Iseedu
of Islamic Education Thoughts and Practices
Vol. 8 No. 1 May 2024 : 87 - 99
E- ISSN : 2615 – 3335
https://journals.ums.ac.id/index.php/iseedu/index

Science Education Data Analysis to Improve Student Achievement: An In-Depth Bibliometric Analysis

Ahmadi Martha Habibi[1*], Fahmi Ulum Al Mubarok[2], Chusniatun[3], Mahasri Shobahiya[4], Zaim Hilmi Musyaffa[5]
[1234]Faculty of Islamic Religion, Universitas Muhammadiyah Surakarta, Indonesia
[5]Department of International Relations, Erciyes Üniversitesi, Turkey

*Correspondence: Ahmadi Martha Habibi
Email: fahrime09@gmail.com
DOI:
https://doi.org/10.23917/iseedu.v8i1.8849

Received: 11-03-2024
Accepted: 12-04-2024
Published: 19-05-2024

Abstract:

This study was conducted to map and analyze the development of research on Science Education Data in publications that have been indexed in Scopus. The method applied is bibliometric analysis, which explores information from publications related to "Data of Science Education" from 2015 to 2024 in the Scopus database. The analysis process involved the use of tools such as Excel and R/R-Studio. The analysis results were visualized using VOS viewer to display keyword occurrences and document citations simultaneously. Within this study's framework, 1092 publications that met the set criteria were identified. The findings of this study indicate an annual growth of 21.32%, peaking in 2021. The United States was identified as the primary contributor in the number of publications, with the University of Nigeria as the most productive affiliate. The author who achieved the highest productivity in the Data of Science Education theme was Ali, N. By detailing these statistics and trends, this study provides a comprehensive overview of the literature in the education domain, while conveying recommendations for research directions that could be undertaken in the future. It is hoped that the results of this study make a valuable contribution to the in-depth understanding of data in the context of Science Education.

Keywords: Education, Data, Science, Human Experience, Bibliometric.

Introduction

Science education is one of the important pillars in the formation of an informed and innovative society (Artyukhov et al., 2022). In this era of technology and information, data collection and analysis have become critical elements in understanding and improving the quality of science education (Chen & Schintler, 2023). Moreover, in the 4.0 era, science education data is beneficial for evaluating students' progress, teacher development, resource management, learning program development as happened during the COVID-19 pandemic (Wang & Uecker, 2017). Education policy, research and innovation, and much more. Industry 4.0, characterized by the integration of digital technology and automation in various industrial sectors, demands a higher understanding and expertise in the fields of science and technology (Abad-Segura et al., 2020). In the face of Industry 4.0 data, science education has an important role to play in ensuring that students are equipped with relevant skills and knowledge (Sezgin et al., 2022).

The importance of science education data is supported by references and facts that are collected and then used to understand the state of progress and characteristics of science learning at various levels, whether at primary, secondary, or tertiary level public schools (García et al., 2020).

Science education data involves various variables that include aspects such as student achievement (Shoaib et al., 2021). Science education data involves various variables that include aspects such as student achievement, teacher engagement, science curriculum, laboratory facilities, and various other factors that affect the process and outcomes of science learning (Kumari, 2024). Science education data collection is not just about measuring achievement, but also about providing in-depth insights to support informed decision-making to improve the overall quality of science education (Munasi, 2024). Data provide a solid basis for intervention, innovation and the development of appropriate strategies for science education to have a positive impact on student development and societal progress (Papa et al., 2025).

Bibliometric indicators play a very significant role in evaluating scientific research results. Their function is not only limited to measuring the impact of publications, but also includes analysing the interaction between science and technology (Ayowumi & Cln, 2021; Das & Parabhoi, 2020; de Paulo et al., 2023; Dias et al., 2023; Fan et al., 2022). In addition, bibliometric indicators play a role in forming a mapping of various fields of science and tracking the development of new knowledge in a particular discipline. Furthermore, bibliometric indicators are a valuable basis for future strategic planning, helping research and academic institutions to direct their resources more effectively and formulate policies based on data and evidence. Therefore, it can be concluded that bibliometric indicators play a key role in the research sphere, assisting in the understanding, measurement, and planning of scientific developments.

This research aims to see the development and trends of research related to the theme "Data of Science Education" in publications registered in the Scopus database from the period 2015 to 2024. The use of 2015 as the starting point of the research was chosen based on the last ten years of this publication in the Scopus database.

Previous Research

In the student domain, research was conducted by Weisen et al (2024) by measuring the ability and level of development of current students in using electronic devices, especially in classroom learning. On the teacher's domain, this study sought to identify student profiles of ICT use based on data from the 2013 Norwegian International Computer and Information Literacy Study (ICILS) (N = 2426). To explore the profile of ICT use at school and at home for various purposes such as recreation, learning purposes, information exchange and social communication, we took an individual-centred approach and applied latent profile analysis (González-Zamar et al., 2020). Among them, we also discussed the personal web usage for each teacher in relation to their work intensity.

The academic realm of research conducted by Kyei-Akuoko et al (2025) examined home resources and public school performance. This research is descriptive cross sectional which involved 48 children. Children were evaluated in listening, writing, counting, and reading performance. In the realm of research technology by Hamzah in Malaysia, which aims to advance technology-based education globally, massive open online courses (MOOC) (Jacqmin, 2019). It also aims to measure the level of effectiveness in teaching. Research that also discusses learning with blinded learning models which aims to assess the extent of technological developments that can be used and mastered by teachers and students.

From previous research on the theme of Data of Science Education, bibliometric analysis research methods have yet to be used as an effort to map scientific publications in various fields.

Method

The bibliometric analysis method is the cornerstone of this research (Hallinger, 2023; Handayani et al., 2023; Luque-Martínez & Luque-Raya, 2023; Şimşek & Kalıpçı, 2023; Thamaree & Zaby, 2023). Data was collected through a search engine that applies a Boolean approach, exploring the Scopus database from 2015 to 2024. The date 10 December 2023, at 09.30 WIB, was set as the search time. A thorough analysis was conducted by the researcher using tools such as R and Rstudio, VosViewer, and Microsoft Excel. This approach was used to parse information about citations, document content, and relationships between data. The researcher went through three stages in processing the collected dataset.

In the initial stage, the researcher will conduct a literature analysis to check the relevance of the topic to the bibliometric research to be conducted. This step is not only useful to ensure the relevance of the research to the theme, but also to identify appropriate keywords that are considered to cover the research area well.

In the second stage, researchers conducted a search in the Scopus database using the boolean operator TITLE-ABS-KEY (data AND science AND education). From the results of this search, a total of 18,442 documents were found. Furthermore, the filtering process was carried out by applying the Boolean operators TITLE-ABS-KEY (data AND science AND education) AND PUBYEAR> 2014 AND PUBYEAR < 2025 and (LIMIT-TO-KEY).2025 and (LIMIT-TO (SUBJAREA, "ARTS") OR LIMIT-TO (SUBJAREA, "ENVI")) and (LIMIT-TO (DOCTYPE, "ar")) and (LIMIT-TO (LANGUAGE, "English")) and (LIMIT-TO (SRCTYPE, "j")) to narrow down the documents, then there are restrictions on the year of publication ranging from 2015 to 2024 and on the scope of human and art, with the type of document being an article, originating from a journal, and written in English. After this filtering process was completed, 1092 relevant documents were obtained.

In the third stage, an in-depth analysis of the selected documents was conducted. Various tools such as Scopus Analyzer, R and RStudio were used for the purpose of identifying the number of documents published each year, analyzing the distribution pattern of documents by year, extracting information about the authors, the authors' institutional affiliation, and the authors' country of origin. In addition, this stage involved analyzing the level of connectedness between documents with the help of visualizations presented by VOSviewer. All data obtained was further processed and analyzed using Microsoft Excel. The steps of this research are visually presented in the attached figure:

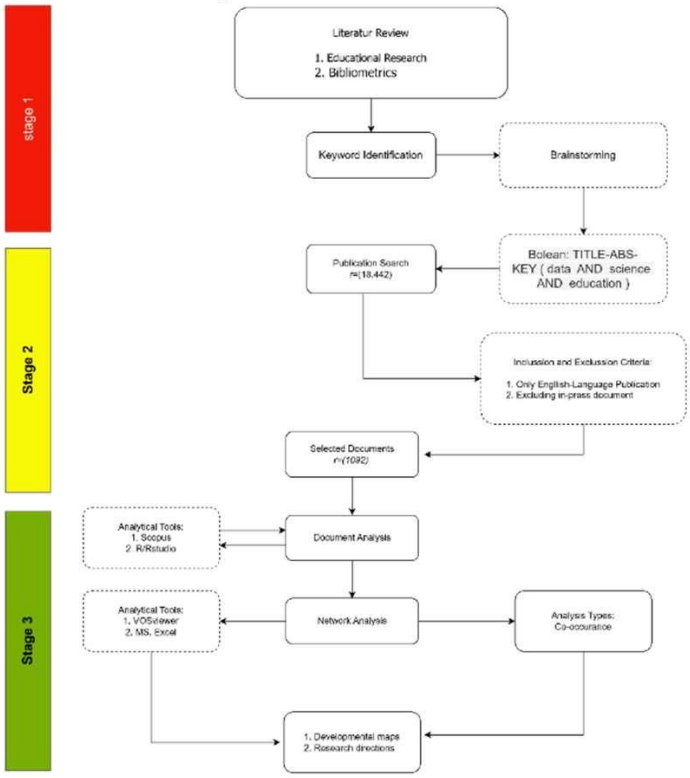

Figure 1. Research Flow

Result and Discussion

Key Information About the Data

Table 1 presents a summary of 4865 records accumulated during a span of 10 years. Encompassing 14669 authors 304 single authors, international authorship collaboration was observed at a rate of 32.74%. The references amounted to 254345, with the average citations per document totalling 20.59.

Table 1. Main Data Information

Description	Results
MAIN INFORMATION ABOUT DATA	
Timespan	2015:2023
Sources (Journals, Books, etc)	782
Documents	4865
Annual Growth Rate %	21,32
Document Average Age	2,59
Average citations per doc	20,59
References	254345
DOCUMENT CONTENTS	
Keywords Plus (ID)	21714
Author's Keywords (DE)	14669
AUTHORS	
Authors	16654

Authors of single-authored docs	304
AUTHORS COLLABORATION	
Single-authored docs	312
Co-Authors per Doc	4,67
International co-authorships %	32,74
DOCUMENT TYPES	
Article	4865

Most Relevant Authors

The following figure shows the ten most influential authors in publications on Data of Science Education. Ali, N, Miralles- Martinez, P, Surif, J., and Yuksel D, lead with the most publications with 4 documents followed by Altay, M. , Archila, P.A., Curle, S., Dastani, M., Ibrahim, N.H. and Panditas, R. 3 documents each

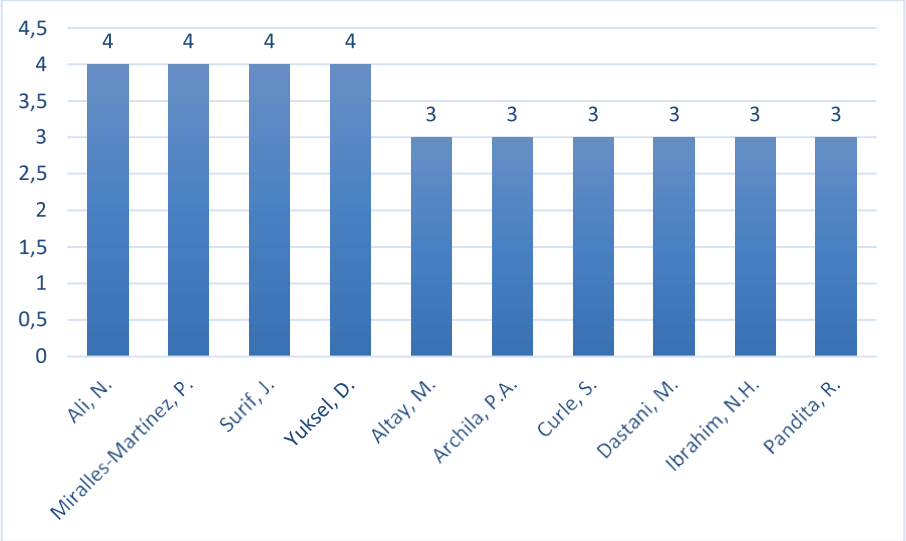

Figure 2. Graph of the number of author publications

Documents by Affiliation

The following figure shows the 10 most influential affiliations in publications on Data of Science Education. The analysis of this theme was mostly carried out by researchers at the University of Nigeria, with a total of 15 articles. Researchers from The Universiti Kebangsaan Malaysia followed this with 13 articles.

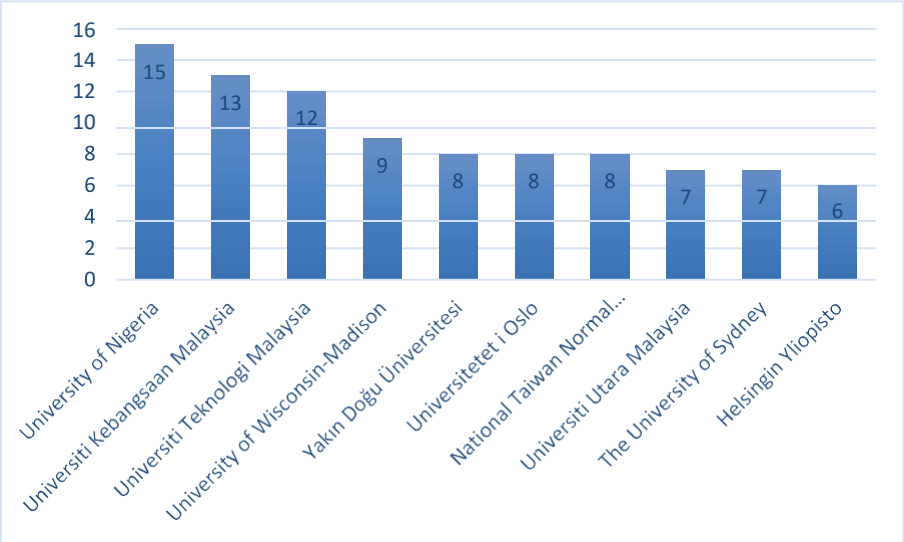

Figure 3. Graph of the number of publications per affiliation

Documents by Country

The following figure illustrates the distribution of publications by country in the Data of Science Education theme. The leader in the number of publications is the United States with 284 documents, followed by Malaysia with 59 documents.

The dominance of Asian countries can be seen with five countries contributing, while countries in other continents such as America, Europe, the UK, Australia and Africa occupy the next dominant position. This shows that research on the theme of Data of Science Education is conducted mainly by countries in the Asian continent.

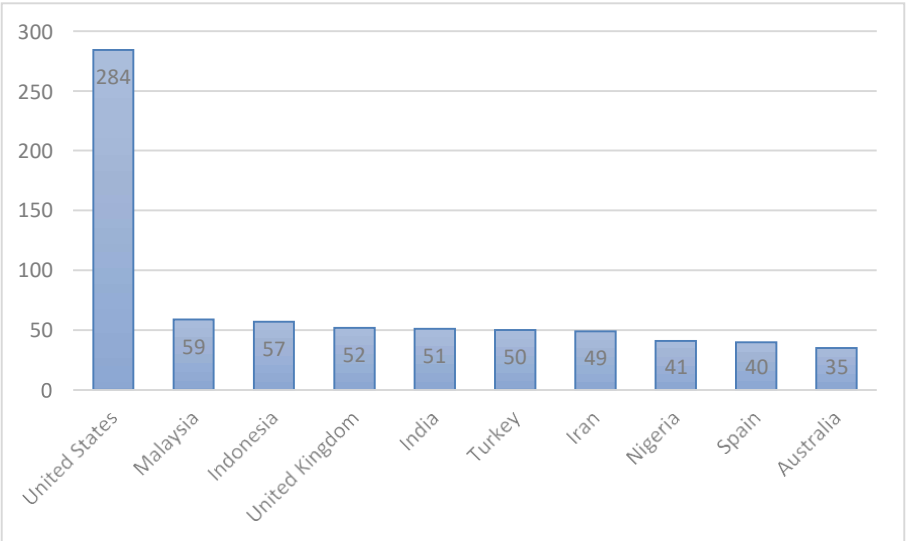

Figure 4. Graph of the number of publications per country

Three-Field Plot

Figure 6 bellow contains 3 observed elements: cited references, author names and keywords. The three elements are then connected by grey plot lines that are related to each other. The first element contains references cited by the authors when researching related themes. Based on the

picture above, there are 4 references. The most cited reference in research with the Data of Science Education theme is Statistical Power Analysis for the behaviour societies, which is displayed in dark red and is associated with the authors Truscott De Mejia A-M and Archila, P.A.

Based on the figure above, in the second element there are 16 authors. The size of the bar chart shows how many research publications from each author. Among the 9 authors who wrote the most articles on the theme of Data of Science Education were Yuksel, D., Soruc, A., Altay, M., Curle, S. and Ali, N.

In the third element, each research topic has connections with authors who are active in producing works related to the Data of Science Education. The analysis identified 8 keywords, with "Higher Education" and "Stem Education" ranking highest. This illustrates that these two words have a strong relationship with Data of Science Education research.

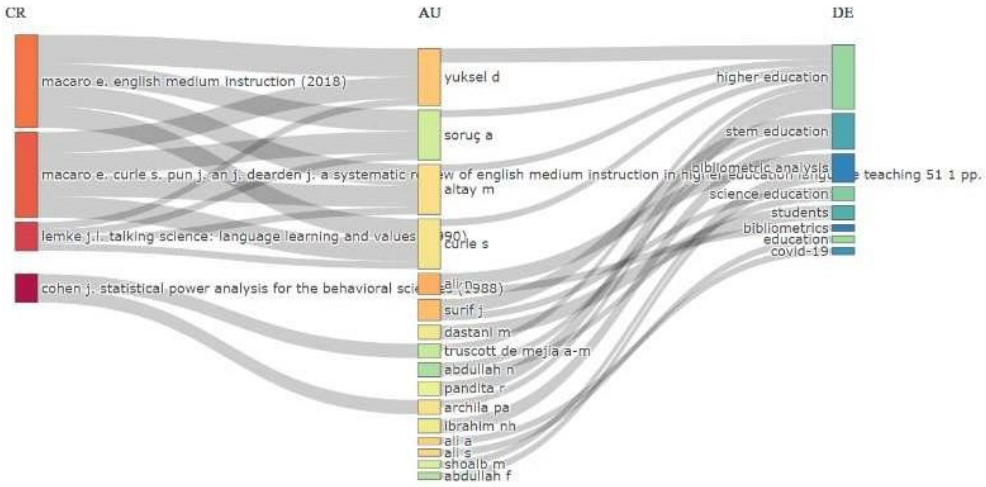

Figure 5. Three-Field Plot (references, author, and keywords)

Corresponding Author's Countries

From the perspectives of "Multiple Country Publication" (MCP) and "Single Country Publication" (SCP), the number of SCPs is greater than that of MCP. In terms of MCPs, the United States (US) achieved the highest number of publications, with 12 documents, followed by Indonesia with 10 documents, and the UK with 9 documents. However, in terms of SCPs, the United States also leads with a much larger number of publications, at 205 documents. Iran ranked second with 37 documents, while India ranked third with 34 documents. Based on continental data, the Asian continent dominates with a total of 8 countries, while the European continent holds the second dominance with a total of 6 countries.

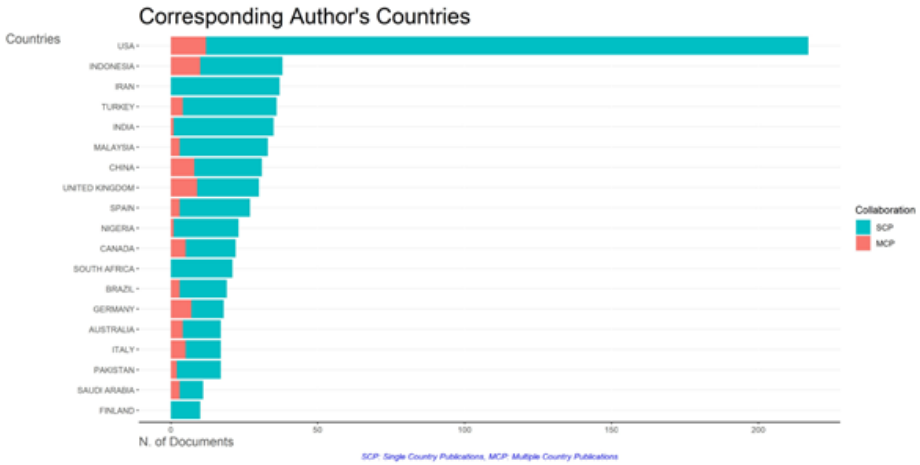

Figure 6. Graph of corresponding author's countries

Most Globally Cited Document

The paper with the largest total citations (TC) is "Cain MK, 2017, Behav Res Methods" with 516 citations. Likewise, the paper with the largest TC per Year "Cain MK, 2017, Behav Res Methods" has citations per year with a total of 73.71 TC per Year.

Based on the available information, it can be concluded that TC (total citations) has a tendency to affect TC per Year (citations per year). In general, papers with higher TC also tend to have significant TC per Year. However, it is only sometimes sure that papers with a large TC always have a higher TC per Year than others, and vice versa.

However, it should be noted that the year of publication of a paper does not significantly affect the TC. While some papers with older years of publication may have lower TC than papers with more recent years of publication, there are also exceptions where papers with older years of publication may have high TC. Therefore, from the available data, there is no consistent relationship between year of publication and number of citations (TC).

Table 2. Most globally cited documents

Paper	Total Citations	TC per Year	Normalised TC
Cain MK, 2017, Behav Res Methods	516	73,71	23,97
Bonney R, 2016, Public Underst Sci	491	61,38	16,42
Simis MJ, 2016, Public Underst Sci	363	45,38	12,14
Akçayir M, 2016, Comput Hum Behav	340	42,50	11,37
Chandler J, 2019, Behav Res Methods	297	59,40	26,46
Maldonado-Mahauad J, 2018, Comput Hum Behav	199	33,17	12,14
Charron N, 2015, Soc Indic Res	174	19,33	7,83
Duffy MC, 2015, Comput Hum Behav	158	17,56	7,11
Harrati N, 2016, Comput Hum Behav	155	19,38	5,18
Mayer RE, 2019, Appl Cogn Psychol	146	29,20	13,01

Thematic Map

The theme mapping shown in Figure 8 consists of a Basic Theme, motor theme, niche theme, and emerging and declining theme. The basic theme is a category that includes the basic theme of existing research. In the given data, the cluster "students" is an example of a basic theme. Basic themes are often the main focus of research that receives extensive attention from researchers.

Motor Themes: This category includes themes that are the driving force behind research development. The "Human" cluster (Cluster 8) is an example of a motor theme in this data. Human themes are actively researched and influence the direction and trends of research related to science education data. Motor themes often influence how research and publications evolve.

Niche Themes: This category includes more specific and less general themes than the basic and motor themes. An example of a niche theme in this data is the "Attitude" cluster (Cluster 6). Themes in the niche theme category have fewer studies or are only of interest to a specific group of researchers.

Emerging or Declining Themes: This category includes themes that are emerging or declining in popularity in research or publications. In this data, the theme "Iran" (Cluster 5) can be categorized as an emerging theme because its number of Occurrences and relevance in research is still significant, but not as much as the themes "Human" (Cluster 8) or Attitude (Cluster 6).

By conducting analyses based on these groupings, we can explore the distribution and trend patterns related to themes in available research data or publications. This classification proves its usefulness in identifying key themes that dominate a research field, emerging themes, and themes that require further attention and in-depth research.

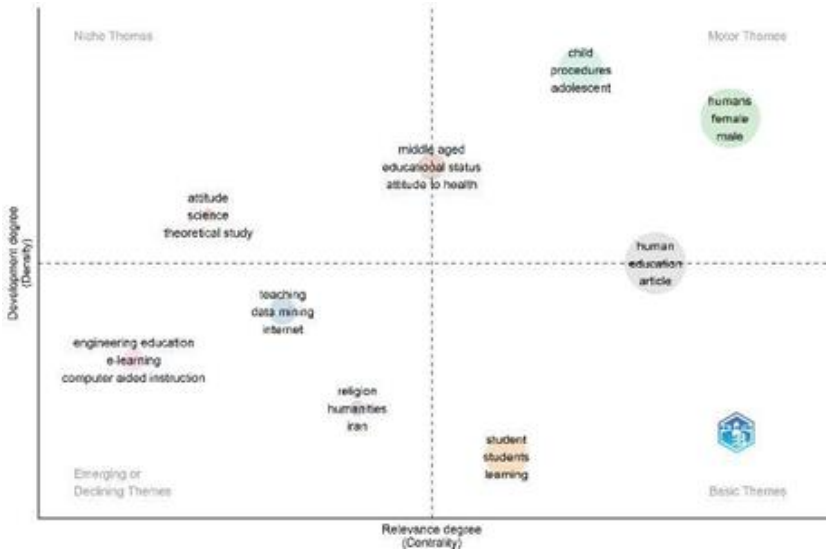

Figure 7. The thematic map

Network Analysis

In the figure 8, the number of occurrences of research on Data of Science Education was analyzed, with a minimum number of clusters of 15. Three clusters were formed, with the number of items in each cluster being: cluster 1 with 44 items, cluster 2 with 28 items and cluster 3 with 25 items. The dominating keyword is "Education" with a total occurrences of 174 and link strength of 972.

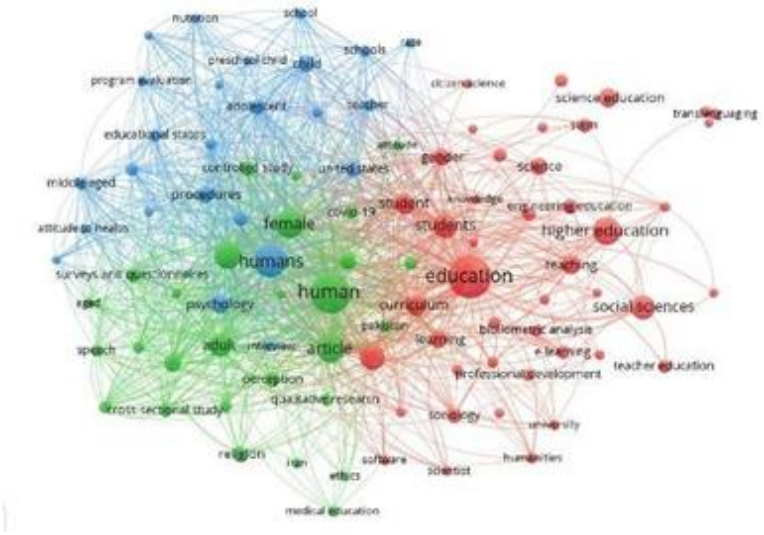

Figure 8. Network distribution of keywords based on the number of occurrences

The figure 9 shows the keyword network analysis based on overlay. It can be seen that the keywords translanguaging, science education, covid-19, and bibliometric analysis are keywords that have been used in the latest year, namely around 2021. Meanwhile, the keywords surveys and questionnaires, middle age and psychology are keywords with relatively old usage around 2018.

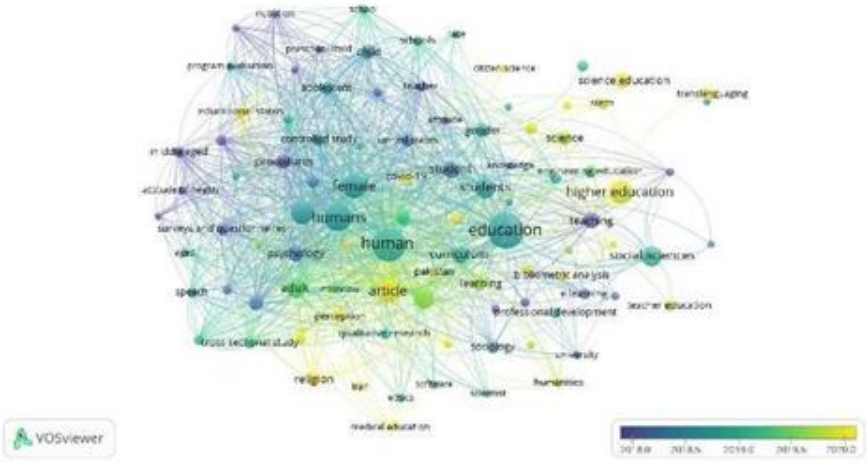

Figure 9. Image of keyword network distribution based on overlay

The research density analysis revealed that the keywords "Human" and "Education" have a high density in this theme research. This suggests that focusing on these aspects plays a central role in the context of Data Science Education. On the other hand, keywords such as "teacher education," "statistical and numerical data," and "educational research" are relatively under-researched. Nonetheless, further research on these sparse keywords can provide valuable insights to unearth elements that may not have been fully explored before.

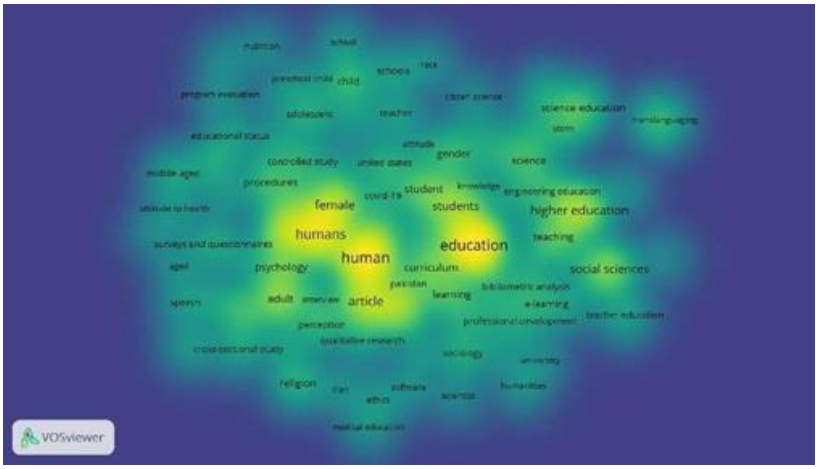

Figure 10. Image of keyword network distribution based on density

The table below shows the occurrence in each cluster that represents the main themes in the Data of Science Education research field. The theme in the first cluster is Social Science Education and Human Experience, and the theme in the second cluster is Questionnaire and Adult Human. At the same time, the theme in the third cluster is Child psychology and Adolescent.

Table 3. Keywords and the number of occurrences in each cluster

	Keywords	Occurrences	Cluster
1.	Education	174	
2.	Higher education	73	
3.	Human experiment	62	1
4.	Students	58	
5.	Social science	55	
1.	Human	149	
2.	Article	74	
3.	Adult	53	2
4.	Questionnaire	27	
5.	Major clinical study	25	
1.	Humans	104	
2.	Psychology	34	
3.	Procedures	28	3
4.	child	29	
5.	adolescent	21	

Conclusion

Publications with the theme Data of Science Education peaked in 2021 with a total publication of 187 documents. The most relevant author is Ali, N., who has four publications. The affiliation with the most publications is the University of Nigeria, which has a total of 15 publications. The country with the most publications is the US, with 284 documents. The largest MCP is the US. Park D-B, 2009, "Cain MK, 2017, Behav Res Methods" is the document with the highest number of citations, namely, 516 citations. The basic theme in research on Dof Science Education is Human; education is

the dominating keyword with a link strengh of 972. Translanguaging, Science Education, Covid-19, and Bibliometric Analysis are keywords with usage in the latest year, which is around 2021.

References

Abad-Segura, E., González-Zamar, M.-D., Infante-Moro, J. C., & García, G. R. (2020). Sustainable management of digital transformation in higher education: Global research trends. *Sustainability (Switzerland)*, *12*(5). https://doi.org/10.3390/su12052107

Artyukhov, A., Volk, I., Surowiec, A., Skrzypek-Ahmed, S., Bliumska-Danko, K., Dluhopolskyi, O., & Shablystyi, V. (2022). Quality of Education and Science in the Context of Sustainable Development Goals—From Millennium Goals to Agenda 2030: Factors of Innovation Activity and Socio-Economic Impact. *Sustainability (Switzerland)*, *14*(18). https://doi.org/10.3390/su141811468

Ayowumi, O., & Cln, O. (2021). Bibliometric analysis of research productivity of librarians in college library, federal college of education (special), Oyo, Oyo, State. Nigeria. *Library Philosophy and Practice*, *2021*, 1–11. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85108214917&partnerID=40&md5=b89f3123387f91714a45878efb9539ec

Chen, Z., & Schintler, L. A. (2023). Rediscovering regional science: Positioning the field's evolving location in science and society. *Journal of Regional Science*, *63*(3), 617–642. https://doi.org/10.1111/jors.12634

Das, J. M., & Parabhoi, L. (2020). Research Productivity of LIS Women Faculty in India: A Bibliometric Study during 1988-2018. *Library Philosophy and Practice*, *2020*, 1–17. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85098144770&partnerID=40&md5=b43d5de16d0a030d5562bca61cab0b5d

de Paulo, R. C., de Araújo Rabelo, T. T., de Oliveira Dias, F., Vilela, D. C., & Mafra, R. Z. (2023). ST&I PUBLIC POLICIES: A BIBLIOMETRIC STUDY OF SCIENTIFIC PRODUCTION. *Revista de Gestao Social e Ambiental*, *18*(1). https://doi.org/10.24857/rgsa.v18n01-023

Dias, F. D. O., Bonfim, H. D. S., Bombardelli, J., & Matos, G. B. D. C. (2023). BIBLIOMETRIC ANALYSIS OF SCIENTIFIC PRODUCTION ON CARBON CREDITS (2012-2021). *Revista de Gestao Social e Ambiental*, *17*(1). https://doi.org/10.24857/RGSA.V17N1-026

Fan, S., Jiang, J., Li, F., Zeng, G., Gu, Y., & Guo, W. (2022). A Bibliometric Analysis of the Literature on Postgraduate Teaching. *Sustainability (Switzerland)*, *14*(24). https://doi.org/10.3390/su142417047

García, E. G., Magaña, E. C., & Ariza, A. C. (2020). Quality education as a sustainable development goal in the context of 2030 agenda: Bibliometric approach. *Sustainability (Switzerland)*, *12*(15). https://doi.org/10.3390/SU12155884

González-Zamar, M.-D., Abad-Segura, E., López-Meneses, E., & Gómez-Galán, J. (2020). Managing ICT for sustainable education: Research analysis in the context of higher education. *Sustainability (Switzerland)*, *12*(19). https://doi.org/10.3390/su12198254

Hallinger, P. (2023). Mapping the Contributions of the Review of Educational Research to Education, 1931–2020. *Review of Educational Research*, *93*(6), 791–824. https://doi.org/10.3102/00346543221132771

Handayani, D. I., Vanany, I., & Ciptomulyono, U. (2023). Blockchain Application in Food Supply Chains: Bibliometric Analysis and Future Research. *International Journal on Food System Dynamics*, *14*(2), 146–165. https://doi.org/10.18461/ijfsd.v14i2.F2

Jacqmin, J. (2019). Providing MOOCs: A FUN way to enroll students? *Information Economics and Policy*, *48*, 32–39. https://doi.org/10.1016/j.infoecopol.2018.10.002

Kumari, R. K. (2024). Revolutionizing Education: Harnessing Graph Machine Learning for Enhanced Problem-Solving in Environmental Science and Pollution Technology. *Nature*

Environment and Pollution Technology, *23*(4), 2247–2254. https://doi.org/10.46488/NEPT.2024.v23i04.038

Kyei-Akuoko, C., Mensah, R. O., Kuusongno, D. D., Ebow Yalley, C., & Amponsah, K. D. (2025). Evaluation of blended learning: challenges, academic performance shifts, and the pros and cons in a selected Technical University. *Cogent Arts and Humanities*, *12*(1). https://doi.org/10.1080/23311983.2024.2435713

Luque-Martínez, T., & Luque-Raya, I. M. (2023). Strategic analysis of Spanish scientific production by field: Natural, Medical and Life Sciences. *Revista Espanola de Documentacion Cientifica*, *46*(4). https://doi.org/10.3989/REDC.2023.4.1409

Munasi, K. R. (2024). Evaluating the Integration of Education for Sustainable Development in the Life Sciences Curriculum: A Case Study of Grade 11 Teachers in Vhembe East District. *Sustainability (Switzerland)*, *16*(24). https://doi.org/10.3390/su162410918

Papa, C. C., Clay, K., Cooper, L. T., & Stark, S. C. (2025). Science-based communication and education needed to improve forest carbon science, policy, and management outcomes. *Environmental Research Letters* , *20*(2). https://doi.org/10.1088/1748-9326/ada8c4

Sezgin, A., Orbay, K., & Orbay, M. (2022). Educational Research Review From Diverse Perspectives: A Bibliometric Analysis of Web of Science (2011–2020). *SAGE Open*, *12*(4). https://doi.org/10.1177/21582440221141628

Shoaib, M., Abdullah, F., & Ali, N. (2021). A Research Visualization of Academic Learning Skills among Students in Higher Education Institutions: A Bibliometric Evidence from 1981 to 2020. *Library Philosophy and Practice*, *2021*, 1–34. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85108548479&partnerID=40&md5=93d787b0e25d4d6cb6eb2acffa6435e9

Şimşek, E. K., & Kalıpçı, M. B. (2023). A bibliometric study on higher tourism education and curriculum. *Journal of Hospitality, Leisure, Sport and Tourism Education*, *33*. https://doi.org/10.1016/j.jhlste.2023.100442

Thamaree, A., & Zaby, S. (2023). BIBLIOMETRIC REVIEW OF RESEARCH ON CORPORATE GOVERNANCE AND FIRM VALUE. *Journal of Governance and Regulation*, *12*(1), 42–52. https://doi.org/10.22495/jgrv12i1art4

Wang, X., & Uecker, J. E. (2017). Education, Religious Commitment, and Religious Tolerance in Contemporary China. *Review of Religious Research*, *59*(2), 157–182. https://doi.org/10.1007/s13644-017-0286-5

Weisen, S., Do, T., Peczuh, M. C., Hufnagle, A. S., & Maruyama, G. (2024). How are first-generation students doing throughout their college years? An examination of academic success, retention, and completion rates. *Analyses of Social Issues and Public Policy*, *24*(3), 1274–1287. https://doi.org/10.1111/asap.12413