Interdisciplinary Research in Artificial Intelligence: Lessons from COVID-19

Abbonato, Diletta; Bianchini, Stefano; Gargiulo, Floriana; Venturini, Tommaso

How to cite

ABBONATO, Diletta et al. Interdisciplinary Research in Artificial Intelligence: Lessons from COVID-19. In: Quantitative science studies, 2024, p. 1–20. doi: 10.1162/qss_a_00329

This publication URL: https://archive-ouverte.unige.ch/unige:182671
Publication DOI: 10.1162/qss_a_00329

Check for updates

Abbonato, D., Bianchini, S., Gargiulo, F., Venturini, T. (2024) Interdisciplinary Research in Artificial Intelligence: Lessons from COVID-19. Quantitative Science Studies. Advance Publication. https://doi.org/10.1162/qss_a_00329

Interdisciplinary Research in Artificial Intelligence: Lessons from COVID-19

Diletta Abbonato[1,2], Stefano Bianchini[*1], Floriana Gargiulo[3], and Tommaso Venturini[4]

[1] *BETA – University Strasbourg, France*

[2] *DEMS, University of Milan – Bicocca, Italy*

[3] *GEMASS, CNRS – Sorbonne Université, France*

[4] *MEDIALAB – Université de Genève, Swiss*

Abstract

Artificial intelligence (AI) is widely regarded as one of the most promising technologies for advancing science, fostering innovation, and solving global challenges. Recent years have seen a push for teamwork between experts from different fields and AI specialists, but the outcomes of these collaborations have yet to be studied. We focus on approximately 15,000 papers at the intersection of AI and COVID-19 – arguably one of the major challenges of recent decades – and show that interdisciplinary collaborations between medical professionals and AI specialists have largely resulted in publications with low visibility and impact. Our findings suggest that impactful research depends less on the overall interdisciplinary of author teams and more on the diversity of knowledge they actually harness in their research. We conclude that team composition significantly influences the successful integration of new computational technologies into science and that obstacles still exist to effective interdisciplinary collaborations in the realm of AI.

Keywords— Artificial intelligence, COVID-19, interdisciplinarity, team science, science of science

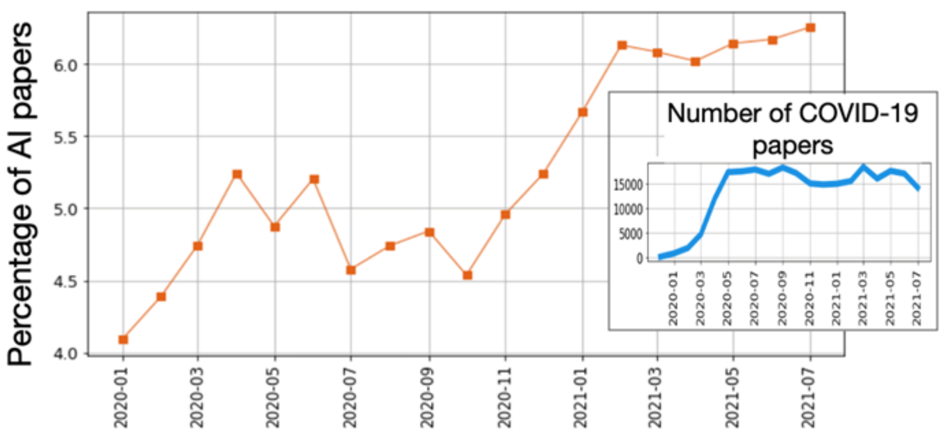
*Corresponding author: `s.bianchini@unistra.fr`

Peer Review: https://www.webofscience.com/api/gateway/wos/peer-review/10.1162/qss_a_00329

1 Introduction

Interdisciplinarity has become a buzzword in science policy. And with good reason. Disciplines have for decades – in some cases, centuries – facilitated scientific progress by providing scholars with the scaffolding of a coherent paradigm and the possibility of standing on the shoulders of their predecessors. However, disciplinary boundaries have often proved to be a stumbling block to innovation, as growing specialization makes it ever harder (though ever more necessary) to venture into unexplored research territories and combine intellectual tools originating from different traditions (Jones, 2009). These entrenched boundaries are especially problematic when facing unprecedented research challenges that require fresh thinking and unrestrained experimentation. Such a situation presented itself recently with the outbreak of the COVID-19 pandemic. The urgency and gravity of the situation prompted researchers in epidemiology and medical science not only to mobilize all the resources available within their disciplines, but to look beyond them for new ideas and external collaborations. Among them, the alliance with artificial intelligence (AI) emerged as one of the most promising (Fig. 1).

Figure 1: COVID-19 publications mentioning AI technology



Notes: Fraction of COVID-19 papers mentioning AI technologies. Inset: Total number of COVID-19 papers. After an initial period of exponential growth, scientific production related to the COVID-19 virus stabilized in May 2020. At the same time, AI research dedicated to COVID-19 virus remained relatively marginal until summer 2020 when it began to record constant linear growth, so that by July 2021 it accounted for nearly 7% of total COVID-19 scientific production. Source: Own elaboration on CORD-19 data.

Although AI techniques have a long history, the field has recently been revived by the escalating power of computational technologies and the growing availability of data on social and natural phenomena. This has led to the development of new machine learning approaches, which have yielded remarkable results within and beyond data science (Cardon et al., 2018; Frank et al., 2019). Recent studies have shown that AI/ML techniques are indeed changing the "way of doing science", from agenda setting and hypothesis formulation to experimentation, knowledge sharing, and public involvement, with a considerable impact on scientific practices (Cockburn et al., 2018; Agrawal et al., 2018; Xu et al., 2021; Bianchini et al., 2022; Birhane et al., 2023; Van Noorden and Perkel, 2023; Koehler and Sauermann, 2024).[1]

The coronavirus pandemic hits at the peak of this cycle of AI hype and, unsurprisingly, many scholars quickly embraced the idea of adopting AI techniques to tackle the challenges presented by COVID-19 (DeGrave et al., 2021; Khan et al., 2021; Roberts et al., 2021).[2] Opportunities for collaborative funding have emerged globally to bring various scientific communities together, and researchers from different backgrounds have come together to try to harness the potential of AI in COVID-19 research (Ahuja et al., 2020; Luengo-Oroz et al., 2020). Some of these collaborations offered substantial contributions to the fight against the pandemic. By reviewing some of the most cited papers in our dataset, we found some interesting uses of AI for COVID-19 research, particularly for making sense of large literature or data archives (cf. for example, Mistry et al., 2021; Salari et al., 2020; Wynants et al., 2020), and some thoughtful assessments of the effectiveness of AI and big data in medicine (cf. for example, Wang et al., 2020a; Agbehadji et al., 2020). Many other publications at the AI/COVID-19 intersection, however, never gained much visibility or scientific traction. What can explain these contrasting outcomes?

Previous research shows that (large) interdisciplinary teams produce more cited research and high-impact papers (Wuchty et al., 2007; Fortunato et al., 2018), and that diversity – not only epistemic, but also institutional and ethnic – is beneficial for producing novel, valuable ideas (Taylor and Greve, 2006). Teams comprising researchers with different backgrounds, methodological approaches, and experience have access to a broader pool of knowledge, which allows them to produce more creative outputs than those produced by less collaborative science (Stephan, 2012; Uzzi et al., 2013; Gargiulo et al., 2022). This can be explained by the functional diversity of teams, that is, differences in the way scientists encode problems and attempt to solve them; as Hong and Page (2004) put it succinctly: "diversity trumps ability". Collaborative projects also serve to boost visibility by exposing scientific findings to a wider and more diverse readership (Leahey, 2016). In the case of COVID-19 research, this suggests that collaborations between AI experts and clinicians may result in successful research outcomes, as domain specialists could provide their "on-the-ground" knowledge to identify promising areas for investigation, while technology experts could offer access to the latest computational methods.

[1] The number of AI publications has surged nearly five-fold in the past decade, constituting around 5 percent of total scientific publications; most of this research has gradually shifted from AI development to its application, which currently represents some 70 percent of scientific activity (Arranz et al., 2023).

[2] It is worth noting that AI is seen by many as a technological solution to meet contemporary global challenges, such as sustainable development, green transition, global health and others (see, e.g., Schwalbe and Wahl, 2020; Vinuesa et al., 2020). Yet, it is equally important to note that the benefits brought by technology are such only under proper AI governance frameworks (Truby, 2020).

Team diversity, however, is not without its disadvantages. Teams that are too large and heterogeneous often suffer from lower consensus-building, cognitive diversity, higher coordination costs, and emotional conflict. As diversity increases, it may become more difficult to convert specialized expertise into scientific outputs (Lee et al., 2015). Studies show that team performances depend more on how the team interacts than on the characteristics of its members (Woolley et al., 2010), and that most successful collaborations seem to be achieved through efforts that, while interdisciplinary, combine relatively close fields (Yegros-Yegros et al., 2015).[3] Furthermore, some physicians were and continue to be suspicious of AI, for it threatens their professional expertise and challenges their authority over clinical truth and, more generally, the place of medicine within society (see Hanemaayer (2021) for a discussion on the historical resistance to AI and algorithmic computer technologies in medicine). Recent empirical evidence suggests indeed that it is difficult for AI-engaged research to mix well with more "traditional" domain-specific research (Duede et al., 2024). Consequently, difficulties may have arisen in collaborations between AI and COVID-19 experts, potentially leading to less impactful and visible scientific outcomes compared to teams composed solely of AI or clinical specialists.[4]

In this paper, we examine the impact of interdisciplinarity by investigating a large corpus of scientific publications at the intersection of COVID-19 and AI (about 15,000 papers retrieved from the COVID-19 Open Research Dataset, CORD-19 – version 2021-08-09 – and supplemented by other metadata from Altmetric and OpenAlex), and studying which forms of interdisciplinarity are more strongly associated with scientific impact. In the remainder, we first describe the metrics of interdisciplinarity used in our study, and then link these metrics to three indicators of scientific "success", namely the number of citations, online visibility, and outreach to other disciplines.

2 Material and methods

2.1 Data

Our analysis combines data from three different databases – CORD-19, OpenAlex, and Altmetric – and is based on the pre-processing protocol illustrated in the Supplementary Material (Fig. S1). The COVID-19

[3] A comprehensive review of the rich literature on the impact of interdisciplinary research is beyond the scope of this article. However, it is interesting to note that the question is still open and debated in the scientific community. For instance, while some studies on environmental sciences and biomedicine suggest long-term benefits of interdisciplinary approaches, particularly in terms of introducing novel ideas (Steele and Stier, 2000; Schilling and Green, 2011; Wang et al., 2015; Larivière et al., 2015; Okamura, 2019), others indicate that interdisciplinary research may reduce both scientific productivity (Leahey et al., 2017) and impact (Levitt and Thelwall, 2008).

[4] One of the main arguments for skepticism toward computational technologies and more recently AI/ML is that medicine is first and foremost the "art" of dealing with the patient as a whole person, which involves knowledge of social relationships and normativity. As Goldhahn et al. (2018) explicitly put it: "*The physician-patient relationship is a relationship between mortal beings vulnerable to illness and death. Computers [...] are not people and do not care about anything*" (p.2).

Open Research Dataset (CORD-19) is a growing corpus of publications on COVID-19 and other coronavirus infections (Wang et al., 2020b). It includes, in the period that we considered (from 01/12/2019 to 31/08/2021), around 600K documents from different sources, including WHO, PubMed central, bioRxiv and medRxiv. Within this large corpus, we focused specifically on a subset of publications that included, in their abstract or title, at least one keyword related to AI. Our list of around 300 AI keywords (see Table S1 in the Supplementary Material) was created by merging the terms mentioned in the Wikipedia AI Glossary for AI with other 'AI vocabularies' (Baruffaldi et al., 2020; Bianchini et al., 2022; Gargiulo et al., 2023).

For each paper in this subset, we retrieved additional metadata from OpenAlex. We discarded all documents with missing information and obtained a final corpus of 16,148 AI publications on COVID-19 (COVID-19+AI dataset). We retrieved the metadata for all the references cited by the publications in our corpus (circa 300K unique papers) and for all the papers that cite them (c. 200K papers). OpenAlex metadata included the DOI, which we used for retrieving the 'attention score' for each paper in the COVID-19+AI dataset from the website Altmetric.com. The score provides a measure of online visibility for scholarly contents (e.g., mentions on the news, in blogs, and on Twitter; article page-views and downloads; GitHub repository watchers). Finally, we used the author identifier in OpenAlex to retrieve the previous publications of all 87,552 authors present in our corpus (around 150K papers) and the institutions to which they are affiliated.

2.2 Measuring interdisciplinarity

The concept of interdisciplinarity is multifaceted, often ambiguous, and there is no consensus on the definition and operationalization of interdisciplinary research (cf. for example, Porter et al., 2007; Huutoniemi et al., 2010; Leydesdorff and Rafols, 2011; Yegros-Yegros et al., 2015; Wang and Schneider, 2020; Fontana et al., 2020, 2022). Here, we use different measures of interdisciplinarity that consider the diversity of team members and references cited in a paper.

Each document, i, in our data is characterized by a set of authors (A_i), a set of references and citations (R_i, C_i), a set of AI keywords, if any, (W_i), the journal where it is published (J_i), and its altmetric score (M_i). Each author, a, in our corpus is associated with his/her list of papers (P_a) and with his/her three most recent papers (P_a^3).

Using a measure inspired by pairwise mutual information and based on the co-occurrence of journals in the reference lists of all articles, we compute a matrix, D, of distances between all journals in the dataset (the more two journals are regularly cited together, the smaller is their distance). To build the distance matrix, we first calculate the mutual co–citation network among journals (where two journals are linked if they appear simultaneously in a reference list). Self-loops are removed. The network is weighted and the weights, w_{ij}, correspond to the number of co–occurrences. Normalizing these weights, we define a connection probability among journals in the following way:

$$p_{ij} = \frac{w_{ij}}{\sum_{i>j} w_{ij}} \tag{1}$$

The structure of this network, however, is biased by the heterogeneity in terms of the number of publications among the journals: some important relationships among small journals could be hidden by their relative size compared to large journals. For this reason, instead of using the weighted adjacency matrix of this network for calculating journal similarity, we introduce a measure based on point–wise mutual information (PMI), that is:

$$pmi_{ij} = max\{0, \frac{1}{log_2 w_{ij}} log_2(\frac{w_{ij}}{p_i p_j})\} \tag{2}$$

where $p_i = \sum_j w_{ij}$. This measure is a similarity ranging between 0 and 1. Hence, we obtain the distance as $D_{ij} = 1 - pmi_{ij}$.

Using this notion of distance, we define two types of interdisciplinarity metrics: the first is related to team composition (measuring the disciplinary span of the previous papers by the contributors of a paper); the second is related to the knowledge mobilized in the paper (measuring the disciplinary span in papers' references). For each dimension (team and knowledge), we introduce a further distinction between interdisciplinarity metrics specifically related to AI, and the more general interdisciplinarity, providing us with four main different metrics:

- *AI Team Expertise* is the fraction of previous AI publications for each author, averaged over the entire team:

$$AI\ Team\ Expertise_i = \frac{1}{\#\mathcal{A}_i} \sum_{a \in \mathcal{A}_i} \frac{\#\{j \in \mathcal{P}_a | \mathcal{W}(j) \neq \{\}\}}{\#\mathcal{P}_a}$$

- *Share AI References* is the fraction of cited references related to AI:

$$Share\ AI\ References_i = \frac{\#\{j \in \mathcal{R}_i | \mathcal{W}(j) \neq \{\}\}}{\#\mathcal{R}_i}$$

- *PMI (Team)* is the average disciplinary dispersion (in term of journal distances) of team authors:

$$PMI\ (Team)_i = \frac{1}{\#\mathcal{A}_i} \sum_{a \in \mathcal{A}_i} \left(\frac{1}{3} \sum_{k \neq l \in \mathcal{P}_a^3} \mathbf{D}_{J(k)J(l)} \right)$$

- *PMI (References)* is the average distance among all the journals cited in the references:

$$PMI\ (References)_i = \frac{1}{\#(\mathcal{R}_i \times \mathcal{R}_i)} \sum_{(u,v)\in(\mathcal{R}_i \times \mathcal{R}_i)} \mathbf{D}_{J(u)J(v)}$$

The first two metrics measure the share of AI in the author teams and knowledge mobilized by the publications, respectively. The last two measure levels of general interdisciplinarity in the teams and the knowledge mobilized by the publications.

In the scientometric literature, the indicator we use to characterize interdisciplinarity is similar to the disparity measure known as the Rao-Stirling indicator (Stirling, 2007). This measure uses another way to manage the bias of a distance based on the matrix defined by 1, inserting the relative frequencies of the journals in the calculation of the index:

$$\Delta(i) = \sum_{ij} w_{ij} p_i p_j \tag{3}$$

Thus, for the sake of completeness, we also calculate the Rao-Stirling *disparity*, D_i, for teams and knowledge composition. Our metric, as well as the Rao-Stirling disparity, takes into account the relative distance among the journals/disciplines, avoiding treating two journals/disciplines that are very similar in content as truly different.

We also extend the analysis to two other dimensions traditionally used to define interdisciplinarity: *variety* and *balance*. The variety, V_i, is the count of the number of different journals where the authors previously published (for team) and of the different journals cited in the references (for knowledge). The balance, B_i, is the Gini index of the frequency associated to each journal for authors previous publications (teams) and for the references (knowledge).

Finally, we define three indicators of "success" for the publications in our corpus, namely: the *number of citations*, N_i, the *altmetric score*, M_i, and the *interdisciplinary spread*, I_i – i.e., how a paper is cited in a diverse set of disciplines - defined as:

$$\mathcal{I}(i) = \frac{1}{\#(\mathcal{C}_i \times \mathcal{C}_i)} \sum_{(u,v)\in(\mathcal{C}_i \times \mathcal{C}_i)} \mathbf{D}_{J(u)J(v)}$$

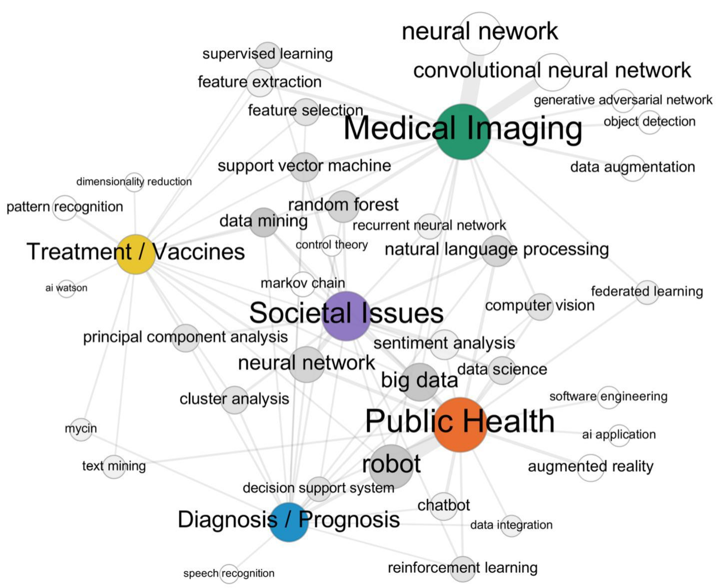
Descriptive statistics of the variables used for this study are reported in Table S2 in the Supplementary Material.

2.3 AI applications

By running a LDA (Latent Dirichlet Allocation) topic modelling on the abstracts of the papers in our corpus, we obtained five distinct areas in which AI/ML techniques have been applied (Fig. 2):

- *Societal Issues* (including epidemiology and infodemics), with some recurrent terms such as social medium, infectious disease, mental health, reproduction number, social distance, etc.;
- *Medical Imaging*: chest X-ray, chest scan, tomography, etc.;
- *Diagnosis and Prognosis*: clinical trials, risk factors, mechanical ventilation, etc.;
- *Treatments and Vaccines*: molecular docking, spike protein, gene expression, drug discovery, etc.;
- *Public Health*: public health, contact tracing, health system, face mask, etc..

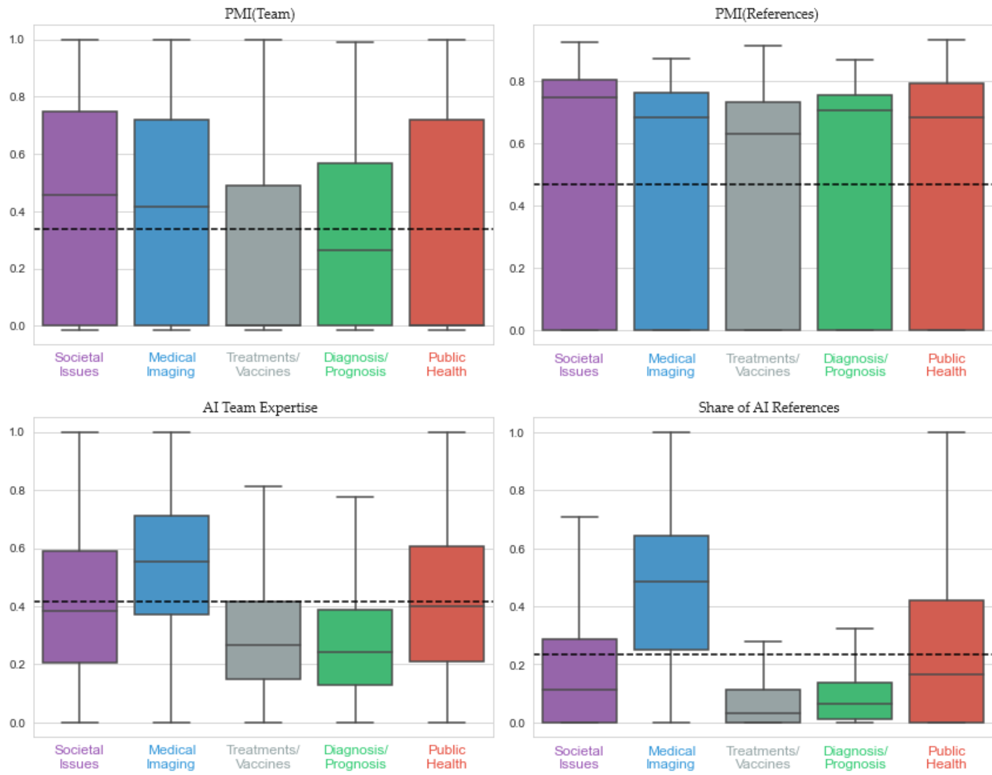
Figure 2: AI application areas for COVID-19 research



Notes: Co-occurrence of AI keywords (gray nodes) and COVID-19 topics (colored nodes). Edges are weighted by the number of articles using each keyword in each topic. Nodes are sized according to their popularity (number of articles). Keywords are colored according to their degree, from white keywords specific to a single topic to dark gray keywords used in multiple topics. The consistency score of the LDA model is 0.53.

A closer reading of the terms characterizing each topic suggests that AI has found a multitude of applications (Bullock et al., 2020; Naudé, 2021; Yang et al., 2020; Piccialli et al., 2021) – see Table S3 in the Supplementary Material. In the case of societal issues, AI seems to have been used mainly for predicting the

Figure 3: Interdisciplinarity metrics in the different axes of COVID-19 research



Notes: General (top) and AI-related interdisciplinarity (bottom). The dotted line represents the mean.

spread of disease over time and space, modeling public policy interventions (e.g., social distancing) and risk assessment, and fighting misinformation and disinformation on social media. In the case of medical imaging, we see the deployment of deep learning models (e.g., CNN) to detect signs of COVID-19 from X-ray images and computed tomography (CT) scans. Another area of application, particularly of machine learning and deep learning, is the identification of possible treatments and vaccines, as well as the re-purposing of existing drugs. Finally, AI appears to support the management of the public health system, for example, robotics providing assistance in the delivery of healthcare tasks.

Each application area may have required specific skills and know-how from researchers with diverse backgrounds and experiences, as well as the (re)combination of different types of knowledge. Our corpus reveals a high level of general interdisciplinarity both in the teams and in the knowledge mobilized by the publications across all research topics – with a slightly higher knowledge heterogeneity in societal issues and diagnosis/prognosis (Fig. 3 top).

In the case of AI, we observe very different scenarios at the topic level. Indeed, the share of teams with more AI experts is markedly higher in medical imaging and public health research, whereas teams working on vaccines, treatments, and prognosis seem to rely very little on AI knowledge (Fig. 3 bottom).

3 Results

3.1 What determines 'success'

We model the various impact measures – i.e., the number of citations received by the publication, the Altmetric attention score, and the interdisciplinary spread – as a function of the different interdisciplinarity metrics discussed earlier and a set of control variables, namely: *AI Collaborator* (=1 if the team includes at least one AI researcher); *Top AI Collaborator* (=1 if the team includes an AI researcher with past number of citations in the top 10° percentile of the citation distribution); *Academic Age* (average academic age of team members, in logs); *Past Impact* (average H-Index of team members based on past publications, in logs); *Nb. Countries* (number of participating countries within a team, in logs); and *Nb. References* (number of cited references, in logs). We also included a complete set of fixed effects for the month of publication and the dominant topic. The number of citations is a count variable and was modeled using a negative binomial regression. The continuous variables – attention score and interdisciplinarity spread – were modeled using ordinary least square regressions.

As shown in Table 1 and 2, the most notable result to emerge from our model is that collaborations with researchers experienced in AI (*AI Collaborator*) do not have a significant impact, and those involving a high share of researchers with established track records of AI publications (*AI Team Expertise*) receive, ceteris paribus, fewer citations, have less online visibility, and struggle to reach distant disciplines. Only those teams that include a top AI researcher (*Top AI Collaborator*) present a positive impact on citations received by their publication, albeit this impact is not strong. Similarly, the ratio of AI-related references (*Share AI References*) has a null or negative impact on the Altmetric attention score and interdisciplinary spread. All in all, research interdisciplinarity limited to AI does not seem to have *any* influence on the impact of COVID-19 publications, and when it does, this influence is negative.

What appears to ensure the impact of a publication is, above all else, the interdisciplinarity of the knowledge mobilized via its references, that is the actual epistemological diversity of the research conducted by a team. Regardless of how we operationalize this diversity, we find a systematic positive effect on all impact measures (except for disparity in the models on the number of citations and attention score). The effect is consistently higher than that of more classic features, such as past impact or the number of affiliated countries. The overall diversity of team members generally has a much less strong, significant and in many cases negative effect.

Table 1: Determinants of 'success' – Nb. Citations and Attention Score

	Nb. Citations				*Attention Score*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
AI Team Expertise	-0.219***	-0.217***	-0.199***	-0.201***	-0.498***	-0.509***	-0.479***	-0.481***
	(0.062)	(0.062)	(0.062)	(0.061)	(0.057)	(0.057)	(0.057)	(0.057)
Share AI References	0.268***	0.274***	0.271***	0.363***	-0.352***	-0.345***	-0.365***	-0.304***
	(0.057)	(0.057)	(0.057)	(0.057)	(0.053)	(0.053)	(0.053)	(0.053)
PMI (Team)	-0.386***				-0.065			
	(0.067)				(0.062)			
PMI (References)	0.482***				0.287***			
	(0.045)				(0.042)			
Balance (Team)		-0.014				-0.123***		
		(0.047)				(0.044)		
Balance (References)		0.220***				0.285***		
		(0.044)				(0.041)		
Disparity (Team)			0.474***				0.318***	
			(0.081)				(0.075)	
Disparity (References)			-0.182				-0.375***	
			(0.113)				(0.104)	
Variety (Team)				-0.052***				0.018*
				(0.012)				(0.011)
Variety (References)				0.014***				0.004***
				(0.001)				(0.001)
AI Collaborator	0.026	0.013	0.009	-0.024	-0.304**	-0.304**	-0.293**	-0.302**
	(0.134)	(0.134)	(0.134)	(0.133)	(0.123)	(0.123)	(0.123)	(0.123)
Top AI Collaborator	0.471***	0.474***	0.481***	0.458***	0.109	0.109	0.105	0.106
	(0.075)	(0.075)	(0.075)	(0.074)	(0.070)	(0.070)	(0.070)	(0.070)
Past Impact [log]	0.189***	0.184***	0.185***	0.186***	0.193***	0.194***	0.193***	0.193***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)
Academic Age [log]	-0.327***	-0.326***	-0.333***	-0.318***	-0.278***	-0.276***	-0.281***	-0.280***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.020)	(0.021)	(0.021)	(0.020)
Nb. Countries [log]	0.722***	0.726***	0.731***	0.678***	0.212***	0.213***	0.214***	0.197***
	(0.031)	(0.031)	(0.031)	(0.031)	(0.029)	(0.029)	(0.029)	(0.029)
Nb. References [log]	0.195***	0.200***	0.208***	0.154***	0.086***	0.088***	0.110***	0.081***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.009)
Log Likelihood	-55,085	-55,109	-55,121	-54,983				
AIK	110,250	110,299	110,322	110,046				
Adjusted R^2					0.192	0.191	0.188	0.190
F Statistic					86.540***	86.027***	84.352***	85.419***
# Observations	14,019	14,019	14,019	14,019	14,019	14,019	14,019	14,019

Notes: The statistical model for evaluating the relationship of different interdisciplinary metrics on two indicators of 'success': the number of citations received by the publication (Columns 1–4) and the Altmetric attention score (Column 5–8). Coefficient estimates of time and topic fixed effects have been omitted from the table. The asterisks ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

Table 2: Determinants of 'success' – Interdisciplinarity Spread

	Interd. Spread			
	(1)	(2)	(3)	(4)
AI Team Expertise	-0.010	-0.042***	-0.068***	-0.089***
	(0.013)	(0.013)	(0.014)	(0.013)
Share AI References	-0.010	0.050***	0.094***	0.097***
	(0.012)	(0.012)	(0.013)	(0.012)
PMI (Team)	0.281***			
	(0.014)			
PMI (References)	0.594***			
	(0.010)			
Balance (Team)		0.281***		
		(0.010)		
Balance (References)		0.427***		
		(0.009)		
Disparity (Team)			1.029***	
			(0.018)	
Disparity (References)			0.740***	
			(0.025)	
Variety (Team)				0.206***
				(0.003)
Variety (References)				0.004***
				(0.0002)
AI Collaborator	0.021	0.002	0.001	0.008
	(0.028)	(0.028)	(0.030)	(0.029)
Top AI Collaborator	0.018	0.024	0.030*	0.022
	(0.016)	(0.016)	(0.016)	(0.016)
Past Impact [log]	-0.002*	0.006***	0.015***	0.003**
	(0.001)	(0.001)	(0.001)	(0.001)
Academic Age [log]	-0.004	-0.017***	-0.035***	-0.022***
	(0.005)	(0.005)	(0.005)	(0.005)
Nb. Countries [log]	0.118***	0.109***	0.094***	0.093***
	(0.007)	(0.007)	(0.007)	(0.007)
Nb. References [log]	0.007***	0.008***	0.014***	0.007***
	(0.002)	(0.002)	(0.002)	(0.002)
Adjusted R^2	0.551	0.554	0.476	0.515
F Statistic	441.300***	446.700***	327.200***	383.400***
# Observations	14,019	14,019	14,019	14,019

Notes: The statistical model for evaluating the relationship of different interdisciplinary metrics on the interdisciplinary spread. Coefficient estimates of time and topic fixed effects have been omitted from the table. The asterisks $***$, $**$ and $*$ indicate significance at the 1%, 5%, and 10% levels, respectively.

3.2 Robustness checks

We made sure that our main findings are robust to alternative specifications and arbitrary choices. First, we performed the same modeling exercise using OpenAlex 'concepts' instead of journals. We replicated the models by considering level-0 concepts (e.g., computer science) associated to each journal and level-1 concepts (e.g., machine learning), which are more granular. A journal can be associated with more than one concept; in this case, we considered the concept with the highest 'confidence score' provided by OpenAlex. Second, we considered the 5 most recent papers by each author instead of 3. Third, we re-estimated the models for the number of citations with a quasi-Poisson instead of a negative binomial regression. Finally, we excluded all publications that are still pre-print as of December 31, 2023.

4 Discussion

The COVID-19 pandemic sparked a global research effort to address this unprecedented event. The scientific system responded promptly to the early stages of the virus and the international scientific community called upon its diverse expertise to assess the clinical and pathogenic characteristics of the disease and to formulate therapeutic and epidemiological strategies to cope with it. Policymakers were also quick to seek advice from ethicists, sociologists, and economists on how best to deal with the crisis (Fry et al., 2020; Chahrour et al., 2020). Against this backdrop, AI applications represented a promising approach to face many of the challenges posed by the pandemic. Yet a number of studies focusing on the application of AI-based approach to COVID-19 research have identified various barriers and shortcomings. They include poor data quality and flow, as well as the lack of global standards and database interoperability (e.g., genetic sequences, protein structures, medical imagery and epidemiological data); the inability of algorithms to work without sufficient knowledge of the domain; overly exacting computational, architectural, and infrastructural requirements; and the legal and ethical opacity associated with privacy and intellectual property (Bullock et al., 2020; Luengo-Oroz et al., 2020; Naudé, 2020; Khan et al., 2021; Piccialli et al., 2021).

In this paper, we have analyzed the role played by different forms of interdisciplinarity, both at the team level and in the research conducted, and their repercussions on various measures of scientific impact. Our research was, in part, motivated by the fact that policy initiatives around the world have emerged – and continue to emerge – aimed at encouraging collaboration between the AI community and specialists in various domains. However, we have no direct evidence of the effectiveness of these initiatives.

Our study provides an important takeaway message for policy-makers and science administrators: collaborations involving AI researchers did not necessarily result in more impactful science. As our analysis revealed, the visibility, relevance and spread of the publications we considered all seem to be linked to the diversity of references rather than that of authors. What generates high-impact science, in other words, is not the *possible* interdisciplinarity associated with team diversity, but the *actual* epistemological diversity hardwired into a paper.

Author contributions. Conceptualization (DA, SB, FG, TV); Data curation (DA); Methodology (DA, SB, FG, TV); Formal analysis (DA); Visualization (DA, FG, TV); Writing (DA, SB, FG, TV); Supervision (SB); Funding acquisition (SB, FG, TV).

Data availability. Data and codes to reproduce the main analyses and robustness tests are available here: https://github.com/zabbonat/AI-COVID19-interdisciplinarity

Funding. This work was supported by the La Mission pour les Initiatives Transverses et Interdisciplinaires (MITI) of the Centre National de la Recherche Scientifique (CNRS). This work was also supported by the European Union – Horizon 2020 Program under the scheme "INFRAIA-01-2018-2019 – Integrating Activities for Advanced Communities", Grant Agreement n.871042, "SoBigData++: European Integrated Infrastructure for Social Mining and Big Data Analytics" (http://www.sobigdata.eu). Stefano Bianchini received financial support through the SEED project – Grant agreement ANR-22-CE26-0013. Floriana Gargiulo received financial support through the ScientIA project ANR-21-CE38-0020.

Competing interests. The authors declare no competing interests.

References

Agbehadji, I. E., Awuzie, B. O., Ngowi, A. B., and Millham, R. C. (2020). Review of big data analytics, artificial intelligence and nature-inspired computing models towards accurate detection of COVID-19 pandemic cases and contact tracing. *International Journal of Environmental Research and Public Health*, 17(15):5330.

Agrawal, A., McHale, J., and Oettl, A. (2018). Finding needles in haystacks: Artificial intelligence and recombinant growth. In *The economics of Artificial Intelligence: An Agenda*, pages 149–174. University of Chicago Press.

Ahuja, A. S., Reddy, V. P., and Marques, O. (2020). Artificial intelligence and COVID-19: A multidisciplinary approach. *Integrative Medicine Research*, 9(3).

Arranz, D., Bianchini, S., Di Girolamo, V., and Ravet, J. (2023). *Trends in the use of AI in science : A bibliometric analysis*. Publications Office of the European Union.

Baruffaldi, S., van Beuzekom, B., Dernis, H., Harhoff, D., Rao, N., Rosenfeld, D., and Squicciarini, M. (2020). Identifying and measuring developments in artificial intelligence: Making the impossible possible.

Bianchini, S., Müller, M., and Pelletier, P. (2022). Artificial intelligence in science: An emerging general method of invention. *Research Policy*, 51(10):104604.

Birhane, A., Kasirzadeh, A., Leslie, D., and Wachter, S. (2023). Science in the age of large language models. *Nature Reviews Physics*, 5(5):277–280.

Bullock, J., Luccioni, A., Pham, K. H., Lam, C. S. N., and Luengo-Oroz, M. (2020). Mapping the landscape of artificial intelligence applications against COVID-19. *Journal of Artificial Intelligence Research*, 69:807–845.

Cardon, D., Cointet, J.-P., Mazières, A., and Carey-Libbrecht, L. (2018). Neurons spike back. *Réseaux*, 211(5):173–220.

Chahrour, M., Assi, S., Bejjani, M., Nasrallah, A. A., Salhab, H., Fares, M., Khachfe, H. H., Salhab, H. A., and Fares, M. Y. (2020). A bibliometric analysis of COVID-19 research activity: A call for increased output. *Cureus*, 12(3).

Cockburn, I. M., Henderson, R., and Stern, S. (2018). The impact of artificial intelligence on innovation: An exploratory analysis. In *The Economics of Artificial Intelligence: An agenda*, pages 115–146. University of Chicago Press.

DeGrave, A. J., Janizek, J. D., and Lee, S.-I. (2021). AI for radiographic COVID-19 detection selects shortcuts over signal. *Nature Machine Intelligence*, 3(7):610–619.

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

Fontana, M., Iori, M., Montobbio, F., and Sinatra, R. (2020). New and atypical combinations: An assessment of novelty and interdisciplinarity. *Research Policy*, 49(7):104063.

Fontana, M., Iori, M., Sciabolazza, V. L., and Souza, D. (2022). The interdisciplinarity dilemma: Public versus private interests. *Research Policy*, 51(7):104553.

Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., Petersen, A. M., Radicchi, F., Sinatra, R., Uzzi, B., et al. (2018). Science of science. *Science*, 359(6379):eaao0185.

Frank, M. R., Wang, D., Cebrian, M., and Rahwan, I. (2019). The evolution of citation graphs in artificial intelligence research. *Nature Machine Intelligence*, 1(2):79–85.

Fry, C. V., Cai, X., Zhang, Y., and Wagner, C. S. (2020). Consolidation in a crisis: Patterns of international collaboration in early COVID-19 research. *PloS One*, 15(7):e0236307.

Gargiulo, F., Castaldo, M., Venturini, T., and Frasca, P. (2022). Distribution of labor, productivity and innovation in collaborative science. *Applied Network Science*, 7(1):19.

Gargiulo, F., Fontaine, S., Dubois, M., and Tubaro, P. (2023). A meso-scale cartography of the AI ecosystem. *Quantitative Science Studies*, 4(3):574–593.

Goldhahn, J., Rampton, V., and Spinas, G. A. (2018). Could artificial intelligence make doctors obsolete? *BMJ*, 363.

Hanemaayer, A. (2021). Don't touch my stuff: Historicising resistance to AI and algorithmic computer technologies in medicine. *Interdisciplinary Science Reviews*, 46(1-2):126–137.

Hong, L. and Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences*, 101(46):16385–16389.

Huutoniemi, K., Klein, J. T., Bruun, H., and Hukkinen, J. (2010). Analyzing interdisciplinarity: Typology and indicators. *Research Policy*, 39(1):79–88.

Jones, B. F. (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *The Review of Economic Studies*, 76(1):283–317.

Khan, M., Mehran, M. T., Haq, Z. U., Ullah, Z., Naqvi, S. R., Ihsan, M., and Abbass, H. (2021). Applications of artificial intelligence in COVID-19 pandemic: A comprehensive review. *Expert Systems with Applications*, 185:115695.

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

Larivière, V., Gingras, Y., Sugimoto, C. R., and Tsou, A. (2015). Team size matters: Collaboration and scientific impact since 1900. *Journal of the Association for Information Science and Technology*, 66(7):1323–1332.

Leahey, E. (2016). From sole investigator to team scientist: Trends in the practice and study of research collaboration. *Annual Review of Sociology*, 42:81–100.

Leahey, E., Beckman, C. M., and Stanko, T. L. (2017). Prominent but less productive: The impact of interdisciplinarity on scientists' research. *Administrative Science Quarterly*, 62(1):105–139.

Lee, Y.-N., Walsh, J. P., and Wang, J. (2015). Creativity in scientific teams: Unpacking novelty and impact. *Research Policy*, 44(3):684–697.

Levitt, J. M. and Thelwall, M. (2008). Is multidisciplinary research more highly cited? A macrolevel study. *Journal of the American Society for Information Science and Technology*, 59(12):1973–1984.

Leydesdorff, L. and Rafols, I. (2011). Indicators of the interdisciplinarity of journals: Diversity, centrality, and citations. *Journal of Informetrics*, 5(1):87–100.

Luengo-Oroz, M., Hoffmann Pham, K., Bullock, J., Kirkpatrick, R., Luccioni, A., Rubel, S., Wachholz, C., Chakchouk, M., Biggs, P., Nguyen, T., et al. (2020). Artificial intelligence cooperation to support the global response to covid-19. *Nature Machine Intelligence*, 2(6):295–297.

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., Finn, R. D., and Bateman, A. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Research*, 49(D1):D412–D419.

Naudé, W. (2020). Artificial intelligence vs COVID-19: Limitations, constraints and pitfalls. *AI & Society*, 35(3):761–765.

Naudé, W. (2021). Artificial intelligence: Neither utopian nor apocalyptic impacts soon. *Economics of Innovation and New Technology*, 30(1):1–23.

Okamura, K. (2019). Interdisciplinarity revisited: Evidence for research impact and dynamism. *Palgrave Communications*, 5(1).

Piccialli, F., Di Cola, V. S., Giampaolo, F., and Cuomo, S. (2021). The role of artificial intelligence in fighting the COVID-19 pandemic. *Information Systems Frontiers*, 23(6):1467–1497.

Porter, A., Cohen, A., David Roessner, J., and Perreault, M. (2007). Measuring researcher interdisciplinarity. *Scientometrics*, 72(1):117–147.

Roberts, M., Driggs, D., Thorpe, M., Gilbey, J., Yeung, M., Ursprung, S., Aviles-Rivero, A. I., Etmann, C., McCague, C., Beer, L., et al. (2021). Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nature Machine Intelligence*, 3(3):199–217.

Salari, N., Hosseinian-Far, A., Jalali, R., Vaisi-Raygani, A., Rasoulpoor, S., Mohammadi, M., Rasoulpoor, S., and Khaledi-Paveh, B. (2020). Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: A systematic review and meta-analysis. *Globalization and Health*, 16(1):57.

Schilling, M. A. and Green, E. (2011). Recombinant search and breakthrough idea generation: An analysis of high impact papers in the social sciences. *Research Policy*, 40(10):1321–1331.

Schwalbe, N. and Wahl, B. (2020). Artificial intelligence and the future of global health. *The Lancet*, 395(10236):1579–1586.

Steele, T. W. and Stier, J. C. (2000). The impact of interdisciplinary research in the environmental sciences: A forestry case study. *Journal of the American Society for Information Science*, 51(5):476–484.

Stephan, P. (2012). *How economics shapes science*. Harvard University Press.

Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society Interface*, 4(15):707–719.

Taylor, A. and Greve, H. R. (2006). Superman or the fantastic four? Knowledge combination and experience in innovative teams. *Academy of Management Journal*, 49(4):723–740.

Truby, J. (2020). Governing artificial intelligence to benefit the un sustainable development goals. *Sustainable Development*, 28(4):946–959.

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

Van Noorden, R. and Perkel, J. M. (2023). AI and science: What 1,600 researchers think. *Nature*, 621(7980):672–675.

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S. D., Tegmark, M., and Fuso Nerini, F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature Communications*, 11(1):233.

Wang, C. J., Ng, C. Y., and Brook, R. H. (2020a). Response to COVID-19 in Taiwan. *JAMA*, 323(14):1341.

Wang, J., Thijs, B., and Glänzel, W. (2015). Interdisciplinarity and impact: Distinct effects of variety, balance, and disparity. *PloS One*, 10(5):e0127298.

Wang, L. L., Lo, K., Chandrasekhar, Y., Reas, R., Yang, J., Burdick, D., Eide, D., Funk, K., Katsis, Y., Kinney, R., et al. (2020b). Cord-19: The COVID-19 open research dataset. *ArXiv*.

Wang, Q. and Schneider, J. W. (2020). Consistency and validity of interdisciplinarity measures. *Quantitative Science Studies*, 1(1):239–263.

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., and Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science*, 330(6004):686–688.

Wuchty, S., Jones, B. F., and Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, 316(5827):1036–1039.

Wynants, L., Van Calster, B., Collins, G. S., Riley, R. D., Heinze, G., Schuit, E., Albu, E., Arshi, B., Bellou, V., Bonten, M. M. J., Dahly, D. L., Damen, J. A., Debray, T. P. A., de Jong, V. M. T., De Vos, M., Dhiman, P., Ensor, J., Gao, S., Haller, M. C., Harhay, M. O., Henckaerts, L., Heus, P., Hoogland, J., Hudda, M., Jenniskens, K., Kammer, M., Kreuzberger, N., Lohmann, A., Levis, B., Luijken, K., Ma, J., Martin, G. P., McLernon, D. J., Navarro, C. L. A., Reitsma, J. B., Sergeant, J. C., Shi, C., Skoetz, N., Smits, L. J. M., Snell, K. I. E., Sperrin, M., Spijker, R., Steyerberg, E. W., Takada, T., Tzoulaki, I., van Kuijk, S. M. J., van Bussel, B. C. T., van der Horst, I. C. C., Reeve, K., van Royen, F. S., Verbakel, J. Y., Wallisch, C., Wilkinson, J., Wolff, R., Hooft, L., Moons, K. G. M., and van Smeden, M. (2020). Prediction models for diagnosis and prognosis of COVID-19: Systematic review and critical appraisal. *BMJ*, page m1328.

Xu, Y., Liu, X., Cao, X., Huang, C., Liu, E., Qian, S., Liu, X., Wu, Y., Dong, F., Qiu, C.-W., et al. (2021). Artificial intelligence: A powerful paradigm for scientific research. *The Innovation*, 2(4).

Yang, G.-Z., J. Nelson, B., Murphy, R. R., Choset, H., Christensen, H., H. Collins, S., Dario, P., Goldberg, K., Ikuta, K., Jacobstein, N., et al. (2020). Combating COVID-19—The role of robotics in managing public health and infectious diseases.

Yegros-Yegros, A., Rafols, I., and D'este, P. (2015). Does interdisciplinary research lead to higher citation impact? The different effect of proximal and distal interdisciplinarity. *PloS One*, 10(8):e0135095.