



# Research Trends of Latent Class Analysis in Education: Bibliometric Analysis Based on Scopus Data

Sa'adatul Ulwiyah, Jumriani Sultan, and Heri Retnawati

Universitas Negeri Yogyakarta, Yogyakarta, Indonesia  
saadatul.2022@student.uny.ac.id

**Abstract.** Latent Class Analysis (LCA) is one of the important statistical tools used in educational research. Latent Class Analysis is used to identify unobserved groups based on patterns in the data. This study aims to capture the landscape of relevant previous research related to Latent Class Analysis in education using bibliometric analysis. Data was taken from the Scopus database which was filtered into 130 publications. The number of articles related to Latent Class Analysis in Education has increased every year, especially in recent years, such as 2019, 2021, and potentially in 2023. The citation trend related to Latent Class Analysis in the scope of educational research occurred in 2019 with a total of 159 citations. Wang Y and Garcia-Fernandes are the authors with the highest number of publications, while Boscardin CK is the author with the highest number of citations. The article that had the most significant impact was the article by Bradshaw CP, Buckley JA, and Ialongo NS entitled "School-Based Service Utilization Among Urban Children with Early Onset Educational and Mental Health Problems: The Squeaky Wheel Phenomenon". The article was published by School Psychology Quarterly and has received 108 citations. The United States, China and Spain are the most influential countries in this field. The keywords "mental health", "structural equation modeling", "surveys and questionnaires", and "adults" are some of the new research areas related to latent class analysis in education.

**Keywords:** Bibliometric, Latent Class Analysis, Education, Grouping Students.

## 1 Introduction

Education is an important aspect of human life. Education is the main foundation for achieving balance and perfection in the development of individuals and society [1]. Through education, individuals can acquire the knowledge, skills, and values necessary to reach their maximum potential and contribute to Society [2]. Thus, education is not only a learning process, but is also the main key in the formation of individuals capable of bringing positive change to society.

As a field that continues to develop, education requires a basic scientific approach to understand the dynamics and changes that occur in it. One of the important statistical tools used in educational research is Latent Class Analysis (LCA). Latent class analysis is one of several techniques for reducing data into parts LCA is a statistical method used to identify hidden patterns or classes in data [3]. LCA used to identify qualitatively distinct subgroups within a population that often share certain

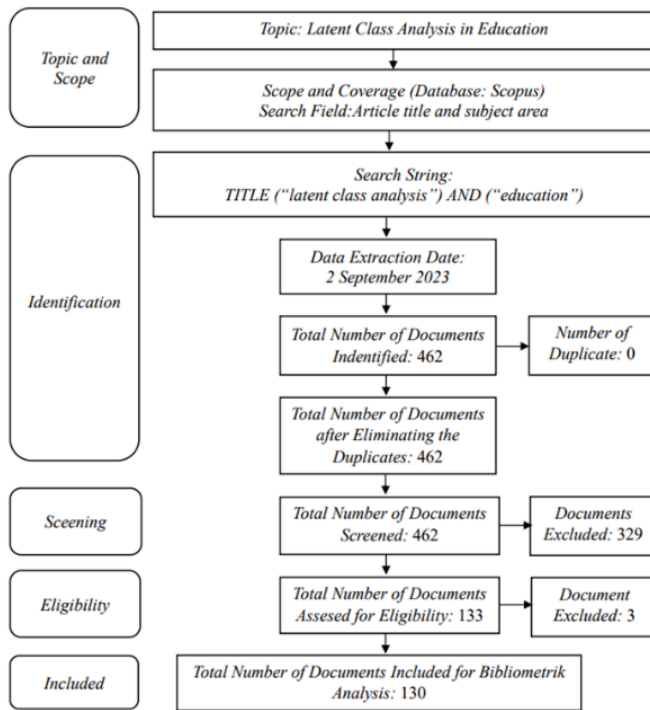
external characteristics [4]. In the educational context, LCA aims to identify groups or categories that may exist among students or respondents based on similarities in certain characteristics or behavior [5], [6]. LCA allows grouping individuals based on certain characteristics or behaviors that may not be directly visible through conventional analysis. In other words, LCA can provide valuable insights into educational development and increase understanding of individual differences in it.

Over time, research in the field of LCA in the educational sphere has developed rapidly. Scientific papers, journals and related research have appeared in various journals and academic platforms. However, no research has yet comprehensively investigated the development and trends in the use of latent class analysis in educational research using a bibliometric approach. Bibliometric analysis is an effective method for analyzing scientific developments in a field. Bibliometric analysis is a method systematic review which is used to identify current trends and issues in previous research literature, with the aim of providing a comprehensive picture of a particular research field [7]. Recently, many research reports on bibliometric analysis have been published with the aim of understanding the trend or direction of research development in a particular topic [8]–[13].

By collecting and analyzing bibliometric data, it will help provide a comprehensive picture of how Latent Class Analysis research in education has developed from its early years to the present. This information will be useful to researchers, educators, and policy makers in the field of education, assisting them in determining future research directions and identifying potential collaborations. Therefore, this research aims to conduct a bibliometric analysis of Latent Class Analysis research in the world of education to identify trends, the most productive authors, the most contributing journals, and dominant research topics.

## 2 Method

In looking for data sources related to "Latent Class Analysis in the field of education", researchers used the comprehensive Scopus database, encompassing data from Q1 to Q4. The process of refining the data that has been collected consists of several stages as shown in Figure 1. The first stage is identification, then continued with the screening, feasibility and inclusion stages as the final stage [11].



**Fig. 1.** Data collection process.

The first stage is identification, at this stage the researcher enters keywords in a search on the Scopus database. The keywords entered are "Latent Class Analysis" and "Education". The results of this identification step produced 462 documents. The next stage is the screening process, where the researcher makes a selection based on certain criteria, namely: field area (social sciences, psychology, multidisciplinary), type of document (article), language (English), and type of source (journal). After the screening process, 133 publications were found that met these criteria, meaning that there were 329 publications that had been discarded and were not continued at the next stage.

Publication of the screening results, then a manual feasibility evaluation is carried out. In this stage, researchers looked at the abstracts and titles of 133 publications, and assessed whether the publications included or discussed latent class analysis in the field of education. At this stage, 130 publications were obtained that were worthy of being included in the next stage.

Metadata from 130 documents was then downloaded in CSV and RIS format. The next step is to conduct a bibliometric analysis to determine the research landscape on latent class analysis in the field of education. The author uses analytical tools such as VOSviewer, Microsoft Excel, and Biblioshiny to carry out this bibliometric analysis. VOSviewer and Biblioshiny are two very useful software in bibliometric data analysis and visualization of scientific information. VOSviewer is software used to visualize and analyze networks and clusters of scientific publications based on co-citation,

co-authorship, or keyword co-occurrence [14]. Meanwhile, Biblioshiny is a software package running in the R statistical programming environment that facilitates bibliometric analysis and the creation of interactive visualizations of reference data [15].

### 3 Results and Discussion

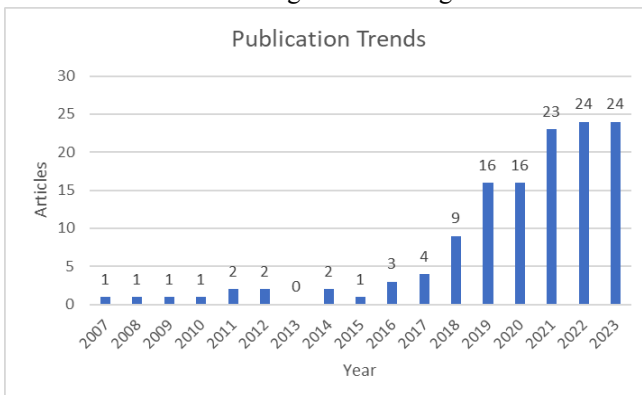
#### 3.1 Results

**Trends in LCA Research Publications in Education.** In this research, publication characteristics analysis is used to examine research topics related to latent class analysis in global student groupings. The results of bibliometric analysis show that there are 130 articles originating from the Scopus database with publication years ranging from 2007 to 2023 used in this research. Overall, Figure 2 presents the main information of the database used.



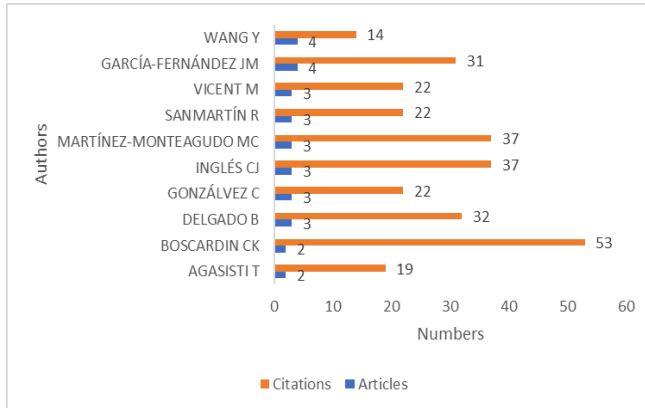
Fig. 2. Main information database.

Furthermore, the trend of growth and development of publications based on year of publication is depicted in Figure 3. It can be seen that in 2022 and 2023 the most articles were published compared to previous years. In publication trends, it can be seen that publications have increased from year to year, although in the period between 2007 and 2017 there was no significant change.



**Fig. 3.** Number of publication from 2007 to 2023.

**Publication Trends Based on Author Productivity.** According to the database, there are a total of 474 authors who have published articles regarding the Use of Latent Class Analysis in the field of education.



**Fig. 4.** Top 10 authors with the largest number of articles.

Based on Figure 4, it can be seen that the two authors who have the highest number of publications, namely Wang Y and Garcia-Fernandes, have both published 4 publications. However, Garcia Fernandes received a higher number of citations, namely 31, while Wang Y only received 14 citations. Then, among the 10 authors listed above, Boscardin CK received the highest number of citations, namely 53 citations from 2 published articles.

**Publication Trends Based on Documents with the Highest Number of Citations. In**

Table 1, the top 5 articles with the highest number of citations in this research field are displayed.

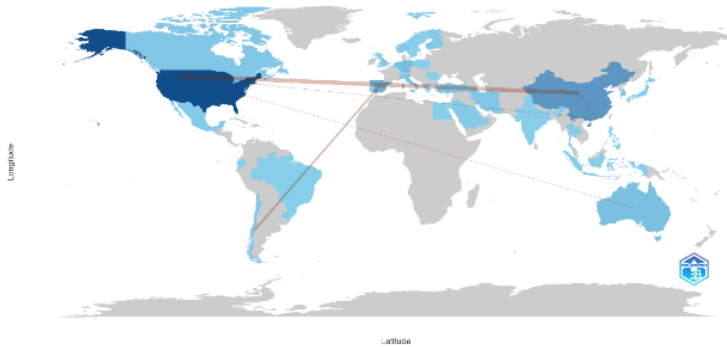
**Table 1.** Top 5 article with the highest number of citations.

No	Authors	Title	Cited by
1	Bradshaw C.P.; Buckley J.A.; Ialongo N.S.	School-Based Service Utilization Among Urban Children With Early Onset Educational and Mental Health Problems: The Squeaky Wheel Phenomenon	108
2	Penuel W.R.; Boscardin C.K.; Masyn K.; Crawford V.M.	Teaching with student response systems in elementary and secondary education settings: A survey study	44
3	Chow A.; Salmela-Aro K.	Task-values across subject domains: A gender comparison using a person-centered approach	42
4	Krell M.; Upmeier in Belzen A.; Krüger D.	Students' Levels of Understanding Models and Modelling in Biology: Global or Aspect-Dependent?	39
5	Bozick R.; DeLuca S.	Not making the transition to college: School, work, and opportunities in the lives of American youth	31

In Table 1 above, it can be seen that the articles that have the most significant influence are those written by Bradshaw C.P., Buckley J.A., and Ialongo N.S. with the title "School-Based Service Utilization Among Urban Children with Early Onset

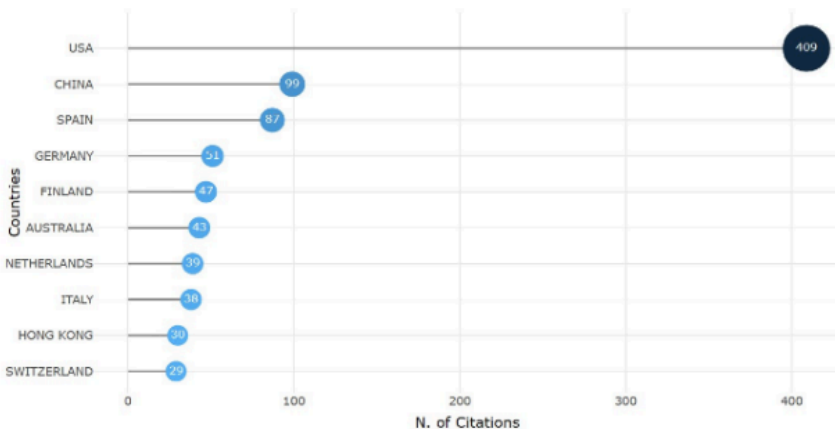
Educational and Mental Health Problems: The Squeaky Wheel Phenomenon". This article was published by School Psychology Quarterly and has received 108 citations.

**Geographic Distribution Based on Collaboration Between Countries and Number of Citations.** Figure 5 below depicts the distribution of the authors' countries of origin as well as collaboration between countries in publications related to latent class analysis in the field of education.



**Fig. 5.** Geographical distribution of publications.

It can be seen that there are several significant collaborations, such as the collaboration between the USA and China, the USA and Australia, as well as the collaboration between Chile and Spain. Furthermore, the geographical distribution can also be seen based on the number of citations from various countries, which can be seen in Figure 6 below.



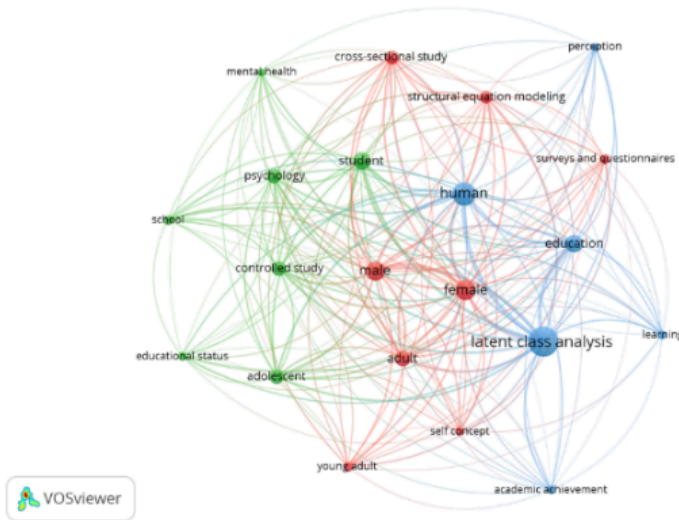
**Fig. 6.** Table of top 10 countries with the highest number of citations.

Based on Figure 6, 7 countries were identified with the number of citations in research regarding latent class analysis in the field of education. It can be seen that the

United States (USA), China (China), and Spain are the most influential countries in this research based on the number of citations received. The United States dominates with 409 citations, indicating a very significant contribution to the development of science in this context. Meanwhile, China follows with 99 citations, and Spain with 87 citations.

### Publication Characteristics Based on Co-occurrence Analysis

*Author's Keyword Analysis.* At this stage, researchers determine the minimum threshold for keyword occurrences, namely 10. This means that only keywords that have been used together in at least 10 publications or more will be displayed in the analysis using VOSviewer. From the initial 1204 keywords, after applying this threshold, only 21 keywords remained displayed.



**Fig. 7.** Appearance with Keywords (Threshold of Appearance  $\geq 10$ ).

The research focus can be seen through the clusters shown in Figure 7. These clusters are differentiated through various colors that represent research related to latent class analysis in the field of education. There are three main clusters that can be identified, namely the red, green and blue clusters.

*Keyword Trends.* Keyword trend analysis is used to determine the novelty of a research. Figure 8 keyword trends in Latent Class Analysis research publications in the field of education. the colors show publication trends by year. The bluish color indicates keywords used in the early years, while yellow indicates keywords that appear in recent publications.





Publication trends based on documents with the highest number of citations, the analysis results show that the article written by Bradshaw et al., (2008) had the most significant impact, the article entitled "School-Based Service Utilization Among Urban Children with Early Onset Educational and Mental Health Problems: The Squeaky Wheel Phenomenon". This article was published by *School Psychology Quarterly* and has received 108 citations. The significant impact of this article shows that the findings or methodology presented in this article is highly relevant and valuable to the research community. This is in line with [19], which states that the citation level shows how much research has been referred to as a reference for other research, the higher the citation level shows the research has become a strong reference for other researchers. The high number of citations indicates that this article has become a key reference in the literature related to latent class analysis in educational contexts.

The geographic distribution of citations in this study identifies the 10 countries with the highest number of citations in research on latent class analysis in education. The increasing number of collaborations indicates that there are advantages and benefits obtained [20]. The United States (USA), China (China), and Spain emerged as the most influential countries based on the number of citations received. Apart from the top three countries, seven other countries also made important contributions to this research. This illustrates the diversity of collaboration and global contributions to the understanding of latent class analysis in education. This information can be used to identify the role of different countries in the development of science and understand the extent to which this research influences the development of knowledge around the world. It also provides important insights for educational researchers and practitioners about collaborations that have the potential to support future research developments.

Publication characteristics based on co-occurrence analysis, the results of this cluster analysis provide valuable insight into research trends in latent class analysis in education. Cluster colors reflect changes in publication trends over time. A number of keywords, such as "mental health," "structural equation modeling," "surveys and questionnaires," and "adult," indicate new developments in this research, indicating a shift in research focus and researchers' interests. This information is very useful for researchers and academics in choosing appropriate and relevant research topics in latent class analysis in education.

#### 4 Conclusions

Based on the research results and discussion, the trends observed from the analysis of 130 documents related to Latent Class Analysis in the field of education from 2007 to 2023 show a significant increase in research focus. Publications in 2019 recorded the highest achievement in terms of citations. The analysis also revealed that some authors, such as Wang Y and Garcia-Fernandes, are major contributors in publications related to this topic. Meanwhile, Boscardin CK was one of the authors whose work was most cited, demonstrating his important contribution to developing an understanding of Latent Class Analysis. One of the most influential articles in this domain is Bradshaw CP, Buckley JA, and Ialongo NS's "School-Based Service

Utilization Among Urban Children with Early Education and Mental Health Problems: The Squeaky Wheel Phenomenon". The publication, published in *School Psychology Quarterly*, has received 108 citations. The cross-national collaborations were between the United States and China, the United States and Australia, and between Chile and Spain. Countries such as the United States, China, and Spain were listed as the countries with the highest number of citations. Analysis of keyword trends showed a shift in research focus in new directions, with increased interest in topics such as "mental health", "structural equation modeling", "surveys and questionnaires", and "adults". This suggests an evolution in the way Latent Class Analysis is approached in the field of education, reflecting efforts to understand more about the psychological and social implications in educational contexts. The four parts of the research focus can be used as a guideline for future research that wants to take the theme of Latent Class Analysis, especially in the field of education.

The limitations of this research are; 1) the data analyzed comes from the Scopus database, so there are still many other databases that can be used such as Wos and others; 2) this research only discusses Latent Class Analysis in the field of education in general, so there are still many other fields that can be further researched; 3) the data in this study taken on September 2, 2023 cannot reflect research after the research is carried out further; 3) the data in this study taken in this study cannot reflect research after that time, so there may be slight differences.

## References

1. Nurkholis, "Education in Efforts to Advance Technology" *Jurnal Kependidikan*, vol. 1, no. 1, pp. 24–44, Jan. 2013, doi: 10.24090/JK.V1I1.530.
2. N. Hayati *et al.*, "Exploring the Potential of Creativity and Innovation: The Role of Character Education at MTS Miftahus Sudur Campor Proppo," *Dewantara: Jurnal Pendidikan Sosial Humaniora*, vol. 2, no. 3, pp. 111–128, Jul. 2023, doi: 10.30640/DEWANTARA.V2I3.1351.
3. T. Pardede, "Latent Class Analysis for Categorical Data Grouping," in *Prosiding Seminar Nasional Matematika, Sains, dan Teknologi*, Universitas Terbuka, 2013, p. A.1-A.16.
4. B. E. Weller, N. K. Bowen, and S. J. Faubert, "Latent Class Analysis: A Guide to Best Practice," *Journal of Black Psychology*, vol. 46, no. 4, pp. 287–311, May 2020, doi: 10.1177/00957984209320932/ASSET/IMAGES/LARGE/10.1177\_00957984209320932-FIG1.JPEG.
5. W. Widhiarso and M. N. R. Hadjam, "Application of latent class analysis to detect unique characteristics in the construct of teacher efficacy in teaching," *Jurnal Penelitian dan Evaluasi Pendidikan*, vol. 20, no. 2, pp. 244–254, Nov. 2016, doi: 10.21831/PEP.V20I2.7698.
6. Riswan, "Grouping Mathematics Achievement of Indonesian Students Based on TIMSS Survey Results Using Latent Class Logistic Analysis," vol. 13, no. 1, 2013.
7. A. B. D. Nandiyanto, R. Ragadhita, N. D. Al Husaeni, and W. C. Nugraha, "Research trend on the use of mercury in gold mining: Literature review and bibliometric analysis," *Moroccan Journal of Chemistry*, vol. 11, no. 1, pp. 11-1 (2023) 01-19, 2023, doi: 10.48317/IMIST.PRSM/MORJCHEM-V11I1.36576.
8. Zafrullah *et al.*, "Item Response Theory In Education: A Biblioshiny Analysis (1987-2023)," *Journal of Education Global*, vol. 1, no. 1, pp. 101–114, Jun. 2023.

9. N. Sanusi, F. Agus Triansyah, I. Muhammad, S. Susanti, K. Komunikasi, and P. Tinggi, "Bibliometric Analysis: Communication Skills Research in Higher Education," *JIIP - Jurnal Ilmiah Ilmu Pendidikan*, vol. 6, no. 3, pp. 1694–1701, Mar. 2023, doi: 10.54371/JIIP.V6I3.1763.
10. E. Kusumawati, "Digital Leadership in Education: A Bibliometric Analysis," *Journal of Education and Teaching (JET)*, vol. 4, no. 2, pp. 252–260, Mar. 2023, doi: 10.51454/JET.V4I2.254.
11. I. Muhammad, F. Agus Triansyah, A. Fahri, and B. Lizein, "Bibliometric Analysis: Self-Efficacy Research in High Schools (1987-2023)," *Edukatif: Jurnal Ilmu Pendidikan*, vol. 5, no. 1, pp. 519–532, Mar. 2023, doi: 10.31004/EDUKATIF.V5I1.4713.
12. S. Ulwiyah, Z. Zafrullah, R. T. Ayuni, and A. Wahyuni, "View of The Use of SPSS in Mathematics Education: Biblioshiny & Bibliometric Analysis (1997-2023)," *Journal of Technology Global*, vol. 1, no. 1, pp. 26–33, Jun. 2023.
13. Zafrullah, S. Ulwiyah, and Nofriyandi, "Rasch Model Analysis On Mathematics Test Instruments: Biblioshiny (1983-2023)," *Mathematics Research and Education Journal*, vol. 7, no. 2, pp. 1–13, Nov. 2023, doi: 10.25299/MREJ.2023.VOL7(2).14550.
14. Van Eck, Nees Jan and Waltman, Ludo, "Vosviewer: A Computer Program for Bibliometric Mapping (November 2009 2)," ERIM Report Series Reference No. ERS-2009-005-LIS, Available at SSRN: <https://ssrn.com/abstract=1346848>
15. M. Aria and C. Cuccurullo, "bibliometrix : An R-tool for comprehensive science mapping analysis," *Journal of Informetrics*, vol. 11, no. 4, pp. 959–975, Nov. 2017, doi: 10.1016/j.joi.2017.08.007.
16. I. Muhammad and D. Juandi, "Discovery Learning Model in Junior High School Mathematics Learning: A Bibliometric Review," *Euler : Jurnal Ilmiah Matematika, Sains dan Teknologi*, vol. 11, no. 1, pp. 74–88, Jun. 2023, doi: 10.34312/euler.v11i1.20042.
17. D. R. Aulianto, P. M. Yusup, and Y. Setianti, "Utilization of the 'Publish Or Perish' Application as a Citation Analysis Tool in the Journal of Communication Studies, Padjadjaran University," in *Seminar Nasional MACOM III Universitas Padjadjaran 2019 "Communication and Information Beyond Boundaries,"* 2019, pp. 873–880.
18. C. P. Bradshaw, J. A. Buckley, and N. S. Ialongo, "School-Based Service Utilization Among Urban Children With Early Onset Educational and Mental Health Problems: The Squeaky Wheel Phenomenon," *School Psychology Quarterly*, vol. 23, no. 2, pp. 169–186, Jun. 2008, doi: 10.1037/1045-3830.23.2.169.
19. Y. R. Fiandari, "Bibliometric Analysis of Research Development Publications on Digital Marketing of Halal Cosmetics," *Jurnal Dokumentasi Dan Informasi*, vol. 43, no. 2, pp. 95–111, Dec. 2022, doi: 10.14203/J.BACA.V43I2.843.
20. M. Amelia and R. Rahmaida, "Scientific Productivity of Indonesian Researchers on Indonesian Biodiversity Research Based on Scopus Database 1990-2015," *J Biologi Indonesia*, vol. 13, no. 2, Nov. 2017, doi: 10.14203/jbi.v13i2.3398

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

