Visualization Analysis of Artificial Intelligence Technology in Higher Education Based on SSCI and SCI Journals from 2009 to 2019

https://doi.org/10.3991/ijet.v16i08.18447

Ji Wang (✉), Qinglong Zhan
Tianjin University of Technology and Education, Tianjin, China
jiwang@163.com

Abstract—Artificial intelligence (AI) in higher education has received much attention in recent years. The article mainly uses CiteSpace as a tool to analyze the data based on SSCI and SCI journals between 2009 to 2019, which are from the WOS database. The retrievals of topics include “artificial intelligence” and “higher education”. Besides, it uses the keyword co-occurrence function and cluster analysis function of CiteSpace for visualization analysis of AI technologies in higher education. The article elaborated on AI technologies in higher education, such as machine learning, neural networks, and briefly discussed the limitations of AI technologies and the future development trend of AI in higher education.

Keywords—Higher education; artificial intelligence; CiteSpace; Visualization

1 Introduction

Artificial intelligence (AI) has exerted an enormous impact on various fields in society. In the Horizon Report 2019 Higher Education Edition, it pointed out that AI is one of the important developments in educational technology for higher education [1]. AI applications related to teaching and learning in higher education are expected to grow significantly. AI has exerted a considerable impact on higher education and is undoubtedly one of the leading technologies to promote both the overall quality and instructional innovation of universities. In this context, higher education is influenced by AI in many aspects, and the two major areas are curricula and enrollment [2]. Besides, AI technologies have driven the reform of higher education, including teaching philosophy and teaching methods. AI applied in higher education has become a reality [3,4]. However, it is not developed enough, and its application in higher education is not widespread [5]. Accelerating the deep integration of these two fields requires multifaceted efforts that technology promotion is one way. Although AI in higher education has received plenty of attention, it is still unclear about relevant issues, such as the future development trend of AI in higher education. In order to further discuss, the article aims to review the current AI technologies in higher education in the ten years from 2009 to 2019. The specific aims to conduct the article are:
• To know the main AI technologies in higher education in ten years
• To find out the impact that AI technologies have on higher education
• To analyze the limitations of AI technologies found in this article
• To discuss the future development trend of AI in higher education.

To achieve the above aims, this study used CiteSpace, the knowledge mapping software to visually analyze AI technologies in higher education based on the WOS database. The results can draw the future development trend of AI technologies and the influence of AI on higher education. We found that AI technologies currently in use include: machine learning, neural networks, AI-based data mining, and virtual reality, and that intelligence inspired by the brain is the development direction of AI, which is likely to apply to the higher education field in the future.

2 Materials and Methods

2.1 Data source

For this article, we analyzed the data based on SSCI and SCI journals which are collected from the core data collection of the Web of Science (WOS) database, and the retrieval of the topics include: “artificial intelligence” and “higher education”. The search span is 2009-2019. The documents type is set to “article”. This article mapped 90 documents as sample documents in WOS. The process is to define the samples with the application of the inclusion criteria, which is shown in Figure 1.

2.2 Tool

CiteSpace is used as the knowledge map analysis tool to visually analyze the developments of AI technologies in higher education in this article. CiteSpace is a scientific knowledge graph analysis software developed by Professor Chaomei Chen of the School of Computing and Information of Drexel University (Drexel University) based on JAVA, and it uses the principles of bibliometrics to analyze keywords in literature, which can explore and mine the dynamic process of scientific research and can show the trend of scientific development [6].
2.3 Procedure

In this article, documents retrieved from WOS were exported in Reworks format and were named with the prefix “download_”, and then were imported into the CiteSpace software. In the software interface, the time span of the “Time Slicing” module was set to 2009-2019, and the time slice was set to 1. We selected TOP N in the “Selection Criteria” module and set N to 50, that is, select the 50 keywords with the highest frequency of citations per year to make the generated network more comprehensive. The rest were set as the default. After setting the conditions, execute GO. The software can generate the keyword co-occurrence map. Then use keyword clustering and logarithmic likelihood algorithm LLR to generate the keyword clustering knowledge map. The brief flow chart is shown in Figure 2.

Fig. 1. PRISMA diagram (referenced [7] and [8])
3 Results

3.1 Time map

The time map of changes in the number of documents can visually reflect the development of the subject area. Figure 3 is the time distribution of the number of papers and citations of papers published in 2009-2019 for sample documents. From the left of Figure 3, it can be observed that the number of documents has increased significantly over time, while the number of articles published in 2009-2014 is relatively small. Especially in 2014, the number of documents is the least. The number of documents has rebounded and increased significantly after 2015. The number of articles in 2016 and 2018 has decreased compared with the previous year, but the overall trend is rising. The number of articles published in 2019 is the largest, showing that in recent years, AI technologies in higher education have received more and more attention. The number of citations on the right of Figure 3 also shows a significant increasing trend, which means that the study of AI in higher education also has upside potential.
3.2 **Keyword knowledge map**

Keywords can represent a summary of the content of the article. The keyword knowledge map is conducive to keyword co-occurrence analysis, which can intuitively describe the relationship among keywords in the documents. Keyword co-occurrence analysis is utilized to analyze the hotspots of the application of AI technologies in higher education in this article. By observing and analyzing the frequency of keywords and the centrality of keywords, it can reflect what the hotspots in the field are. As shown in Figure 4, nodes represent keywords, and the connections of nodes represent the co-occurrence relationship among keywords. The size of a node describes the frequency that the keywords represented by the node appear. Keywords with higher frequency or higher centrality are the hotspots in the field, and keywords with high frequency and high centrality are the most representative. Table 1 lists the frequency and centrality of the top ten hot keywords.

The top ten hot keywords, according to the following Figure 4 and Tables 1 listed, are AI, education, design, machine learning, higher education, deep learning, future, neural network, and model. The frequency of other keywords including children, computer, perspective, evolutionary algorithm, future, and competence, is less than or equal to 3.
Table 1. Keyword frequency and centrality

<table>
<thead>
<tr>
<th>Rank</th>
<th>Keywords</th>
<th>Frequency</th>
<th>Centrality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Artificial intelligence</td>
<td>37</td>
<td>0.37</td>
</tr>
<tr>
<td>2</td>
<td>Education</td>
<td>44</td>
<td>0.44</td>
</tr>
<tr>
<td>3</td>
<td>Design</td>
<td>7</td>
<td>0.16</td>
</tr>
<tr>
<td>4</td>
<td>Machine learning</td>
<td>6</td>
<td>0.01</td>
</tr>
<tr>
<td>5</td>
<td>Higher education</td>
<td>6</td>
<td>0.10</td>
</tr>
<tr>
<td>6</td>
<td>Big Data</td>
<td>5</td>
<td>0.04</td>
</tr>
<tr>
<td>7</td>
<td>Deep Learning</td>
<td>5</td>
<td>0.07</td>
</tr>
<tr>
<td>8</td>
<td>Future</td>
<td>5</td>
<td>0.01</td>
</tr>
<tr>
<td>9</td>
<td>Neural network</td>
<td>4</td>
<td>0.10</td>
</tr>
<tr>
<td>10</td>
<td>Model</td>
<td>4</td>
<td>0.05</td>
</tr>
</tbody>
</table>

The frequency and centrality of AI (F=37, C=0.37), and the frequency and centrality of education (F=44, C=0.44) are in the top rank, which describes that artificial intelligence has a deep connection with education and that AI technology can vastly promote education. The keyword design (F=7, C=0.16) also has a very high research interest. Machine learning (F=6, C=0.01), deep learning (F=5, C=0.07) and neural network (F=4, C=0.10) are AI technologies in higher education (F=6, C=0.10). Model (F=4, C=0.05) refers to the AI algorithms.
3.3 Keyword clustering knowledge map

Cluster analysis of keywords aims to further explore the hotspots of the study on AI technologies in higher education. First, import sample documents into CiteSpace. Second, make keyword co-occurrence maps. Third, use keyword clustering and logarithmic likelihood algorithm LLR after clustering co-occurring keywords. A cluster visualization map is obtained and shown in Figure 5. The correspondence between cluster tags and the high-frequency keywords under their tags is shown in Table 2.

![Fig. 5. Keyword clustering knowledge map](http://www.i-jet.org)

<table>
<thead>
<tr>
<th>Cluster Label</th>
<th>High-frequency keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>#0 (activities and operations)</td>
<td>artificial intelligence</td>
</tr>
<tr>
<td>#1 virtual reality</td>
<td>augmented reality, children</td>
</tr>
<tr>
<td>#2 discipline development</td>
<td>machine learning, deep learning, education</td>
</tr>
<tr>
<td>#3 upgrade pathways</td>
<td>model</td>
</tr>
<tr>
<td>#4 assessment</td>
<td>evolutionary algorithm, competence</td>
</tr>
<tr>
<td>#5 neural networks</td>
<td>deep learning, neural network</td>
</tr>
<tr>
<td>#6 scientometric data mining</td>
<td>higher education</td>
</tr>
<tr>
<td>#7 case-based learning</td>
<td>future, design</td>
</tr>
<tr>
<td>#9 cloud computing</td>
<td>cloud computing</td>
</tr>
</tbody>
</table>

**Note:** Cluster #8 and Cluster #1 are both virtual reality, so cluster #8 is not shown.
According to the analysis of the keyword co-occurrence map and the keyword clustering map, the contents in the keyword clustering knowledge map involve machine learning, virtual reality, neural networks, subject development, scientometric data mining, evaluation, Case-based learning, and cloud computing.

Four areas of applications in AIEd [8] include: a) profiling and prediction; b) assessment and evaluation; c) adaptive systems and personalization; d) intelligent tutoring systems. We mainly focus on AI technologies in these four areas. Besides, we further discussed intelligent virtual reality.

**In the aspect of prediction:** Machine learning, educational data mining, and neural networks can get valuable results in prediction. In learner performance prediction, these technologies show the ability to learn based on historical experience, which can predict results of learners’ final grades [9], learners who have difficulties in learning [10], the potential of learners’ performance, and can help improve learners’ learning quality [11-15] as well as instructors’ performance [16]. However, using historical data as predictive variables has a certain landmark but lacks accuracy and adaptability. Algorithms perceiving changes in the environment to perform calculations are forward-looking compared to training based on historical data. In addition to this, different algorithms have different prediction effects for complex and diverse data. Selecting suitable AI algorithms has a positive impact on the prediction. Comparing and analyzing the prediction results of several types of algorithms trained by reasonably designed data can successfully select the appropriate algorithm for prediction [17,18]. In dropout prediction, these technologies can take advantage of algorithms and data to identify factors affecting dropout and learners who are at risk of dropping out in online learning (such as MOOC or e-learning courses) [19-21]. With the increase in the types of data that can be calculated, the amount of data has extended to a large extent, increasing the pressure of calculation. The single algorithm is weak in the face of increasingly complex data. Multiple algorithms may show valuable performance in this context, but they also add to the complexity and difficulty of the calculation, making the calculation process opaque.

**In the aspect of assessment and evaluation:** AI is a powerful tool to open the “black box of learning”, which expands the dimensions of education evaluation, and the effectiveness of evaluation are also realized. The AI-based evaluation system is used to accurately evaluate the learning situation of learners [22,23]. Related technologies such as machine learning and educational data mining have made considerable contributions to the evaluation and are of great help to provide accurate guidance, which makes the evaluation method more flexible. They can identify learners’ learning patterns even difficulties in the learning process by analyzing the correlation between learners’ personal information[24]. Learners’ learning conditions, including learners’ mastery of subject knowledge, weaknesses, learning styles, learners’ attitudes towards subjects, can be utilized as variables for data analysis. AIAssess is an intelligent assessment software designed for learners to study science and mathematics, which can assess learners’ understanding and mastery of knowledge while they are studying. Argument maps supported by AI are a way to assist learners in higher education in demonstrating interaction and potential learning and can assess learners’ weaknesses. However, the complex process of information processing may lead to errors, which will affect the final
processing results and will lead to privacy issues and errors. Although the availability of data can enable technological breakthroughs, the limitations of reliance on data are gradually exposed.

In the aspect of adaptive learning systems and personalization: The blended learning model is widely used in the field of higher education. With the continuous development of diverse technologies, the environment as an important variable has become a source that technologies and algorithms can learn from. AI algorithms realize the adaptability of their functions by constructing adaptability models. Machine learning and neural networks can be used as adaptable technologies in adaptable education systems, successfully building the learner model [25]. In higher education, the degree of interaction between instructors and learners is relatively weak. The learning management system (LMS) as the adaptive teaching system can effectively understand the needs of learners [26]. In the online learning environment, learners can use the learning management system to carry out relevant learning operations. LMS under the blended learning model can adequately support the online learning environment (OLE) of higher education institutions (HEI) as a medium to mobilize the interaction between instructors and students to meet the learning needs of learners. For the purpose that instructors and relevant decision-makers understand the needs of learners, adaptive models provide educators with an environment where online course materials are placed and received by learners to interact with other learners or instructors. However, the adaptability of models built by machine learning and neural networks is not enough, and the adaptive model does not have mature general capabilities.

In the aspect of Intelligence Tutor Systems (ITS): With big data, deep learning, neural networks, and algorithms constantly being strengthened, the intelligent tutor system is progressing in an adaptive, personalized, and scientific direction. Technologies such as data mining and machine learning provide emotional intervention for encouraging learning motivation and improving learners’ performance in intelligent tutor systems [27-29]. The data in ITS is representative to a certain extent in reflecting emotional information. Intelligent tutor systems perform previous training on these technologies to obtain complete and accurate information from low-level data, effectively predicting learners’ emotions. These technologies successfully help systems collect emotional information, intervene in learners’ learning, and incorporate emotional dialogue and politeness principles to stimulate learners’ learning motivation as well as learning performance, which also serve to solve relevant problems. Intelligent virtual assistants have an outstanding performance in problem-solving [30,31]. Georgia Institute of Technology uses an AI-based virtual teaching assistant called Jill Watson to answer learners’ questions, which can tailor solutions to different learners’ problems based on learners’ questions. This type of AI relies on previous experience to recognize and answer questions raised by learners, and solve problems in specific fields. AI technologies have been trained by previous data and artificial labeling. When faced with atypical problems, the performance of the system may be unsatisfactory.

In the aspect of intelligent virtual reality: Currently, AI and virtual reality (VR) are showing a trend of mutual integration [32]. AI technology has the potential to apply in the field of virtual reality. The integrated products of virtual reality and AI are gradually being applied in the field of higher education. VR Job is an application that
combines the technical advantages of chatbots and virtual reality for training employees to improve interview results [33], which aims to help learners successfully employed. Intelligent virtual agents and virtual learning systems belong to the field of combining AI and virtual reality (VR). The combination of AI and VR is still in the development stage. In the future, the integration of AI and VR in higher education would produce deeper research.

4 Discussion and Conclusion

In this article, we mainly explored AI technologies in higher education. We concluded that AI technologies in higher education include machine learning, neural networks, AI-based educational data mining, and the relationship between AI and VR, furthermore, introduced these technologies applications. AI technologies show powerful performance in prediction as well as evaluation. Moreover, they can be used to build adaptable models to provide better learning services. In the intelligent tutor system, they provide emotional strategies to analyze learning and improve performance. The two fields of virtual reality and AI show a trend of convergence and intersection, which greatly improves the degree of interaction in education.

More importantly, AI technologies have exerted a considerable impact on higher education. AI was proposed at the Dartmouth University seminar in 1956 [34]. Many vital advances in artificial intelligence originated in the university environment. This positive upward trend is likely to continue to develop in the future [30]. The combination of these two areas complements each other and promotes mutual progress. According to the development of the times, there is no doubt that AI technology is continuously improving. Similarly, the interest in studying AI applications in higher education has also increased in recent years.

According to the results of the study, we summarized the limitations of related AI technologies in higher education, which include [35]: a) AI lacks adaptability and learning abilities comparable to humans, such as transferring learning, learning by analogy, which is caused by the strong dependence of intelligent computing on data. b) Calculation consumes a lot of energy, and the calculation process is opaque and lacks interpretability to users. c) AI can not understand the meaning of the data processed and easily involves privacy issues. Therefore, the development of AI technology is still in the progress. Although the predictive performance of different AI algorithms is different, the well-structured data is an important condition for the success of AI algorithms’ prediction in learners’ performance. Likewise, machine learning and neural networks can obtain good prediction results in a good context. The computational performance of algorithms trained on historical data can cope with typical situations but not non-typical situations, which may result in inaccurate results in prediction. Furthermore, the training of algorithms and models takes plenty of time and energy consumption. Due to the large-scale amount and the complexity of the data, the calculation process is opaque, lacking interpretation capabilities and existing certain biases and other related restrictions. As relevant technologies become more and more mature, prediction depends heavily on the training and completion of the data. The redundancy of the data would
hinder the calculation rate of the prediction. Although AI is a powerful driving force for VR and AR, the diversity of interaction is limited by AI technology. The integration of AI and VR for promoting the development of higher education is still in its infancy. The shortcomings of AI technologies are likely to limit the development of AI in higher education. So it is necessary to deepen the integration of AI and higher education from the perspective of enhancing technology.

According to the limitations concluded by the results, we concluded that the future development trend of AI is developing toward brain-like intelligence [36]. In higher education, the limitations of AI technologies have caused a large amount of energy loss and issues concerning privacy and ethics in the education system. Intelligence inspired by the brain can bring adaptive services to higher education, vastly reducing ethical implications and risks. With breakthroughs in brain science and neuroscience, AI is developing toward the direction of intelligence inspired by the brain. The brain has the characteristics of low energy consumption and fault tolerance [37], which is due to its structure and function. The brain is composed of many neurons linked together, and different brain regions are responsible for different functions. The structure of neuron communication and plasticity creates the characteristics of low energy consumption and the high response of the brain. Therefore, Large-scale brain simulations and research on computational models inspired by the brain are significant. Promoting the future development of AI, in this regard, is a necessary and feasible approach to realize the close connection between AI and higher education. Brain-inspired AI would get rid of the dependence on data in the future. Similarly, the environment would be an essential variable to stimulate AI learning, thus providing intelligence comparable to humans. With the maturity of AI in brain-like aspects in the future, it is possible to apply brain-inspired educational robots and brain-inspired teaching systems in higher education.

The importance of this article is to achieve the aims proposed at the beginning of the article. We discussed AI technologies in higher education by using CiteSpace software, explained the limitations of the technology based on the findings, and explored the future development trends of AI in higher education. It is an important implication of this review that we encourage technological updates to accelerate the integration of AI and higher education. Also, there may be some possible limitations in this article. The article used WOS as a core database. Therefore, the views represented by the results are limited. Co-word analysis and cluster analysis produce visual images, but they vary with different classification criteria so that the use of methods and the interpretation of results are subjective. Besides, this article mainly focused on the discussion of technologies, but it still needs to explore AI in the higher education field from other perspectives. In the future, we will conduct in-depth discussions on artificial intelligence and higher education topics from a future perspective.

5 References


6 Authors

Ji Wang is a postgraduate at the School of Information Technology, Tianjin University of Technology and Education.

Dr. Qinglong Zhan is a professor, associate dean graduate tutor at the School of Information Technology, Tianjin University of Technology and Education.