

Peeking into the AI Hype: Investigating Research Trends and Collaboration Dynamics in Artificial Intelligence in Education

Yipu Zheng, Jie Chen, Zhuqian Zhou, Zach Friedman, Paulo Blikstein
yipu.zheng@tc.columbia.edu, jc5230@tc.columbia.edu, zz2404@tc.columbia.edu,
zjf2003@tc.columbia.edu, paulob@tc.columbia.edu
Teachers College, Columbia University

Abstract: This project presents an in-depth analysis of the dynamics of the International Artificial Intelligence in Education (IAIED) Society through co-authorship networks and keyword analysis, discussing its implications to the larger research community. In addition, this study provides methodological contributions to the analysis of author and topic networks within multidisciplinary communities such as the Learning Sciences.

Introduction

With the growing attention to Artificial Intelligence (AI) in recent years, the application of AI technologies to education followed: AIED has been developing rapidly as an interdisciplinary research area (Feng & Kirkley, 2020). But the development of AI-powered content delivery systems, personalized learning management software, and adaptive testing platforms has been much faster than the growth of the research community (Zawacki-Richter et al., 2019). Now propelled by the global pandemic and remote learning, the AIED technologies might impact millions of students, teachers and schools all over the world at an unprecedented speed. Such significant implications of AI in Education (AIED) leave researchers and research communities with greater responsibility than ever to examine strengths and weaknesses, risks and opportunities, and costs and benefits of implementing AI in various learning scenarios. As the AIED field was growing, so were the calls for closer examination of issues of ethics and social justice around AI (Zuboff, 2015; Berendt et al., 2020). In particular, the International Artificial Intelligence in Education (IAIED) Society, as one of the earliest and more well-defined research communities in this area, has been viewed as an important leader in conversations around AIED. Given such a context, this paper aims to present a bird's-eye view of the recent research in this field through social network analysis on co-authorship networks in the community. The study is guided by the following research questions, which attempt to examine how the field is prepared to navigate itself in a world in which it has a critical social responsibility:

1. What are the major topics and interests in the IAIED research community (2013-2020)?
2. Is the IAIED research community open to new topics and new members, and attuned to new societal concerns?

Methods

To address our research questions, we analyzed time-evolving co-authorship networks and keyword usage trends of conference full papers from the International Conference of Artificial Intelligence in Education (ICAIED) and journal papers from the International Journal of Artificial Intelligence in Education (IJAIED) published from January 2013 to October 2020. The data were retrieved from Springer and the IJAIED official website. We collected the links of the web pages of target papers, automatically accessed the website's source code, stored the source code as text files, and then used *BeautifulSoup* library to extract data from the source code. The dataset contained 949 articles from 1915 authors from 2013-2020, and included the following fields: article title, author names and affiliations, publication year, keywords, abstract, and references. Full body texts of articles were also retrieved for further research.

In our data analysis, we examined the overall trends of topics from 2013 to 2020 through the most frequently used keywords, ascending keywords, and descending keywords over time. Additionally, we built time-evolving coauthorship networks with a two-year sliding window and utilized them as a proxy for the evolving collaboration network of the IAIED community. Core members of the community were identified using degree centrality. Keywords produced by core members and big sub-groups in the co-authorship network were further classified into distinct conceptual groups to reveal different ideas and research foci of those major contributors.

In the end, newcomers in the research community were identified through publishing patterns, and their keyword usage was analyzed.

Preliminary findings and implications

There are 2136 keyword stems presented in a total of 949 publications. The ten most frequently mentioned keywords are Intelligent Tutoring Systems (n=167), Natural Language Processing (n=46), Machine Learning (n=39), Student Model (n=35), Education Data Mining (n=32), Learning Analytics (n=31), Metacognition (n=26), Affect (n=24), Learning (n=24), and Collaborative Learning (n=23). The number of keywords grows from 507 in 2013-2014 to 593 in 2015-2016, 825 in 2017-2018, and 804 in 2019-2020, possibly indicating novel ideas being regularly introduced to the IAIED society.

The IAIED co-authorship network demonstrated a right-skewed distribution in the number of authors by degree of centrality. More than 50% of the authors (1046 out of 1915) in our IAIED dataset possessed a degree of zero to four (i.e., they have zero to four co-authorship ties in total in the period of 2013-2020). We identified 35 core members, who had more than 20 total co-authorship ties and analyzed their collaboration and publications. A majority of them (n=29, 83%) have published work on Intelligent Tutoring System. For the rest of the core members (n=6, 22.9%), all of them study Educational Games. The analysis of core members' publications reveals Intelligent Tutoring Systems and Educational Games being two major research topics in the IAIED society. In addition, the integration of new community members is vital to the growth and vitality of a scientific community. In attempting to measure such integration, we proposed five categories of authors in terms of their publishing patterns from January 2013-October 2020. The large size of the Joined Recently group (n = 542) shows the growth of the IAIED society from the perspective of membership. Additionally, the keyword analysis shows an increase of more than two hundred keywords during this same period of time (2019-2020).

Research topics of the IAIED society can also be investigated by analyzing the most frequently used keywords in large co-authorship cliques. The largest sixteen cliques were inspected, and they consist of nearly half (49.3%) of the authors in the IAIED society from 2013 to 2020. Intelligent Tutoring Systems (n=10, 62.5%), Game-Based Learning (n=3, 18.8%), Collaborative Learning (n=4, 25%) and Conversational Agents (n=3, 18.8%) are the major topics in these cliques.

One unexpected finding from our analysis is the use of synonymous but distinct terminologies to describe the same concepts in different cliques. This phenomenon may hinder inter-group communication since multiple terminologies need to be learned and used accordingly. In further research, we would like to investigate how different cliques developed over time.

Besides an in-depth analysis of the IAIED research community, an important contribution of this work, with applicability to the larger Learning Sciences community, is methodological. IAIED (and LS) are new, interdisciplinary, and still under construction. Most of the literature we found on bibliometrics was focused on well-defined communities. Many of the methodologies we developed in this paper were created for these types of novel research communities with “fuzzier,” “in-flux” keywords. We also developed a novel way of categorizing author participation in a research community based on publishing patterns. In this way, newcomers in the community can be identified, and the integration of new community members can be studied. Our future research will apply the method to other closely related research communities (e.g. Educational Data Mining, Learning Analytics, and Learning at Scale) to validate current findings and explore more emerging trends.

References

- Berendt, B., Littlejohn, A., & Blakemore, M. (2020). AI in education: learner choice and fundamental rights. *Learning, Media and Technology*, 45(3), 312-324.
- Feng, S., & Kirkley, A. (2020). Mixing patterns in interdisciplinary co-authorship networks at multiple scales. *Scientific reports*, 10(1), 1-11.
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—where are the educators?. *International Journal of Educational Technology in Higher Education*, 16(1), 39.
- Zuboff, S. (2015). Big other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, 30(1), 75-89.