated in all model specifications.

However, when examining Table 4.7 in the Appendix, we can see that the effects are quite heterogeneous across labels. We chose the four labels for which more than

1000 papers had been classified to perform the regressions. The effect of cognitive distance between team members is visible in the "Technical Advance" category but not significant for the remaining labels. Conversely, in Table 4.9, we can see that the results in terms of exploratory profiles are mainly driven by the 'Interesting Hypothesis' label. Results here are a bit different since we observe a U shape, meaning that highly specialized or highly diverse teams most often publish articles labelled as "Interesting hypotheses". Results are quite similar when using Poisson regression and modelling the number of times a paper is labelled in a given category as visible in Table 4.8 and Table 4.10 in the Appendix.

4.4.2 Cognitive dimension and impact

This subsection examines the relationship between the team's cognitive dimension and impact measures. To this end, we report the results of an OLS to identify the joint impact of authors' intra-diversity and inter-diversity on the indicators. The outcomes of these models are presented in 4.5 and 4.6.

Our analysis emphasises the need to differentiate the forms of impact to understand better how the cognitive aspect influences scientific recognition. Indeed, we use the traditional indicator of the number of citations and indicators of disruption and consolidation. The composition of the teams has a significant influence on the type of impact of the studies conducted.

The Table 4.5 regression tables indicate a double inverse U-shaped relationship between the cognitive dimension and the number of citations. Table 4.13 shows that both turning points are around 45%. Following Uzzi, a too-conventional work might not be as impactful as the contribution is more marginal. Conversely, peers may not sufficiently consider a too-novel study. This phenomenon is reflected in the composition of the teams as we can see in the differences between consolidation and disruption indicators. Indeed, to consolidate, it is necessary to have a team with a low average exploratory profile and low average cognitive distance between members. The relationship is negative for consolidation indicators (DeIn and Depth) for both intra and inter-individual levels; the effect is sometimes captured via quadratic terms. This means that cognitive diversity is negatively related to the fact that papers citing the focal paper also cite each other or cite many of the references from the focal article. Specialized teams are the ones who consolidate the science.

For disruptive indicators, the picture is rather different (DI1, DI5, DI1nok and Breadth). Cognitive distance still seems to be globally favorable for disruption.

Table 4.5: Scientific recognition: cognitive diversity and average exploratory profile (Field-Weighted)

	Dependent variable:						
	# cit.	DI1	DI5	DI1nok	DeIn	Breadth	Depth
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Author inter $_{abs}$ (FW)	0.031***	0.021***	0.034***	0.047***	-0.067***	-0.010*	0.002
	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)	(0.006)	(0.007)
Author inter $_{abs}$ˆ2 (FW)	-0.036***	0.012**	0.005	0.002	0.008	0.015***	-0.012**
	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)
Author intra $_{abs}$ (FW)	0.070***	-0.057***	0.026***	-0.008	0.009	0.014**	-0.004
	(0.008)	(0.007)	(0.008)	(0.008)	(0.009)	(0.006)	(0.008)
Author intra $_{abs}$ˆ2 (FW)	-0.072***	0.038***	0.009	0.024***	-0.030***	0.021***	-0.038***
	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.007)
# References	0.003***	-0.001***	-0.003***	-0.002***	0.004***	-0.0001*	0.001***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.00005)	(0.0001)
# Meshterms	0.008***	-0.002***	-0.003***	-0.003***	0.005***	-0.003***	0.006***
	(0.0004)	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0002)	(0.0004)
# Authors	0.012***	-0.006***	-0.002***	-0.005***	0.006***	-0.009***	0.012***
	(0.0004)	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.0003)	(0.0004)
SJR	0.039***	-0.019***	0.002	-0.006***	0.008***	-0.026***	0.030***
	(0.005)	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)	(0.004)
Year	Yes	Yes	Yes	Yes	Yes	Yes	
Journal Cat.	Yes	Yes	Yes	Yes	Yes	Yes	
Observations	1,826,207	1,826,207	1,826,207	1,826,207	1,826,207	1,826,207	1,826,207
R^2	0.173	0.029	0.069	0.034	0.137	0.051	0.075
Adjusted R^2	0.173	0.029	0.069	0.034	0.137	0.051	0.075
Residual Std. Error	0.266	0.281	0.281	0.280	0.270	0.270	0.291
F Statistic	1,621.946***	233.770***	575.115***	269.625***	1,227.699***	413.321***	629.396***

Notes: This table reports coefficients of the effect of cognitive diversity and average exploratory profile on scientific recognition using PKG. Standard errors are cluster robust at the journal level: ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively. The effects are estimated with an OLS. Variables are field-weighted and constant term, scientific field (Scimago Journal Category), and time-fixed effects are incorporated in all model specifications.

Then, the Breadth disruption indicator, which examines how often articles citing the focal paper also cite each other, seems to indicate a U-shaped relationship with a turning point at 0.33, i.e. if the individuals are very distant or if they are very close, this produces the most disruptive articles in the sense that the citations will be concentrated towards the focal paper.

Although not always significant, the intra-individual effect is more mixed; teams with higher average explorative profiles globally appear to have a higher disruption potential, but this does not hold for DI1. The DI1NOK index follows the same pattern as DI5, with the exception that it is the quadratic term that takes over.

The articles that are consolidating science are articles with low team diversity and low average exploratory profiles. Here we can observe the notion of highly specialized individuals who conduct more confirmatory and therefore consolidating research. The opposite is true for disruption. The teams' diversity always seems beneficial for proposing disruptive ideas. Articles receiving the most citation are again a matter of a trade-off between a cognitively not-too-distant team and a somewhat reasonable average level of exploration[8].

In Table 4.6, we specify the team's composition in terms of exploratory/exploitative profile and found that the relationship of the cognitive distance with the impact measures remains almost similar. For consolidation metrics and citation counts, the share of exploitative individuals is clearly beneficial. The exploitative profile reduces the risk of failure as researchers learn from experience and combinations that have failed [Vincenti, 1990]. Whereas too exploratory profiles seem to affect the expected number of citations negatively, the effect appears mixed for consolidation since it is positive DeIn and insignificant for Depth. In both cases, combining the two types of profiles is harmful. At the same time, the share of exploitative individuals is positive, suggesting that combining these two types of profiles is not optimal for consolidating research. To achieve disruption, it is better to minimize the number of individuals who are too exploratory or too specialized, but combining both types of profiles seems once again essential. We can see how the impact of highly explorative profiles is always negative, and the impact of exploitative profiles is also negative. Still, the interaction between the two is always positive for all disruptiveness measures.

In conclusion, the analysis shows how teams with a high share of specialized individuals or low average exploratory profiles are teams that consolidate science. In contrast, teams that get the most recognition in terms of disruption combine highly exploitative and highly exploratory individuals and have cognitively more distant members[9].

4.5 Conclusion

This paper examines the effect of exploratory scholars and, in a broader way, team composition on creativity. Our findings suggest that the cognitive dimension plays

[8]For regressions without field-year normalization as presented in Table 4.16, the results are more mixed and less clear. The cognitive aspect seems to follow a U-shaped pattern, with teams that are very close or distant being the most disruptive. The results are more robust for breadth, with diversity consistently appearing to be beneficial.

[9]For regressions without field-year normalization (see Table 4.19), the results are less homogeneous for the cognitive distance aspect, but the combination of explorative and exploitative is robust. The interaction of the two consistently leads to disruption.

Table 4.6: Scientific recognition: cognitive diversity, highly exploratory and exploitative profile (Field-Weighted)

	Dependent variable:						
	# cit.	DI1	DI5	DI1nok	DeIn	Breadth	Depth
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Author inter $_{abs}$ (FW)	0.088***	-0.058***	0.020*	0.004	-0.019	-0.004	0.003
	(0.010)	(0.009)	(0.011)	(0.010)	(0.012)	(0.007)	(0.009)
Author inter $_{abs}$ˆ2 (FW)	-0.073***	0.067***	0.026***	0.042***	-0.037***	0.025***	-0.028***
	(0.010)	(0.009)	(0.009)	(0.008)	(0.009)	(0.007)	(0.008)
Share exploratory	-0.023***	-0.055***	-0.041***	-0.056***	0.058***	-0.006	-0.003
	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.004)	(0.005)
Share exploitative	0.029***	-0.033***	-0.056***	-0.049***	0.058***	-0.024***	0.032***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
Share exploratory * Share exploitative	-0.023**	0.132***	0.047***	0.096***	-0.087***	0.059***	-0.034***
	(0.012)	(0.010)	(0.011)	(0.010)	(0.011)	(0.011)	(0.012)
# References	0.003***	-0.001***	-0.003***	-0.002***	0.004***	-0.0001*	0.001***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.00005)	(0.0001)
# Meshterms	0.008***	-0.002***	-0.003***	-0.003***	0.005***	-0.003***	0.006***
	(0.0004)	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0002)	(0.0004)
# Authors	0.012***	-0.007***	-0.003***	-0.005***	0.006***	-0.010***	0.013***
	(0.0004)	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.0003)	(0.0004)
SJR	0.038***	-0.018***	0.003	-0.006***	0.007***	-0.026***	0.030***
	(0.005)	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)	(0.004)
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Journal Cat.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,826,207	1,826,207	1,826,207	1,826,207	1,826,207	1,826,207	1,826,207
R^2	0.174	0.030	0.071	0.035	0.139	0.051	0.075
Adjusted R^2	0.174	0.030	0.071	0.035	0.139	0.050	0.075
Residual Std. Error	0.266	0.281	0.280	0.280	0.270	0.270	0.291
F Statistic	1,619.636***	239.244***	586.510***	281.922***	1,243.711***	410.608***	626.296***

Notes: This table reports coefficients of the effect of cognitive diversity and highly exploratory and exploitative profiles on scientific recognition using PKG. Standard errors are cluster robust at the journal level: ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively. The effects are estimated with an OLS. Variables are field-weighted and constant term, scientific field (Scimago Journal Category), and time-fixed effects are incorporated in all model specifications.

a crucial role in the creative process, and significantly influences the two pillars of creativity: originality and success. We first show that the team's cognitive diversity strongly influences novelty (*realized* and *perceived*) of the research conducted. We also show that a double-inversed U-shaped relationship exists between cognitive dimensions (intra and inter) and the impact in terms of citations. Our study also highlights the strong connection between the cognitive dimension and the nature of these citations. Teams with more exploitative profiles tend to consolidate science, while those with more exploratory individuals disrupt it and propose more distant knowledge combinations, only when associated with exploitative ones. Our research underscores how team composition in terms of profiles lies at the heart of scientific creativity.

Multiple limitations arise in our study. First, concerning data used, PKG is based on advanced heuristics and algorithms to disambiguate authors using affiliation and additional metadata Xu et al. [2020]. While there is a considerable amount of research on addressing noise in Knowledge Graphs [Fasoulis et al., 2020] and improvements in these methods may increase their reliability in the future, we cannot guarantee that errors or inconsistencies will not occur when dealing with author-level information in PKG.

Other shortcomings are directly related to the creation of our indicator. First, many methods and hyper-parameters were chosen for the simplicity of computation. The embedding is a pre-trained model from SpaCy and is not state-of-the-art. One should compare the behavior of different embedding techniques but also on what kind of text they are applied and the distance measure used. We suspect that the two papers might be close given a specific embedding and distance measure but highly distant given other parameters. In addition, the distance between the two papers would vary depending on whether the distance metric is applied to the paper's title, abstract, or full text. The semantic distances between researchers can be influenced by biases inherent in the fields and journal practices. For example, if researchers publish in different journals, the structure and format of their abstracts may be affected even if their research topic or area of expertise remains unchanged. Another hyper-parameter we used is the time window for an author's past publication. We considered a time window of 5 years. This suggests that any paper published by the author before this point would not be captured. One could argue that past behavior influences current behavior, and a highly diverse background can be proxied by recent publications. Yet no evidence supports this hypothesis. Another issue is how we define authors' cognitive aspect by considering only past publications. Although we do not try to approximate the skills of a researcher but only their disposition to do diverse research, we are not sure how working on a topic is enough to understand then and manage this new knowledge. This raises the question of the exact competencies of a transformational leader and if the past paper is sufficient to proxy it. Also, a specialized author could have previously worked on distant papers but only on his topic/methodology. Our measure defines it as diverse, yet is it true? Although solo publications can be used to construct an author's profile, the increasing significance of teamwork in scientific research makes it uncertain whether a complete and precise profile can be established solely on this basis. Another option could be to incorporate external information, such as educational background, and assign greater weight to papers that align with the author's education. However, obtaining this information can be challenging as it often requires web scraping, which is not easily scalable. The last issue in our mind about using past publications is ghost and honorary authorship as it is common that some authors contributed very little to the production of the article. [Sugimoto and Larivière, 2018, Pruschak and Hopp, 2022]. Both are problems to consider while defining a coauthored paper as part of your knowledge space.

In our analysis, we solely focused on the cognitive diversity of researchers, but diversity encompasses various aspects as highlighted by prior research studies [Medin and Lee, 2012, Hofstra et al., 2020]. According to Koopmann et al. [2021], there are four proximity dimensions among researchers, namely cognitive, institutional, social, and geographical. Relying solely on PKG to approximate all of these dimensions could be challenging. Still, alternative sources such as OpenAlex could provide more comprehensive information on a researcher's institutions, past institutions, and authors' characteristics. For instance, relying on PKG to construct a researcher's seniority could be biased because of the restriction on health sciences papers. Exploring these additional channels could lead to developing supplementary measures that complement cognitive diversity.

Another area worth exploring is the temporal dynamic between exploring new ideas and exploiting existing ones. As we discussed earlier, discovering new concepts is essential for addressing major challenges. However, there is often a pattern of moving through cycles of exploration and exploitation within a particular field. Similarly, authors may initially focus on a particular subject and then switch to a different area to gain a fresh perspective on the first one once they have developed sufficient expertise.

To increase the efficiency of the scientific system, it is necessary to conduct further research on the composition of research teams and their impact on creativity. Our preliminary results indicate that policymakers and grant evaluators should consider both individual and team-level characteristics and not only citations when making decisions about research funding and support. We have explored some research avenues to deepen our understanding of this phenomenon, and we encourage other researchers to build upon our work in this area. By continuing to investigate these factors, we can develop more effective strategies for supporting and fostering creativity within research teams, ultimately leading to more impactful and innovative scientific outcomes.

4.6 Appendix

Figures

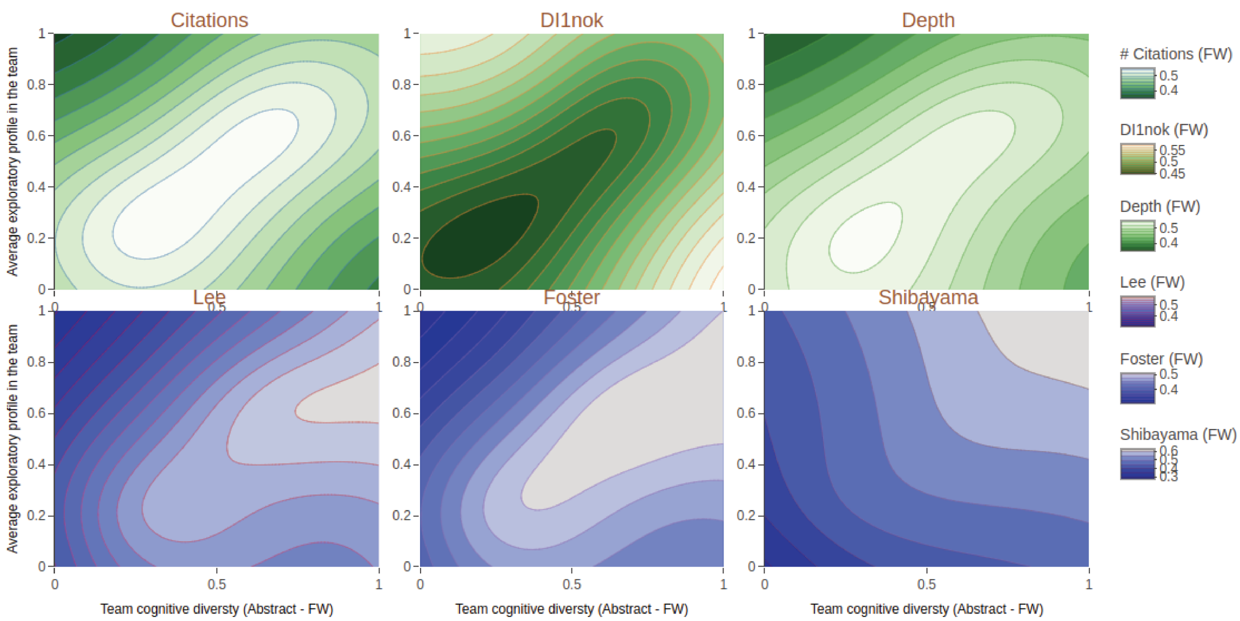


Figure 4.6: Relation between cognitive diversity, average exploratory profile and Novelty/ Scientific Impact

Regressions

Novelty indicators and Faculty Opinion

Table 4.7: Faculty Opinions: Cognitive diversity and average exploratory profile (Field-Weighted)

	Dependent variable:			
	Logit Model			
	Interesting Hyp.	Technical Adv.	Confirmation	Controversial
Author inter $_{abs}$ (FW)	-0.625	1.485***	-0.427	-0.757
	(0.387)	(0.338)	(0.386)	(0.491)
Author inter $_{abs}$ˆ2 (FW)	0.414	-1.101***	0.310	0.543
	(0.382)	(0.328)	(0.381)	(0.516)
Author intra $_{abs}$ (FW)	-0.191	0.209	0.001	0.278
	(0.388)	(0.336)	(0.384)	(0.580)
Author intra $_{abs}$ˆ2 (FW)	-0.016	-0.465	0.199	0.016
	(0.383)	(0.324)	(0.365)	(0.602)
# References	0.006***	-0.012***	-0.0002	-0.001
	(0.001)	(0.002)	(0.001)	(0.002)
# Meshterms	0.019***	-0.040***	0.011***	0.004
	(0.004)	(0.006)	(0.004)	(0.005)
# Authors	-0.026***	0.018***	0.005	-0.017*
	(0.006)	(0.005)	(0.004)	(0.009)
SJR	0.065***	-0.019**	-0.008	0.006
	(0.011)	(0.010)	(0.005)	(0.008)
Year	Yes	Yes	Yes	Yes
Journal Cat.	Yes	Yes	Yes	Yes
Observations	12,555	12,555	12,555	12,555
Log Likelihood	-7,919.383	-7,202.326	-7,657.496	-3,866.333
Akaike Inf. Crit.	16,112.770	14,678.650	15,588.990	8,006.667

Notes: This table reports coefficients of the effect of cognitive diversity and average exploratory profile on perceived novelty from Faculty Opinions. Standard errors are cluster robust at the journal-level: ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The effects are estimated with a Logit model. Variables are field-weighted and constant term, scientific field (Scimago Journal Category) and time fixed effects are incorporated in all model specifications.

Table 4.8: Faculty Opinions: Cognitive diversity and average exploratory profile (Field-Weighted)

	Dependent variable:			
	Poisson Model			
	Interesting Hyp.	Technical Adv.	Confirmation	Controversial
Author inter $_{abs}$ (FW)	-0.410*	1.225***	-0.330	-0.828**
	(0.209)	(0.204)	(0.287)	(0.413)
Author inter $_{abs}$ˆ2 (FW)	0.291	-0.918***	0.259	0.689
	(0.210)	(0.192)	(0.285)	(0.440)
Author intra $_{abs}$ (FW)	-0.009	0.320	0.129	0.271
	(0.184)	(0.217)	(0.248)	(0.495)
Author intra $_{abs}$ˆ2 (FW)	-0.119	-0.492**	-0.041	-0.108
	(0.181)	(0.219)	(0.223)	(0.521)
# References	0.003***	-0.006***	-0.0001	0.0005
	(0.001)	(0.002)	(0.001)	(0.001)
# Meshterms	0.013***	-0.024***	0.007*	0.002
	(0.002)	(0.005)	(0.003)	(0.005)
# Authors	-0.017***	0.009***	0.003	-0.012
	(0.004)	(0.003)	(0.003)	(0.009)
SJR	0.039***	-0.0004	0.002	0.014*
	(0.007)	(0.007)	(0.004)	(0.007)
Year	Yes	Yes	Yes	Yes
Journal Cat.	Yes	Yes	Yes	Yes
Observations	12,555	12,555	12,555	12,555
Log Likelihood	-10,420.250	-9,880.221	-8,978.963	-4,358.803
Akaike Inf. Crit.	21,114.510	20,034.440	18,231.920	8,991.606

Notes: This table reports coefficients of the effect of cognitive diversity and average exploratory profile on perceived novelty from Faculty Opinions. Standard errors are cluster robust at the journal level: ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The effects are estimated with a Poisson model. Variables are field-weighted and constant term, scientific field (Scimago Journal Category) and time fixed effects are incorporated in all model specifications.

Table 4.9: Faculty Opinions: Cognitive diversity, highly exploratory and exploitative profile (Field-Weighted)

	Dependent variable:			
	Logit Model			
	Interesting Hyp.	Technical Adv.	Confirmation	Controversial
Author inter $_{abs}$ (FW)	-0.859***	1.500***	-0.289	-0.476
	(0.277)	(0.317)	(0.309)	(0.377)
Author inter $_{abs}$ˆ2 (FW)	0.590*	-1.244***	0.228	0.421
	(0.308)	(0.338)	(0.333)	(0.371)
Share exploratory	-0.754*	-0.644	0.868**	-0.193
	(0.450)	(0.443)	(0.441)	(0.607)
Share exploitative	0.304***	-0.014	-0.070	-0.097
	(0.083)	(0.118)	(0.097)	(0.160)
Share exploratory * Share exploitative	2.911***	0.015	-1.069	1.822
	(1.073)	(1.132)	(1.048)	(1.650)
# References	0.006***	-0.012***	-0.0001	-0.001
	(0.001)	(0.002)	(0.001)	(0.002)
# Meshterms	0.019***	-0.041***	0.012***	0.004
	(0.004)	(0.006)	(0.004)	(0.005)
# Authors	-0.024***	0.018***	0.005	-0.018*
	(0.006)	(0.005)	(0.004)	(0.009)
SJR	0.065***	-0.019*	-0.008	0.005
	(0.010)	(0.010)	(0.005)	(0.008)
Year	Yes	Yes	Yes	Yes
Journal Cat.	Yes	Yes	Yes	Yes
Observations	12,555	12,555	12,555	12,555
Log Likelihood	-7,910.400	-7,202.819	-7,655.698	-3,866.431
Akaike Inf. Crit.	16,096.800	14,681.640	15,587.400	8,008.863

Notes: This table reports coefficients of the effect of cognitive diversity and highly exploratory and exploitative profiles on perceived novelty from Faculty Opinions. Standard errors are cluster robust at the journal-level: ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The effects are estimated with a Logit model. Variables are field-weighted and constant term, scientific field (Scimago Journal Category) and time fixed effects are incorporated in all model specifications.

Table 4.10: Faculty Opinions: Cognitive diversity, highly exploratory and exploitative profile (Field-Weighted)

	Dependent variable:			
	Poisson Model			
	Interesting Hyp.	Technical Adv.	Confirmation	Controversial
Author inter $_{abs}$ (FW)	-0.485***	1.336***	-0.171	-0.560*
	(0.157)	(0.224)	(0.248)	(0.322)
Author inter $_{abs}$ˆ2 (FW)	0.318*	-1.106***	0.121	0.472
	(0.163)	(0.224)	(0.261)	(0.332)
Share exploratory	-0.718**	-0.550*	0.478*	-0.083
	(0.334)	(0.312)	(0.246)	(0.513)
Share exploitative	0.135***	-0.026	-0.020	-0.084
	(0.049)	(0.077)	(0.066)	(0.171)
Share exploratory * Share exploitative	2.112***	-0.106	-0.564	1.674
	(0.662)	(0.762)	(0.640)	(1.504)
# References	0.003***	-0.006***	-0.0001	0.0005
	(0.001)	(0.002)	(0.001)	(0.001)
# Meshterms	0.013***	-0.024***	0.007**	0.002
	(0.002)	(0.005)	(0.003)	(0.005)
# Authors	-0.016***	0.009***	0.003	-0.013
	(0.004)	(0.003)	(0.003)	(0.009)
SJR	0.039***	0.00003	0.002	0.014*
	(0.007)	(0.007)	(0.004)	(0.007)
Year	Yes	Yes	Yes	Yes
Journal Cat.	Yes	Yes	Yes	Yes
Observations	12,555	12,555	12,555	12,555
Log Likelihood	-10,415.010	-9,880.661	-8,977.912	-4,358.090
Akaike Inf. Crit.	21,106.010	20,037.320	18,231.830	8,992.180

Notes: This table reports coefficients of the effect of cognitive diversity and highly exploratory and exploitative profiles on perceived novelty from Faculty Opinions. Standard errors are cluster robust at the journal level: ***, ** and * indicate significance at the 1%, 5% and 10% level, respectively. The effects are estimated with a Poisson model. Variables are field-weighted and constant term, scientific field (Scimago Journal Category) and time fixed effects are incorporated in all model specifications.

Novelty indicators with Mesh Terms

Cognitive diversity and average exploratory profile effect on Novelty

Table 4.11: Combinatorial Novelty: cognitive diversity and average exploratory profile (Field-Weighted/ Meshterms)

	Dependent variable:			
	Uzzi	Lee	Foster	Wang
	(1)	(2)	(3)	(4)
Author inter $_{abs}$ (FW)	0.067***	0.114***	0.056***	0.062***
	(0.009)	(0.007)	(0.008)	(0.006)
Author inter $_{abs}$ˆ2 (FW)	-0.016**	-0.050***	-0.025***	-0.008
	(0.008)	(0.006)	(0.008)	(0.006)
Author intra $_{abs}$ (FW)	-0.020*	0.025***	-0.029**	-0.055***
	(0.012)	(0.009)	(0.013)	(0.009)
Author intra $_{abs}$ˆ2 (FW)	-0.047***	-0.062***	-0.042***	-0.010
	(0.010)	(0.008)	(0.011)	(0.007)
# References	0.001***	0.001***	0.001***	0.0002***
	(0.0001)	(0.00005)	(0.0001)	(0.00003)
# Meshterms	0.007***	0.014***	0.0004	0.029***
	(0.001)	(0.001)	(0.0004)	(0.0004)
# Authors	0.002***	0.008***	0.005***	0.001***
	(0.0004)	(0.0004)	(0.0004)	(0.0003)
SJR	-0.004***	-0.006***	-0.005***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Year	Yes	Yes	Yes	Yes
Journal Cat.	Yes	Yes	Yes	Yes
Observations	661,821	1,823,859	1,823,859	1,823,859
R^2	0.029	0.083	0.015	0.153
Adjusted R^2	0.029	0.083	0.015	0.152
Residual Std. Error	0.285	0.276	0.300	0.360
F Statistic	86.982***	699.050***	121.183***	1,390.452***

Notes: This table reports coefficients of the effect of cognitive diversity and average exploratory profile on combinatorial novelty using PKG. Standard errors are cluster robust at the journal level: ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively. The effects are estimated with an OLS. Variables are field-weighted and constant term, scientific field (Scimago Journal Category), and time-fixed effects are incorporated in all model specifications.

Share of Highly Exploratory Profile

Table 4.12: Combinatorial Novelty: Cognitive diversity, highly exploratory and exploitative profile (Field-Weighted/ Meshterms)

	Dependent variable:			
	Uzzi	Lee	Foster	Wang
	(1)	(2)	(3)	(4)
Author inter $_{abs}$ (FW)	0.045***	0.115***	0.007	0.002
	(0.011)	(0.010)	(0.013)	(0.008)
Author inter $_{abs}$ˆ2 (FW)	0.013	-0.032***	0.030**	0.027***
	(0.010)	(0.010)	(0.012)	(0.007)
Share exploratory	-0.107***	-0.113***	-0.150***	-0.063***
	(0.006)	(0.006)	(0.007)	(0.005)
Share exploitative	0.068***	0.045***	0.056***	0.026***
	(0.003)	(0.003)	(0.004)	(0.002)
Share exploratory * Share exploitative	0.185***	0.189***	0.254***	0.106***
	(0.016)	(0.013)	(0.016)	(0.012)
# References	0.001***	0.001***	0.001***	0.0002***
	(0.00005)	(0.00005)	(0.0001)	(0.00003)
# Meshterms	0.007***	0.014***	0.0004	0.029***
	(0.001)	(0.001)	(0.0004)	(0.0004)
# Authors	0.003***	0.008***	0.005***	0.001***
	(0.0004)	(0.0004)	(0.0004)	(0.0003)
SJR	-0.004***	-0.006***	-0.004***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Year	Yes	Yes	Yes	Yes
Journal Cat.	Yes	Yes	Yes	Yes
Observations	661,821	1,823,859	1,823,859	1,823,859
R^2	0.033	0.086	0.019	0.152
Adjusted R^2	0.032	0.086	0.019	0.152
Residual Std. Error	0.284	0.275	0.299	0.360
F Statistic	97.014***	721.442***	149.772***	1,383.049***

Notes: This table reports coefficients of the effect of cognitive diversity and highly exploratory and exploitative profiles on combinatorial novelty using PKG. Standard errors are cluster robust at the journal level: ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively. The effects are estimated with an OLS. Variables are field-weighted and constant term, scientific field (Scimago Journal Category), and time-fixed effects are incorporated in all model specifications.

Turning points

Table 4.13: Turning Points for Combinatorial Novelty and Scientific Impact

Regression	Author intra *abs* (FW)	Author inter *abs* (FW)
Uzzi	0.318	2.725
Lee	0.229	2.441
Foster	0.244	2.521
Wang	0.038	1.75
Shibayama	2	1.203
# Cit.	0.486	0.43
DI1	0.75	0.875
DI5	-1.44	-3.4
DI1nok	0.166	-11.75
DeIn	0.15	-4.187
Breadth	-0.33	0.33
Depth	-0.052	0.083

Notes: This table reports the turning points of the effect of cognitive diversity and average exploratory profiles on combinatorial novelty and scientific recognition in Table 4.2 and 4.5.

Regression without field-year weighting

Table 4.14: Combinatorial Novelty: cognitive diversity and average exploratory profile (References)

	Dependent variable:				
	Uzzi	Lee	Foster	Wang	Shibayama
	(1)	(2)	(3)	(4)	(5)
Author inter $_{abs}$	183.520***	4.377***	0.940***	3.061***	0.268***
	(23.173)	(0.204)	(0.033)	(0.252)	(0.008)
Author inter $_{abs}$ˆ2	-176.966***	-3.915***	-1.005***	-1.936***	-0.195***
	(31.732)	(0.270)	(0.043)	(0.335)	(0.012)
Author intra $_{abs}$	198.281***	3.825***	1.052***	0.095	0.226***
	(22.235)	(0.222)	(0.074)	(0.365)	(0.011)
Author intra $_{abs}$ˆ2	-403.151***	-8.090***	-2.057***	0.130	-0.153***
	(38.759)	(0.381)	(0.107)	(0.619)	(0.018)
# References	0.518***	0.009***	0.001***	0.076***	0.0004***
	(0.072)	(0.0004)	(0.00004)	(0.007)	(0.00002)
# Meshterms	1.287***	0.025***	0.003***	-0.043***	0.001***
	(0.119)	(0.002)	(0.0003)	(0.003)	(0.0001)
# Authors	1.371***	0.025***	0.004***	0.005	0.002***
	(0.113)	(0.001)	(0.0004)	(0.004)	(0.0001)
SJR	-1.151***	-0.020***	-0.011***	-0.093***	-0.002***
	(0.264)	(0.004)	(0.002)	(0.021)	(0.0003)
Year	Yes	Yes	Yes	Yes	Yes
Journal Cat.	Yes	Yes	Yes	Yes	Yes
Observations	1,647,446	1,815,631	1,815,631	1,815,631	1,809,185
R^2	0.020	0.168	0.151	0.158	0.253
Adjusted R^2	0.020	0.168	0.151	0.158	0.253
Residual Std. Error	192.756	1.258	0.235	4.341	0.066
F Statistic	139.319***	1,504.472***	1,328.955***	1,399.846***	2,523.818***

Notes: This table reports coefficients of the effect of cognitive diversity and average exploratory profile on combinatorial novelty using PKG. Standard errors are cluster robust at the journal level: ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively. The effects are estimated with an OLS. The constant term, scientific field (Scimago Journal Category), and time-fixed effects are incorporated in all model specifications.

Table 4.15: Combinatorial Novelty: cognitive diversity and average exploratory profile (Meshterms)

	Dependent variable:			
	Uzzi	Lee	Foster	Wang
	(1)	(2)	(3)	(4)
Author inter $_{abs}$	14.010***	1.829***	0.399***	0.951***
	(1.109)	(0.067)	(0.023)	(0.070)
Author inter $_{abs}$ˆ2	-16.049***	-2.002***	-0.495***	-0.915***
	(1.477)	(0.087)	(0.034)	(0.086)
Author intra $_{abs}$	11.644***	1.408***	0.405***	-0.177*
	(1.403)	(0.095)	(0.041)	(0.105)
Author intra $_{abs}$ˆ2	-28.595***	-2.603***	-1.066***	-0.578***
	(2.138)	(0.140)	(0.063)	(0.154)
# References	0.038***	0.002***	0.001***	0.001***
	(0.002)	(0.0001)	(0.00004)	(0.0001)
# Meshterms	-0.022*	0.028***	0.001***	0.058***
	(0.013)	(0.001)	(0.0003)	(0.001)
# Authors	0.012	0.011***	0.002***	-0.0001
	(0.008)	(0.001)	(0.0003)	(0.001)
SJR	-0.052	-0.011***	-0.002***	0.015***
	(0.032)	(0.002)	(0.001)	(0.003)
Year	Yes	Yes	Yes	Yes
Journal Cat.	Yes	Yes	Yes	Yes
Observations	661,832	1,823,889	1,823,889	1,823,889
R^2	0.064	0.179	0.120	0.174
Adjusted R^2	0.063	0.179	0.120	0.174
Residual Std. Error	7.929	0.536	0.206	0.716
F Statistic	193.801***	1,638.907***	1,020.027***	1,586.294***

Notes: This table reports coefficients of the effect of cognitive diversity and average exploratory profile on combinatorial novelty using PKG. Standard errors are cluster robust at the journal level: ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively. The effects are estimated with an OLS. The constant term, scientific field (Scimago Journal Category), and time-fixed effects are incorporated in all model specifications.

Table 4.16: Scientific recognition: cognitive diversity and average exploratory profile

	Dependent variable:						
	# cit.	DI1	DI5	DI1nok	DeIn	Breadth	Depth
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Author inter $_{abs}$	39.622***	-0.019***	-0.043***	0.640***	-3.961***	0.044*	-0.080***
	(12.175)	(0.006)	(0.008)	(0.050)	(0.306)	(0.023)	(0.024)
Author inter $_{abs}$^2	-55.784***	0.034***	0.068***	-0.485***	3.848***	-0.023	0.057*
	(15.267)	(0.008)	(0.010)	(0.066)	(0.379)	(0.032)	(0.033)
Author intra $_{abs}$	99.833***	-0.064***	-0.067***	-0.090	-2.059***	0.111***	-0.033
	(12.635)	(0.007)	(0.008)	(0.073)	(0.385)	(0.032)	(0.033)
Author intra $_{abs}$^2	-130.168***	0.069***	0.094***	0.138	2.499***	-0.021	-0.141***
	(16.543)	(0.010)	(0.012)	(0.105)	(0.541)	(0.047)	(0.049)
# References	0.681***	-0.0002***	-0.0004***	-0.003***	0.023***	-0.0002***	0.001***
	(0.027)	(0.00001)	(0.00002)	(0.0001)	(0.001)	(0.00004)	(0.0001)
# Meshterms	0.338***	-0.0004***	-0.001***	-0.006***	0.019***	-0.003***	0.005***
	(0.080)	(0.00004)	(0.0001)	(0.0005)	(0.002)	(0.0002)	(0.0003)
# Authors	3.405***	-0.0004***	-0.0002***	-0.009***	0.023***	-0.008***	0.010***
	(0.365)	(0.00003)	(0.00005)	(0.0005)	(0.002)	(0.0003)	(0.0004)
SJR	16.482***	-0.001***	0.001***	-0.013***	0.025***	-0.023***	0.025***
	(1.269)	(0.0001)	(0.0002)	(0.002)	(0.009)	(0.003)	(0.003)
Year	Yes	Yes	Yes	Yes	Yes	Yes	
Journal Cat.	Yes	Yes	Yes	Yes	Yes	Yes	
Observations	1,826,237	1,826,237	1,826,237	1,826,237	1,826,237	1,826,237	1,826,237
R^2	0.116	0.042	0.077	0.133	0.238	0.107	0.159
Adjusted R^2	0.116	0.042	0.077	0.133	0.238	0.107	0.158
Residual Std. Error	126.203	0.056	0.061	0.467	1.591	0.250	0.241
F Statistic	984.300***	328.932***	626.917***	1,151.082***	2,343.227***	904.045***	1,416.198***

Notes: This table reports coefficients of the effect of cognitive diversity and average exploratory profile on scientific recognition using PKG. Standard errors are cluster robust at the journal level: ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively. The effects are estimated with an OLS. The constant term, scientific field (Scimago Journal Category), and time-fixed effects are incorporated in all model specifications.

Table 4.17: Combinatorial Novelty: Cognitive diversity, highly exploratory and exploitative profile (References)

	Dependent variable:				
	Uzzi	Lee	Foster	Wang	Shibayama
	(1)	(2)	(3)	(4)	(5)
Author inter $_{abs}$	280.445***	6.383***	1.533***	1.696***	0.430***
	(30.366)	(0.225)	(0.066)	(0.347)	(0.011)
Author inter $_{abs}$ˆ2	-311.743***	-6.771***	-1.834***	-0.650	-0.350***
	(39.387)	(0.298)	(0.082)	(0.493)	(0.015)
Share exploratory	-22.978***	-0.416***	-0.087***	-0.328***	0.009***
	(3.028)	(0.029)	(0.005)	(0.034)	(0.001)
Share exploitative	8.714***	0.234***	0.048***	-0.468***	-0.021***
	(1.809)	(0.015)	(0.004)	(0.082)	(0.001)
Share exploratory * Share exploitative	29.023***	0.541***	0.186***	0.208	-0.052***
	(8.047)	(0.084)	(0.013)	(0.129)	(0.004)
# References	0.514***	0.009***	0.001***	0.076***	0.0004***
	(0.073)	(0.0004)	(0.00004)	(0.007)	(0.00002)
# Meshterms	1.292***	0.025***	0.004***	-0.042***	0.001***
	(0.119)	(0.002)	(0.0003)	(0.003)	(0.0001)
# Authors	1.472***	0.027***	0.005***	0.002	0.001***
	(0.118)	(0.001)	(0.0004)	(0.003)	(0.0001)
SJR	-1.206***	-0.022***	-0.011***	-0.090***	-0.002***
	(0.260)	(0.004)	(0.002)	(0.020)	(0.0003)
Year	Yes	Yes	Yes	Yes	Yes
Journal Cat.	Yes	Yes	Yes	Yes	Yes
Observations	1,647,446	1,815,631	1,815,631	1,815,631	1,809,185
R^2	0.020	0.167	0.150	0.158	0.252
Adjusted R^2	0.020	0.167	0.150	0.158	0.252
Residual Std. Error	192.763	1.258	0.235	4.340	0.067
F Statistic	138.223***	1,493.115***	1,310.908***	1,399.600***	2,503.790***

Notes: This table reports coefficients of the effect of cognitive diversity and highly exploratory and exploitative profiles on combinatorial novelty using PKG. Standard errors are cluster robust at the journal level: ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively. The effects are estimated with an OLS. The constant term, scientific field (Scimago Journal Category), and time-fixed effects are incorporated in all model specifications.

Table 4.18: Combinatorial Novelty: Cognitive diversity, highly exploratory and exploitative profile (Meshterms)

	Dependent variable:			
	Uzzi	Lee	Foster	Wang
	(1)	(2)	(3)	(4)
Author inter $_{abs}$	23.097***	2.721***	0.573***	0.687***
	(1.222)	(0.095)	(0.036)	(0.103)
Author inter $_{abs}$ˆ2	-28.625***	-3.128***	-0.798***	-0.852***
	(1.611)	(0.118)	(0.047)	(0.125)
Share exploratory	-0.852***	-0.091***	-0.068***	-0.073***
	(0.100)	(0.007)	(0.004)	(0.007)
Share exploitative	1.977***	0.078***	0.045***	0.057***
	(0.092)	(0.007)	(0.003)	(0.006)
Share exploratory * Share exploitative	1.251***	0.083***	0.130***	0.135***
	(0.393)	(0.023)	(0.010)	(0.025)
# References	0.038***	0.002***	0.001***	0.001***
	(0.001)	(0.0001)	(0.00004)	(0.0001)
# Meshterms	-0.022*	0.028***	0.001***	0.058***
	(0.013)	(0.001)	(0.0003)	(0.001)
# Authors	0.028***	0.012***	0.003***	0.001
	(0.008)	(0.001)	(0.0003)	(0.001)
SJR	-0.064**	-0.012***	-0.003***	0.015***
	(0.031)	(0.002)	(0.001)	(0.003)
Year	Yes	Yes	Yes	Yes
Journal Cat.	Yes	Yes	Yes	Yes
Observations	661,832	1,823,889	1,823,889	1,823,889
R^2	0.065	0.179	0.119	0.174
Adjusted R^2	0.065	0.179	0.119	0.174
Residual Std. Error	7.923	0.537	0.206	0.716
F Statistic	197.576***	1,631.871***	1,014.063***	1,574.819***

Notes: This table reports coefficients of the effect of cognitive diversity and highly exploratory and exploitative profiles on combinatorial novelty using PKG. Standard errors are cluster robust at the journal level: ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively. The effects are estimated with an OLS. The constant term, scientific field (Scimago Journal Category), and time-fixed effects are incorporated in all model specifications.

Table 4.19: Scientific recognition: cognitive diversity, highly exploratory and exploitative profile

	Dependent variable:						
	# cit.	DI1	DI5	DI1nok	DeIn	Breadth	Depth
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Author inter $_{abs}$	114.316***	-0.088***	-0.115***	0.313***	-4.628***	0.114***	-0.095***
	(13.961)	(0.009)	(0.011)	(0.084)	(0.556)	(0.030)	(0.033)
Author inter $_{abs}$^2	-122.966***	0.101***	0.143***	-0.193*	4.793***	-0.093**	0.041
	(16.903)	(0.011)	(0.013)	(0.103)	(0.669)	(0.039)	(0.043)
Share exploratory	-2.597**	-0.006***	-0.005***	-0.057***	0.176***	0.007**	-0.009***
	(1.100)	(0.001)	(0.001)	(0.005)	(0.018)	(0.003)	(0.003)
Share exploitative	3.759***	-0.005***	-0.008***	-0.085***	0.284***	-0.021***	0.027***
	(1.085)	(0.0005)	(0.001)	(0.004)	(0.015)	(0.002)	(0.002)
Share exploratory * Share exploitative	-23.770***	0.019***	0.014***	0.115***	-0.181***	0.031***	-0.017*
	(3.529)	(0.002)	(0.002)	(0.017)	(0.064)	(0.010)	(0.010)
# References	0.679***	-0.0002***	-0.0004***	-0.003***	0.023***	-0.0002***	0.001***
	(0.027)	(0.00001)	(0.00002)	(0.0001)	(0.001)	(0.00004)	(0.0001)
# Meshterms	0.337***	-0.0004***	-0.001***	-0.006***	0.019***	-0.003***	0.005***
	(0.080)	(0.00004)	(0.0001)	(0.0005)	(0.002)	(0.0002)	(0.0003)
# Authors	3.430***	-0.0004***	-0.0003***	-0.010***	0.025***	-0.009***	0.011***
	(0.364)	(0.00003)	(0.00005)	(0.0005)	(0.002)	(0.0003)	(0.0004)
SJR	16.427***	-0.001***	0.001***	-0.012***	0.024**	-0.023***	0.025***
	(1.268)	(0.0001)	(0.0002)	(0.002)	(0.009)	(0.003)	(0.003)
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Journal Cat.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,826,237	1,826,237	1,826,237	1,826,237	1,826,237	1,826,237	1,826,237
R^2	0.116	0.042	0.078	0.134	0.239	0.107	0.159
Adjusted R^2	0.116	0.042	0.078	0.134	0.239	0.107	0.159
Residual Std. Error	126.204	0.056	0.061	0.467	1.590	0.250	0.241
F Statistic	980.132***	328.824***	630.089***	1,160.693***	2,346.784***	901.029***	1,411.683***

Notes: This table reports coefficients of the effect of cognitive diversity and highly exploratory and exploitative profiles on scientific recognition using PKG. Standard errors are cluster robust at the journal level: ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively. The effects are estimated with an OLS. The constant term, scientific field (Scimago Journal Category), and time-fixed effects are incorporated in all model specifications.

General Conclusion

This thesis provides several insights concerning AI implications in science. It is organized into four chapters: the first two chapters focus on AI's impact on scientific discovery and the drivers of AI adoption for domain scientists. The last two chapters emphasize the relationship between team composition and scientific discovery. This conclusion briefly summarizes the chapters and proposes an extension of this work in light of recent changes in the landscape of artificial intelligence.

Chapter 1 examines the dissemination and impact of artificial intelligence, specifically neural networks, in science. We show that while neural network methods do not serve as an autopilot for knowledge navigation, they represent a powerful and versatile research tool that impacts knowledge creation in tangible ways. The chapter proposes that AI be considered an *emerging general method of invention* and provides a comprehensive perspective on AI's role in fostering knowledge creation.

Chapter 2 explores the drivers of AI adoption for domain scientists. It uses Scientific & Technical Human Capital (STHC) as a valuable framework for understanding these incentives, revealing that institutional factors and social environment composition are strongly related to AI adoption. Furthermore, we show the importance of collaborating with early-career researchers and that individuals with diverse backgrounds adopt AI easily. This perspective shifts the focus from aggregate trends in AI adoption to individual researchers' characteristics and their social and institutional environments, enabling a more nuanced understanding of how AI is integrated into research practices.

Chapter 3 introduces *Novelpy*, an open-source Python package designed to compute novelty and disruption indicators for scientific documents or patents. This tool offers the scientometrics community a centralized module to analyze and compare various measures of novelty and disruptiveness. The creation of this tool addresses a gap in the scientometrics community and sets the groundwork for future studies to explore the relationship between these indicators systematically.

Chapter 4 develops an indicator at the author level. It explores the relationship between the team's composition, the article's novelty and its scientific recognition. We highlight the essential role of the cognitive dimension in the creative process, as it significantly affects originality and success. The chapter emphasizes the importance of team composition in terms of cognitive profiles for scientific creativity, showing that teams that combine both highly exploratory and exploitative individuals are more able to disrupt science and propose more novel knowledge combinations.

However, this work can be further expanded, particularly in how AI involvement in research can be considered. This thesis views AI as a tool explicitly mentioned in scientific articles. As this technology spreads, its use will be less explicit, and it will become increasingly difficult to see the submerged part of the iceberg, i.e., the implicit use of AI. When we began this work, AI was still extremely task-specific, and our analyses showed that AI could be considered a new super microscope with the flexibility to analyze any content. Now, generative textual AI completely changes the game. GPT-like models are also trained on specific tasks, predicting word sequences in a text. However, to make this prediction, it is necessary to understand all the entities present in a text and be able to combine them. GPT-4 does not answer 2 to 1+1 only because it read it but because it understood the concept of counting and addition. This subtle difference is significant because it implies that the exploration of the entire knowledge space can be achieved within the same entity.

With the centralization of almost all information from the internet in a single model, capable of understanding its meaning and relation with other information, we witness a potential second shift in how AI will impact science. Today, discussions mainly revolve around considering AI as a potential author (see recent debates on *Nature*). Still, this method allows for much more than simply transcription of ideas into academic language. It truly enables a "Human + Machine" experience for researchers, as they can use this chatbot during the conceptualization and data-processing phases, but also to suggest articles suitable for a research question, and point out inconsistencies and shortcomings in a text. It opens the door to a more systematic exploration of the knowledge space in an informal manner for users since it takes place through interposed messages. In fact, GPT-4 is the AI entity the closer to the competent colleague you meet at the coffee machine, with whom a quick informal exchange on a given problem can quickly solve it. These interactions with these hypothetical colleagues will mostly be mentioned in the acknowledgements of an article; their contribution to the knowledge-creation process remains almost invisible. The same goes for generative AI. Today, individuals using generative AI in their research do not explicitly mention it, which makes it increasingly difficult to understand how it can modify research. Therefore it would seem necessary to consider a qualitative approach to this phenomenon. While quantitative methods offer large-scale results, the lack of detailed personal narratives or experiences deprives us of an in-depth understanding of the use of AI, the actors involved, and its precise role in the knowledge-creation process.

Conclusion Générale

Cette thèse fournit plusieurs enseignement sur les implications de l'utilisation de l'IA dans la science. Elle est organisée en quatre chapitres : les deux premiers se concentrent sur l'impact de l'IA sur la découverte scientifique et les moteurs de l'adoption de l'IA pour les scientifiques de domaine d'application. Les deux derniers chapitres de la thèse met en avant la relation entre la composition de l'équipe et la découverte scientifique. Cette conclusion résume brièvement les chapitres et propose une extension de ce travail au vu des changements récents dans le secteur de l'intelligence artificielle.

Le chapitre 1 donne un aperçu de la diffusion et de l'impact de l'intelligence artificielle, en particulier des réseaux neuronaux, dans la science. Nous montrons que si les réseaux neuronaux ne constituent pas des systèmes capable naviguer automatiquement dans un espace de connaissances, ils représentent néanmoins un outil de recherche puissant et polyvalent qui a un impact tangible sur la création de connaissances. Le chapitre suggère que l'IA doit être considérée comme une *méthode générale d'invention émergente* et fournit une perspective globale sur le rôle de l'IA dans la stimulation de la création de connaissances.

Le chapitre 2 explore les moteurs de l'adoption de l'IA par les scientifiques de domaine d'application. Il fait appel au capital humain scientifique et technique (STHC) comme cadre de référence pour comprendre ces incitations, et révèle que les facteurs institutionnels et la composition de l'environnement social sont étroitement liés à l'adoption de l'IA. Nous montrons l'importance des chercheurs en début de carrière et des personnes ayant des expériences diverses pour favoriser l'adoption et la réutilisation de l'IA. Cette perspective met l'accent non plus sur les tendances globales de l'adoption de l'IA, mais sur les caractéristiques individuelles des chercheurs et sur leur environnement social et institutionnel. Cela permet une compréhension plus nuancée de la manière dont l'IA est intégrée dans les pratiques scientifiques.

Le chapitre 3 présente *Novelpy*, un module Python open-source conçu pour calculer des indicateurs de nouveauté et de disruption sur des documents scientifiques ou des brevets. Cet outil offre à la communauté scientométrique un moyen centralisé d'analyser et de comparer systématiquement ces mesures. La création de cet outil comble une lacune dans la communauté scientométrique et jette les bases de futures études visant à explorer la relation entre ces indicateurs de manière systématique.

Le chapitre 4 développe un indicateur au niveau auteur. Il explore la relation entre la composition de l'équipe, la nouveauté d'un article et sa reconnaissance scientifique. Nous montrons le rôle essentiel de la dimension cognitive dans le processus créatif, dans la mesure où elle influe considérablement sur l'originalité et la réussite. Le chapitre souligne l'importance de la composition de l'équipe en termes de profils cognitifs pour la créativité scientifique, en montrant que les équipes qui combinent des individus hautement exploratifs et exploitatifs sont plus à même de bouleverser la science et de proposer des combinaisons de connaissances plus nouvelles.

Toutefois, ce travail peut être développé davantage, en particulier en ce qui concerne la manière dont l'implication de l'IA dans la recherche est envisagée. Cette thèse considère l'IA comme un outil nécessairement mentionné explicitement dans un article scientifique. Au fur et à mesure que cette technologie se répandra, son utilisation sera moins explicite et il deviendra de plus en plus difficile de voir la partie immergée de l'iceberg, c'est-à-dire l'utilisation implicite de l'IA. Lorsque nous avons commencé ce travail, l'IA restait extrêmement ciblée, et nos analyses ont montré que l'IA pouvait être considérée comme un nouveau super microscope ayant la flexibilité d'analyser n'importe quel contenu. Aujourd'hui, l'IA textuelle générative change complètement la donne. Les modèles de type GPT sont également formés à des tâches spécifiques, en prédisant les mots suivants dans un texte. Cependant, pour faire cette prédiction, il est nécessaire de comprendre toutes les entités présentes dans un texte et d'être capable de les combiner. GPT-4 ne répond pas 2 à 1+1 uniquement parce qu'il l'a lu, mais parce qu'il a compris le concept de dénombrement et d'addition. Cette différence subtile est considérable car elle implique que l'exploration de l'ensemble de l'espace de connaissances peut être réalisée au sein d'une même entité.

Avec la centralisation de la quasi-totalité des informations provenant d'Internet dans un modèle unique, capable de comprendre leur signification et leur relation avec d'autres informations, nous assistons à un second changement potentiel dans la manière dont l'IA influencera la science. Aujourd'hui, les discussions tournent principalement autour du fait de considérer l'IA comme un auteur potentiel (voir les récents débats sur Nature). Pourtant, cette méthode permet bien plus qu'une simple retranscription d'idées dans un langage académique. En effet, les chercheurs peuvent utiliser ce chatbot pendant les phases de conceptualisation et de traitement des données, se faire suggérer des articles adaptés à une question de recherche et signaler les incohérences et les lacunes dans un texte. Cela ouvre la porte à une exploration plus systématique de l'espace de connaissances de manière informelle, puisque cela se fait par messages interposés. En fait, GPT-4 est l'entité IA la plus proche du collègue compétent que l'on rencontre à la machine à café, avec lequel un échange informel rapide sur un problème donné permet de le résoudre rapidement. Ces interactions avec ces collègues hypothétiques seront le plus souvent mentionnées dans les remerciements d'un article ; leur contribution au processus de création de connaissances reste presque invisible. Il en va de même pour l'IA générative. Aujourd'hui, les personnes qui utilisent l'IA générative dans leur recherche ne la mentionnent pas explicitement, ce qui rend de plus en plus difficile la compréhension de la manière dont elle peut modifier la recherche. Il semble donc nécessaire d'envisager une approche qualitative de ce phénomène. Bien que les méthodes quantitatives offrent des résultats à grande échelle, l'absence de récits ou d'expériences personnelles détaillés nous prive d'une compréhension approfondie de l'utilisation de l'IA, des acteurs impliqués et de son rôle précis dans le processus de création de connaissances.

Bibliography

G. Abramo, C. A. D'Angelo, and L. Zhang. A comparison of two approaches for measuring interdisciplinary research output: The disciplinary diversity of authors vs the disciplinary diversity of the reference list. *Journal of Informetrics*, 12(4): 1182–1193, 2018.

H. A. Abt and E. Garfield. Is the relationship between numbers of references and paper lengths the same for all sciences? *Journal of the American Society for Information Science and Technology*, 53(13):1106–1112, 2002.

D. Acemoglu and P. Restrepo. Robots and jobs: Evidence from us labor markets. *Journal of Political Economy*, 128(6):2188–2244, 2020.

J. Adams. The fourth age of research. *Nature*, 497(7451):557–560, 2013.

R. Adner and D. A. Levinthal. The emergence of emerging technologies. *California Management Review*, 45(1):50–66, 2002.

P. Aghion, B. F. Jones, and C. I. Jones. Artificial intelligence and economic growth. In *The economics of artificial intelligence: An agenda*, pages 237–282. University of Chicago Press, 2018.

A. Agrawal, J. McHale, and A. Oettl. Finding needles in haystacks: Artificial intelligence and recombinant growth. In *The economics of artificial intelligence: An agenda*, pages 149–174. University of Chicago Press, 2018.

N. Ahmed and M. Wahed. The de-democratization of ai: Deep learning and the compute divide in artificial intelligence research. *arXiv Preprint arXiv:2010.15581*, 2020.

B. Alberts, M. W. Kirschner, S. Tilghman, and H. Varmus. Rescuing us biomedical research from its systemic flaws. *Proceedings of the National Academy of Sciences*, 111(16):5773–5777, 2014.

M. Amendola, J.-L. Gaffard, et al. Novelty, hysteresis, and growth. Technical report, Observatoire Francais des Conjonctures Economiques (OFCE), 2014.

D. Amodei and D. Hernandez. Review of ai and compute, 2018. URL `https://openai.com/blog/ai-and-compute/`. OpenAI Blog.

C. Angermueller, T. Pärnamaa, L. Parts, and O. Stegle. Deep learning for computational biology. *Molecular Systems Biology*, 12(7):878, 2016.

A. Annoni, P. Benczur, P. Bertoldi, B. Delipetrev, G. De Prato, C. Feijoo, E. F. Macias, E. G. Gutierrez, M. I. Portela, H. Junklewitz, et al. Artificial intelligence: A european perspective. Technical report, Joint Research Centre (Seville site), 2018.

D. Archibugi. Choosing your mentor: A letter to creative minds. *Journal of Innovation Economics & Management*, pages I99–13, 2021.

L. Argote, S. L. Beckman, and D. Epple. The persistence and transfer of learning in industrial settings. *Management Science*, 36(2):140–154, 1990.

M. Aria and C. Cuccurullo. Bibliometrix: An r tool for comprehensive analysis of scientific literature. *Journal of Informetrics*, 11:959–975, 2017.

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

S. Arts, J. Hou, and J. C. Gomez. Natural language processing to identify the creation and impact of new technologies in patent text: Code, data, and new measures. *Research Policy*, 50(2):104144, 2021.

C. Ayoubi, M. Pezzoni, and F. Visentin. At the origins of learning: Absorbing knowledge flows from within the team. *Journal of Economic Behavior & Organization*, 134:374–387, 2017.

C. Ayoubi, M. Pezzoni, and F. Visentin. Does it pay to do novel science? the selectivity patterns in science funding. *Science and Public Policy*, 48(5):635–648, 2021.

P. Azoulay. Small research teams 'disrupt'science more radically than large ones, 2019.

P. Azoulay, J. S. Graff Zivin, and G. Manso. Incentives and creativity: evidence from the academic life sciences. *The RAND Journal of Economics*, 42(3):527–554, 2011.

R. Baradaran, R. Ghiasi, and H. Amirkhani. A survey on machine reading comprehension systems. *Natural Language Engineering*, 28(6):683–732, 2022.

S. Baruffaldi, B. van Beuzekom, H. Dernis, D. Harhoff, N. Rao, D. Rosenfeld, and M. Squicciarini. Identifying and measuring developments in artificial intelligence. 2020.

D. J. Berndt and J. Clifford. Using dynamic time warping to find patterns in time series. In *KDD Workshop*, volume 10, pages 359–370. Seattle, WA, USA:, 1994.

S. Bianchini, M. Müller, and P. Pelletier. Artificial intelligence in science: An emerging general method of invention. *Research Policy*, 51(10):104604, 2022.

M. Bikard, F. Murray, and J. S. Gans. Exploring trade-offs in the organization of scientific work: Collaboration and scientific reward. *Management Science*, 61(7): 1473–1495, 2015.

N. Bloom, C. I. Jones, J. Van Reenen, and M. Webb. Are ideas getting harder to find? *American Economic Review*, 110(4):1104–44, 2020.

M. A. Boden. Creativity and knowledge. *Creativity in Education*, pages 95–102, 2001.

M. A. Boden. *The creative mind: Myths and mechanisms*. Psychology Press, 2004.

M. A. Boden. Computer models of creativity. *AI Magazine*, 30(3):23–23, 2009.

E. Bonabeau. Decisions 2.0: The power of collective intelligence. *MIT Sloan Management Review*, 50(2):45, 2009.

R. Bonney, J. L. Shirk, T. B. Phillips, A. Wiggins, H. L. Ballard, A. J. Miller-Rushing, and J. K. Parrish. Next steps for citizen science. *Science*, 343(6178): 1436–1437, 2014.

L. Bornmann and H.-D. Daniel. What do citation counts measure? a review of studies on citing behavior. *Journal of Documentation*, 2008.

L. Bornmann, S. Devarakonda, A. Tekles, and G. Chacko. Do disruption index indicators measure what they propose to measure? the comparison of several indicator variants with assessments by peers. retrieved december 6, 2019, 2019a.

L. Bornmann, A. Tekles, H. H. Zhang, and Y. Y. Fred. Do we measure novelty when we analyze unusual combinations of cited references? a validation study of bibliometric novelty indicators based on f1000prime data. *Journal of Informetrics*, 13(4):100979, 2019b.

L. Bornmann, S. Devarakonda, A. Tekles, and G. Chacko. Disruptive papers published in scientometrics: meaningful results by using an improved variant of the disruption index originally proposed by wu, wang, and evans (2019). *Scientometrics*, pages 1–7, 2020.

K. J. Boudreau, E. C. Guinan, K. R. Lakhani, and C. Riedl. Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. *Management Science*, 62(10):2765–2783, 2016.

P. Bourdieu. The specificity of the scientific field and the social conditions of the progress of reason. *Social Science Information*, 14(6):19–47, 1975.

K. W. Boyack, R. Klavans, and K. Börner. Mapping the backbone of science. *Scientometrics*, 64(3):351–374, 2005.

B. Bozeman and E. Corley. Scientists' collaboration strategies: implications for scientific and technical human capital. *Research Policy*, 33(4):599–616, 2004.

B. Bozeman, J. S. Dietz, and M. Gaughan. Scientific and technical human capital: an alternative model for research evaluation. *International Journal of Technology Management*, 22(7-8):716–740, 2001.

T. F. Bresnahan and M. Trajtenberg. General purpose technologies 'engines of growth'? *Journal of Econometrics*, 65(1):83–108, 1995.

H. Brooks. The relationship between science and technology. *Research Policy*, 23(5): 477–486, 1994.

E. Brynjolfsson and A. McAfee. *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & Company, 2014.

E. Brynjolfsson and T. Mitchell. What can machine learning do? workforce implications. *Science*, 358(6370):1530–1534, 2017.

Y. Bu, L. Waltman, and Y. Huang. A multi-dimensional framework for characterizing the citation impact of scientific publications. *arXiv Preprint arXiv:1901.09663*, 2019.

R. S. Burt. Structural holes and good ideas. *American Journal of Sociology*, 110(2): 349–399, 2004.

A. Cabrera and E. F. Cabrera. Knowledge-sharing dilemmas. *Organization Studies*, 23(5):687–710, 2002.

N. Carayol and J.-M. Dalle. Sequential problem choice and the reward system in open science. *Structural Change and Economic Dynamics*, 18(2):167–191, 2007.

N. Carayol, M. Lanoë, et al. The impact of project-based funding in science: Lessons from the anr experience. Technical report, Groupe de Recherche en Economie Théorique et Appliquée (GREThA), 2017.

N. Carayol, L. Agenor, and L. Oscar. The right job and the job right: Novelty, impact and journal stratification in science. *Impact and Journal Stratification in Science (March 5, 2019)*, 2019.

P. Chellappandi and C. Vijayakumar. Bibliometrics, scientometrics, webometrics/-cybermetrics, informetrics and altmetrics–an emerging field in library and information science research. *Shanlax International Journal of Education*, 7(1):5–8, 2018.

C. M. Christensen and C. M. Christensen. *The innovator's dilemma: The revolutionary book that will change the way you do business*. HarperBusiness Essentials New York, NY, 2003.

I. M. Cockburn, R. Henderson, and S. Stern. The impact of artificial intelligence on innovation. *National Bureau of Economic Research*, No. w24449, 2018.

J. S. Coleman. Social capital in the creation of human capital. *American Journal of Sociology*, 94:S95–S120, 1988.

R. Costas and T. Franssen. Reflections around 'the cautionary use' of the h-index: response to teixeira da silva and dobránszki. *Scientometrics*, 115(2):1125–1130, 2018.

J. N. Cummings. Work groups, structural diversity, and knowledge sharing in a global organization. *Management Science*, 50(3):352–364, 2004.

K. B. Dahlin and D. M. Behrens. When is an invention really radical?: Defining and measuring technological radicalness. *Research Policy*, 34(5):717–737, 2005.

P. R. Daugherty and H. J. Wilson. *Human + machine: Reimagining work in the age of AI.* Harvard Business Press, 2018.

P. A. David. The dynamo and the computer: an historical perspective on the modern productivity paradox. *American Economic Review*, 80(2):355–361, 1990.

D. J. de Solla Price. Little science, big science – and beyond. 1963.

D. J. de Solla Price. Networks of scientific papers. *Science*, 149(3683):510–515, 1965.

J. Degrave, F. Felici, J. Buchli, M. Neunert, B. Tracey, F. Carpanese, T. Ewalds, R. Hafner, A. Abdolmaleki, D. de Las Casas, et al. Magnetic control of tokamak plasmas through deep reinforcement learning. *Nature*, 602(7897):414–419, 2022.

P. J. DiMaggio and W. W. Powell. The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, pages 147–160, 1983.

T. Eloundou, S. Manning, P. Mishkin, and D. Rock. Gpts are gpts: An early look at the labor market impact potential of large language models. *arXiv Preprint arXiv:2303.10130*, 2023.

T. C. Erren, D. M. Shaw, and P. Lewis. Small groups, open doors: Fostering individual and group creativity within research communities, 2017.

European-Commission. On artificial intelligence - a european approach to excellence and trust. White Paper COM(2020) 65 final, Brussels, 2020.

H. J. Falk-Krzesinski, N. Contractor, S. M. Fiore, K. L. Hall, C. Kane, J. Keyton, J. T. Klein, B. Spring, D. Stokols, and W. Trochim. Mapping a research agenda for the science of team science. *Research Evaluation*, 20(2):145–158, 2011.

R. Fasoulis, K. Bougiatiotis, F. Aisopos, A. Nentidis, and G. Paliouras. Error detection in knowledge graphs: Path ranking, embeddings or both. *arXiv Preprint arXiv:2002.08762*, 2020.

L. Fleck. *Genesis and development of a scientific fact*. University of Chicago Press, 2012.

L. Fleming. Recombinant uncertainty in technological search. *Management Science*, 47(1):117–132, 2001.

L. Fleming, S. Mingo, and D. Chen. Collaborative brokerage, generative creativity, and creative success. *Administrative Science Quarterly*, 52(3):443–475, 2007.

M. Fontana, M. Iori, F. Montobbio, and R. Sinatra. New and atypical combinations: An assessment of novelty and interdisciplinarity. *Research Policy*, 49(7):104063, 2020.

S. Fortunato, C. T. Bergstrom, K. Börner, J. A. Evans, D. Helbing, S. Milojević, A. M. Petersen, F. Radicchi, R. Sinatra, B. Uzzi, et al. Science of science. *Science*, 359(6379), 2018.

J. G. Foster, A. Rzhetsky, and J. A. Evans. Tradition and innovation in scientists' research strategies. *American Sociological Review*, 80(5):875–908, 2015.

J. G. Foster, F. Shi, and J. Evans. Surprise! measuring novelty as expectation violation. 2021.

M. F. Fox. Gender, environmental milieu, and productivity in science. 1991.

C. Franzoni, P. Stephan, and R. Veugelers. Funding risky research. *Entrepreneurship and Innovation Policy and the Economy*, 1(1):103–133, 2022.

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

J. Furman and R. Seamans. Ai and the economy. *Innovation Policy and the Economy*, 19(1):161–191, 2019.

J. L. Furman and F. Teodoridis. Automation, research technology, and researchers' trajectories: Evidence from computer science and electrical engineering. *Organization Science*, 31(2):330–354, 2020.

A. Geuna. *The economics of knowledge production: funding and the structure of university research*. Edward Elgar Publishing, 1999.

W. Glänzel and A. Schubert. Double effort= double impact? a critical view at international co-authorship in chemistry. *Scientometrics*, 50(2):199–214, 2001.

B. González-Pereira, V. Guerrero-Bote, and F. Moya-Anegon. The sjr indicator: A new indicator of journals' scientific prestige. *arXiv Preprint arXiv:0912.4141*, 2009.

I. Goodfellow, Y. Bengio, and A. Courville. *Deep learning*. MIT press, 2016.

C. A. E. Goodhart. *Monetary theory and practice: The UK experience*. Springer, 1984.

R. J. Gordon. Perspectives on the rise and fall of american growth. *American Economic Review*, 106(5):72–76, 2016.

G. Graetz and G. Michaels. Robots at work. *Review of Economics and Statistics*, 100(5):753–768, 2018.

Z. Griliches. Hybrid corn: An exploration in the economics of technological change. *Econometrica, Journal of the Econometric Society*, pages 501–522, 1957.

A. K. Gupta, K. G. Smith, and C. E. Shalley. The interplay between exploration and exploitation. *Academy of Management Journal*, 49(4):693–706, 2006.

D. Hain, R. Jurowetzki, T. Buchmann, and P. Wolf. Text-based technological signatures and similarities: how to create them and what to do with them. *arXiv Preprint arXiv:2003.12303*, 2020.

D. Hassabis, D. Kumaran, C. Summerfield, and M. Botvinick. Neuroscience-inspired artificial intelligence. *Neuron*, 95(2):245–258, 2017.

K. He, X. Zhang, S. Ren, and J. Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*, pages 1026–1034, 2015.

S. He, Y. Li, Y. Feng, S. Ho, S. Ravanbakhsh, W. Chen, and B. Póczos. Learning to predict the cosmological structure formation. *Proceedings of the National Academy of Sciences*, 116(28):13825–13832, 2019.

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

S. Hemlin, C. M. Allwood, B. R. Martin, and M. D. Mumford. Why is leadership important for creativity in science, technology, and innovation? introduction. In *Creativity and leadership in science, technology, and innovation*, pages 1–26. Routledge, 2013.

R. M. Henderson and K. B. Clark. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly*, pages 9–30, 1990.

A. J. Hey, S. Tansley, K. M. Tolle, et al. *The fourth paradigm: data-intensive scientific discovery*, volume 1. Microsoft research Redmond, WA, 2009.

B. Hofstra, V. V. Kulkarni, S. Munoz-Najar Galvez, B. He, D. Jurafsky, and D. A. McFarland. The diversity–innovation paradox in science. *Proceedings of the National Academy of Sciences*, 117(17):9284–9291, 2020.

F. Jacob. Evolution and tinkering. *Science*, 196(4295):1161–1166, 1977.

L. B. Jeppesen and K. R. Lakhani. Marginality and problem-solving effectiveness in broadcast search. *Organization Science*, 21(5):1016–1033, 2010.

B. F. Jones. The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *The Review of Economic Studies*, 76(1):283–317, 2009.

B. F. Jones, S. Wuchty, and B. Uzzi. Multi-university research teams: Shifting impact, geography, and stratification in science. *Science*, 322(5905):1259–1262, 2008.

J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021.

Q. Ke, E. Ferrara, F. Radicchi, and A. Flammini. Defining and identifying sleeping beauties in science. *Proceedings of the National Academy of Sciences*, 112(24): 7426–7431, 2015.

S. Kelchtermans, D. Neicu, and R. Veugelers. Off the beaten path: What drives scientists' entry into new fields? *FEB Research Report MSI_2008*, pages 1–42, 2020.

R. D. King, J. Rowland, S. G. Oliver, M. Young, W. Aubrey, E. Byrne, M. Liakata, M. Markham, P. Pir, L. N. Soldatova, et al. The automation of science. *Science*, 324(5923):85–89, 2009.

J. Klinger, J. Mateos-Garcia, and K. Stathoulopoulos. A narrowing of ai research? *arXiv Preprint arXiv:2009.10385*, 2020.

J. Klinger, J. Mateos-Garcia, and K. Stathoulopoulos. Deep learning, deep change? mapping the evolution and geography of a general purpose technology. *Scientometrics*, 126:5589–5621, 2021.

T. Koopmann, M. Stubbemann, M. Kapa, M. Paris, G. Buenstorf, T. Hanika, A. Hotho, R. Jäschke, and G. Stumme. Proximity dimensions and the emergence of collaboration: a hyptrails study on german ai research. *Scientometrics*, pages 1–22, 2021.

A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, 2017.

T. S. Kuhn. *The structure of scientific revolutions*. University of Chicago press, 1962.

S. Kumar. Co-authorship networks: a review of the literature. *Aslib Journal of Information Management*, 2015.

V. Kumar, S. Sendhilkumar, and G. Mahalakshmi. Author similarity identification using citation context and proximity. In *2017 Second International Conference on Recent Trends and Challenges in Computational Models (ICRTCCM)*, pages 217–221. IEEE, 2017.

B. M. Lake, T. D. Ullman, J. B. Tenenbaum, and S. J. Gershman. Building machines that learn and think like people. *Behavioral and Brain Sciences*, 40:e253, 2017.

K. R. Lakhani, K. J. Boudreau, P.-R. Loh, L. Backstrom, C. Baldwin, E. Lonstein, M. Lydon, A. MacCormack, R. A. Arnaout, and E. C. Guinan. Prize-based contests can provide solutions to computational biology problems. *Nature Biotechnology*, 31(2):108–111, 2013.

J. Lanier. *You are not a gadget: A manifesto*. Vintage, 2011.

Q. V. Le. Building high-level features using large scale unsupervised learning. In *2013 IEEE international conference on acoustics, speech and signal processing*, pages 8595–8598. IEEE, 2013.

E. Leahey. From sole investigator to team scientist: Trends in the practice and study of research collaboration. *Annual Review of Sociology*, 42:81–100, 2016.

Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.

Y.-N. Lee, J. P. Walsh, and J. Wang. Creativity in scientific teams: Unpacking novelty and impact. *Research Policy*, 44(3):684–697, 2015.

M. K. Leung, A. Delong, B. Alipanahi, and B. J. Frey. Machine learning in genomic medicine: a review of computational problems and data sets. *Proceedings of the IEEE*, 104(1):176–197, 2015.

C. H. Liao. How to improve research quality? examining the impacts of collaboration intensity and member diversity in collaboration networks. *Scientometrics*, 86(3): 747–761, 2011.

Y. Lin. Where do new ideas come from: New directions in science emerge from disconnection and discord. Technical report, University of Chicago, 2021.

Y. Lin, J. Evans, and L. Wu. Novelty, disruption, and the evolution of scientific impact. 2021.

G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, and C. I. Sánchez. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42:60–88, 2017.

G. Loewenstein. The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, 116(1):75, 1994.

J. Ma, R. P. Sheridan, A. Liaw, G. E. Dahl, and V. Svetnik. Deep neural nets as a method for quantitative structure–activity relationships. *Journal of Chemical Information and Modeling*, 55(2):263–274, 2015.

J. Mairesse, M. Pezzoni, et al. The impact of novelty in scientific articles: The case of french physicists. *Revue d'Économie Industrielle*, 174(2e), 2021.

A. H. Marblestone, G. Wayne, and K. P. Kording. Toward an integration of deep learning and neuroscience. *Frontiers in Computational Neuroscience*, page 94, 2016.

J. G. March. Exploration and exploitation in organizational learning. *Organization Science*, 2(1):71–87, 1991.

V. Marx. The big challenges of big data. *Nature*, 498(7453):255–260, 2013.

K. Matsumoto, S. Shibayama, B. Kang, and M. Igami. Introducing a novelty indicator for scientific research: validating the knowledge-based combinatorial approach. *Scientometrics*, 2021.

J. McCarthy, M. Minsky, N. Rochester, and C. Shannon. Proposal for the 1956 dartmouth summer research project on artificial intelligence. *Dartmouth College, Hanover, NH, USA*, 1955.

M. A. McFadyen and A. A. Cannella Jr. Social capital and knowledge creation: Diminishing returns of the number and strength of exchange relationships. *Academy of Management Journal*, 47(5):735–746, 2004.

J. McLevey and R. McIlroy-Young. Introducing metaknowledge: Software for computational research in information science, network analysis, and science of science. *Journal of Informetrics*, 11(1):176–197, 2017.

D. L. Medin and C. D. Lee. Diversity makes better science. *APS Observer*, 25, 2012.

R. K. Merton. Priorities in scientific discovery: a chapter in the sociology of science. *American Sociological Review*, 22(6):635–659, 1957.

T. Mikolov, K. Chen, G. Corrado, and J. Dean. Efficient estimation of word representations in vector space. *arXiv Preprint arXiv:1301.3781*, 2013.

T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems*, pages 3111–3119, 2013b.

S. Milojević. Quantifying the cognitive extent of science. *Journal of Informetrics*, 9 (4):962–973, 2015.

J. Mingers and L. Leydesdorff. A review of theory and practice in scientometrics. *European Journal of Operational Research*, 246(1):1–19, 2015.

R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley. Deep learning for healthcare: review, opportunities and challenges. *Briefings in Bioinformatics*, 19(6): 1236–1246, 2018.

M. S. Mizruchi and L. C. Fein. The social construction of organizational knowledge: A study of the uses of coercive, mimetic, and normative isomorphism. *Administrative Science Quarterly*, 44(4):653–683, 1999.

J. A. Moral Muñoz, E. Herrera Viedma, A. Santisteban Espejo, M. J. Cobo, et al. Software tools for conducting bibliometric analysis in science: An up-to-date review. 2020.

B. Mueller. The building blocks of creativity and new ideas. *RAUSP Management Journal*, 54:242–246, 2019.

J. Nahapiet and S. Ghoshal. Social capital, intellectual capital, and the organizational advantage. *Academy of Management Review*, 23(2):242–266, 1998.

Nature-Editorial. The scientific events that shaped the decade. *Nature*, 576:337–338, 2019.

R. R. Nelson. *An evolutionary theory of economic change*. Harvard University Press, 1985.

N. J. Nilsson. *The quest for artificial intelligence*. Cambridge University Press, 2009.

I. Nonaka. A dynamic theory of organizational knowledge creation. *Organization Science*, 5(1):14–37, 1994.

B. Nooteboom. *Learning and innovation in organizations and economies*. OUP Oxford, 2000.

B. Nooteboom, W. Van Haverbeke, G. Duysters, V. Gilsing, and A. Van den Oord. Optimal cognitive distance and absorptive capacity. *Research Policy*, 36(7):1016–1034, 2007.

E. Noyons, H. Moed, and A. Van Raan. Integrating research performance analysis and science mapping. *Scientometrics*, 46(3):591–604, 1999.

OECD. Artificial intelligence in society. *OECD Publishing*, 2019.

OECD. Effective policies to foster high-risk/high-reward research. (112), 2021. doi: https://doi.org/https://doi.org/10.1787/06913b3b-en. URL `https://www.oecd-ilibrary.org/content/paper/06913b3b-en`.

OECD. A blueprint for building national compute capacity for artificial intelligence. (350), 2023. doi: https://doi.org/https://doi.org/10.1787/876367e3-en. URL `https://www.oecd-ilibrary.org/content/paper/876367e3-en`.

I. Olier, O. I. Orhobor, T. Dash, A. M. Davis, L. N. Soldatova, J. Vanschoren, and R. D. King. Transformational machine learning: Learning how to learn from many related scientific problems. *Proceedings of the National Academy of Sciences*, 118 (49):e2108013118, 2021.

C. O'neil. *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown, 2017.

P. Otlet. *Traité de documentation: le livre sur le livre, theéorie et pratique*. Editiones Mundaneum, 1934.

S. Page. The difference. In *The Difference*. Princeton University Press, 2008.

P. B. Paulus and B. A. Nijstad. *Group creativity: Innovation through collaboration*. Oxford University Press, 2003.

P. B. Paulus, N. W. Kohn, L. E. Arditti, and R. M. Korde. Understanding the group size effect in electronic brainstorming. *Small Group Research*, 44(3):332–352, 2013.

P. Pelletier and K. Wirtz. Novelpy: A python package to measure novelty and disruptiveness of bibliometric and patent data. *arXiv Preprint arXiv:2211.10346*, 2022.

A. M. Petersen, M. E. Ahmed, and I. Pavlidis. Grand challenges and emergent modes of convergence science. *Humanities and Social Sciences Communications*, 8(1):1–15, 2021.

J. M. Phillips and B. Malone. Increasing racial/ethnic diversity in nursing to reduce health disparities and achieve health equity. *Public Health Reports*, 129(1_suppl2): 45–50, 2014.

H. Poincaré. Mathematical creation. *The Monist*, pages 321–335, 1910.

N. Priaulx and M. Weinel. Connective knowledge: what we need to know about other fields to 'envision'cross-disciplinary collaboration. *European Journal of Futures Research*, 6(1):1–18, 2018.

J. Priem, H. Piwowar, and R. Orr. Openalex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. *arXiv preprint arXiv:2205.01833*, 2022.

A. Pritchard et al. Statistical bibliography or bibliometrics. *Journal of Documentation*, 25(4):348–349, 1969.

G. Pruschak and C. Hopp. And the credit goes to…-ghost and honorary authorship among social scientists. *Plos One*, 17(5):e0267312, 2022.

A. Purkayastha, E. Palmaro, H. J. Falk-Krzesinski, and J. Baas. Comparison of two article-level, field-independent citation metrics: Field-weighted citation impact (fwci) and relative citation ratio (rcr). *Journal of Informetrics*, 13(2):635–642, 2019.

M. Raghu and E. Schmidt. A survey of deep learning for scientific discovery. *arXiv Preprint arXiv:2003.11755*, 2020.

W. Raghupathi and V. Raghupathi. Big data analytics in healthcare: promise and potential. *Health Information Science and Systems*, 2:1–10, 2014.

A. Rajkomar, E. Oren, K. Chen, H. N. Dai AM, M. Hardt, P. Liu, X. Liu, J. Marcus, M. Sun, P. Sundberg, et al. Scalable and accurate deep learning with electronic health records. npj dig med 1 (1), 2018.

K. Ram and K. Broman. arxiv: Interface to the arxiv api. *https://CRAN.R-project.org/package=aRxiv*, 2019.

D. Ravì, C. Wong, F. Deligianni, M. Berthelot, J. Andreu-Perez, B. Lo, and G.-Z. Yang. Deep learning for health informatics. *IEEE journal of Biomedical and Health Informatics*, 21(1):4–21, 2016.

C. M. Rawlings and D. A. McFarland. Influence flows in the academy: Using affiliation networks to assess peer effects among researchers. *Social Science Research*, 40(3):1001–1017, 2011.

A. Rodríguez-Navarro. Research assessment based on infrequent achievements: A comparison of the united states and europe in terms of highly cited papers and nobel prizes. *Journal of the Association for Information Science and Technology*, 67(3):731–740, 2016.

N. Rosenberg. Scientific instrumentation and university research. *Research Policy*, 21(4):381–390, 1992.

D. Rotolo, D. Hicks, and B. R. Martin. What is an emerging technology? *Research Policy*, 44(10):1827–1843, 2015.

J. Ruiz-Rosero, G. Ramírez-González, and J. Viveros-Delgado. Software survey: Scientopy, a scientometric tool for topics trend analysis in scientific publications. *Scientometrics*, 121(2):1165–1188, 2019.

M. A. Runco and G. J. Jaeger. The standard definition of creativity. *Creativity Research Journal*, 24(1):92–96, 2012.

S. J. Russell. *Artificial intelligence a modern approach*. Pearson Education, Inc., 2010.

P. Savadjiev, J. Chong, A. Dohan, M. Vakalopoulou, C. Reinhold, N. Paragios, and B. Gallix. Demystification of ai-driven medical image interpretation: past, present and future. *European Radiology*, 29:1616–1624, 2019.

M. Savona. The value of data: Towards a framework to redistribute it. 2019.

J. Schmidhuber. Deep learning in neural networks: An overview. *Neural Networks*, 61:85–117, 2015.

U. Schmoch. Double-boom cycles and the comeback of science-push and market-pull. *Research Policy*, 36(7):1000–1015, 2007.

E. W. Schneider. Course modularization applied: The interface system and its implications for sequence control and data analysis. 1973.

J. A. Schumpeter et al. *Business Cycles*, volume 1. McGraw-Hill New York, 1939.

I. Seeber, E. Bittner, R. O. Briggs, T. De Vreede, G.-J. De Vreede, A. Elkins, R. Maier, A. B. Merz, S. Oeste-Reiß, N. Randrup, et al. Machines as teammates: A research agenda on ai in team collaboration. *Information & Management*, 57 (2):103174, 2020.

J. Sevilla, L. Heim, A. Ho, T. Besiroglu, M. Hobbhahn, and P. Villalobos. Compute trends across three eras of machine learning. In *2022 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2022.

D. Shen, G. Wu, and H.-I. Suk. Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19:221–248, 2017.

F. Shi, J. G. Foster, and J. A. Evans. Weaving the fabric of science: Dynamic network models of science's unfolding structure. *Social Networks*, 43:73–85, 2015.

S. Shibayama, D. Yin, and K. Matsumoto. Measuring novelty in science with word embedding. *PloS One*, 16(7):e0254034, 2021.

S. J. Shin and J. Zhou. When is educational specialization heterogeneity related to creativity in research and development teams? transformational leadership as a moderator. *Journal of Applied Psychology*, 92(6):1709, 2007.

P. E. Stephan. The economics of science. *Journal of Economic Literature*, 34(3): 1199–1235, 1996.

C. R. Sugimoto and V. Larivière. *Measuring research: what everyone needs to know*. Oxford University Press, 2018.

D. P. Sullivan, C. F. Winsnes, L. Åkesson, M. Hjelmare, M. Wiking, R. Schutten, L. Campbell, H. Leifsson, S. Rhodes, A. Nordgren, et al. Deep learning is combined with massive-scale citizen science to improve large-scale image classification. *Nature Biotechnology*, 36(9):820–828, 2018.

C. Sutton and L. Gong. Popularity of arxiv. org within computer science. *arXiv Preprint arXiv:1710.05225*, 2017.

I. Tahamtan and L. Bornmann. Creativity in science and the link to cited references: Is the creative potential of papers reflected in their cited references? *Journal of Informetrics*, 12(3):906–930, 2018.

A. Taylor and H. R. Greve. Superman or the fantastic four? knowledge combination and experience in innovative teams. *Academy of Management Journal*, 49(4):723–740, 2006.

M. Trajtenberg. Ai as the next gpt: a political-economy perspective. Technical report, National Bureau of Economic Research, 2018.

D. Trapido. How novelty in knowledge earns recognition: The role of consistent identities. *Research Policy*, 44(8):1488–1500, 2015.

W. Tsai and S. Ghoshal. Social capital and value creation: The role of intrafirm networks. *Academy of Management Journal*, 41(4):464–476, 1998.

A. Uddin, J. Bhoosreddy, M. Tiwari, and V. K. Singh. A sciento-text framework to characterize research strength of institutions at fine-grained thematic area level. *Scientometrics*, 106(3):1135–1150, 2016.

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

F. van der Wouden and H. Youn. The impact of geographical distance on learning through collaboration. *Research Policy*, 52(2):104698, 2023.

T. Van Leeuwen, H. Moed, R. Tijssen, M. Visser, and A. Van Raan. Language biases in the coverage of the science citation index and its consequencesfor international comparisons of national research performance. *Scientometrics*, 51(1): 335–346, 2001.

V. Van Roy, D. Vertesy, and G. Damioli. Ai and robotics innovation. *Handbook of Labor, Human Resources and Population Economics*, pages 1–35, 2020.

S. Vannuccini and E. Prytkova. Artificial intelligence's new clothes? from general purpose technology to large technical system. *From General Purpose Technology to Large Technical System (April 7, 2021). SWPS*, 2, 2021.

A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin. Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 2017.

R. Veugelers and J. Wang. Scientific novelty and technological impact. *Research Policy*, 48(6):1362–1372, 2019.

W. G. Vincenti. What engineers know and how they know it analytical studies from aeronautical history. 1990.

D. Wadden, S. Lin, K. Lo, L. L. Wang, M. van Zuylen, A. Cohan, and H. Hajishirzi. Fact or fiction: Verifying scientific claims. *arXiv Preprint arXiv:2004.14974*, 2020.

C. S. Wagner, T. A. Whetsell, and S. Mukherjee. International research collaboration: Novelty, conventionality, and atypicality in knowledge recombination. *Research Policy*, 48(5):1260–1270, 2019.

M. F. Wagstaff, G. L. Flores, R. Ahmed, and S. Villanueva. Measures of curiosity: A literature review. *Human Resource Development Quarterly*, 32(3):363–389, 2021.

M. L. Wallace, V. Larivière, and Y. Gingras. A small world of citations? the influence of collaboration networks on citation practices. *PloS One*, 7(3):e33339, 2012.

L. Waltman. A review of the literature on citation impact indicators. *Journal of Informetrics*, 10(2):365–391, 2016.

D. Wang and A.-L. Barabási. *The Science of Science*. Cambridge University Press, 2021.

J. Wang, R. Veugelers, and P. Stephan. Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy*, 46(8):1416–1436, 2017.

S. Wang and R. A. Noe. Knowledge sharing: A review and directions for future research. *Human Resource Management Review*, 20(2):115–131, 2010.

M. L. Weitzman. Recombinant growth. *The Quarterly Journal of Economics*, 113 (2):331–360, 1998.

J. Whitfield. Group theory; what makes a successful team? john whitfield looks at research that uses massive online databases and network analysis to come up with some rules of thumb for productive collaborations. *Nature*, 455(7214):720–724, 2008.

M. Winding, B. D. Pedigo, C. L. Barnes, H. G. Patsolic, Y. Park, T. Kazimiers, A. Fushiki, I. V. Andrade, A. Khandelwal, J. Valdes-Aleman, et al. The connectome of an insect brain. *Science*, 379(6636):eadd9330, 2023.

S. G. Winter and R. R. Nelson. An evolutionary theory of economic change. *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship*, 1982.

WIPO. Wipo technology trends 2019: Artificial intelligence. *Geneva: World Intellectual Property Organization*, 2019.

U. Witt. Propositions about novelty. In *Rethinking Economic Evolution*. Edward Elgar Publishing, 2016.

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

L. Wu, A. Kittur, H. Youn, S. Milojević, E. Leahey, S. M. Fiore, and Y.-Y. Ahn. Metrics and mechanisms: Measuring the unmeasurable in the science of science. *Journal of Informetrics*, 16(2):101290, 2022.

Q. Wu and Z. Yan. Solo citations, duet citations, and prelude citations: New measures of the disruption of academic papers. *arXiv Preprint arXiv:1905.03461*, 2019.

S. Wu and Q. Wu. A confusing definition of disruption, Apr 2019.

S. Wuchty, B. F. Jones, and B. Uzzi. The increasing dominance of teams in production of knowledge. *Science*, 316(5827):1036–1039, 2007.

Y. Xie and A. A. Killewald. *Is American science in decline?* Harvard University Press, 2012.

F. Xu, L. Wu, and J. Evans. Flat teams drive scientific innovation. *Proceedings of the National Academy of Sciences*, 119(23):e2200927119, 2022.

J. Xu, S. Kim, M. Song, M. Jeong, D. Kim, J. Kang, J. F. Rousseau, X. Li, W. Xu, V. I. Torvik, et al. Building a pubmed knowledge graph. *Scientific Data*, 7(1): 1–15, 2020.

Q. Zhang, L. T. Yang, Z. Chen, and P. Li. A survey on deep learning for big data. *Information Fusion*, 42:146–157, 2018.

S. Zhang, L. Yao, A. Sun, and Y. Tay. Deep learning based recommender system: A survey and new perspectives. *ACM Computing Surveys (CSUR)*, 52(1):1–38, 2019.

S. Zuboff. *The age of surveillance capitalism: The fight for a human future at the new frontier of power: Barack Obama's books of 2019*. Profile Books, 2019.

List of Figures

List of Tables

Pierre Pelletier

ARTIFICIAL INTELLIGENCE IN SCIENCE: DIFFUSION AND IMPACT

RÉSUMÉ

Cette thèse étudie l'impact de l'intelligence artificielle (IA) sur le système scientifique, en se concentrant sur la création de connaissances et les facteurs d'adoption. Le Chapitre 1 explore la manière dont l'IA affecte la production de connaissances, c'est à dire son originalité et l'impact scientifique associé. Le chapitre 2 examine les facteurs qui favorisent l'adoption de la technologie de l'IA dans les domaines d'application scientifiques. Le chapitre 3 présente *Novelpy*, un outil open-source basé sur Python qui calcule diverses mesures de nouveauté et de disruption. Enfin, le chapitre 4 vise à comprendre les sources de ces indicateurs de nouveauté en intégrant une dimension cognitive dans l'étude de la créativité en science.

Mots clefs: Intelligence Artificielle; Equipes Scientifiques; Adoption de l'IA; Nouveauté Combinatoire; Impact Scientifique

RÉSUMÉ EN ANGLAIS

This thesis studies the impact of artificial intelligence (AI) on the scientific system, focusing on knowledge creation and adoption factors. Chapter 1 explores how AI affects knowledge production, namely its originality and associated scientific impact. Chapter 2 examines the factors driving AI adoption in application domains. Chapter 3 introduces *Novelpy*, a Python-based open-source tool that computes various measures of novelty and disruptiveness. Finally, Chapter 4 aims to understand the sources of these novelty indicators by integrating a cognitive dimension into the study of creativity in science.

Keywords: Artificial Intelligence; Scientific Teams; AI Adoption; Combinatorial Novelty; Scientific Impact