**Dr. Mohit Kumar**
PhD in Management, Assistant Professor, Department of Business Management,
Chhatrapati Shahu Ji Maharaj University, Kanpur, Uttar Pradesh, India.
ORCID ID 0000-0003-2532-4012
*drmohitkumar@csjmu.ac.in*

**Dr. Pravin Kumar Agrawal**
PhD in Management, Assistant Professor, Department of Business Management,
Chhatrapati Shahu Ji Maharaj University, Kanpur, Uttar Pradesh, India.
ORCID ID 0000-0003-4331-8767
*drpravinagrawal@csjmu.ac.in*

**Dr. Supriya Agrawal**
PhD in Management, Assistant Professor, Amity Business School.
Amity University Lucknow, Lucknow, India.
ORCID ID 0000-0001-7187-1354
*sagrawal@lko.amity.edu*

**Vibha Kushwaha**
Research Scholar, Department of Business Management,
Chhatrapati Shahu Ji Maharaj University, Kanpur, Uttar Pradesh, India.
ORCID ID 0000-0003-2313-2391
*vibhak15@gmail.com*

# MASSIVE OPEN ONLINE COURSES IN EDUCATION: A BIBLIOMETRIC STUDY OF GLOBAL TRENDS

**Abstract.** The past decade has witnessed a revolution in the learning landscape worldwide, driven by the rapid expansion of digital technologies and open-access learning platforms. Among them, Massive Open Online Courses (MOOCs) have been at the centre stage of learning innovation, promising flexible, low-cost, and borderless learning. This study aims to map the global research landscape of Massive Open Online Courses (MOOCs) using bibliometric techniques to understand the evolution, intellectual structure, and thematic trends of the field. Drawing from a dataset of 10,059 publications indexed in Scopus, the analysis was conducted using Bibliometrix (R package) and VOSviewer to perform co-citation analysis, bibliographic coupling, and thematic mapping. The results reveal that the United States, China, and the United Kingdom are the leading contributors in terms of research output. Key institutions such as MIT and Stanford dominate citation impact, while authors like Yuan L. and Alario-Hoyos S. show high scholarly influence. The analysis reveals a mismatch between research output volume and scholarly influence, particularly in certain regions. Thematic mapping highlights core clusters in pedagogy, learner engagement, and assessment tools, while underexplored areas include inclusivity, multilingual access, and international policy coordination. A key contribution of this study lies in uncovering structures in the MOOC literature which reveal hidden intersections between technical, pedagogical, and policy-related domains. The findings underscore the growing maturity of MOOCs as an interdisciplinary field, while also pointing to fragmentation in research efforts. This study not only charts the growth and diversification of MOOC research but also highlights neglected areas crucial for educational equity and global collaboration. The findings have applied value for educators, platform developers, and policymakers seeking to expand the reach, quality, and impacts of MOOCs in an evolving digital learning environment.

**Keywords:** Massive Open Online Courses (MOOCs); Bibliometric Analysis; Online Education; Thematic Mapping; Co-citation Analysis.

## 1. INTRODUCTION

The emergence of Massive Open Online Courses (MOOCs) is perhaps the most revolutionary in the history of learning. MOOCs broke into headlines with the appearance in 2012 of high-profile brands such as Coursera, edX, and Udacity that were widely dubbed as the

© Mohit Kumar, Pravin Kumar Agrawal, Supriya Agrawal, Vibha Kushwaha, 2025. 

"Year of the MOOC" [1]. They declared the democratization of education through free, scalable, access to high-quality learning opportunities provided by best-league institutions for a global marketplace.

MOOCs originally rose to fame as being capable of bridging learning gaps, largely for students from disadvantaged regions who were unable to pursue mainstream higher education. The theorized framework underlying MOOCs – was open-access by definition, modular pedagogy, and mass digital delivery [2]. Eventually, however, their pedagogic flaws, dropout rates, and unreliable learner interaction patterns began attracting criticism [3], [4]. Nevertheless, MOOCs have endured – with variations of adaptive technology, interactive assessment prototypes, and continue to be an active force in the education sector.

The COVID-19 pandemic also sparked greater demand for MOOCs as universities globally were compelled to adopt remote or hybrid learning models. MOOCs evolved from optional supplements to essential tools for large-scale remote education. Consequently, MOOC literature has spread explosively across disciplines as varied as education, computer science, psychology, instructional design, to public policy.

While there have been a number of systematic reviews and meta-analyses of some sub-topics – e.g., learner motivation [5], instructional video production [6] or test design methods [7] – few studies have attempted to charter global MOOC literature.

In this context, bibliometric analysis provides a robust methodological framework to examine the trajectory, impact, and clustering of scholarly knowledge within a large body of academic work [8].

Bibliometrics enables researchers to quantitatively assess trends in publication output, authorship, collaboration networks, citation patterns, keyword co-occurrences, and the conceptual and social structures of a field. This study applies an advanced bibliometric approach – utilizing tools such as Bibliometrix in R, VOSviewer – to analyze a dataset of 10,059 MOOC-related publications indexed in Scopus between 2009 and 2024. In addition to traditional performance analysis, this study incorporates co-citation analysis**,** bibliographic coupling**,** keyword co-occurrence mapping to uncover both visible and latent themes in the literature. The goal of the research is to evaluate the emergence, structure, and impact of MOOC-related research between 2009 and 2024 with a comprehensive bibliometric approach. It aims to trace how the body of knowledge has evolved conceptually and intellectually, identify the leading researchers and global collaboration networks including the dominant and missing research threads.

This study aims to systematically explore the evolution, thematic composition, and intellectual structure of global research on Massive Open Online Courses (MOOCs). By identifying influential authors, institutions, citation patterns, and emerging themes, the study seeks to uncover how MOOC-related scholarship has developed over time and to highlight key areas for future academic and policy attention.

To achieve this, the study explores the following research questions:

- RQ1: What are publication trends, leading authors, and main sources of MOOC literature between 2009 and 2024?
- RQ2: How is the global research in context of MOOC in terms of collaborative institutional engagement, and citation power?
- RQ3: What are the overarching thematic clusters and intellectual structures in MOOC research, and how have they evolved over time?
- RQ4: What is the alignment between research output volume and scholarly impact?

By providing a multidimensional analysis of MOOC research over 15 years, this paper contributes to the meta-scientific understanding of how the field has matured and reoriented itself in response to both technological innovation and global educational challenges. The findings will be of value not only to scholars in education and learning technologies but also to

institutional policymakers, platform developers, and global education stakeholders aiming to harness the full potential of MOOCs in shaping the future of learning.

## 2. LITERATURE REVIEW

In the past decade, MOOCs (Massive Open Online Courses) transformed the global landscape of higher education and turned into a strong force in online learning. Developed as an innovative solution for open, scalable, and accessible education, MOOCs were hailed as democratizing agents who would end persistent issues of educational inequity and access. Early enthusiasm for MOOCs was based on their perceived potential to disrupt pedagogical paradigms, particularly in higher education, by allowing institutions to reach thousands – if not millions – of students across geographic and socio-economic boundaries. This buzz has been supported with the launch of best-in-class platforms such as Coursera, edX, Udacity, and FutureLearn, which made courses developed by the best universities and scholars accessible to the world for free or at a small fee. Pappano's (2012) pronouncement that 2012 would be the "Year of the MOOC" encapsulated the international fervor surrounding the innovation and the excitement that it was about to transform the manner in which teaching and learning were being delivered.

Initial scholarly research over MOOCs was largely conceptual and exploratory in nature, and it revolved around understanding distinctions between cMOOCs (connectivist MOOCs) that emphasized networked, learner-driven learning, and xMOOCs (extended MOOCs), which were more structured, teacher-driven like conventional online courses.

This was soon followed by a great amount of empirical research, wherein learners' demographics, motivation, engagement, and completion rates were studied by researchers. Jordan's (2014) data analysis of MOOCs demonstrated that student dropout rates often exceeded 90 percent, prompting concerns about learner persistence and also course design. Subsequent studies explored how course structure, peer-to-peer interaction, assessment type, and video length influenced learners' results. For instance, Guo, Kim, and Rubin (2014) found that shorter, more engaging videos significantly improved learner retention, suggesting that traditional lecture formats were ill-suited to the MOOC environment. These findings emphasized the need for instructional design that was specifically adapted to the online, large-scale nature of MOOCs.

Studies, like Jordan's (2014), showed student dropout rates often exceeding 90%, raising concerns about learner persistence and course design. Research by Guo, Kim, and Rubin (2014) suggested that shorter, more engaging videos improved learner retention, indicating the need for instructional design specifically adapted to the large-scale online environment of MOOCs. Margaryan, Bianco, and Littlejohn (2015) evaluated MOOCs against instructional design benchmarks, finding that many failed to align with basic principles of feedback and assessment. Issues such as language barriers, low digital literacy, and poor internet connectivity also restricted fair participation, leading to a demand for localized, translated, and culturally responsive MOOC materials, reflecting a move towards "inclusive MOOCs".

While much of the early research centered on learner behavior, another category of literature emerged around instructional quality and design principles. Margaryan, Bianco, and Littlejohn (2015) evaluated MOOCs against established instructional design benchmarks and concluded that many failed to align with basic principles of feedback, and assessment.

Language barriers, low levels of digital literacy, and poor quality internet connectivity restricted fair participation. In response, there has been growing demand for MOOC study materials to be localized, translated into multiple languages, and designed as culturally responsive courses that reflect diverse populations' demands. All of these efforts reflect a

broader movement in the literature towards so-called "inclusive MOOCs" that attempt to rethink the model with international learners in mind.

But sustainability of MOOC remains a concern. It has been challenging for most universities to generate revenue or academic credit through MOOCs, creating doubts about long-term sustainability and strategic alignment. Although certain platforms have created monetization options – such as paid certificates, company training units, and subscription-based models – the literature remains divided arguing about the possibility for MOOCs to transition from the experimental phase into a financially and pedagogically sustainable segment of higher education ecosystems.

More recent progress in MOOC research is represented through the uptake of learning analytics, artificial intelligence, and adaptive learning technologies. Literature review studies such as those by Veletsianos and Shepherdson (2016) and Zawacki-Richter et al. (2018) have attempted to make sense of broad-scale trends in MOOC research, even though fewer have employed advanced bibliometric techniques to characterize the intellectual landscape of the field, collaboration networks, and conceptual evolution.

In order to respond to this limitation, bibliometric analysis offers a strong, objective, and scalable means of analysing how the field of MOOC has developed over time, who are its leading contributors, what topics underlie its dialogue, and where the emerging frontiers are being created. The current research positions itself in this new methodological shift, seeking to build a rich and multi-faceted picture of research in MOOC in education.

## 3. RESEARCH METHODOLOGY

Bibliometric analysis is a quantitative approach to examining, evaluating, and mapping the structural and dynamic character of research scholarly publications in a particular field. It allows researchers to trace intellectual evolution, uncover hidden patterns in the literature, identify dominant authors or institutions, and trace emerging themes.

Although some studies have opted for the Web of Science (WOS) for bibliometric analysis, for this study, Scopus database was used for the following two reasons: (1) Scopus holds a wider coverage of research on MOOCs in areas like education, computer science, and social sciences, particularly after 2010; and (2) Scopus provides more comprehensive export of citation and bibliographic information, including abstracts, affiliations, and references, which is critical for advanced analysis like co-citation networks and topic modelling. The search was limited to documents published between 2009 and 2024.

### 3.1 Study Design

The bibliometric study employed a systematic three-stage research design as presented in Figure 1. The identification of keywords and development of search strings were conducted in the first stage. Following the dual strategy presented by Chen and Xiao (2016), high-frequency words as well as conceptually significant terms were considered. The terms "MOOC," "massive open online courses," "online learning," and "e-learning" were chosen because they occur most frequently in the literature and are used in the context of educational technology. Keywords were searched within the title, abstract, and keywords of Scopus for comprehensive coverage and thematic relevance. 10,059 records were retrieved from Scopus using the refined search query.

Search Query: ("massive open online course*" OR "MOOC*" OR "massive open online learning") AND ("education" OR "learning" OR "teaching" OR "instructional design" OR "higher education")

Only documents published in English and indexed in the appropriate subject areas (Education, Computer Science, Social Sciences, and Engineering) were included. Incomplete metadata, duplicate records, and unrelated documents were removed in data cleaning to end up with a good-quality dataset ready for more detailed analysis.

### 3.2 Data Selection and Preparation

During the second step, bibliographic information of the downloaded documents was retrieved in .csv formats using Scopus. The data set had document metadata such as titles, abstracts, keywords, year of publication, authorship, affiliation, citations, and references. The files were then imported into R Studio, and the Bibliometrix R package [8] was used for preliminary data treatment and descriptive statistical analysis. These main indicators such as yearly scientific production, most productive authors, country contributions, institutional affiliations, and document types were computed.

In addition, co-citation networks, bibliographic coupling, and keyword co-occurrence matrices were also generated using VOSviewer, which supported high-resolution visualizations of thematic clusters and intellectual structures of the data set.

### 3.3 Data Analysis and Visualization

During the third phase, the research utilized different levels of analysis and visualization. Descriptive bibliometric analysis was carried out using Bibliometrix through its web interface Biblioshiny to generate tables and plots of publication evolution, most cited journals, frequency of citation, and contributions of contributors. The conceptual structure of the literature was examined through keyword co-occurrence networks, thematic maps and bibliographic coupling which helped in revealing implicit relationships between education ideas and electronic learning models. Bibliographic coupling is a bibliometric technique used in the measurement of document similarity based on shared references cited. Two papers can be said to be bibliographically coupled if they cite one or more shared sources. The higher the number of shared references cited, the stronger the coupling. Bibliographic coupling detects research papers most likely writing on similar topics or belonging to the same thematic group although they might not cite each other.

In bibliometric studies, intellectual structure and conceptual structure are two complementary elements of academic research. The conceptual structure refers to the underlying topics, themes, and ideas that make up the subject matter. It is likely to be investigated via techniques such as keyword co-occurrence, thematic mapping, and topic modeling, which identify how concepts cluster across documents. Thematic maps categorize research themes by centrality and density, classifying them into basic, motor, niche, and emerging themes. The intellectual structure, however, mirrors the underlying knowledge base and influence between scholars or publications. It is investigated through techniques such as co-citation analysis, which traces frequency of pairing authors or work being cited together and identifies most influential contributors and schools of thought within the field. Along with these aspects, they tend to discover what the field is all about (conceptual) and who has influenced it (intellectual).

Combining graph-based visualizations, network analysis, and topic modelling produced a multi-dimensional representation of MOOC research's knowledge structure.

## 4. DATA ANALYSIS

*Table 1.*

**Main Information**

| Description | Results |
|---|---|
| **MAIN INFORMATION ABOUT DATA** | |
| Timespan | 2009:2024 |
| Sources (Journals, Books, etc) | 3069 |
| Documents | 10059 |
| Annual Growth Rate % | 56.03 |
| Document Average Age | 5.86 |
| Average citations per doc | 13.37 |
| References | 248666 |
| **DOCUMENT CONTENTS** | |
| Keywords Plus (ID) | 18362 |
| Author's Keywords (DE) | 15997 |
| **AUTHORS** | |
| Authors | 18112 |
| Authors of single-authored docs | 1313 |
| **AUTHORS COLLABORATION** | |
| Single-authored docs | 1603 |
| Co-Authors per Doc | 3.26 |
| International co-authorships % | 17.59 |
| **DOCUMENT TYPES** | |
| article | 4207 |
| book | 89 |
| book chapter | 728 |
| conference paper | 5035 |

The bibliometric data extracted from Scopus ranges between the years 2009 and 2024 and comprises 10,059 documents in 3,069 different channels of publication, including journals, books, book chapters, and conference papers. This scope reflects the multidisciplinary nature of Massive Open Online Courses (MOOCs) research that crosses over into educational theory, digital pedagogy, instructional design, and learning analytics. The rate of publication growth per year in this area is a staggering 56.03%, reflecting the rapidly developing academic interest in MOOCs, particularly after the milestone year of 2012, also referred to as the "Year of the MOOC."

Each document, on average, has been cited 13.37 times, suggesting that MOOC-related research has not only proliferated but has also been widely referenced, reflecting its academic relevance and cross-disciplinary appeal. Moreover, the total number of references cited in these documents exceeds 248,000, suggesting a densely interconnected body of scholarship.

In terms of authorship, the MOOC literature includes contributions from 18,112 unique authors, indicating a broad base of academic involvement. Out of these, 1,313 authors have published single-authored papers, resulting in 1,603 such documents, pointing to a modest but meaningful corpus of work that is likely conceptual, theoretical, or reflective in nature. The average number of co-authors per document is 3.26, which signals a high degree of collaboration within the field. However, the rate of international co-authorship remains relatively modest at 17.59%. This suggests that while collaboration is common, much of it may be occurring within national or regional contexts.

The composition of document types reveals further insights into the research landscape. Out of the total documents, 4,207 are journal articles and 5,035 are conference papers. In

addition to these, there are 728 book chapters and 89 books, indicating ongoing efforts to consolidate knowledge and develop comprehensive treatments of MOOC pedagogy, design, and policy.

**Annual Scientific Production**

*Table 2.*

**Annual Scientific Production**

| Year | Articles | Year | Articles |
|---|---|---|---|
| 2009 | 1 | 2017 | 903 |
| 2010 | 0 | 2018 | 1011 |
| 2011 | 8 | 2019 | 1040 |
| 2012 | 14 | 2020 | 1003 |
| 2013 | 207 | 2021 | 1098 |
| 2014 | 514 | 2022 | 963 |
| 2015 | 739 | 2023 | 914 |
| 2016 | 853 | 2024 | 791 |

The above table reflects pattern of publications in Massive Open Online Courses (MOOCs) over time. From a few articles prior to 2012—with only 1 in 2009 and none in 2010—the field experienced an abrupt surge beginning in 2013 with their 207 articles. The later years experienced continued momentum, and publication numbers increased steadily to a peak in 2021 with 1,098 articles.

This peak corresponds with the post-pandemic academic climate, in which online learning models, including MOOCs, became not just alternatives but essential components of educational continuity. The slight decline observed in 2022 (963 articles), 2023 (914 articles), and 2024 (791 articles, till date) may reflect a phase of saturation or shifting focus toward hybrid or micro-credential-based systems that evolved from the original MOOC model.

*Table 3.*

**Most Relevant Sources**

| Sources | Articles |
|---|---|
| ACM International Conference Proceeding Series | 456 |
| Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) | 424 |
| CEUR Workshop Proceedings | 234 |
| Communications In Computer and Information Science | 172 |
| Advances In Intelligent Systems and Computing | 109 |
| International Review of Research in Open and Distributed Learning | 109 |
| Lecture Notes in Networks and Systems | 108 |
| IEEE Global Engineering Education Conference, Educon | 107 |
| International Journal of Emerging Technologies in Learning | 107 |
| Computers and Education | 89 |

The analysis of source-wise distribution highlights the interdisciplinary nature and strong technical orientation of MOOC-related research. The ACM International Conference Proceeding Series emerges as the most prolific outlet, contributing 456 documents, closely followed by Lecture Notes in Computer Science (including subseries in Artificial Intelligence and Bioinformatics) with 424 publications. This dominance of computer science-oriented proceedings underscores the technological foundations of MOOC platforms and the active involvement of computer science researchers in advancing MOOC infrastructure, learning

analytics, and intelligent systems in education. Interestingly, while technical venues dominate the top positions, certain education-focused journals are also prominently featured. The International Review of Research in Open and Distributed Learning (IRRODL) and the International Journal of Emerging Technologies in Learning (iJET) both contributed 109 and 107 documents respectively, affirming the field's grounding in open education, pedagogical innovation, and distance learning methodologies.

**Most Relevant Authors**

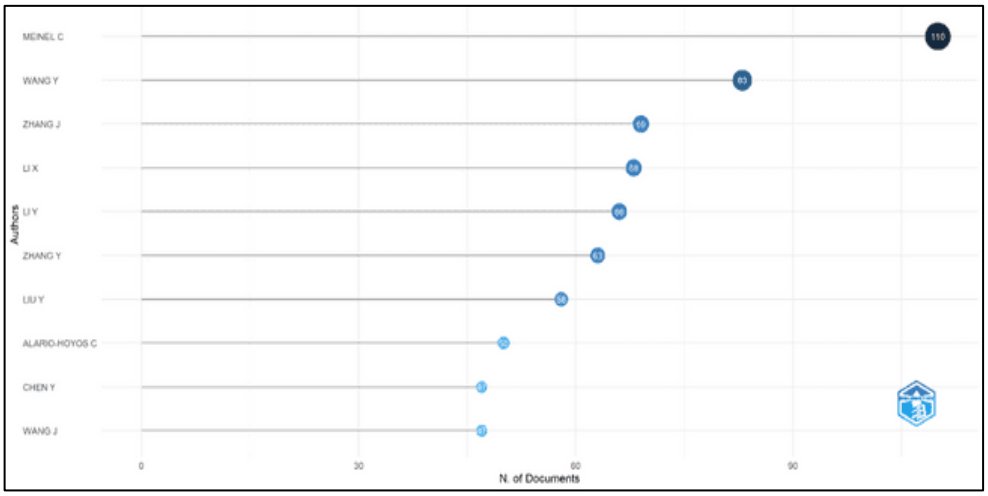


*Figure 1. Most Relevant Authors*

Authorship analysis sees a tight cluster of very productive authors with a high prevalence of MOOC-related work. At the forefront is Menel C, with an outstanding achievement of 110 published articles, representing a consistent and even activity in the area. Second comes Wang Y, with 83 papers, and a group of authors, including Zhang J, Li X, and Lu Y, with each having between 68 to 69 publications.

This form of academic involvement suggests not only individual dedication but probabl y membership in collective cohorts of researchers or institutional research programs targeted toward MOOCs.

The trend also indicates potential for new researchers to co-author alongside these thought leaders, or critically engage with their work through citation and co-authorship network studies.

**Most Relevant Affiliations**

*Table 4.*

**Most Relevant Affiliations**

| Affiliation | Articles |
|---|---|
| Central China Normal University | 156 |
| Beijing Normal University | 124 |
| University of Tasmania | 117 |
| Massachusetts Institute of Technology | 100 |
| Universidad Carlos Iii De Madrid | 99 |
| Tsinghua University | 90 |
| Universidad De Valladolid | 85 |
| South China Normal University | 84 |
| Durham University | 83 |
| Open University of the Netherlands | 81 |

Analysis of the most relevant affiliations in international literature on MOOCs in education establishes a number of top institutions who have made important contributions in this area. Central China Normal University takes the lead with 156 publications, followed by Beijing Normal University at 124. This is further evidence of Chinese universities' leadership in MOOC study and development of dissemination, aligning with the overall policy-level support of the country for digital studies and initiatives and its global educational technology discourse positioning strategy.

Outside of China, some internationally well-known institutions appear as principal contributors such as the University of Tasmania (Australia) and Durham University (UK), perhaps as a result of research into pedagogic innovations, learner activity, and policy studies of MOOCs in developed education systems. Massachusetts Institute of Technology (MIT), with 100 publications, is a classic catalyst of the MOOC phenomenon, especially as it pioneered leadership in creating edX, one of the very first large MOOC platforms.

**Countries' Scientific Production**

*Table 5.*

**Most Relevant Affiliations**

| Country | Freq | Country | Freq |
|---|---|---|---|
| China | 6250 | Germany | 1200 |
| USA | 4856 | Australia | 1109 |
| Spain | 2387 | Malaysia | 780 |
| India | 1698 | Italy | 687 |
| UK | 1414 | France | 631 |

The analysis of countries' science output in the field of MOOCs in education reveals stark regional domination and research intensity. China is by far the leading contributor with 6,250 publications, leaving all other nations far behind. This commanding dominance reflects China's strategic emphasis on digital schooling, e-learning infrastructure, and mainstream takeup of MOOCs in policy and practice. The country's heavy investment in higher education reforms and platforms has likely spurred the research outputs of universities. The United States takes the second position with 4,856 publications, which reflects its role in the MOOC movement via adoption of platforms like Coursera, edX, and Udacity. Spain is the forefront European nation in this sector with 2,387 works. Universidad Carlos III de Madrid and Universidad de Valladolid are examples of Spanish institutions that excel in research into learner analytics, MOOC adoption in tertiary education, and providing multi-lingual courses.

**Most Global Cited Documents**

*Table 6.*

**Most Global Cited Documents**

| Paper | DOI | Total Citations | TC per Year | Normalized TC |
|---|---|---|---|---|
| GUO PJ, 2014, LS - PROC ACM CONF LEARN SCALE-a | 10.1145/2556325.2566239 | 1371 | 114.25 | 48.30 |
| LIYANAGUNAWARDENA TR, 2013, INT REV RES OPEN DISTANCE LEARN | 10.19173/irrodl.v14i3.1455 | 833 | 64.08 | 24.89 |

| | | | | |
|---|---|---|---|---|
| KIZILCEC RF, 2013, ACM INT CONF PROC SER | 10.1145/2460296.2460330 | 779 | 59.92 | 23.27 |
| KIZILCEC RF, 2017, COMPUT EDUC | 10.1016/j.compedu.2016.10.001 | 714 | 79.33 | 47.67 |
| JORDAN K, 2014, INT REV RES OPEN DISTANCE LEARN | 10.19173/irrodl.v15i1.1651 | 672 | 56.00 | 23.67 |
| ZHANG J, 2017, INT WORLD WIDE WEB CONF, WWW | 10.1145/3038912.3052580 | 615 | 68.33 | 41.06 |
| MARGARYAN A, 2015, COMPUT EDUC | 10.1016/j.compedu.2014.08.005 | 548 | 49.82 | 23.39 |
| HONE KS, 2016, COMPUT EDUC | 10.1016/j.compedu.2016.03.016 | 543 | 54.30 | 29.27 |
| ALRAIMI KM, 2015, COMPUT EDUC | 10.1016/j.compedu.2014.08.006 | 538 | 48.91 | 22.97 |
| KAPLAN AM, 2016, BUS HORIZ | 10.1016/j.bushor.2016.03.008 | 522 | 52.20 | 28.14 |

The highest-cited articles in the MOOCs field outline the central and ground breaking contributions that have shaped theoretical and practical domains within this subject matter. The highest-cited paper by Guo PJ (2014), presented at the ACM Conference on Learning at Scale, has 1,371 citations with an impressive 114.25 per year and a total normalized citation (TC) score of 48.30.

Normalized Total Citations (TC) is a metric which standardizes raw citations to be variation-adjusted for publication year, allowing for more equitable comparisons across time. It is generated by dividing total citations by years since publication. It is particularly valuable in choosing influential papers regardless of publication year.

The supremacy of this research in all citation indices is a reflection of its ongoing influence, most likely due to its empirically based proof on video use and learning activity, which have become essential in the development of MOOC platforms. The second most highly cited article by Liyanagunawardena et al. (2013) in IRRODL has 833 total citations and a normalized TC of 24.89. As one of the initial systematic reviews of MOOCs, this article helped to define research agendas and induce significant thematic domains and hence contribute enormously to conceptual knowledge in the emerging field.

Kizilcec et al. appear twice on the list—once in 2013 and again in 2017—indicating sustained academic impact over time. Their 2013 ACM work was cited 779 times, while their 2017 work in Computers & Education was cited 714 times and had an extremely high normalized TC of 47.67, indicating its broader dissemination and lasting influence.

**Most Frequent Words & Word Cloud**

*Table 7.*

**Most Frequent Words**

| **Words** | **Occurrences** |
|---|---|
| e-learning | 3521 |
| massive open online course | 2641 |
| students | 2182 |
| teaching | 1935 |
| curricula | 1763 |
| mooc | 1432 |
| education | 1405 |
| learning systems | 1280 |
| moocs | 763 |
| education computing | 742 |

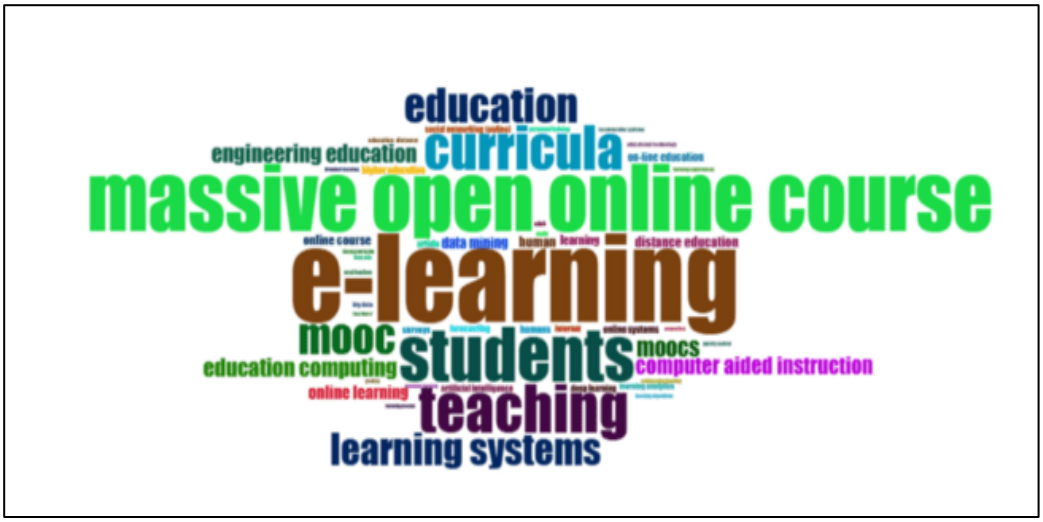


*Figure 2. Most Relevant words*

The analysis of the most frequently occurring words in the corpus of 10,059 MOOC-related articles published between 2009 and 2024 reveals clear thematic trends that define the research context in the field. Most frequently used is the term "e-learning" at 3,521 occurrences, highlighting its central position within the broader digital education context.

The phrase “massive open online course” occurs 2,641 times, while its acronym “MOOC” and plural form “MOOCs” appear 1,432 and 763 times respectively. This variation reflects both keyword indexing habits and the evolution of the discourse—where early literature may have used the full form and later adopted the abbreviated form as the concept gained familiarity.

Terms like “students” (2,182) and “teaching” (1,935) emphasize the learner-centered and pedagogical focus of the literature. These frequencies suggest that much of the research has revolved around how MOOCs affect learners’ experiences, engagement, and outcomes, as well as how educators adapt to or design for these platforms.

**Conceptual Structure**

**Thematic Map**

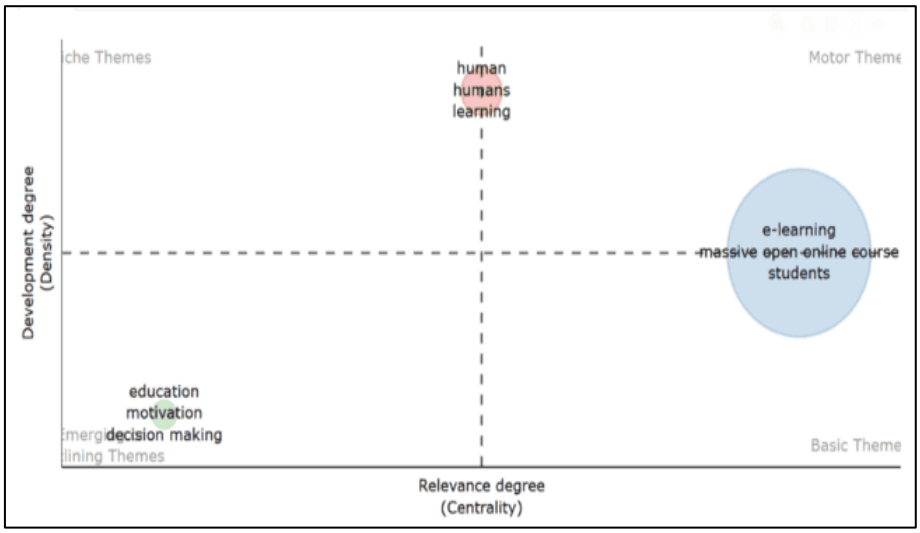


*Figure 3. Thematic Map*

The Figure 4 presents the thematic map. The thematic map plots clusters of keywords on two axes: centrality (horizontal axis) indicating the centrality of the theme to the field, and

density (vertical axis) indicating internal development of the theme. Based on this, themes are classified in four quadrants:

- Upper-right (Major Themes): Central and well-developed
- Upper-left (Niche Themes): Specialized but marginal
- Lower-left (Emerging or Declining Themes): Peripherally located and weakly developed
- Lower-right (Basic Themes): Highly central but underdeveloped areas

These quadrants shed light on both established and new frontiers of MOOC research. In the bottom-right quadrant, which represents Basic Themes, the terms *"e-learning," "massive open online course,"* and *"students"* appear prominently. These are foundational topics that are central to the field but not particularly well developed internally suggesting they are widely discussed but often in general or introductory ways. Their presence here confirms their importance as the core conceptual framework of MOOC research, especially regarding learner-centered studies and broad online education practices.

In the upper-central region**,** bordering the quadrant, we see terms like *"human," "humans,"* and *"learning."* These are part of more specialized bottom-left quadrant**,** research trajectories that, while not as central, are Niche Theme relatively well-developed. In the bottom-left quadrant**,** representing Emerging or Declining Themes**,** we find keywords such as *"education," "motivation,"* and *"decision making."* These are currently less central and underdeveloped, which might suggest either (a) these are new areas of inquiry that haven't been fully explored yet or they are declining in scholarly attention in this domain, possibly due to earlier saturation or shifts in focus.

**Intellectual Structure**

**Co-citation Network**

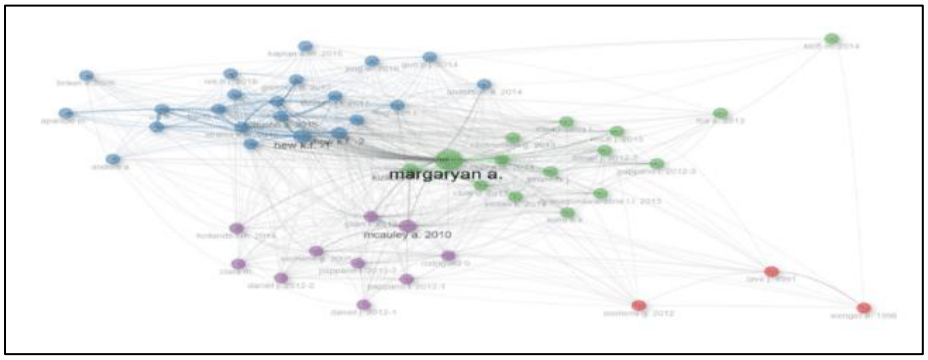


*Figure 4. Co-citation Network*

The co-citation network displayed in the image illustrates how often pairs of authors are cited together across the MOOC literature, revealing the intellectual topography which is also structural layout of influential authors, schools of thought, and research clusters within a scholarly domain, often revealed through co-citation analysis. At the center of the network, Margaryan A. emerges as the most influential and highly co-cited author, indicating their pivotal role in shaping core theoretical or empirical understandings within MOOC studies. Surrounding Margaryan are multiple clusters of co-cited authors, differentiated by colors, suggesting distinct thematic orientations**:**

- The green cluster**,** which includes authors like *Christensen G.*, *Jordan K.*, and *Guo P.J.*, appears to focus on empirical studies of learner behavior, engagement, and outcomes in MOOCs**.** Their work often combines large-scale learner data with evaluation of pedagogical design and course effectiveness.

- The blue cluster**,** containing *Kaplan A.M.*, *Reich J.*, and *Kizilcec R.F.*, reflects a technological and strategic focus**,** covering MOOC platform design, learner analytics, and institutional strategy for online education.
- The purple cluster**,** with *McAuley A.*, *Daniel J.*, and *Hollands F.M.*, is likely rooted in policy and systemic discourse**,** often addressing scalability, educational access, and the role of MOOCs in reforming higher education ecosystems.
- The red cluster**,** featuring foundational figures like *Wenger E.* and *Lave J.*, connects to theoretical frameworks like communities of practice and situated learning. Their continued citation suggests that MOOC research frequently draws from broader educational theory to interpret findings.

**Social Structure**

**Collaboration Network**

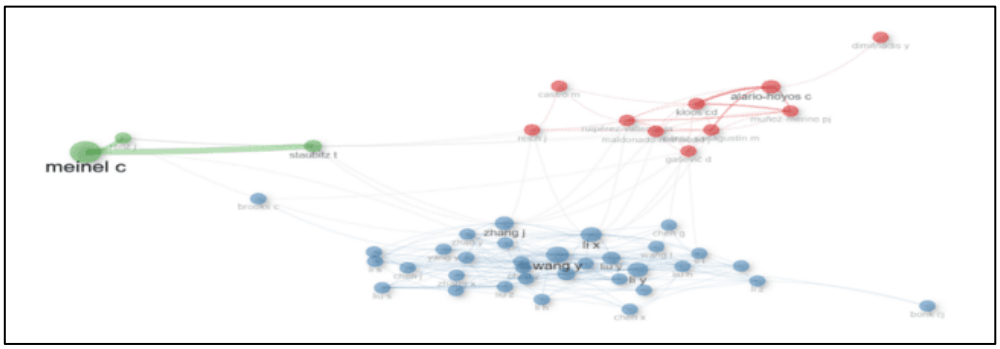


*Figure 5. Co-authorship Network*

Figure 6 shows the co-authorship network of researchers within the field of MOOCs in education. This illustrates author groups that dominate research publications in this research area. Three prominent clusters emerge from the network:

- The Blue Cluster centers around prolific authors such as Wang Y, Zhang J, and Liu Y, indicating a tightly knit and highly collaborative research group. This cluster demonstrates dense intra-group connectivity, suggesting sustained partnerships likely focused on data-driven analysis of MOOCs, including user behavior, learning outcomes, and platform analytics.
- The Red Cluster, led by Alario-Hoyos C and collaborators like Muñoz-Merino PJ and Kloos CD, shows moderate internal connectivity and links to other researchers such as Ruipérez-Valiente JA and Gasevic D. This group may be more involved in the pedagogical and technological design aspects of MOOCs, possibly with a European research focus given the regional affiliations of many authors.
- The Green Cluster, dominated by Meinel C and associates like Staubitz T and Ternier **S**, appears more isolated from the rest of the network. This suggests a self-contained yet influential research group, possibly affiliated with a specific MOOC platform or institution focusing on platform development, scalability, or technical implementation.

**Co-citation analysis by cited authors**

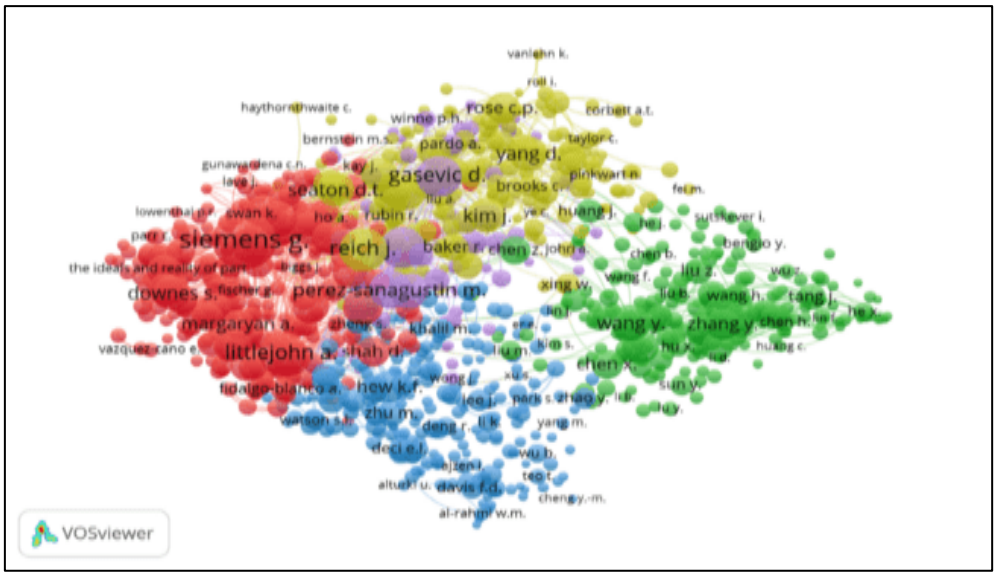


*Figure 6. Co-citation analysis by cited authors*

The Co-citation analysis of citing authors plots intellectual domains and influential contributors towards MOOC. Several distinct clusters emerge, each representing a thematic research community. One prominent cluster includes George Siemens, Stephen Downes, and Allison Littlejohn, who are central to the development of foundational concepts in open online education. The second largest cluster includes researchers such as Dragan Gasevic, Jihyun Kim, and Diyi Yang whose work mainly focuses on learning analytics and educational data mining. The third cluster represented by Yongchuan Wang and Yanyan Zhang has the technical inclination towards MOOC deployment and infrastructure, and this is best supplemented by artificial intelligence. The fourth cluster includes researchers such as K.F. Hew and M. Zhu who is interested in encouraging students, stimulating participation, and retaining drivers.

**Co-citation analysis by cited sources**

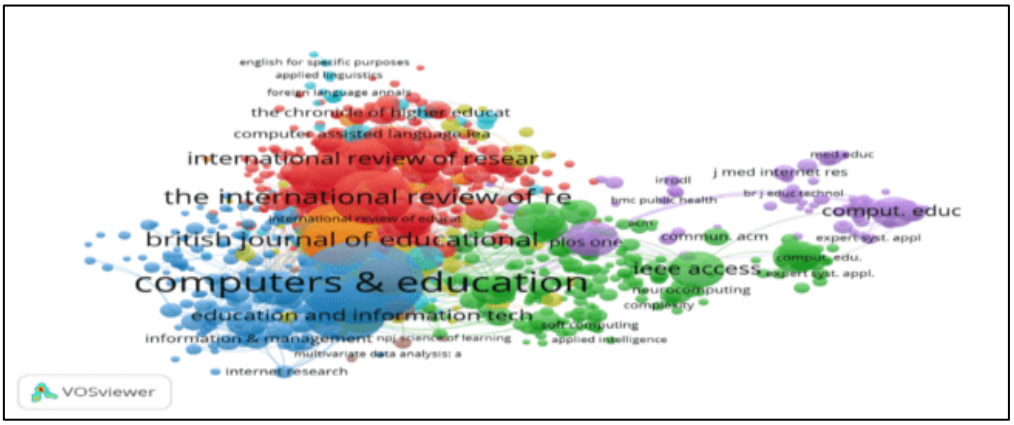


*Figure 7. Co-citation analysis by cited sources*

Co-citation of sources provides a glimpse into journals and publication venues that comprise the academic landscape of MOOC studies. The most central and influential node in the network is occupied by the journal "Computers & Education," which indicates that it is the premier platform for publishing studies on the interface between technology and pedagogy. Some other core journals include the "British Journal of Educational Technology," "The

International Review of Research in Open and Distributed Learning (IRRODL)," and "IEEE Access." These journals are frequently co-cited and hence the centers of debate in the field. The network also shows the interdisciplinarity of MOOC research in clusters of journals from disciplines such as computer science and public health, covered by titles like "J Med Internet Res" and "BMC Public Health." This suggests that the field is not strictly in the purview of education research but somehow overflows into broader arguments regarding accessibility, health education, and digital inclusion.

**Bibliographic Coupling by Countries**

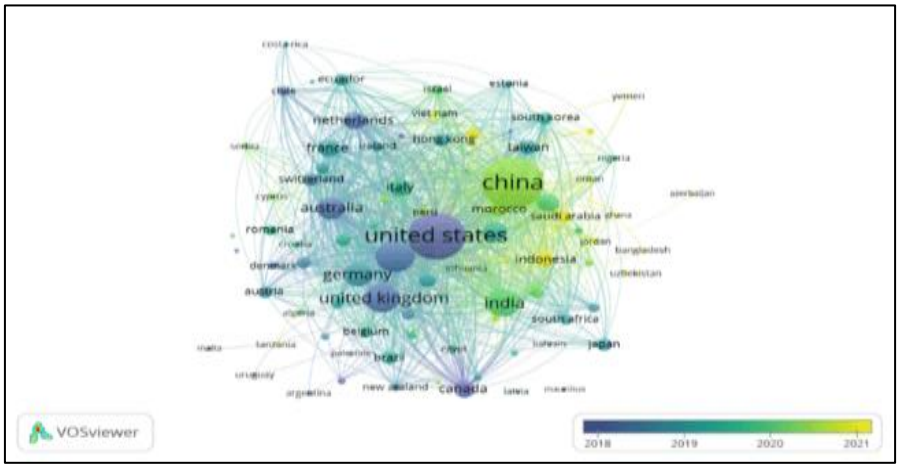


*Figure 8. Bibliographic Coupling by Countries*

The country map of bibliographic coupling also shows the interconnection and intensity of research collaboration among countries in the discipline. The United States, China, and the United Kingdom are the major contributors, as indicated by their large node sizes and dense linkages. These countries have, in turn, referred to similar collections of publications, showing a shared foundation of knowledge and overlapping research agendas. Particularly salient is that India, Germany, Australia, and Canada also exhibit high levels of bibliographic coupling, reflecting active engagement in international scholar debate. Blue to yellow color trend defines the average year of publication, highlighting that countries like India and China have been increasingly making contributions to modern literature.

**Bibliographic Coupling by Organizations**

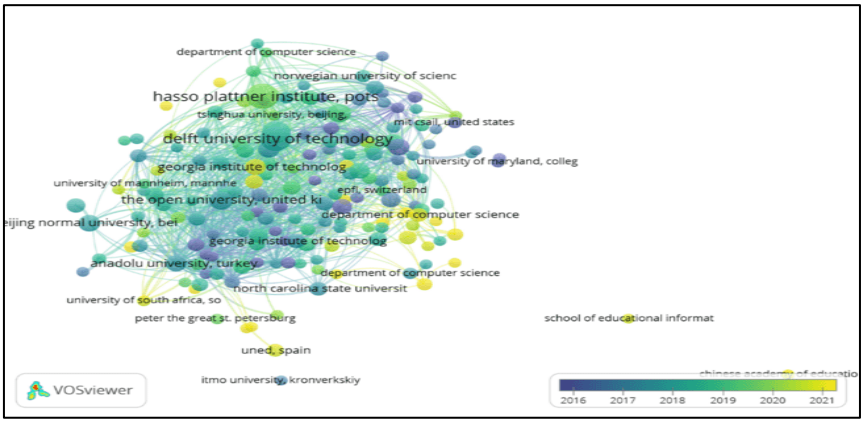


*Figure 9. Bibliographic Coupling by Organizations*

Network structure visualization for bibliographic coupling between institutions suggests clusters of institutions most linked with each other based on shared references within publications. The Hasso Plattner Institute, Delft University of Technology, and The Open University (UK) appear as central hubs, indicating high citation intersection with a range of other institutions. This indicates their influence and centrality in the research network of the field. Several other universities, such as the Georgia Institute of Technology, Tsinghua University, and Anadolu University, are part of dense clusters, indicating thematic and methodological homogeneity.

**Bibliographic Coupling by Authors**

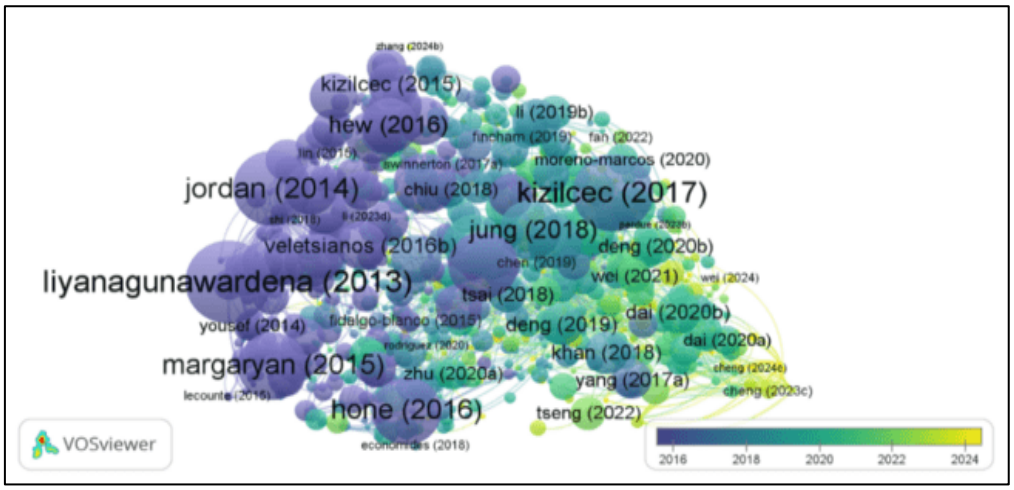


*Figure 10. Bibliographic Coupling by Authors*

Author's bibliographic coupling visualization indicates individual scholars who commonly cite the same papers, suggesting the same lines of research or theoretical lines. The sizes of the nodes and connections of authors like Liyanagunawardena (2013), Jordan (2014), and Kizilcec (2015, 2017) indicate that they are located at the center of this network. These authors have laid groundwork research that continues to inform existing research. The high interconnection density shows a mature and very active research base.

## 5. DISCUSSION

The bibliometric analysis of 10,059 publications that refer to MOOC for the period between 2009 and 2024 confirms a highly dynamic, inter- and transdisciplinary, and globally expanding domain. The fast growth rate of over 56% yearly and steady increase in scholarly output confirm that MOOCs evolved from a disruptive innovation to become a full-fledged area of research into education and technology studies.

International spread at the country level indicates the extent and intensity of international scholarship. China, the United States, Spain, India, and the United Kingdom were the most contributing countries. But citation analysis reveals the more nuanced truth—while China leads all others in sheer quantity, countries like Australia, Hong Kong, and the Netherlands offer significantly higher citation averages, indicative of a more globally influential and effective body of work. This reflects a disparity between the amount of research and its quality that may be indicative of varying levels of institutional support, research grants, and availability of English-language publication.

Keyword co-occurrence analysis and thematic mapping confirm the dominance of pedagogical, technology, and learner-centered themes. Cornerstone concepts such as curriculum, instruction, students, and e-learning continue to shape the conversation, and they reside primarily in the lower-right quadrant of the thematic map – i.e., important but theoretically underdeveloped. These emerging subfields are making their own intellectual niches, which would suggest a trend toward increased specialization and even the formation of subdisciplines within the overall MOOC research universe. Co-citation and bibliographic coupling analysis further solidify the intellectual center of the discipline. Works by authors such as Guo, Kizilcec, Jordan, Liyanagunawardena, and Margaryan are frequently co-cited and are therefore implied to be seminal in character. However, weak integration between pedagogically focused and technologically focused central thematic clusters is an indication where technical and pedagogical research streams are developing in parallel but not necessarily in dialogue with one another.

The second significant point is the under-representation of policy-focused, equity-oriented studies within the central thematic clusters. Despite early discussions of MOOCs as democratizing forces, inclusion, diversity, and issues of access – specifically for students in non-English-speaking and less resourced environments – are sideline concerns, based on the thematic analysis. This indicates a possible split between MOOC discourse and scholarly research agendas.

It validates that MOOC research is evolving into a mature, interdisciplinary research field with a high productivity, thematic variety, and methodological complexity. However, gaps remain – particularly in global cooperation, policy relevance, and interdisciplinary integration. These gaps might maybe be closed not only with the aim of further strengthening the scientific foundation of MOOC research, but also to enhance its practical relevance in reshaping the education of the future.

## 6. CONCLUSION

This study presents a comprehensive bibliometric overview of global MOOC-related research, analyzing 10,059 documents indexed in Scopus. The quantitative results show a consistent upward trend in publication output, with leading contributions from countries such as the United States, China, and the United Kingdom. Institutions like MIT and Stanford emerged as high-impact research hubs, while authors including Yuan L. and Alario-Hoyos S. played key roles in shaping the field. Thematic analysis identified established areas of focus – such as learner engagement, instructional design, and assessment practices – alongside newer themes gaining momentum, including inclusivity, multilingual delivery, and adaptive learning.

In terms of analytical insights, the study mapped how authors, keywords, and citation patterns are interrelated, revealing clusters of research that have grown independently as well as those that are beginning to converge. The co-citation analysis helped identify prominent bodies of work, while thematic mapping illuminated which research areas are central and well-developed versus those that are still emerging and underexplored. One important observation was the mismatch between the volume of research output and its academic impact in certain regions, emphasizing the need for not just producing more studies but ensuring relevance and rigor.

While the study offers a broad and data-rich view of the MOOC research landscape, it is constrained by its reliance on the Scopus database and English-language literature. Future investigations could benefit from integrating data from other sources such as Web of Science, or from focusing on region-specific developments, particularly in low- and middle-income countries where MOOC adoption may follow unique trajectories.

The findings have actionable implications for stakeholders. Platform designers can draw on the most cited themes to improve learner engagement; policymakers can use the insights to address access and equity gaps in digital education; and researchers can focus efforts on underrepresented topics that show potential for impact. The observed imbalance between technological and pedagogical research also suggests the need for more integrated, interdisciplinary approaches going forward.

## 7. Future Directions of Research

While this study has mapped the present landscape of research on MOOCs, there are several areas that need further research. To begin with, more research must probe intensely on the socio-political and policy dimensions of MOOCs, particularly in low- and middle-income countries where access and regulation vary significantly. The role of inclusion, multilingual content, and culturally adaptive learning design remains underrepresented in current literature and deserves targeted investigation.

Also, comparative and longitudinal bibliometric studies exploring shifts in research agendas across regions, languages, and policy settings are required. These themes can potentially enhance the academic rigor and applied relevance of research in context of MOOCs.

## REFERENCES (TRANSLATED AND TRANSLITERATED)


[1] L. Pappano, “The year of the MOOC,” *The New York Times*, Nov. 2, 2012. [Online]. Available: https://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html (in English)

[2] T. R. Liyanagunawardena, A. A. Adams, and S. A. Williams, “MOOCs: A systematic study of the published literature 2008–2012,” *Int. Rev. Res. Open Distrib. Learn.*, vol. 14, no. 3, pp. 202–227, 2013. (in English)

[3] K. Jordan, “Initial trends in enrolment and completion of massive open online courses,” *Int. Rev. Res. Open Distrib. Learn.*, vol. 15, no. 1, pp. 133–160, 2014. (in English)

[4] J. Reich, *Failure to Disrupt: Why Technology Alone Can't Transform Education*. Cambridge, MA, USA: Harvard Univ. Press, 2020. (in English)

[5] K. F. Hew and W. S. Cheung, “Students’ and instructors’ use of massive open online courses (MOOCs): Motivations and challenges,” *Educ. Res. Rev.*, vol. 12, pp. 45–58, 2014. (in English)

[6] P. J. Guo, J. Kim, and R. Rubin, “How video production affects student engagement: An empirical study of MOOC videos,” in *Proc. 1st ACM Conf. Learn. @ Scale*, 2014, pp. 41–50. (in English)

[7] S. Zheng, M. B. Rosson, P. C. Shih, and J. M. Carroll, “Understanding student motivation, behaviors and perceptions in MOOCs,” *Comput. Educ.*, vol. 91, pp. 83–93, 2015. (in English)

[8] M. Aria and C. Cuccurullo, “Bibliometrix: An R-tool for comprehensive science mapping analysis,” *J. Informetrics*, vol. 11, no. 4, pp. 959–975, 2017. (in English)

[9] K. M. Alraimi, H. Zo, and A. P. Ciganek, “Understanding the MOOCs continuance: The role of openness and reputation,” *Comput. Educ.*, vol. 80, pp. 28–38, 2015. (in English)

[10] C. Chen and Y. Xiao, “Selecting publication keywords for domain analysis in bibliometrics: A comparison between natural language processing and bibliographic coupling,” *J. Informetrics*, vol. 10, no. 3, pp. 927–940, 2016. (in English)

[11] S. Downes, *Connectivism and Connective Knowledge: Essays on Meaning and Learning Networks*. Ottawa, Canada: Nat. Res. Council Canada, 2012. (in English)

[12] F. M. Hollands and D. Tirthali, *MOOCs: Expectations and Reality*. New York, NY, USA: Center for Benefit-Cost Stud. Educ., Teachers College, Columbia Univ., 2014. (in English)

[13] A. M. Kaplan and M. Haenlein, “Higher education and the digital revolution: About MOOCs, SPOCs, social media, and the Cookie Monster,” *Bus. Horizons*, vol. 59, no. 4, pp. 441–450, 2016. (in English)

[14] R. F. Kizilcec, C. Piech, and E. Schneider, “Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses,” in *Proc. 3rd Int. Conf. Learn. Analytics Knowl.*, 2013, pp. 170–179. (in English)

[15] R. F. Kizilcec, M. Pérez-Sanagustín, and J. J. Maldonado, "Self-regulated learning strategies predict learner behavior and goal attainment in massive open online courses," *Comput. Educ.*, vol. 104, pp. 18–33, 2017. (in English)
[16] A. Margaryan, M. Bianco, and A. Littlejohn, "Instructional quality of massive open online courses (MOOCs)," *Comput. Educ.*, vol. 80, pp. 77–83, 2015. (in English)
[17] R. Pranckutė, "Web of Science (WoS) and Scopus: The titans of bibliographic information in today's academic world," *Publications*, vol. 9, no. 1, p. 12, 2021. (in English)
[18] G. Siemens, "Massive open online courses: Innovation in education?," in *Open Educational Resources: Innovation, Research and Practice*, R. McGreal, W. Kinuthia, and S. Marshall, Eds. Vancouver, Canada: Commonwealth of Learning and Athabasca Univ., 2013, pp. 5–15. (in English)
[19] A. F. J. Van Raan, "Advances in bibliometric analysis: Research performance assessment and science mapping," in *Measuring Scholarly Impact*, Y. Ding, R. Rousseau, and D. Wolfram, Eds. Cham, Switzerland: Springer, 2014, pp. 17–33. (in English)
[20] G. Veletsianos and P. Shepherdson, "A systematic analysis and synthesis of the empirical MOOC literature published in 2013–2015," *Int. Rev. Res. Open Distrib. Learn.*, vol. 17, no. 2, pp. 198–221, 2016. (in English)
[21] O. Zawacki-Richter, A. Bozkurt, U. Alturki, and A. Aldraiweesh, "What research says about MOOCs – An explorative content analysis," *Int. Rev. Res. Open Distrib. Learn.*, vol. 19, no. 1, pp. 242–259, 2018. (in English)
[22] J. Zhang, K. C. Almeroth, S. Knight, and S. Niemann, "Moodog: Tracking student behaviors in MOOCs," in *Proc. 26th Int. Conf. World Wide Web*, 2017, pp. 353–362. (in English)
[23] P. J. Guo, J. Kim, and R. Rubin, "How video production affects student engagement: An empirical study of MOOC videos," in *Proc. 1st ACM Conf. Learning @ Scale*, New York, NY, USA: ACM, 2014, pp. 41–50. (in English)
[24] K. Jordan, "Initial trends in enrolment and completion of massive open online courses," *Int. Rev. Res. Open Distrib. Learn.*, vol. 15, no. 1, pp. 133–160, 2014. (in English)
[25] A. Margaryan, M. Bianco, and A. Littlejohn, "Instructional quality of massive open online courses (MOOCs)," *Comput. Educ.*, vol. 80, pp. 77–83, 2015. (in English)

*Text of the article was accepted by Editorial Team 02.06.2025*

## МАСОВІ ВІДКРИТІ ОНЛАЙН-КУРСИ В ОСВІТІ: БІБЛІОМЕТРИЧНЕ ДОСЛІДЖЕННЯ ГЛОБАЛЬНИХ ТЕНДЕНЦІЙ

**Д-р Мохіт Кумар,**
кандидат наук з менеджменту, доцент, кафедра бізнес-менеджменту,
Університет Чхатрапаті Шаху Джі Махарадж, Канпур, Уттар-Прадеш, Індія.
ORCID ID 0000-0003-2532-4012
*drmohitkumar@csjmu.ac.in*

**Д-р Правін Кумар Агравал,**
кандидат наук з менеджменту, доцент, кафедра бізнес-менеджменту,
Університет Чхатрапаті Шаху Джі Махарадж, Канпур,
Уттар-Прадеш, Індія.
ORCID ID 0000-0003-4331-8767
*drpravinagrawal@csjmu.ac.in*

**Д-р Супрія Агравал,**
кандидат наук з менеджменту, доцент, Бізнес-школа Аміті.
Університет Аміті Лакхнау, Лакхнау, Індія.
ORCID ID 0000-0001-7187-1354
*sagrawal@lko.amity.edu*

**Вібха Кушваха**
науковий співробітник, кафедра бізнес-менеджменту,
Університет Чхатрапаті Шаху Джі Махарадж, Канпур, Уттар-Прадеш, Індія.
ORCID ID 0000-0003-2313-2391
*vibhak15@gmail.com*

**Анотація.** Останнє десятиліття в усьому світі засвідчило революцію в освітньому ландшафті, спричинену швидким поширенням цифрових технологій і навчальних платформ з відкритим доступом. Серед них Масові відкриті онлайн-курси (МВОК) опинилися в центрі інновацій у навчанні, обіцяючи гнучке, недороге навчання без кордонів. Представлене дослідження має на меті скласти карту глобального дослідницького ландшафту Масових відкритих онлайн-курсів (МВОК) за допомогою бібліометричних методів, щоб зрозуміти еволюцію, інтелектуальну структуру та тематичні тенденції в цій галузі. На основі набору даних з 10 059 публікацій, індексованих в Scopus, аналіз проводився за допомогою Bibliometrix (пакет R) і VOSviewer для аналізу цитування, бібліографічних зв'язків і тематичного картографування. Результати показують, що Сполучені Штати, Китай та Велика Британія є провідними країнами з точки зору результатів досліджень. Ключові інституції, такі як Массачусетський технологічний інститут та Стенфорд, домінують за впливом на цитування, а такі автори, як Юань Л. та Аларіо-Хойос С., демонструють високий науковий вплив. Аналіз виявляє невідповідність між обсягом наукової продукції та науковим впливом, особливо в певних регіонах. Тематичне картування висвітлює основні кластери в педагогіці, залучені учнів та інструментах оцінювання, тоді як недостатньо дослідженими сферами є інклюзивність, багатомовний доступ та міжнародна координація політики. Ключовий внесок цього дослідження полягає у виявленні структур у літературі про МВОК, які розкривають приховані перетини між технічними, педагогічними та політичними сферами. Результати дослідження показують зростаючу зрілість МВОК як міждисциплінарної галузі та фрагментарність дослідницьких зусиль. Це дослідження не лише демонструє зростання та диверсифікацію досліджень МВОК, але й висвітлює ігноровані сфери, що мають вирішальне значення для освітньої рівності та глобальної співпраці. Результати дослідження мають прикладне значення для освітян, розробників платформ і політиків, які прагнуть розширити охоплення, якість і вплив МВОК у цифровому навчальному середовищі, що розвивається.

**Ключові слова:** Масові відкриті онлайн-курси (МВОК); бібліометричний аналіз; онлайн-освіта; тематичне мапування; аналіз цитування.

This work is licensed under Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.