Bibliometric research in the field of artificial intelligence

Wenyan Tu¹, Ji Fang¹, Fangyu, Jiao¹, Yutong Ci*¹

¹Library, Huazhong University of Science and Technology, Hubei, 430074, Wuhan, China

*Corresponding author’s e-mail: cba@hust.edu.cn

Abstract. This paper analyzes the development trend of global artificial intelligence from the perspective of bibliometrics. From the publishing trend, the number of papers published in the field of artificial intelligence is growing rapidly. China and the United States are among the top countries in the output of scientific research papers, and India has the fastest cumulative growth rate. In the research direction, the research on computer vision, image, deep learning and neural network is the research hotspot in recent years.

1. Introduction

Artificial intelligence, as a comprehensive high-tech discipline was born in the 1950. It is a multidisciplinary integrated psychology, cognitive science, information science, and biological science.[1]

The concept of artificial intelligence has gone through a long and brutal initial stage of growth since it was proposed at the Dartmouth Conference in 1956. Hawking argues that "artificial intelligence could also be the end of the history of human civilization unless we learn how to avoid danger"; Musk argues that artificial intelligence is "much more risky" and that it is a "fundamental risk that threatens human civilization". In recent years, with the widespread application of artificial intelligence technologies such as unmanned driving, face recognition, and intelligent voice. The development of a new generation of artificial intelligence technology, open group intelligence and autonomous operation, the discussion of artificial intelligence governance has once again been widely mentioned by all walks of life.[2]

In October 2016, the National Science and Technology Council released two national strategy reports on artificial intelligence, the National Artificial Intelligence Research and Research Strategic Plan and the Prepare for the Future of Artificial Intelligence Future of The Artificial Intelligence marks the beginning of a national effort for the era of artificial intelligence in the United States, the most influential and developed country in the world.[3]

As the U.S. education sector's AI market points out, AI's use in U.S. education will grow by 47.5 percent by 2021. In order to cope with the rapid development of artificial intelligence, the education field in the United States should accelerate and strengthen the training of AI-related talents.[4]

Artificial intelligence has developed rapidly in various countries in recent years. The overall development of Europe in this regard is better, and the development situation of each country has its own characteristics. In 2018, the German federal cabinet adopted the "key points of the artificial intelligence strategy of the federal government" jointly prepared by the Federal Ministry of economy and energy, the Federal Ministry of education and research and the Federal Ministry of labor and society.

AI has emerged as an area of importance for national competitiveness, yet also sees growing...
international collaboration. Several national and international AI policies and strategies have been put forth in recent years, as both causes and consequences of growing AI research ecosystems. This has led to increased scientific output through a variety of dissemination modes, including publications, preprints, conferences, competitions, and software.\[5\]

2. **Data sources and research methods**

**2.1. Data sources**

Based on all types of scientific papers published from 2010 to 2019 in Elsevier scival database, according to the asjc discipline classification system, under the computer science category, this paper defines the sub discipline as artificial intelligence, and obtains nearly 500 thousand relevant scientific papers in the field of artificial intelligence around the world. Then, according to the direct citation relationship between the literatures, we cluster the literatures to get the global research topics in the field of artificial intelligence. (Data retrieval time: March 1, 2021).

**2.2. Research methods**

In this paper, using the method of direct citation between literatures, we first generate research topic clusters from the collection of research literatures according to the citation relationship, and then judge and identify the research topic clusters with high degree of attention according to the cited frequency, browsed times and journal impact factors of the research topic literatures.

In the process of theme creation, Elsevier scival adopted a global micro model to cluster 70 million papers and references in all scientific fields from 1996 to 2019 in Scopus, and identified and formed nearly 9,600 research topics.\[6\]

Publications are clustered into Topics based upon a direct citation analysis. Where there is a weak citation link, there is a break and a new Topic is formed. Topic Clusters are formed by aggregating Topics with similar research interest together to form a broader, higher-level area of research. These Topic Clusters can be used to get a broader understanding of the research being done by a country, institution (or group) or researcher (or group), before drilling into the more niche underlying Topics.

![Figure 1. Direct citation model\[7\]](image)

The topics created by scival based on direct citation relationship will exist forever once they are generated, and a small number of new topics will be generated every year. The later literatures are added to the topics according to the citation relationship. The more the latest literatures are, the newer the topics are. The old topics will not disappear, but they may be in dormant state.

3. **Quantitative analysis of global research in the field of artificial intelligence**

**3.1. Posting trend**

From 2010 to 2019, a total of 498,006 research papers were produced in the field of artificial intelligence, which increased from 38,515 in 2010 to 95,577 in 2019, with a cumulative increase of 1.48 times (Figure 1). 2010-2015 is in a period of stable development, 2016-2019 is in a period of
rapid development, it can be seen that the development of artificial intelligence is more and more rapid.

From the overall trend of publishing, the annual growth of papers in the world is nearly twenty thousand after 2016, and nearly one hundred thousand papers published in 2019. It is expected that the global volume of papers published will exceed one hundred and twenty thousand in 2020.

![Figure 2. Global Posting Trend](image)

The publications in selected year range and identify the journals in which they are published. From the journals we can determine in which quartiles the publications belong. Scival divides the journal into four regions according to citescore value, which are Q1, Q2, Q3 and Q4. The journal partition can be used to judge the influence of the journal. The higher the influence of the journal, the higher the quality of the papers published in these better regions.

From 2010 to 2019, there are more and more high-quality research papers in the field of artificial intelligence, among which the papers published in Q1 journals account for the largest proportion, reaching 33.1%, and showing an increasing trend year by year.

![Figure 3. Trends in the publication of journals by district](image)

3.2. Country analysis

From the analysis of the countries or regions that publish papers, the total number of countries or regions in the field of artificial intelligence worldwide is 180, among which the top five countries are China (124,523), the United States (68,447), India (40,774), Japan (29,695) and the United Kingdom (25,655). From the perspective of the trend of publishing, the volume of papers issued by these five countries is on the rise, of which India has the fastest growth rate, with a cumulative growth of 6.3 times, which is far faster than the total global growth rate. Secondly, China has a cumulative growth of 1.3 times, which is basically consistent with the global average growth rate.

China's total number of papers is much higher than that of other countries. From 2010 to 2014, China's number of papers showed a decreasing trend. From 2015, the number of papers increased significantly. The total number of papers published in 2019 is about 2.5 times that of the United States and India, and 5 times that of Japan and the United Kingdom.
Table 1. The top five countries’ trends in the publication.

| Country | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 |
|---------|------|------|------|------|------|------|------|------|------|------|
| China   | 11413| 9123 | 7401 | 8006 | 7585 | 8662 | 13007| 14538| 18361| 26337|
| USA     | 5025 | 5032 | 4688 | 5273 | 5498 | 6220 | 7243 | 8489 | 9481 | 11498|
| India   | 1385 | 1401 | 1532 | 1943 | 1460 | 1577 | 4146 | 6123 | 10995| 10212|
| Japan   | 2737 | 2421 | 2554 | 2629 | 2431 | 2098 | 2894 | 3316 | 3997 | 4618 |
| UK      | 1957 | 1867 | 1880 | 1959 | 2388 | 2492 | 2787 | 2966 | 3488 | 3871 |

3.3. Institutional analysis

From 2010 to 2019, 10,369 institutions published research papers in the field of artificial intelligence. Among them, the Chinese Academy of Sciences published 7,406 papers. At the same time, the citation frequency of papers published by CAS also ranks first in the world. Among the top ten institutions in the world, seven institutions come from China, accounting for 70%, followed by one from the United States, one from France and one from Singapore. Among the top ten institutions cited, there are six from the United States, accounting for 60%, and only one from the Chinese Academy of Sciences. It can be seen that Chinese institutions have a certain advantage in the number of papers, but the overall influence is not enough. On the contrary, the overall quality of scientific research papers in American institutions is higher, and the degree of attention is also higher.

Table 2. Top 10 organizations in terms of scholarly output and citation frequency.

| Institution                          | Country         | Output | Institution                          | Country         | Citations |
|--------------------------------------|-----------------|--------|--------------------------------------|-----------------|-----------|
| Chinese Academy of Sciences          | China           | 7406   | Chinese Academy of Sciences          | China           | 96036     |
| CNRS                                 | France          | 5976   | Microsoft USA                        | United States   | 89929     |
| Harbin Institute of Technology       | China           | 3817   | Massachusetts Institute of Technology| United States   | 70342     |
| Tsinghua University                  | China           | 3685   | Alphabet Inc.                        | United States   | 66812     |
| Nanyang Technological University     | Singapore       | 3279   | CNRS                                 | France          | 66635     |
| Carnegie Mellon University          | United States   | 3042   | Stanford University                  | United States   | 62310     |
| Beihang University                   | China           | 2985   | Nanyang Technological University     | Singapore       | 61953     |
| Beijing Institute of Technology      | China           | 2865   | Carnegie Mellon University          | United States   | 61924     |
| Zhejiang University                 | China           | 2727   | Institut national de recherche en informatique et en automatique | France | 59926 |
| Shanghai Jiao Tong University        | China           | 2701   | University of California at Berkeley | United States   | 50657     |

3.4. Trend analysis of research topics

Prominence is a new indicator that shows the current momentum of a topic by looking at very recent citations, views and CiteScore values.

- **Prominence = momentum (not the same as importance)**

  Therefore, the higher the prominence value is, it means that more and more researchers are paying attention to this topic, and it also shows that the growth momentum of this topic is more and more fierce. The percentile of prominence of each topic was calculated according to the actual prominence
In the field of artificial intelligence, there are more than four hundred thousand papers in the world, with a total of 3,380 research topics, 39 research topic clusters, and 7 topic clusters with prominence percentile ranking in the top 10% of the world. The highest priority is topic cluster 0, which includes artificial intelligence algorithm, computer vision and computer model. The number of research papers included in this topic cluster reaches 300,336, accounting for 60% of the total number of research papers in the field of artificial intelligence in the world.

| Topic Cluster                                                                 | Number | Output  | FWCI | Prominence |
|------------------------------------------------------------------------------|--------|---------|------|------------|
| Algorithms; Computer Vision; Models                                          | TC.0   | 300336  | 1.58 | 99.799     |
| Semantics; Models; Recommender Systems                                      | TC.37  | 119496  | 1.41 | 97.39      |
| Stroke; Gait; Rehabilitation                                                 | TC.48  | 78799   | 1.06 | 96.386     |
| Decision Making; Fuzzy Sets; Models                                         | TC.211 | 43520   | 1.19 | 93.307     |
| Classification (Of Information); Learning Systems; Algorithms               | TC.195 | 55445   | 1.2  | 92.972     |
| Robots; Robotics; Manipulators                                              | TC.101 | 65354   | 1.05 | 92.838     |
| Models; Social Networking (Online); Algorithms                              | TC.358 | 48448   | 1.5  | 91.767     |

4. Analysis of research hotspots in the field of artificial intelligence

We use text mining and apply a variety of Natural Language Processing techniques to the titles, abstracts and author keywords of the documents in the Topic Cluster in order to identify important keyphrases.

According to the frequency and correlation of phrases, we can get the key phrases in the literature, which also represent the main research content of the literature set. According to the change of the frequency of phrases, we can see the trend of the research topic cluster.

Based on the analysis of topic cluster 0, which has the largest number of papers and highest prominence percentile, this paper concludes that top 50 keywords in this topic cluster by relevance, based on 300,336 publications, and the growth trend of some phrases.

As can be seen from Fig.5, the hot research directions in the field of artificial intelligence include facial recognition, computer vision, neural network, camera, biometry, deep learning, object detection, etc.
From the change trend of phrases, the research on image processing, image recognition, image segmentation and other directions has an obvious growth trend, and the related research on neural network and deep learning has a rapid growth in 2015-2019. From 2010 to 2016, more attention has been paid to the research of computer vision. Since 2016, the research on neural networks has surpassed the research on computer vision. After 2017, the research on deep learning began to surpass the research on computer vision.

5. Conclusion
In a word, since the concept of artificial intelligence was put forward in the 1950s, artificial intelligence has developed into a research subject with great application value. In the past 10 years, the number of scientific research papers in the field of artificial intelligence has increased dramatically. Many countries in the world have promulgated the development strategy of artificial intelligence, especially in the United States, China, India and European countries. The scientific research investment and talent reserve of artificial intelligence have increased significantly. The competition in the field of artificial intelligence has become increasingly fierce. It can be predicted that artificial
intelligence will be widely used and has great commercial value. Speeding up the investment of talents and scientific research in these directions will help to keep a place in the fierce competition in the future.

Reference
[1] Zou L., Zhang X.f. (2012) Artificial Intelligence and Its Development and Application. Information Network Security, 02:11-13.
[2] Ji W.D., Xu Y.C. (2021) The 2020 Global Top Ten Artificial Intelligence Governance Events and Prospects of Reform Trends. National Governance, Z2:2-10.
[3] He Z. (2016) Towards the era of artificial intelligence--Also on the strategic direction of American artificial intelligence and its reference to China's artificial intelligence strategy. E-government, 12: 2-10.
[4] Meghan B. C. (2017) AI in Education Will Grow Exponentially by 2021[EB/OL].
[5] Elsevier.(2018) Artificial Intelligence: How knowledge is created, transferred, and used--Trends in China, Europe, and the United States.
[6] Cui Y.H, Wang S.A., Gao X.W., Yang H., Cao X.W. (2018) Research frontier topic detection and feature analysis based on global micro model. Library and information work, 62 (15): 75-82.
[7] https://www.elsevier.com/__data/assets/pdf_file/0006/548313/Topic-Prominence-Advanced-Webinar.pdf