How Does AI Assist Scientific Research Domains? Evidence Based on 26 Millions Research Articles

Qianqian Xu [1], Jie Meng [2], Jiangen He [3] , Wen Lou[1,*]
[1] *Faculty of Economics and Management, East China Normal University, Shanghai, China, 200062*
[2] *University of Chinese Academy of Sciences, Shanghai, China, 200025*
[3] *School of Information Sciences, The University of Tennessee, Knoxville, TN, USA, 37996-0150*

Abstract
Researchers are paying a lot of attention to artificial intelligence (AI). The development of AI in research and its contributions to other fields remain largely unexplored, nonetheless. We have identified a total of 435994 papers across all disciplines that involve AI methods based on 26 millions articles in the Web of Science database. Using bibliometrics and visualization techniques, the evolution process of AI technology in scientific research is identified. The results of the study show that the growth patterns of AI technology are consistent with the fundamental properties of emerging technologies, characterized by rapid growth, new applications, and changes in size.

Keywords
Artificial intelligence, Bibliometrics, Scientific Fields

1. Introduction

The impact of technology on society is profound. It has transformed not only the way people live, but also affected every branch of science. Significant publicity has been given to AI and its application in science, the renewed scholarly interest in AI, including the bibliometrics community.

The goal of this study is to reveal how does AI assist scientific research domains, and to provide guidance for future research efforts. We conducted a search of 26 millions AI articles from 2000-2019, limiting the study to those that utilized AI technologies. We reveal different perspectives on how AI assists the scientific research domains and the behavioral patterns exhibited by disciplines following AI trends.

2. Data Collection And Methods

Our data source is the Web of Science (WoS) database, since the subsequent search requires search term matching for abstracts, the abstracts were processed for missing values, and after removing the articles with missing abstracts, a total of 26,408,350 articles were obtained. We used AI methods provided by the Papers with Code platform as a search term list to match article abstracts. Papers with Code provides a total of 2060 AI methods. To be conservative, we removed abbreviations as well as one-word terms and ensured that only unambiguous terms were included to prevent errors in matching. In addition, we retained the 50 most common AI terms. In our study, the final list of search terms contains 1960 AI methods.

For our analysis, we used the collection of articles from WOS 2000-2019, and we conducted three search sessions on the dataset. In the initial search, we performed search term matching on abstracts using the findall method in Python, requiring at least one search term to be present in the article abstract. In the second round of search, we removed data that only appeared in methods such as logistic regression and linear regression. In the third round of search, we performed manual verification to ensure that AI methods were mentioned in the article abstracts, and we manually removed some articles with ambiguous terms. Finally, we identified a total of 435,994 articles.

3. Result

We collected a total of 435,994 articles on AI research, accounting for 1.6% of all articles. AI has been a hot topic in recent years, but our data shows that the application of AI technology is not yet common in most fields.

Joint Workshop of the 4th Extraction and Evaluation of Knowledge Entities from Scientific Documents and the 3rd AI + Informetrics (EEKE-AII2023), June 26, 2023, Santa Fe, New Mexico, USA and Online
*Corresponding author.
EMAIL: xuqq0427@gmail.com (Q. Xu); moonjaymengjie@gmail.com (J. Meng); jiangen@utk.edu (J. He); wlou@infor.ecnu.edu.cn (W. Lou);

A total of 449 AI methods appear in the AI research articles we collected, and Figure 1 shows the distribution of AI method vocabulary counts. Most of the AI method vocabulary occurrences are concentrated between 0 and 10. However, the percentage of terms decreases as the number of occurrences increases. Although a large number of AI methods have emerged, the AI methods used in the articles we collected are still few and concentrated in a few common methods.

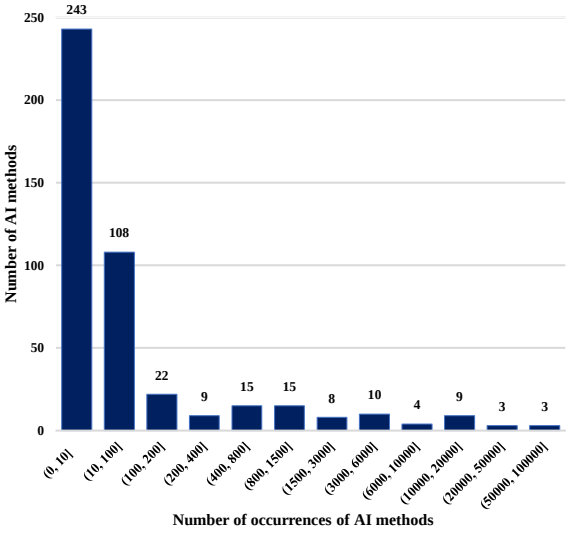

Figure 1: AI method counts distribution

Exploring the popularity of AI methods, we found that bayesian analysis, neural network, and principal component analysis are the most commonly used methods among researchers. Table 1 provides the Top 10 AI methods. Bayesian, neural network and principal component analysis are popular AI methods because they have demonstrated superior performance and have a well-established theoretical and algorithmic basis, which makes them easier for researchers to understand and apply in their respective fields. Furthermore, these methods have been widely studied and applied, leading to a growing body of literature that supports their effectiveness and reliability.

Table 1

Top10 AI methods

AI methods	%
bayesian	19.09
neural network	16.30
principal component analysis	14.16
machine learning	8.42
artificial neural network	7.95
support vector machine	7.36
particle swarm optimization	3.91
linear regression	3.02
random forest	2.90
decision tree	2.77

Research output has increased not only in absolute numbers but also in the total number of papers equivalent to a given scientific field, albeit at a lower level. AI articles account for 1.6% of all papers, meaning that AI articles still represent only a small fraction of the overall research volume. However, recent growth rates in these shares are remarkable.

Our data confirm the explosive period of research activity in all scientific fields (Figure 2), with a high growth rate of AI articles around 2005, around 15%, then experiencing some decline around 2010, then gradually recovering and growing steadily. The growth rates rapidly increased after 2017, reaching 30% by 2019.

Figure 2: Trends in AI publication activity

The proliferation of AI as an emerging technology from initial fields to application areas follows a "double boom" cycle [1], a theory illustrated by the growth patterns of AI in All disciplines. In the initial phase, AI seemed to have a lot of potentials and was given high expectations, triggering considerable development efforts, especially on the theoretical side, which led to the first AI boom. However, during the early development process, some participants found difficulties in translating theory into practice, and most companies failed and stopped their innovation activities, which led to the end of the first boom. But some continued, and over time they overcame some of the more important practical hurdles and showed progress in the application domain, opening the second boom.

4. Conclusion And Future Work

We use bibliometric and visualization methods to study the assistance of AI technologies in different disciplines from 2000-2019. Although AI research has been receiving much attention in recent years, it is still not applied in most research areas. Our study found the same trend in the growth rate of the number of AI articles in all disciplines. AI is an

emerging technology and the growth trend is in line with the 'double-boom' cycle proposed by Schmoch [1]. Future research will focus on the assistance of AI to the scientific field by discipline, from four aspects: growth pattern, temporal diffusion, topic evolution, and research strategy.

5. References

[1] Ulrich Schmoch. 2007. Double-boom cycles and the comeback of science-push and market-pull. Research Policy 36, 7 (September 2007), 1000–1015. DOI:https://doi.org/10.1016/j.respol.2006.11.008.