https://doi.org/10.51408/issi2025_125

How is the Sino-US AI Collaboration Reshaped by the China Initiative?

Dingkang Lin[1], Jiang Li[2]

[1] 652023140004@smail.nju.edu.cn, [2] lijiang@nju.edu.cn
Nanjing University, School of Information Management, 210023 Nanjing (China)

Abstract

The China Initiative, launched by the United States in 2018, has significantly reshaped scientific collaboration patterns between China and the U.S., particularly in the field of artificial intelligence (AI). This study examines the evolving dynamics of Sino-US AI research collaboration, focusing on the post-2018 period marked by geopolitical tensions, by using a comprehensive dataset from DBLP and DBLP-Citation-Network-v16. Our analysis reveals that (1) collaboration between the two nations shows a reversed U-shape where the peak is 2019, (2) China shifts its international collaboration to the EU and the U.S. strengthens ties with Canada, and (3) the AI subfield computer vision experiences the most pronounced impact under the China Initiative, because new collaboration in this field dramatically decreases and existing collaboration is largely suspended, which highlights its vulnerability to geopolitical disruptions.

Introduction

The Launch of the China Initiative in November 2018 in the United States (US) has significantly impacted scientific collaboration patterns of the US. US-based researchers have become increasingly cautious about engaging in collaborations with Chinese counterparts due to perceived risks and potential complications (Lee, 2022). This climate has particularly affected Chinese-American scientists, who have reported experiencing systemic discrimination and targeted scrutiny. The barriers to scientific mobility have become more pronounced, with Chinese scientists facing substantial obstacles in visiting US institutions. These challenges include increased visa denials and heightened bureaucratic hurdles, leading to a noticeable decline in their willingness to engage in US collaborations (Silver et al., 2020). Furthermore, the flow of Chinese students to US institutions has been significantly restricted, with limitations imposed on study fields and a marked decrease in enrollment numbers (Feder, 2019; Tang et al., 2021).

Prior to 2019, Sino-US research collaborations demonstrated consistent growth, predominantly funded by Chinese sources and characterized by a majority of Chinese first authorships (Lee & Haupt, 2020). However, post-2019 data reveals a concerning trend: both the absolute number of Sino-US collaborative publications and their proportion in global collaborative output have declined significantly (Tang, 2024; Wagner & Cai, 2022). At the individual researcher level, the China Initiative's impact is evident in productivity metrics. US scientists collaborating with China have experienced lower research output compared to those collaborating with other countries (Jia et al., 2024). Similarly, Chinese researchers engaged in US collaborations have shown decreased productivity and citation impact, prompting many to redirect their collaborative efforts toward domestic partnerships and

collaborations with other nations (Li & Wang, 2024). The differential impact of these collaboration shifts is particularly noteworthy given the varying degrees of reliance on international partnerships. US scientific innovation demonstrates greater dependence on international collaboration across multiple metrics, including patent filings and research publications (Jang et al., 2022; Wu et al., 2019). This suggests that the decline in Sino-US scientific cooperation may have more substantial implications for US research output and innovation capacity (Wagner & Cai, 2022). The field of AI (artificial intelligence) experienced increasing international collaboration prior to 2019, with the US and France maintaining central positions in global networks, while China emerged as hubs within developing countries' collaboration networks (Hu et al., 2020). However, recent trends indicate challenges. Okamura's (2023) global analysis observed declining multidisciplinary collaboration between China and the US post-2019, including in AI.

This study aims to systematically investigate the evolving dynamics of Sino-US collaboration in AI research, focusing on the question: *How has the China Initiative reshaped Sino-US collaboration in AI*? To address this question, we leverage a comprehensive dataset from DBLP. By categorizing AI research into ten distinct fields and employing robust methods for country attribution, we provide a nuanced analysis of collaboration trends, alternative collaborators of China/US, and potential explanations of the change, offering valuable insights into the broader implications for global AI innovation and scientific collaboration.

Methodology

Data Processing

The primary database used in this study is DBLP (Digital Bibliography & Library Project), an open-source bibliographic information database focused on major computer science publications. We retrieved the DBLP data on November 1, 2024. DBLP was chosen because it offers the most comprehensive collection of research papers published in both journals and conferences within the field of computer science.

The CCF (China Computer Federation) Recommended International Academic Publications Directory (2023 edition) lists 102 AI journals and conferences (CCF, 2023) (available at https://github.com/lindingkang/sino_us_ai_collaboration/blob/main/CCF_ai_conf_joun_2023.csv). Papers published in these venues were classified as AI research papers in our study. All journals and conferences but one are indexed in DBLP, i.e., Journal of Speech, Language, and Hearing Research. Consequently, our initial dataset includes 543,626 papers published in these 101 venues.

DBLP does not provide information on author affiliations. We hence utilized DBLP-Citation-Network-v16, developed by Tang et al. (2008), to augment our dataset in this regard. To address the inconsistencies in the writing of affiliations, we employed four distinct methods to determine the country of each author of the 1,388,182 author pairs from 440,797 articles that had complete records in the dataset: institutional

matching, country matching, manual matching, and AI-assisted matching, as follows,

- Institution-matching: Matching institutions to countries using OpenAlex institution information. We utilized the entire OpenAlex database, encompassing all institutions and their corresponding country information, to perform full-text matching with each author's affiliation texts using all available names, including those in different languages, alternative names, and other variants.
- Country matching: Using country names, aliases, and abbreviations for unmatched cases.
- Manual matching: Manually assigning countries to texts appearing over 30 times.
- AI-assisted matching: Inquiring with DeepSeek regarding the countries associated with the remaining affiliations.

As a result, a total of 1,165,155 texts were successfully matched, and upon conducting a manual verification of a 150-sample subset, we confirmed an accuracy rate of 100%. Following the mapping of all affiliations, we acquired 343,297 papers. By applying a filter for the years 2013 to 2022, we arrived at a final dataset comprising 180,821 articles authored by 237,741 individuals.

We constructed ten subfields of AI, by integrating the subfields from the AI Act by the European Union (https://artificialintelligenceact.com/understanding-ai-types-of-ai/) as well as insights from AI professionals, including machine learning, natural language processing, computer vision, cognitive computing, rule-based AI, robotics, multi-agent systems, expert systems, natural computing, and generative AI. Then, we categorized each journal or conference to one or more of the subfields according to the perspectives derived from large language models (LLMs) and AI professionals (available at https://github.com/lindingkang/sino_us_ai_collaboration/blob/main/CCF_ai_conf_joun_2023.csv). Papers published in a given journal or conference were assigned to the field(s) associated with that venue.

Measures

Okubo et al. (1992) proposed the *Affinity* index, defined as $C_{x,y}/Cx$, where $C_{x,y}$ represents collaborative publications between countries x and y, and C_x is country x's total international collaborations. In this study, we applied its variant to quantify the Sino-US collaboration, i.e.,

$$Affinity = \frac{C_{x,y}}{\sqrt{C_x * C_y}}. \quad (1)$$

We classified Sino-US author pairs in papers into two categories: *existing* collaboration and *new* collaboration. The delineation between old and new collaborations was anchored by the year 2019. Specifically, during the 2019-2022 period, an author pair in a paper was considered to have an *existing* collaboration if they had co-authored an AI-related paper in or before 2018. In contrast, if a pair had

no record of co-authoring any AI-related paper prior to 2018, it was labeled as *new* collaboration. During the 2013-2018 period, for a given year, an author pair in a paper was deemed to have an existing collaboration if they had previously co-authored an AI-related paper before that year. Conversely, if there was no prior record of them co-authoring any AI-related paper before the given year, the pair was classified as a new collaboration.

Based on the definitions above, we classified papers into two groups. Papers where all author pairs were *existing* collaborations were labeled as *existing* collaborations, while papers that included at least one *new* author pair were classified as new collaborations. In the computation of the affinity index, the denominator remained the total number of international collaborative publications between the two countries, while the numerator was the count of either new or existing collaborative papers between them. It should be emphasized that the combined count of *new* and *existing* pairs is not equivalent to the overall number of author pairs from 2019 onwards. This discrepancy signifies author pairs that initially emerged post-2018 but then reoccurred in later years, thereby illustrating the evolving characteristics of collaborative relationships across different time periods.

Furthermore, we characterized *disappeared* collaboration as referring to Sino-US author pairs who were present in papers published in or before 2018 but were absent from publications in any year subsequent to 2018.

Result

China has seen a more pronounced increase in the quantity of AI publications compared to the US, as illustrated in Figure 1. Although the number of US AI publications rose sharply, a downturn emerged in 2022. In terms of international collaboration on AI publications, both China and the US have experienced growth, but this growth rate is slower than that of the overall number of AI publications.

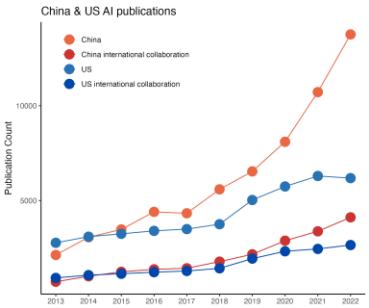

Figure 1. Trends in AI publications of China and US between 2013 and 2022.

Following the initiation of the China Initiative in 2018, there has been a significant rise in Sino-US collaboration in terms of AI publication counts, despite a slight deceleration in the growth rate, as depicted in Figure 2a. In parallel, Sino-EU collaboration in AI publication counts has been on a steady upward trajectory. Regarding the US, its collaborations with both the EU and Canada have increased, albeit at a pace that is not as rapid as that observed with China. To a certain extent, the upward trend in all four of these collaboration curves in Figure 2a can be

attributed to the overall increase in the number of AI publications in China and US, as shown in Figure 1. Accordingly, we leverage the Affinity index to mitigate the effect of publication counts.

The upward trajectory of Sino-US collaboration in AI (measured by the Affinity index) shifted dramatically to a downward trend, forming a reversed U-shape as illustrated in Figure 2b. The peak in the curve in 2019 may likely be due to publication delays. In contrast, China redirected its international collaboration towards the EU among all other countries/territories, while the US shifted its focus to Canada. The EU is not an alternative collaborator for the US, as their collaboration has seen a significant decline since 2019.

Next, we turn our attention to the declining Sino-US collaboration. Figure 2c demonstrates that both *existing* Sino-US collaborations and *new* collaborations (measured by Sino-US author pairs) have been sharply decreasing since 2019. It is important to note that the values in Figure 2c do not represent the number of Sino-US co-authored publications, but rather the Affinity index of Sino-US collaboration, which is divided into *new* and *existing* categories. Clearly, the decline in *existing* collaborations is more pronounced than that of *new* collaborations, suggesting that the overall decrease in Sino-US collaboration is primarily due to the contraction of *existing* collaborative relationships.

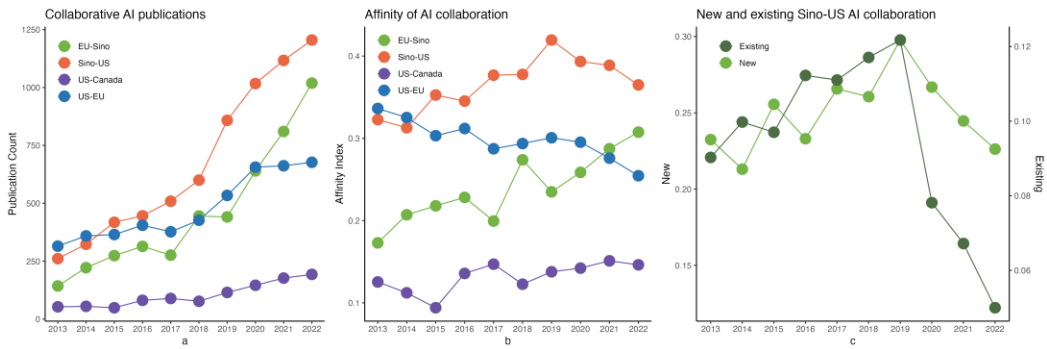

Figure 2. Sino-US collaboration in AI publications. 2a. Trends of Sino-US collaborative AI publications. 2b. Trends of Sino-US AI collaboration measured by the Affinity and their alternative collaborators. 2c. Trends of *new* and *existing* Sino-US AI collaboration.

Upon examining the subfields of AI, we observe that *new* collaborations in the majority of these subfields follow the general trend, with a steady decline since 2019. The most significant decreases are seen in robotics, natural language processing, and cognitive computing, with respective decline rates of 40.1%, 35.9%, and 35.9%, as depicted in Figure 3a. Figure 3b indicates that *existing* collaborations in most subfields also underwent a rapid decrease from 2019 to 2022, with rule-based AI, computer vision, and expert systems being the most impacted, experiencing decline rates of 61.9%, 61.6%, and 57.4% respectively.

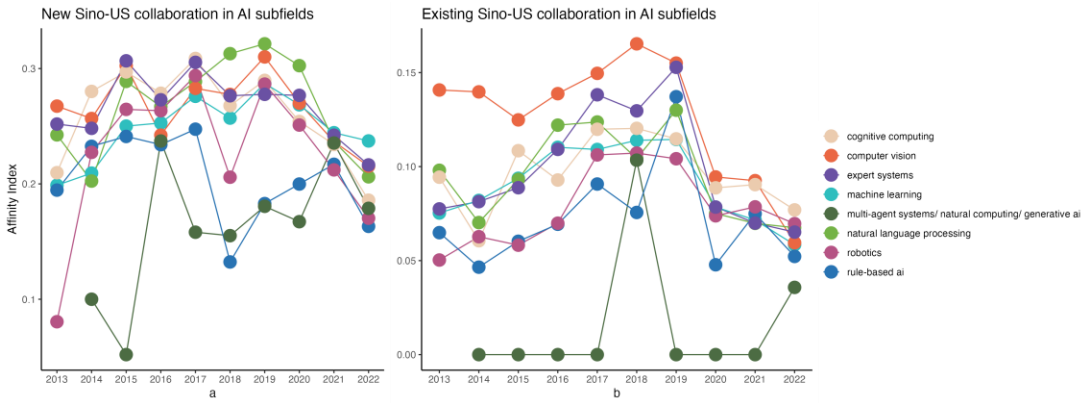

Figure 3. Trends of Sino-US collaboration in ten AI subfields (Due to limited publications, the fields of "multi-agent systems," "natural computing," and "generative AI" are combined into one). 3a. Affinity index of *new* collaboration in each AI subfield. 3b. Affinity index of *existing* collaboration in each AI subfield.

The decline of existing collaborations in Figure 2c indicates the disappearance of Sino-US collaboration. It is verified that a significant number of Sino-US author pairs have vanished since 2019, as indicated in Figure 4a. The count of such *disappeared* pairs skyrocketed to over 5,000 in 2020 and has stayed at a high level since then. Upon examining the subfields, it was found that computer vision was the most heavily impacted area, as illustrated in Figure 4b.

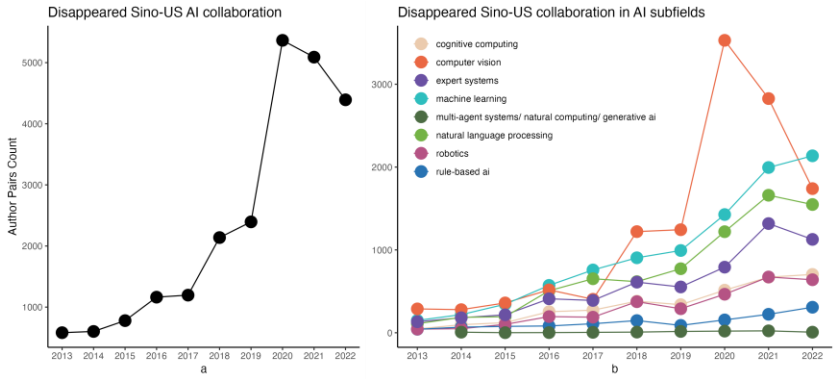

Figure 4. Trends of disappeared Sino-US AI collaboration. 4a. Disappeared Sino-US AI collaboration. 4b. Disappeared Sino-US collaboration is AI subfields.

Preliminary findings

This study provides a comprehensive analysis of the Sino-US collaboration in AI research after the launch of the China Initiative in 2018, leveraging 101 AI-related journals and conferences indexed in DBLP and DBLP-Citation-Network-v16. We delved into the ten distinct AI subfields to explore why changes happened.

The initial findings are as follows: (1) we identified a reversed U-shaped pattern in Sino-US AI collaboration from 2013 to 2022, with the peak occurring in 2019. The significant decline in Sino-US collaboration can be attributed to a sharp reduction in

both *new* and *existing* collaborative efforts. (2) In response to the China Initiative, China has turned to the EU as an alternative partner in AI, while the US has primarily looked to Canada for collaboration. (3) The AI subfields of computer vision has been most heavily affected by the China Initiative. This is due to a steep decrease of *new* collaborations and a near suspension of *existing* collaborations.

This study offers initial statistical insights, with the analysis grounded in observational findings rather than causal inferences. Moving forward, we aim to apply a difference-in-differences approach to rigorously establish causality, validate the current observations, and delve deeper into the underlying factors driving these trends.

Acknowledgments

This study is financially supported by the National Social Science Fund Major Project (24&ZD321). We extend our sincere gratitude to Dr. Yanbo Wang for his constructive comments, and Professor Jie Tang and his team for their invaluable assistance with data access and processing.

References

CCF. (2023). *CCF Recommended International Academic Publications Directory*. https://www.ccf.org.cn/Academic_Evaluation/By_category/

Feder, T. (2019). Trade wars and other geopolitical tensions strain US-China scientific collaborations. *Physics Today, 72*(11), 22-26. https://doi.org/10.1063/Pt.3.4338

Hu, H. T., Wang, D. B., & Deng, S. H. (2020). Global Collaboration in Artificial Intelligence: Bibliometrics and Network Analysis from 1985 to 2019. *Journal of Data and Information Science, 5*(4), 86-115. https://doi.org/10.2478/jdis-2020-0027

Jang, B., Choung, J. Y., & Kang, I. (2022). Knowledge production patterns of China and the US: quantum technology. *Scientometrics, 127*(10), 5691-5719. https://doi.org/10.1007/s11192-022-04478-4

Jia, R. X., Roberts, M. E., Wang, Y., & Yang, E. D. (2024). The impact of US-China tensions on US science: Evidence from the NIH investigations. *Proceedings of the National Academy of Sciences of the United States of America, 121*(19). https://doi.org/10.1073/pnas.2301436121

Lee, J. J., & Haupt, J. P. (2020). Winners and losers in US-China scientific research collaborations. *Higher Education, 80*(1), 57-74. https://doi.org/10.1007/s10734-019-00464-7

Li, M. X., & Wang, Y. (2024). Influence of political tensions on scientific productivity, citation impact, and knowledge combinations. *Scientometrics, 129*(4), 2337-2370. https://doi.org/10.1007/s11192-024-04973-w

Okubo, Y., Miquel, J. F., Frigoletto, L., & Dore, J. C. (1992). Structure of International Collaboration in Science - Typology of Countries through Multivariate Techniques Using a Link Indicator. *Scientometrics, 25*(2), 321-351. https://doi.org/Doi 10.1007/Bf02028090

Tang, J., Zhang, J., Yao, L., Li, J., Zhang, L., & Su, Z. (2008). *ArnetMiner: extraction and mining of academic social networks* Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, Las Vegas, Nevada, USA. https://doi.org/10.1145/1401890.1402008

Tang, L. (2024). Halt the ongoing decoupling and reboot US-China scientific collaboration. *Journal of Informetrics, 18*(2). https://doi.org/10.1016/j.joi.2024.101521

Tang, L., Cao, C., Wang, Z., & Zhou, Z. (2021). Decoupling in science and education: A collateral damage beyond deteriorating US-China relations. *Science and Public Policy, 48*(5), 630-634. https://doi.org/10.1093/scipol/scab035

Wagner, C. S., & Cai, X. (2022). Changes in co-publication patterns among China, the European Union (28) and the United States of America, 2016-2021. *arxiv*. https://doi.org/https://arxiv.org/abs/2202.00453

Wu, L. F., Zhu, H. Y., Chen, H. C., & Roco, M. C. (2019). Comparing nanotechnology landscapes in the US and China: a patent analysis perspective. *Journal of Nanoparticle Research, 21*(8). https://doi.org/10.1007/s11051-019-4608-0