A Bibliometric Analysis of the Papers Published in the Journal of Artificial Intelligence in Education from 2015-2019

https://doi.org/10.3991/ijai.v2i1.14481

Clare Baek, Tenzin Doleck(*)
University of Southern California, Los Angeles, United States
doleck@usc.edu

Abstract—To analyze the current research status and trends of the artificial intelligence in education field, we applied bibliometric methods to examine the articles published in one of the representative journals of the field, International Journal of Artificial Intelligence in Education, from 2015 to 2019. We analyzed 135 articles retrieved from the Web of Science database and examined prolific countries, collaboration networks, prolific authors, keywords, and the citations the articles received. Through examining keywords, we found that the authors largely focused on students and learning. Through examining prolific authors and countries, we found active publication of corresponding authors from United States, United Kingdom, Canada, and Germany. We found international collaboration among some researchers and institutions, such as strong collaboration network between United States and Canada. We suggest reinforcement in building more widespread international partnership and expanding collaboration network by including diverse institutions. International collaboration and expanded institutional network can improve research by incorporating various perspectives and expertise.

Keywords—Artificial Intelligence in Education, Bibliometrics, Information Science, Mapping Research, Research Evaluation, Research Trends

1 Introduction

Artificial Intelligence has been evolving, expanding, and making notable impact on education since the very first conferences of the field started in 1980’s. Artificial Intelligence in Education (AIEd) has created innovative and effective environment for learners and instructors, and transformed the ways of assessment, providing feedback, collaboration, discussion, and delivering instruction. With the drastic and continuous improvement of technology and concomitant increase of computer science education for all k-12 settings as well as higher education, Artificial Intelligence has been increasingly used by diverse educators and learners [1,2]. AIEd plays a crucial role in expanding equitable access to computer science education as well as improving learning experience through technologies [1,3].
In the past five years between 2015-2019, AIED evolved to include a wide spectrum of learning platforms. For example, Intelligent Tutoring System provides personalized learning to meet the diverse needs of students through automated evaluation and personalized feedback [4]. Other examples include on-line learning platforms that support collaborative learning experience with the use of conversational agents and automated essay scoring for free text responses [5,6]. In addition, instructions that involve abstract concepts or require highly visual illustration can benefit from using virtual reality, such as virtual reality surgery training [7]. As educational technology continues to advance and computer science curriculum is increasingly embedded into many educational settings, the role of AIED and its practical impact on both global and intercultural dimensions will increase [7]. Therefore, it is timely and relevant to examine the status of the current literature in AIED.

For this study, we have chosen the International Journal of Artificial Intelligence in Education (IJAIED) to examine the status of the current literature of the artificial intelligence in education. The International Journal of Artificial Intelligence in Education (IJAIED) is one of the representative journals in the field of AIED [3]. Since its publication of the first volume in 1989, IJAIED has been publishing articles on various topics of application and design of artificial intelligence techniques that can support and enhance learning. The International Journal of Artificial Intelligence in Education (IJAIED) is the official journal of the International Artificial in Education Society, which is an interdisciplinary community of computer science, education, and psychology [8].

This article aims to provide in education scholars with a comprehensive understanding of the current research status of artificial intelligence in education by using bibliometrics to analyze the International Journal of Artificial Intelligence of the last 5 years, from 2015-2019. This article explores the most influential articles, prolific authors, scientific collaboration patterns and structure, and the main topics discussed in the journal in the period of 2015-2019. More specifically, first, it provides education scholars with a comprehensive understanding of the current research status of AIED. Second, this article identifies the prolific articles, authors, and collaboration structure presented in the IJAIED. Lastly, this article examines the evolution of the most frequently used keywords across the last 5-year time period. Research questions we attempt to answer are the following:

1. What are the most influential articles published in IJAIED from 2015-2019?
2. Who are the most prolific authors in IJAIED from 2015-2019?
3. What are the collaboration patterns and structure between countries in IJAIED from 2015-2019?
4. What are the most frequently used keywords in articles published in IJAIED from 2015-2019?

2 Literature Review

Bibliometrics is a technique that is useful in examining the current status and trend of a scientific field’s literature, which reveals various patterns of publications such as
productivity, collaboration, and prolific topics through visual and numerical illustration. Numerous studies in diverse fields have used bibliometric method. For example, Firdaus et al. used bibliometrics to examine the current trend and future directions of blockchain research [9]. Arici et al. revealed current research trends of augmented reality in science education by using bibliometric analysis and examined the most used keywords in articles [10]. In addition, they examined the most cited authors, journals, and most used research design types in the field. de Bem Oliveira et al. evaluated the structure of the production of scientific knowledge regarding the use of molecular markers for sugarcane studies [11]. While some studies examined the overall field using bibliometrics, other studies narrowed their focus on analyzing a specific journal in the field.

Analyzing one of the prolific journals that well represents a field using a bibliometric overview can convey the leading trends of the past and present, and reveal likely future developments of the journal [12-14]. Furthermore, such an exercise reveals the progress, evolving interest and themes, quality of research, current collaboration status, and future direction of the field [15]. For example, Martinez-Lopez et al. analyzed the leading trends of an international journal in the field of marketing, the European Journal of Marketing, by identifying the citation structure, prolific articles, and authors [13]. Prieto-Gutiérrez & Segado-Boj used bibliometrics to analyze Annals of Library and Information Studies, and compared its trends with other local journals in the field of library and information science [12]. Dereli et al. analyzed articles in the journal, Total Quality Management and Business Excellence, over 13 years of period on articles’ structure trend, methods used, frequently used keywords, and authors’ countries [16]. English et al. analyzed the journal, Religious Education: An Interfaith Journal of Spirituality, Growth and Transformation, over a 10-year period to evaluate implications for the journal and the field by looking at themes, research methodologies, institutional affiliation, contributor characteristics, and the pattern of research interests evolving in the field [15]. Chen et al. analyzed the journal, Computers & Education, to examine the research status and trends of the educational technology field [17]. In this analysis, Chen et al. discussed evolution of the main characteristics such as statistics of articles, citations, and authors published in Computers & Education and also examined top-ranked keywords along with their evolution across time.

3 Methodology

In the current study, we applied bibliometrics to examine the articles published in the IJAIED [18]. All articles published in the journal were retrieved from the Web of Science (WoS) database on Jun 10, 2019. We searched for the full publication name “The International Journal of Artificial Intelligence in Education” in the publication name field. The publications were retrieved in the bibtext file format for analysis. Our inclusion criterion was that the articles have to be published in IJAIED between the period of 2015-2019. This resulted in a total of 135 unique articles including 3 review articles. All 135 publications were included in our analysis. Information such as bibliographic information of the articles, annual citations, title, abstract, publication year,
authors, authors’ institutions, and countries were extracted for analysis. The retrieved data was analyzed using the Biblioshiny in the Bibliometrix package of R [18], which is commonly used for quantitative research in bibliometrics. The parameter of bibliometric analysis was set to include top 20 such as top 20 most local cited sources.

4 Findings

4.1 Main information about the articles

Table 1 shows the main information about the 135 articles we collected. The articles had 364 authors total and most of the articles had more than one author with 336 multi-authored articles and only 28 single-authored articles. Average authors per article was 2.7, average co-authors per article was 3.13, and average number of articles per author was 0.371. This suggests that most authors collaborated with two to three authors. Average number of articles per author is less than one and the number of author appearances is greater than the total number of authors, which shows that some authors have multiple publications.

<table>
<thead>
<tr>
<th>Description</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Articles</td>
<td>135</td>
</tr>
<tr>
<td>Article</td>
<td>132</td>
</tr>
<tr>
<td>Review article</td>
<td>3</td>
</tr>
<tr>
<td>Keywords Plus</td>
<td>288</td>
</tr>
<tr>
<td>Author’s keywords</td>
<td>540</td>
</tr>
<tr>
<td>Period</td>
<td>2015-2019</td>
</tr>
<tr>
<td>Average citations per articles</td>
<td>4.178</td>
</tr>
<tr>
<td>Authors</td>
<td>364</td>
</tr>
<tr>
<td>Author Appearances</td>
<td>422</td>
</tr>
<tr>
<td>Authors of single authored articles</td>
<td>28</td>
</tr>
<tr>
<td>Authors of multi-authored articles</td>
<td>336</td>
</tr>
<tr>
<td>Articles per author</td>
<td>0.371</td>
</tr>
<tr>
<td>Authors per article</td>
<td>2.7</td>
</tr>
<tr>
<td>Co-authors per articles</td>
<td>3.13</td>
</tr>
<tr>
<td>Collaboration index</td>
<td>3.2</td>
</tr>
</tbody>
</table>

The most local cited sources are listed in Fig. 1. This shows sources such as journals or conference proceedings that were included in at least one of the reference lists of the article set from 2015-2019 in IJAIED. The most cited journals are *International Journal of Artificial Intelligence*, with 315 citations, and *Lecture Notes in Computer Science*, with 284 citations. Following these two most cited journals, *Computers & Education* and *International Journal of Artificial Intelligence in Education* are the third most cited sources.
4.2 Authors

A total of 364 authors published articles between the years of 2015-2019. Fig. 2 depicts Lotka’s Law, which computes the frequency distribution of scientific productivity [19]. Here, Lotka’s Law reveals that about 90% of the authors (329 authors) have one publication and about 6% of the authors have 2 publications. There are outliers, a very few authors with more than 4 publications.

Fig. 1. Most local cited sources in IJAIED from 2015-2019

Fig. 2. Frequency distribution of scientific productivity shown through Lotka’s Law
Fig. 3. Author impact by H-index

Fig. 3 lists the top 20 authors with the highest impact by H-index, and Fig. 4 lists most relevant authors based on the number of articles they published. The Hirsh index, generally referred to as H-index, indicates the minimum number of times an author or a journal has been cited. Koedinger has by far the highest H-index and most number of articles. Koedinger and Aleven ranked in the top 5 for both authors with the highest H-index and most number of articles. This suggests that these authors are the most productive authors with active and frequent publication in IJAIED for the period of 2015-2019.

Fig. 4. Most relevant authors per number of articles
Authors’ collaboration network is presented in Fig. 5. The thickness of the radial represents the number of collaborations. Here, Koedinger and Aleven appear to have a strong collaboration network with each other, represented by a thick radial. Koedinger, McLaren, and Aleven collaborate all together and Koedinger and McLaren collaborate together. Allen and McNamara also have notable number of collaborations together. It is interesting to note that Koedinger, Aleven and McLaren are all listed as relevant authors (Fig. 4). This suggests that there is active collaboration amongst the most productive authors.

![Author's collaboration network](image)

**Fig. 5.** Author’s collaboration network

4.3 Countries

Fig. 6 lists 20 countries of corresponding authors with most number of articles. Authors from United States have the most number of articles, followed by Canada and United Kingdom. MCP indicates multiple countries publication, which is the number of articles with at least one co-author from a different country. SCP indicates single county publication, which is the number of articles with all authors from one country. Australia, Japan, and New Zealand show high international collaboration. While, surprisingly, United States shows low international collaboration.

In terms of collaboration network, United States and Canada appear to have the strongest collaboration network with each other (Fig. 7). United States also has a collaboration network with Germany. Australia has a collaboration network with United Kingdom and, Netherlands has a collaboration network with United Kingdom. This shows that only a few countries have established active collaboration network and collaboration network is limited to countries in North America and Europe. There seems to be no international collaboration with countries in other continents.
Fig. 6. Corresponding author’s country. *MCP: multiple countries publication; SCP: single country publication

Fig. 7. Collaboration network between countries

4.4 Articles

Global citations measure the number of citations an article has received from articles contained in the entire database, in our case, Web of Science. Specifically, global citations measure the impact of an article in the whole bibliographic database. Table 2 shows that the most global cited article was written by Westerfield et al. in 2015,
followed by Burrows et al. in 2015. TC per year indicates the yearly average number of times each document has been cited. It is notable that the article by Rivers & Koedinger published in 2017 received the highest total TC per year, 11.

**Table 2.** 10 Most global cited documents

<table>
<thead>
<tr>
<th>Author</th>
<th>Total citations</th>
<th>TC per year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Westerfield et al., 2015</td>
<td>28</td>
<td>7</td>
</tr>
<tr>
<td>Burrows et al., 2015</td>
<td>23</td>
<td>5.8</td>
</tr>
<tr>
<td>Rivers &amp; Koedinger, 2017</td>
<td>22</td>
<td>11</td>
</tr>
<tr>
<td>Kinshuk et al., 2016</td>
<td>22</td>
<td>7.3</td>
</tr>
<tr>
<td>Johnson &amp; Lester, 2016</td>
<td>21</td>
<td>7.0</td>
</tr>
<tr>
<td>Nye, 2015</td>
<td>20</td>
<td>5.0</td>
</tr>
<tr>
<td>Kim &amp; Baylor, 2016</td>
<td>19</td>
<td>6.3</td>
</tr>
<tr>
<td>Aleven et al., 2016</td>
<td>14</td>
<td>4.7</td>
</tr>
<tr>
<td>Bull &amp; Kay, 2016</td>
<td>14</td>
<td>4.7</td>
</tr>
<tr>
<td>R. Baker, 2016</td>
<td>13</td>
<td>4.3</td>
</tr>
</tbody>
</table>

Local citations measure the number of citations an article has received from articles included in the analyzed collection. Local citations measure the impact of an article in the analyzed collection, in our case, the 135 articles in IJAIED from 2015-2019 (Table 3). The most local cited articles, with 3 citations, are both first-authored by Aleven and published in 2016. Another article with 3 citations was by Biswas et al., published in 2016.

**Table 3.** 10 Most local cited documents

<table>
<thead>
<tr>
<th>Author</th>
<th>Local citations</th>
<th>Global citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aleven et al., 2016</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>Aleven et al., 2016</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>Biswag et al., 2016</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>R. Baker, 2016</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>Gilbert et al., 2015</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Matsuda et al., 2015</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Roll &amp; Wylie, 2016</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>McLaren et al., 2015</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Ogan et al., 2015</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Burrows et al., 2015</td>
<td>1</td>
<td>23</td>
</tr>
</tbody>
</table>

Reference Publication Year Spectroscopy (RPYS) is a quantitative method for identifying the historical origins of research fields and topics [20-21]. RPYS creates a temporal profile of cited references for a set of articles that emphasizes years where relatively significant findings were published. RPYS allows to identify the temporal roots of a discipline [22]. The Reference Publication Year Spectroscopy shows that the articles in the journal tend to cite references from the more recent period (Fig. 8). There seems to be two noticeable peaks before 2000: a peak in 1994 and another peak in 1999. After the peak in 1999, it is evident in the spectroscopy that the number of cited
references increase drastically as the years go up, which again suggests that the articles cite the most recent references over the older ones.

Fig. 8. Reference spectroscopy, references from 1954-2018

4.5 Most frequently used words

We examined the most frequently used words over time in author’s keywords and abstracts in the articles we have collected. Author keywords are a list of terms that were chosen by authors, as words that represent the overall content of the article. For author’s keywords, the most frequently used words have been “intelligent tutoring systems”, followed by “intelligent tutoring system” and “learning analytics.” Fig. 9 shows that “intelligent tutoring system” has been the top word used over the period of 2015-2019, and the conspicuous difference between the frequency of “intelligent tutoring system” and other words is especially evident. In author’s keywords, the use of “learning analytics” seems to grow over the years, especially in the last two years.
For abstract, “learning” and “students” have been the most frequently used words over time (Fig. 10). One notable trend is that for abstracts, the annual frequency growth of “learning” appears to be slowing down between the year 2017 to 2019 (Fig. 10). The use of word “students” seems to be continually growing. This might imply the change of trend in research. Although it is not the most used word in each category, “feedback” has been one of the most frequently used words in author’s keywords and abstract.

From looking at the most frequently used words, we found the following:

1. Intelligent tutoring system is notably one of the most prevalent topics of research in IJAIED during the period of 2015-2019.
2. The topics clearly put weight on educational components, as suggested by frequently used words such as students, feedback, and motivation.
3. As expected, author’s keywords and abstract have overlapping terms.

Fig. 9. Word dynamics for authors’ keywords (cumulative)
A caveat of using Biblioshiny for bibliometric analysis is its inability to recognize keywords that contain more than one word such as “learning analytics” and “intelligent tutoring system” and as a result, reading them as separate words and analyzing them separately. For example, intelligent tutoring system is displayed separately into three different word (“intelligent”, “tutoring”, “system.”) for abstracts of the articles as shown in Fig. 10. However, for author’s keywords, “intelligent tutoring system” and “learning analytics” are analyzed together as depicted in Fig. 9. We speculate that this could be due to how the keywords have an option for the author to enter the whole word in the article. For abstracts, the software seems to read a block of text without the ability to detect multiple words combined to make one word.

4.6 Conceptual structure: Co-occurrence network

Co-occurrence shows words that occur together in an article. As the same pair of words co-occurs increasingly more often, the link between them gets stronger [23]. Keywords Plus of the articles contains words or phrases contained in the titles of an article’s references but not in the title of the article itself. Keywords Plus is as effective...
as author’s keywords in investigating the knowledge structure of scientific fields, although it is less comprehensive in representing an article’s content [24]. In Keywords Plus of the articles, “students” co-occurred with multiple words, such as knowledge, intelligent tutoring systems, and framework (Fig. 11). The word “performance” co-occurred with multiple words as well including motivation, achievement, and system. Also, “motivation” and “achievement” co-occurred, with a strong link. This makes sense as the two topics are often related variables in a study.

Fig. 11. Co-occurrence network (Keywords Plus)

Fig. 12 shows that there is an expansive co-occurrence network in abstracts between the words used. Especially the words “students”, “learning”, and “systems” occur with words such as “study”, “learners”, “and design.” The words “students” and “learning” appear to have a strong link in the co-occurrence network, which makes sense as this network pertains to an educational journal. It is notable that there is a strong co-occurrence link between multiple terms in the abstract.
4.7 Intellectual structure

Fig. 13 provides a multi-field plot between top authors, top affiliations and countries in the top 20 ranking. Fig. 13 depicts that most of top authors and top affiliations are from United States, Germany, Canada, and New Zealand. Most of the top affiliations from other countries have a relationship with the U.S. institutions, for example, Germany with North Carolina State University and Arizona State University. Australia, Brazil, China, United Kingdom, and Mexico do not appear to have a relationship with the U.S. institutions. In addition, some countries have none to very limited international relationship with the institutions and authors. This conveys active international collaboration amongst a group of countries and the concentrated collaboration structure within the several countries and institutions.

Fig. 14 provides a multi-field plot between top 20 words used in titles, abstract, and author’s keywords. Most titles that include the words “intelligent”, “tutoring”, or “system” also include them in abstract, and keywords. Abstracts that used “feedback” included a variety of words in keywords such as collaborative learning, student modeling, intelligent tutoring system, learning analytics, educational data mining, and motivation. Most of the top 20 words in the titles that include “students” and similarly, “students” are used with most of the other top 20 keywords. One interesting word appearing only in the top keywords and not in the titles or abstract is “gift.”
**Fig. 13.** Three-Fields Plot (Relationship among Top Authors, Top Affiliations, and Top Countries.

**Fig. 14.** Three-Fields Plot (Relationship among Top Titles, Top Abstract, and Top Keywords)
5 Discussion

Our bibliometric analysis found that students and learning are two core topics and anchors of the studies in IJAIED. We found that intelligent tutoring system is the most frequently studied and discussed topic in IJAIED. Also, we found that there has been active contribution of the prolific authors in the journal all throughout 2015-2019. These prolific authors have maintained a collaboration network with each other. The country with the most number of relevant authors and institutions is United States. United States also has the most number of published articles, far exceeding the country with the second most number of published articles, Canada. United States and Canada have a strong collaboration link. There has been some international collaboration between countries in North America and Europe. However, the collaboration has been limited to these two continents.

Establishing and maintaining a widespread international and global AIED research community beyond WEIRD (western, education, industrialized, rich and democratic countries) will be crucial in making sure that the research has practical impact [7]. Improving generalizability of the AIED models such that the intervention, system, or new technology are designed for learners from all background would require AIED community to expand its research to include population that is more diverse. More specifically, new research should include samples from developing countries, and historically underserved population groups [25]. This way, AIED can be a unique and powerful tool that could reach remote locations and supplement additional resources through systems such as intelligent tutoring system which could provide personalized learning and automated tutor to underrepresented communities and developing countries. Therefore, more diverse and inclusive local and international collaboration could lead AIED to improve generalizability and expand its perimeter in reaching more population.

Learning Analytics and Educational Data Mining have become increasingly studied topics, especially in the most recent years [26]. Learning Analytics and Educational Data Mining are two related fields that emerged separately in the early 2000’s, both focusing on improving education using learning technology [27]. These two fields are closely related to AIED in ways that focus on developing innovative technological tools and analytical method to enhance learning context for instructors, learners, and administrators in all settings. As these three communities continue to develop and evolve, it is important to investigate each field’s objective to understand the similarities and differences amongst each other. This will lead the three fields to find ways to collaborate, learn from each other, and grow together.

This study has some limitations. For example, we collected data from Web of Science database and the limitations of this database may apply to this study as well. In addition, there are other studies published in 2019 that we were not able to include as our data collection was conducted in June 2019. However, this study makes a timely contribution as it provides a holistic overview of the most recent patterns and trends of IJAIED. As Artificial Intelligence in Education continues to evolve, it would be helpful to investigate the patterns and trends of other journals in the field as well using bibliometric markers on this resource base.
6 References


7 Authors

Clare Baek is with University of Southern California, in United States.

Tenzin Doleck is with University of Southern California, in United States. Email: doleck@usc.edu

Article submitted 2020-03-30. Resubmitted 2020-04-08. Final acceptance 2020-04-16. Final version published as submitted by the authors.