Discover Education

RESEARCH Open Access

A comprehensive investigation utilizing topic modeling to analyze research interests and trends in digital literacy

Erdem Çekmez[1*] and Özcan Özyurt[2]

*Correspondence:
Erdem Çekmez
erdemcekmez@trabzon.edu.tr
[1]Trabzon University, Trabzon, Turkey
[2]Karadeniz Technical University, Trabzon, Turkey

Abstract

The objective of this study is to delineate the existing research landscape and prevalent trends in digital literacy by applying topic modeling methodologies. For this purpose, 2,223 articles published between 2007 and mid-2024 retrieved from the Scopus database were analyzed using the Latent Dirichlet Allocation model. The findings revealed a consistent increase in publications over time, with a significant acceleration from 2019 onwards. Ten distinct topics were identified, with "Digital Literacy & Online Education," "Digital Literacy & Citizenship," and "Digital Literacy & Teaching Profession" being the most prominent themes. Trend analysis showed that "Digital Literacy & Online Education," "Digital Literacy & Health," and "Economy & Older Adults" demonstrated the most significant recent growth. The implications of these findings indicate a dynamic, pandemic-influenced shift in digital literacy research towards practical applications in health and the economy. This underscores the urgent need for policies and educational practices that address emerging digital competencies and divides, particularly concerning older adults and marginalized communities.

Keywords Digital literacy, Research trends, Latent Dirichlet Allocation, Topic modeling

1 Introduction

Digital literacy (DL) in education is a complex concept encompassing various skills crucial for active engagement in the digital world, emphasizing not just tool usage but also understanding social norms. UNESCO and other institutions view DL technically, focusing on skills like data retrieval and comprehension [65], but it goes beyond to include interpretation, communication, and significance attribution in digital settings [3]. Loosely put, DL is the ability to navigate, evaluate, and create digital content, blending technical skills with critical thinking [37, 61], evolving from computer, information, and media literacy with diverse perspectives [37].

The role of DL in education is multifaceted. Educational approaches vary, with some focusing on basic technical skills and others on the socio-cultural aspects of digital participation [50], including challenges like surveillance and the profit motives in digital

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

spaces [43]. DL is now widely seen as a core competence for citizenship in a knowledge society, involving the meaningful use of digital technologies and critical evaluation [25, 30]. Its empirical structure also includes performance and tool-dependent dimensions influenced by digital environments [52]. Proficiency in DL is crucial for students to adapt to technological advancements, utilize them effectively for goals such as enhancing data security and efficiency [53], and engage in safe internet practices to prevent cybercrime [35]. These advancements also include immersive technologies like virtual reality, whose growing importance for providing interactive, hands-on training in vocational education is highlighted by recent bibliometric analyses [42].

Given its importance, DL has become a critical focus for educational policies and curricula designed to enhance student skills [63]. The need for robust pedagogic frameworks that support DL education is evident, with a focus on meaning-making and creation rather than merely operationalizing technology [8]. This is underscored by research on digital transformation in higher education, which highlights the importance of digital pedagogies and continuous adaptation [18]. Ultimately, DL is closely linked with the digital competencies necessary for problem-solving, communication, and collaboration, which are crucial for empowering marginalized communities, fostering socio-economic development [44], and bridging the wider digital divide [57].

As the body of literature on DL expands, synthesizing this vast information presents a significant challenge. Traditional review methods, such as systematic reviews, often have a narrow scope, while bibliometric analyses tend to focus on quantitative metrics like citation counts and may lack in uncovering latent patterns [6]. To overcome these limitations, this study employs topic modeling, a data mining technique that enables a thorough semantic analysis of large textual datasets to uncover implicit details [31]. This approach allows for the identification of underlying research themes and trends with greater depth, thereby clarifying the unique contribution and urgency of the present research.

In summary, DL in education transcends technical skills to include the critical thinking, communication, and ethical considerations essential for navigating the digital age. By integrating DL into curricula and fostering digital citizenship, individuals can be empowered to participate effectively and responsibly in a digital society.

1.1 Literature review

In the previous section, it can be understood from the reported research that DL plays a crucial role in the teaching and learning process. As a result, the number of studies in the field of education that address DL from different perspectives is increasing day by day. Due to this increase, various types of review studies have been conducted with the aim of examining the field with a holistic perspective and finding answers to various questions through the analysis or synthesis of research in the literature. In this section, research reports conducted for this purpose have been summarized.

One line of research is composed of bibliometric reviews aimed at portraying the research landscape. The foci of these studies include: (i) the practical implications and tangible effects of DL on learning outcomes, pedagogical strategies, and student engagement [7]; (ii) scientific studies conducted at the tertiary level with a specific focus on the prominence of studies on tourism [10]; (iii) the conceptual and intellectual structure of the DL research field [15]; (iv) the emergence of new publishing trends and the rise

of burgeoning research domains following the pandemic period [58]; and (v) the current state of research on DL studies and the influence of the pandemic on this area of research [62]. The objectives and main conclusions of these studies are outlined in Table 1.

In sum, these studies highlight the increasing importance of DL across various fields. Integrating DL into education now emphasizes a holistic approach, including responsible technology use and digital citizenship. Despite global interest, tourism studies lag in DL tools, posing challenges for future professionals. Key research themes identified include conceptualization, measurement, and impact of DL. During the pandemic, research focused on application and challenges in higher education, with future trends leaning towards online education and digital equity. Emerging themes like 'ehealth' and 'digital health literacy' are prominent, with major contributions from the US, Spain, and the UK.

Additionally, systematic reviews form another significant area of research, which are conducted to answer specific research questions or to examine a topic in depth by analyzing the literature comprehensively and impartially.

The central themes of these studies include: (i) teachers' professional role, satisfaction, higher educational practice, and sustainable education [38]; (ii) the use and appropriateness of instruments designed to measure DL [46]; (iii) the scope and characteristics of

Table 1 The objectives and main conclusions of bibliometric studies

The study	Objective (The study aims to:)	Main conclusions
Arissaputra et al. [7]	Analyze research trends in DL education, emphasizing practical applications and outcomes.	The study concludes that integrating DL into education now emphasizes a holistic approach, including responsible technology use, critical thinking, and digital citizenship. Educational institutions, policymakers, and educators are urged to prioritize DL initiatives to equip students with essential digital skills.
Caldevilla-Domínguez et al. [10]	Analyze studies on DL in higher education, particularly in tourism, and identify the prominence of tourism studies and challenges of integrating ICTs in the tourism sector.	While there is significant global interest in DL, especially among students, tourism studies receive limited attention, posing challenges for future professionals. Training institutions must better equip tourism students with DL tools to ensure competitiveness. Integrating ICT in higher education is crucial for professional development and workforce readiness.
Choudhary and Bansal [15]	Contribute to advancing DL knowledge by identifying research gaps and emphasizing the existing themes in DL literature.	Key authors, studies, and research themes were discerned via bibliometric and content analysis. The investigation delineated five notable research themes within the realm of DL: conceptualization, antecedents, measurement, training and interventions, and impact and issues.
Tan [58]	Depict the prevailing and subsequent research themes and keywords in literature from 2020 to 2023, focusing on higher education during the pandemic period.	Studies on DL in higher education during the pandemic mainly focused on application, evaluation, and challenges faced by students and teachers. Future research trends are expected to concentrate on online education, digital literacies, pedagogy, learning processes, and teaching strategies, with a growing emphasis on digital equity.
Wang and Si [62]	To perform a comprehensive examination of the scholarly output concerning DL investigation from 1990 to 2022 with a specific emphasis on the influence of the COVID-19 pandemic era.	The study identified 'ehealth', 'mhealth', 'digital health literacy', 'digital health', 'fake news', and 'online learning' as emerging themes in DL research. Additionally, most DL research has been conducted in the United States, Spain, and the United Kingdom. The Journal of Adolescent and Adult Literacy, the Journal of Medical Internet Research, and Computers & Education have been identified as the journals that have published the most research in this field.

DL initiatives carried out to enhance the capabilities of teenage girls [40]; (iv) the process of digital transformation and the acquisition of DL within the intricate framework of higher education establishments [19], (v) the definitions and usages of the terms digital competency and DL [55]; (vi) the empirical evidence presented by international research studies pertains to the DL levels among university students [22]; (vii) the elements of critical digital literacies pertaining to educational settings, as delineated in scholarly works and global policy papers [26]; and (viii) the factors affecting students' DL [49]. The objectives and main conclusions of these studies are outlined in Table 2.

Table 2 The objectives and main conclusions of systematic review studies

The study	Objective (The study aims to)	Main conclusions
Li and Yu [38]	Investigate the impact of the COVID-19 pandemic on teachers' professional roles, career satisfaction levels, and DL.	There is a significant correlation between teachers' DL level, career satisfaction, and professional role, affecting the online teaching and learning process. Teachers require assistance in improving their DL in order to adjust to the evolving complexities of their professional duties amidst the COVID-19 pandemic.
Oh et al. [46]	Evaluate studies on DL in older adults and examine the adequacy of DL assessment tools in addressing age-specific requirements.	The paper indicates that more studies are needed to enhance the measurement of DL among older adults for better resource allocation and support. DL significantly impacts the well-being of older adults through information and communication technologies.
Meherali et al. [40]	Assess the effectiveness of DL interventions in empowering adolescent girls in low- and middle-income countries.	DL interventions, including mobile phones and internet-based strategies, are effective in empowering adolescent girls in LMICs by enhancing access to health services, information, and educational resources. Nevertheless, there is insufficient evidence regarding the efficacy of DL in improving girls' opportunities to access financial services and achieve economic empowerment.
Farias-Gaytan et al. [19]	Ascertain research pertaining to digital transformation and DL in institutions of higher education through a systematic study of literature.	It is crucial for higher education institutions to prioritize the development of DL within the faculty and students in order to comply with the demands of the modern workforce and society. The study underscores the significance of assessing digital proficiencies in both educators and learners to pinpoint areas of excellence and areas needing improvement for tailored training initiatives.
Spante et al. [55]	Understand the referencing strategy to DL and digital competence over time, disciplines, countries, methods, and levels of analysis.	The study delineated distinct referencing strategies for digital competence and DL. Digital competence is frequently delineated in policy documents, while DL is predominantly referenced or defined in scientific research. Consequently, future research on DL should concentrate on investigating the origins of definitions, examining how various definitions interact or conflict, and elucidating the rationale behind specific definitions to enhance comprehension of the research's objectives.
Gutiérrez-Ángel et al. [22]	Analyze the empirical evidence provided by international studies concerning the DL of university students pursuing degrees related to the field of education.	Pre-service teachers exhibit intermediate to higher levels of DL, particularly in communication and collaboration skills, but show an intermediate-low level in digital content creation. There is a need to formulate effective intervention designs and strategies to capacitate pre-service teachers in DL.
Ilomäki et al. [26]	Identify the dimensions of critical digital literacies and their relevance in educational practices and to create a research-based framework for school education.	Critical digital literacies integrate various subject domains and are linked to societal phenomena and rapid technological changes, emphasizing the need for regular revisions in definition and content. Information literacy, digital media literacy practices, technological literacy, participation in the digital society, and digital identity constitute the fundamental elements of critical digital literacies.
Peng and Yu [49]	Define DL, and examine the relationship between students' DL and self-control, technostress, and engagement.	Family socioeconomic status and parents' education level positively correlate with students' DL. DL significantly impacts students' self-control, technostress, and engagement, enhancing learning outcomes and reducing online crime.

To sum up, these studies underscore the crucial role of DL in various contexts. DL is positively influenced by family socioeconomic status and parents' education, enhancing students' self-control, reducing technostress, and improving engagement. Critical digital literacies require regular updates to keep pace with societal changes and technological advancements. Pre-service teachers show strong DL in communication but need improvement in digital content creation. Higher education must prioritize DL development to meet modern workforce demands. DL interventions effectively empower adolescent girls in low- and middle-income countries, though more evidence is needed on financial empowerment. Enhancing DL among older adults is vital for their well-being. There is a significant link between teachers' DL levels and their career satisfaction, necessitating support for DL improvement, especially during the COVID-19 pandemic.

In the conducted literature review, in addition to the aforementioned review studies, two scoping reviews and one integrative literature review were identified. In their integrative literature review study, Reddy et al. [51] aimed to provide a comprehensive review of the field of DL, particularly through the lens of information and communication technology. Their study redefined and identified six components of DL. The authors recommend that these components can be used to develop a DL scale to measure the digital competencies of individuals. In the first study classified as a scoping review, Khanlou et al. [33] aimed to explore the impact of digital technology on education, daily living, community integration, and employment for young adults with developmental disabilities. The study emphasizes the importance of facilitators for promoting digital inclusion of young adults with developmental disabilities, highlighting barriers such as training, affordability, and lack of fit with needs. Moreover, the study identifies the significance of organizational infrastructure, training, and software design features in facilitating technology utilization. Of the other recently published scoping review study, Alowais et al. [4] aimed to explore educational and training approaches for equipping pharmacy students and trainees with DL skills. The authors highlight the importance of theoretical knowledge in health informatics, proficiency in various digital technologies, and practical experience in informatics centered on patient care and interprofessional collaboration as essential skills for DL in pharmacy education. It underscores the necessity of early integration of DL training in undergraduate curricula to establish a robust skill set in these areas.

1.2 Justification for the research and statement of the problem

As outlined in the previous section, the examination of DL has been a central focal point of inquiry across various perspectives within scholarly literature. Several scholars have investigated this particular subject, scrutinizing it from a range of viewpoints and theoretical frameworks. The body of literature regarding DL is consistently growing, resulting in the appearance of novel research inquiries that necessitate further exploration and clarification. An inquiry aimed at outlining contemporary research interests, emerging patterns, and advancements within DL literature would be beneficial for comprehending the present landscape and offering insights for prospective researchers.

As explicated in the literature review section, numerous studies endeavor to articulate the understanding within the field across various thematic frameworks. Each of these studies undeniably brings distinctive contributions to their corresponding areas of investigation. Nevertheless, it is pivotal to acknowledge the constraints intrinsic to these

inquiries. For instance, systematic review studies frequently exhibit a restricted scope concerning particular research inquiries, potentially constraining the incorporation of a wide range of research in the field. As a result, the outcomes of such reviews could fail to capture intricate nuances and details present in the body of research. Regarding bibliometric research methodology, it is employed to quantitatively evaluate bibliographic indicators, which aid in classifying information based on various variables. However, this approach places emphasis on specific factors including the volume of publications, the citation status of authors, the network of co-publications, and popular topics [6]. Consequently, it may lack in uncovering latent patterns within the literature and executing semantic analyses.

In current discussions, there is an increasing inclination towards employing topic modeling methods based on data/text mining to reduce risks and facilitate thorough semantic analysis [23, 31]. The utilization of topic modeling enhances the breadth of bibliometric analysis and systematic review investigations, enabling semantic evaluations to uncover implicit details from extensive textual datasets [31]. Lately, the technique of topic modeling has garnered significant attention in scholarly investigations aimed at achieving a thorough comprehension of diverse domains. Undoubtedly, a plethora of studies of such kind are distinguishable within the current corpus of academic literature, to mention some, investigations on intelligent cities [48], the integration of AI in educational settings [14], the implementation of blended learning approaches [66], and the utilization of technology in math education [12, 24].

By employing semantic analysis, this study seeks to reveal the predominant research focuses and patterns within the field of DL in education. Alongside delineating the characteristics of the discipline, another objective of this study is to illuminate the current landscape of scholarly articles on DL in education, tracing its evolution from its historical beginnings to present advancements. This study, which encompasses a collection of 2223 articles, constitutes one of the most comprehensive explorations of the discourse on DL to date.

The present study delves into the analysis of literature on DL by scrutinizing research papers listed in the Scopus database spanning from the year 2007 to mid-2024. In accordance with this objective, the study aims to address the following research questions:

RQ1. What is the temporal distribution of DL articles?

RQ2. What are the prevailing topics addressed in DL articles?

RQ3. What emerging trends are observed in the topics covered within DL articles?

2 Methodology

2.1 Methodological approach

Topic Modeling Analysis (TMA) constitutes a prevalent methodology utilized across a multitude of fields including data science, and natural language processing. TMA is frequently classified as an "unsupervised" technique due to its ability to autonomously derive latent themes or topics from unstructured textual data without reliance on predetermined classifications [21]. The utilization of computational methodologies such as Latent Dirichlet Allocation (LDA) enables TMA to reveal latent themes and organizational frameworks within extensive compilations of textual documents [16]. In the current study, LDA algorithm [9], which is acknowledged as a generative approach for the development of probabilistic topic models, has been utilized.

2.2 The search process and exclusion-inclusion criteria

The Scopus database is utilized for its broad interdisciplinary coverage in this research. It catalogs titles from more than 7000 global publishers, serving as an extensive index of international research outputs [17]. In alignment with the academic corpus and the viewpoints of three subject matter experts, the criteria for the investigation were determined. Upon establishing the criteria, the following query was executed via Scopus search engine on June 16, 2024, and 2,245 articles were found.

TITLE-ABS ("digital literac*" AND ("education" OR "teaching" OR "learning" OR "instruction" OR "training")) AND (LIMIT-TO (DOCTYPE, "ar")) AND (LIMIT-TO (PUBSTAGE, "final")) AND (LIMIT-TO (LANGUAGE, "English")) AND (LIMIT-TO (SRCTYPE, "j")).

The examination of the articles revealed that only 13 studies, representing less than 1% of the total, were published from 1997 to 2006. As a result, it was decided to integrate the research conducted from the year 2007 onwards. After removing duplicates and those that do not contain abstract, 2,223 articles remained out of the initial 2,245. The criteria for inclusion and exclusion that were established to discern relevant studies upon the completion of this process are shown in Table 3.

2.3 Corpus preparation, the preprocessing of data, and the implementation of the analysis

The textual corpus for the analysis was created by merging the title, abstract, and author keywords for each of the 2,223 articles into a single textual unit. This combined text for each article served as the document input for the subsequent topic modeling analysis.

The efficacy of LDA analysis results is contingent upon corpus quality. Consequently, preparatory actions were undertaken on the corpus prior to topic extraction. The initial preparatory measure involved eliminating duplicate studies and those lacking abstracts. Following the retrieval of articles from the Scopus database, their components were compiled into an Excel file. It was noted that nearly all abstracts concluded with a "©" symbol succeeded by an irrelevant phrase (e.g., © The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023). A macro was developed to expunge this section from each abstract in the Excel file.

In the context of LDA analysis, n-grams denote sequences of 'n' words that collectively convey a coherent meaning. The incorporation of n-grams within LDA enhances topic modeling by facilitating the identification of co-occurring multi-word expressions, which in turn improves the coherence and interpretability of the identified topics [34, 45]. Research conducted by Lau et al., [36] demonstrated that the integration of bigram collocations into document representations yields more parsimonious models and enhances topic coherence within topic modeling frameworks. Consequently, the authors advocate for the treatment of multiword named entities as single tokens to ensure consistent analytical outcomes. In alignment with this guidance, our study undertook a large

Table 3 The criteria for literature selection

Inclusion	-Including a combination of the terms specified in the search string in the title or summary; the document type being an article and the publication stage variable being final; to be written in English; the source type being a journal. -Publications spanning the period from 2007 to mid-2024.
Exclusion	- Publications in a language other than English -The publications categorized as reports, data papers, retractions, letters, short surveys, errata, or notes. - Those that do not contain abstract

number of preliminary LDA analyses by assigning values of 5, 4, 3, and 2 to the n-gram variable, thereby systematically exploring bigrams, trigrams, quadgrams, and pentagrams within the corpus. For each specified n-gram value, we identified the resultant meaningful word clusters and subsequently modified the corpus by eliminating spaces between the constituent words of these clusters within the Excel file, effectively consolidating them into singular tokens. For illustrative purposes, the phrase "small and medium-sized enterprises" was transformed into "smallandmediumsizedenterprises", "technological pedagogical content knowledge" was altered to "technologicalpedagogicalcontentknowledge," "emergency remote teaching" became "emergencyremoteteaching," and "social capital" was converted to "socialcapital". After it was observed that there were no other meaningful groups of words in the corpus, LDA analyses were conducted with the n-gram variable set to 1 to obtain the topics. In the findings section, the word groups are presented as they were input into the analysis.

Before conducting an LDA analysis, data and text mining techniques, along with data pre-processing, were applied in a sequential manner [2]. The implementation of these methodologies facilitated the transformation of raw, unprocessed data into a cleaned and refined dataset. The first step involved executing a data cleaning process. This phase focused on removing any variables related to text language from the dataset to improve its integrity and quality. The main aim was to optimize the machine's ability to process natural language texts effectively. The sequence of procedures included several stages: initially, the text was tokenized to break it down into individual words. Next, all words were converted to lowercase. Following this, special characters and punctuation were eliminated from the words. Additionally, irrelevant and stop words—such as 'a', 'an', 'is', 'the', 'of', 'for', and others—were removed due to their lack of semantic significance within the text. Finally, lemmatization was performed to reduce words to their root forms. Upon completion of these preprocessing steps, the final dataset was assembled with the remaining words, which was then utilized for the LDA analysis.

TMA was carried out using the LDA-Multicore algorithm from the Gensim library [20] within the Python programming environment. The analysis of the corpus was conducted through an iterative process, consistent with the guidance provided in the literature [47]. To determine the optimal number of topics (referred to as K), various K values ranging from 8 to 20 were tested. Throughout the iterative analysis, a coherence score was calculated for each K value, indicating that the model with 10 topics achieved the best performance, yielding a coherence score of 0.43. The final LDA model was trained with the following hyperparameters: workers = 3, random_state = 42, minimum_probability = 0.25, passes = 40, iterations = 500, alpha = 'asymmetric', and eta = 'symmetric'.

Once the number of topics was established, the researchers employed the pyLDAvis tool for topic visualization. PyLDAvis, a web-based application, was utilized to enhance the visual understanding and labeling of the topics. During the visualization phase, a specific parameter known as the λ value, constrained between 0 and 1, was set to 0.6 based on recommendations from the existing literature [54]. This λ value represents the importance of individual words within the topics. Subsequently, the topics were labeled by identifying the most pertinent terms associated with each topic.

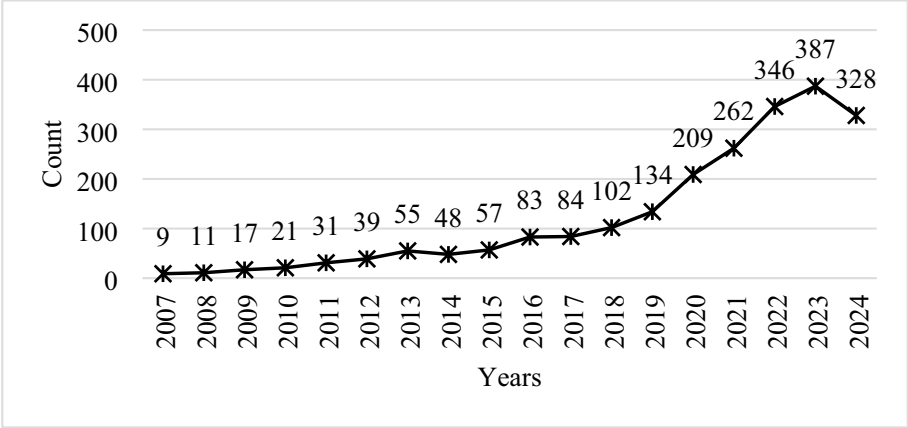


Fig. 1 The distribution of the counts of the studies over the years

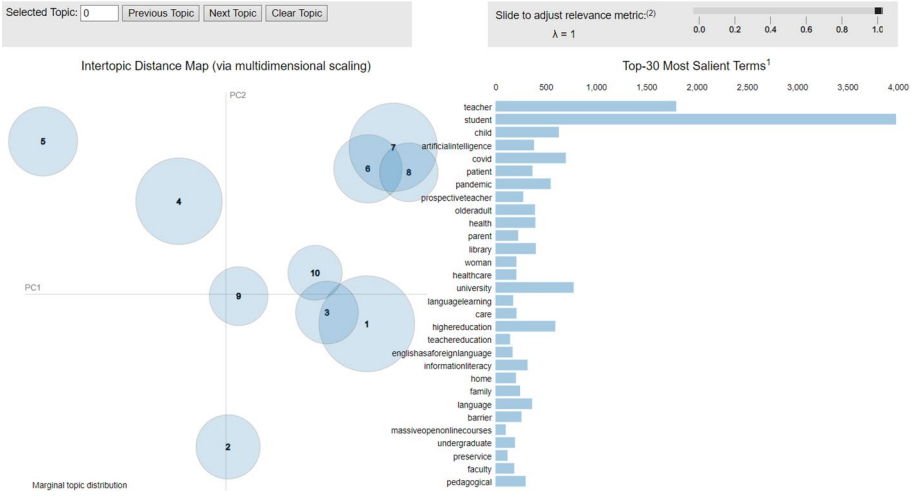


Fig. 2 The distribution of the topics and the most salient terms

3 Findings

3.1 The findings pertaining to the annual distribution of the published articles

The first problem addressed relates to the allocation of research endeavors within the domain of DL across different years. The graphical representation delineating the fluctuations in the number of published studies over time is illustrated in Fig. 1. As depicted in Fig. 1, there has been a persistent rise in the count of published articles, particularly accelerating from 2019 onwards. The insight that the tally of research papers in 2024 seems diminished relative to the preceding year does not intrinsically suggest a reduction in the number of studies carried out last year, as the search for articles was executed on June 16, 2024.

3.2 Findings related to the prevailing topics

In this segment, the results concerning significant research interests and prevailing trends within the domain of DL are presented. Utilizing LDA for topic modeling analysis, ten topics were discerned. These emerging topics are represented in the inter-topic distance map, alongside the top 30 most salient terms within the analyzed corpus, as illustrated in Fig. 2.

The findings illustrated in Fig. 2 indicate that topics 2, 4, 5, 9, and to a certain extent, topic 10 are characterized by their independence. Conversely, topics 6, 7, and 8 exhibit a high degree of interrelation. In addition to this, when considering the intersection areas represented by topics 1, and 3, it becomes evident that they are interrelated. The inferences drawn regarding these relationships have been addressed in the discussion and conclusion section.

To facilitate the categorization of the identified topics, the researchers first conducted a preliminary labeling by analyzing the most salient terms within each topic. Subsequently, these researcher-generated labels, along with their corresponding term clusters, were presented to three subject matter experts. The experts were asked to evaluate whether the proposed labels accurately and appropriately represented the corresponding cluster of terms. A full consensus was achieved in this initial step, as all three experts independently agreed on the suitability of the labels. The labels of the topics, their corresponding percentages, and the associated terms are presented in Table 4.

The figures presented in Table 4 reveals that within the domain of DL, the predominant area of focus among researchers is "DL & online education" (16.46%). This is succeeded by the themes of "DL & citizenship" (14.76%) and "DL & teaching profession" (13.73%). Conversely, the topic receiving the least attention is "Women & marginalized communities" (5.1%). In order to exemplify the issues addressed by the studies within the topics, five research foci for each topic are presented in Table 5.

3.3 Findings related to emerging trends within DL articles

To ascertain the variations in the volume of the topics over time, the period from 2007 to mid-2024 was divided into six distinct intervals, each lasting three years except for the last one. Subsequently, the frequencies of publications pertinent to each topic were ascertained for these intervals. These intervals were quantified using integers ranging from 1 to 6, symbolizing the continuum from historical to contemporary contexts. In addition to the frequencies, a best-fit line is added to the volume charts created according to the frequencies of the articles. The slope of the best-fit line is taken as an indicator of the acceleration (ACC) of the topics. Hence, a positive slope value indicates that the frequencies of articles within a topic are showing an increasing trend, whereas a negative value indicates the opposite scenario. The frequencies of the articles within each time period, the percentage of the number of articles in each period relative to all works related to that topic are presented in Table 6 along with ACC values. The ordering of the topics in Table 6 is based on the size of their volumes, from top to bottom.

Based on the data in Table 6, the charts depicting how the volumes of the topics change over time, along with best-fit lines, have been created and are presented in Fig. 3. Considering the data presented in Table 6; Fig. 3, it is observed that the most balanced distribution in terms of volume across time periods occurs in the topic of "1. DL & citizenship". Additionally, it is noted that, except for this topic, all other topics have more articles in the 6th time period than in the 5th time period. For these reasons, it has the lowest acceleration value among all topics (ACC = 6).

On the other hand, in terms of acceleration, the topics "4. Economy & older adults" (ACC = 11.1) and "5. DL & health" (ACC = 11) share the lead. In both topics, the number of articles published in the 6th time period constitutes approximately 64% of all articles published across all time periods. The topics "7. DL & online education", "10. Emerging

Table 4 The topics, related terms and percentages

No	Topic name	30 most salient terms	%
7	DL & online education	Student, covid, highereducation, digitalliteracyskill, access, challenge, effectiveness, distancelearning, resource, blendedlearning, transition, learningmanagementsystem, readiness, selfregulatedlearning, learningoutcome, academic, digitaldivide, emergencyremoteteaching, informationliteracy, onlineeducation, selfdirectedlearning, massiveopenonlinecourses, assessment, policy, openeducation, feedback, higherorderthinkingskills, digitalscholarship, selfefficacy, digitalresearch	16.46
1	DL & citizenship	Student, civic, fakenews, digitalcitizenship, identity, text, digitalliteracypractice, communication, medialiteracy, multimodel, socialpractice, reading, teacher, digitaltext, youtube, digitalstorytelling, screen, education, criticaldigitalliteracy, language, multiliteracies, ereader, writing, socialmedia, literacy, digitalmedia, socialnetwork, misinformation, informationliteracy, society	14.76
6	DL & teaching profession	Teacher, propectiveteacher, digitalcompetence, teacherducation, teaching, preservice, englishasforeignlanguageteacher, professionaldevelopment, teachertraining, pedagogical, technologicalpedagogicalcontentknowledge, secondaryschool, collaboration, belief, primaryschool, inservice, attitude, digitaltechnology, classroom, curriculum, highereducation, teachereducation, englishasforeignlanguage, teachingpractices, digitalliteracyskill, elementaryschool, technologyintegration, perception, educationaltechnology	13.73
2	Family & child DL practices	Child, parent, family, education, covid, earlychildhood, rural, digitalliterapractice, mother, disability, poverty, digitaldivide, socialmedia, challenge, medium, digitaltechnology, mobilephone, assessment, urban, risk, digitalinclusion, access, cyberbullying, socioeconomic, play, support, digitalmedia, digitalcompotence, onlinerisk, policy	9.21
4	Economy & older adults	Covid, pandemic, olderadult, socialsupport, barrier, wellbeing, ageing, service, society, access, digitalskill, digitaldivide, worker, crisis, labor, digitaleconomy, financial, economic, support, employee, smallandmediumsizedenterprises, digitaltransformation, industry, entrepreneurship, socialisolation, digitalinclusion, business, rural, sector, socioeconomic	8.59
5	DL & health	Telehealth, olderadult, mobileapp, ehealthliteracy, telemedicine, digitalhealth, healthtechnology, delivery, selfmanagement, attitude, healthservice, mobilehealth, hospital, ruralhealth, nursing, publichealth, digitalhealthliteracy, healthinequalities, healthinformation, healthcareproviders, therapy, healtheducation, digitaldivide	8.54
3	DL & libraries	Library, publiclibrary, student, informationliteracy, digitalinclusion, society, informationseeking, access, digitalliteracytraining, visualization, digitaldivide, service, cybersecurity, staff, libraryservice, academiclibraries, curriculum, user, resource, digitaltechnology, education, community, public, government, university, digitalskill, citizen, highereducation, organization	7.92
8	DL & language teaching	Student, languagelearning, perception, englishasforeignlanguage, machinetranslation, englishasforeignlanguagestudent, linguistic, mobilelearning, criticalthinking, englishlanguageteaching, writing, attitude, digitaltool, translation, digitalcompetence, secondlanguage, highereducation, englishasforeignlanguageteacher, foreignlanguage, digitaltechnology, curriculum, pedagogical, vocabulary, digitalstorytelling, englishforacademicpurposes, communication, englishforspecificpurposes, languageteaching, reading, socialmedia	7.87
10	Emerging technologies & DL.	Artificialintelligence, chatgpt, chatbot, academic, ethical, child, programming, language, writing, game, videogame, robotics, problemsolving, stem, digitalliteracyskill, conversationalagent, scratch, privacy, computationalthinking, pedagogical, student, curriculum, challenge, teacher, health, internetofthings, creativity, engineering, pandemic, business	7.83
9	Women & marginalized communities	Woman, empowerment, agricultural, informationliteracy, service, digitalcompetence, farmer, community, internetusage, inequality, refugee, rural, lowincome, digitaldivide, marginalized, household, migrant, digitalliteracytraining, resident, socialcapital, sustainabledevelopment, awareness, socioeconomic, health, family, wellbeing, socialmedia, access, challenge, policy	5.10

Table 5 Examples of research issues addressed within the topics

Topic	Examples of research issues (The foci of the studies include:)
7. DL & online education	(1) The challenges of emergency remote teaching during the COVID-19 pandemic; (2) the role of higher-order thinking skills in enhancing DL and academic success among higher education students, particularly in the context of online learning environments; (3) the implementation and effectiveness of learning management systems in supporting blended learning; (4) the impact of self-regulated learning strategies on students' DL and academic performance, emphasizing the importance of self-regulation in online learning contexts; (5) the development and accessibility of massive open online courses for inclusive education.
1. DL & citizenship	(1) The effectiveness of DL interventions in combating the spread of fake news on social media platforms; (2) the development of digital citizenship skills among university students through engagement with DL practices and social networks; (3) the role of digital storytelling in fostering critical DL; (4) The impact of media literacy education on teachers' abilities to address misinformation and fake news in English as a foreign language classrooms; (5) The influence of digital media use and social practices on the development of digital citizenship skills in young people, enhancing their engagement with social media.
6. DL & teaching profession	(1) Integrating digital competence into teacher preparation to enhance classroom technology use; (2) developing technological pedagogical content knowledge for effective DL in language teaching; (3) assessing DL competence of preservice teachers; (4) enhancing DL and teaching performance through professional development, addressing technostress; (5) exploring DL practices' impact on prospective teachers' professional identity.
2. Family & child DL practices	(1) The impact of DL programs on reducing the digital divide and promoting digital inclusion among rural youth; (2) the role of social media in shaping the socioeconomic interactions and digital practices of young people; (3) the effectiveness of DL workshops in enhancing parenting skills and supporting child development; (4) the challenges and strategies of mothers as they mediate their children's DL practices in English; (5) the influence of DL on cyberbullying experiences among primary school children.
4. Economy & older adults	(1) The impact of DL programs on older adults' digital social behavior and well-being, particularly during the COVID-19 pandemic; (2) the role of digital skills in facilitating financial inclusion for young people in complex scenarios; (3) the barriers and enablers of digital transformation in small and medium-sized enterprises; (4) the influence of social media on users' self-efficacy and loneliness; (5) the challenges of digital inclusion for older adults, emphasizing the importance of technological support services in reducing the digital divide.
5. DL & health	(1) The exploration of telehealth solutions in healthcare, focusing on their utilization, the barriers encountered, and strategies for promotion during the COVID-19 pandemic; (2) the investigation of challenges related to digital access and literacy among older adults; (3) the analysis of various factors that influence users' willingness to adopt mobile health applications; (4) the study of digital health literacy among university students to enhance understanding of their ability to navigate digital health information; (5) examining how digital health literacy tools highlight health inequalities among different population groups.
3. DL & libraries	(1) The exploration of DL skills among students preparing for competitive exams, particularly in the context of library resources and support; (2) the investigation of university students' digital competence and information-seeking behaviors; (3) the analysis of community-based DL training programs and their role in promoting digital inclusion and curriculum development; (4) the examination of the relationship between DL, online privacy concerns, and cybersecurity behaviors; (5) the assessment of public libraries' roles in advancing digital government initiatives and enhancing digital inclusion within communities.
8. DL & language teaching	(1) The impact of DL on technostress, online engagement, autonomy, and academic success among language learners; (2) the role of digital storytelling in enhancing critical thinking and creativity among EFL teacher candidates; (3) the development of machine translation literacy and autonomy through inclusive practices; (4) the influence of DL on the academic productivity and technostress of EFL students; (5) the investigation of the effectiveness of gamified learning environments, such as digital escape rooms, in teaching English for specific purposes.

Table 5 (continued)

Topic	Examples of research issues (The foci of the studies include:)
10. Emerging technologies & DL	(1) The adoption of artificial intelligence in agriculture, focusing on its role in transforming digitalization and productivity in developing countries; (2) the development of DL and computational thinking in young children through educational robotics; (3) the role of conversational agents in promoting physical activity among women; (4) the exploration of young children's perceptions of ubiquitous computing and the internet of things; (5) the impact of DL and artificial intelligence tools on university students' learning experiences, with a focus on AI-assisted learning.
9. Women & marginalized communities	(1) The role of DL in empowering women in rural areas; (2) the impact of DL on farmers' pro-environmental behavior; (3) the influence of DL on the social inclusion of refugees and asylum seekers; (4) the intersection of DL and social connectedness among immigrant seniors; (5) the role of DL training in reducing the digital divide for marginalized populations.

Table 6 The distribution of the frequencies in topics according to time intervals

Topic no	1 2007–2009	2 2010–2012	3 2013–2015	4 2016–2018	5 2019–2021	6 2022–2024	Total	ACC
7	2 (0.5%)	9 (2.5%)	25 (6.8%)	38 (10.3%)	106 (28.8%)	188 (51.1%)	368	9.6
1	8 (2.5%)	26 (8%)	33 (10.2%)	65 (20%)	97 (29.9%)	96 (29.5%)	325	6
6	4 (1.3%)	11 (3.6%)	26 (8.5%)	32 (10.5%)	89 (29.2%)	143 (46.9%)	305	8.8
2	9 (4.4%)	7 (3.4%)	18 (8.8%)	29 (14.2%)	51 (24.9%)	91(44.4%)	205	7.7
4	2 (1.1%)	3 (1.6%)	6 (3.2%)	10 (5.3%)	46 (24.2%)	123 (64.7%)	190	11.1
5	1 (0.5%)	3 (1.6%)	7 (3.7%)	13 (6.8%)	44 (23.2%)	122 (64.2%)	190	11
3	5 (2.8%)	12 (6.8%)	16 (9.1%)	27 (15.3%)	54 (30.7%)	62 (35.2%)	176	6.9
8	2 (1.1%)	10 (5.7%)	18 (10.2%)	25 (14.2%)	50 (28.4%)	71 (40.3%)	176	7.7
10	3 (1.7%)	6 (3.4%)	8 (4.6%)	17 (9.7%)	38 (21.7%)	103 (58.9%)	175	9.9
9	1 (0.9%)	4 (3.5%)	3 (2.7%)	13 (11.5)	30 (26.6%)	62 (54.9%)	113	9.9
Total	37	91	160	269	605	1061	2223	

technologies & DL", and "9. Women & marginalized communities" closely follow the two leading topics in terms of acceleration, exhibiting a similar increase to each other.

In addition to the volumes of the topics and their accelerations within themselves, the changes in the volumes of the topics relative to each other over specified time intervals were examined. For this, the ratio of each topic to all the studies conducted in each time period was calculated. The acceleration of the relevant topic compared to the other topics in the periods was also calculated. In this calculation, the percentages of the relevant topic in each period were taken into account. Graphs depicting how the volumes of topics change compared to all studies are presented in Fig. 4, along with the best-fit line. The acceleration values, which indicate whether the volumes of topics are trending upwards or downwards compared to all topics across different time periods, are given in Fig. 5.

As illustrated in Fig. 4, the ratios of the volumes of the topics exhibit varying changes over time intervals. The topic "DL & online education" demonstrates the most significant upward trend in terms of its ratio compared to others, possessing the highest acceleration value of 2.37 (see Fig. 5). Furthermore, it is evident that the topics "DL & health" and "Economy & older adults" have also captured a notable upward trend in their volume ratios relative to others, achieving acceleration values of 1.61 and 1.25, respectively. In addition, although the calculated acceleration value for the topic of 'Women & marginalized communities' is relatively small, the increase in its proportion over the last two time periods is noteworthy.

On the other hand, "DL & citizenship" has exhibited the most pronounced downward trend, characterized by an acceleration value of -2.77, particularly due to its stagnant

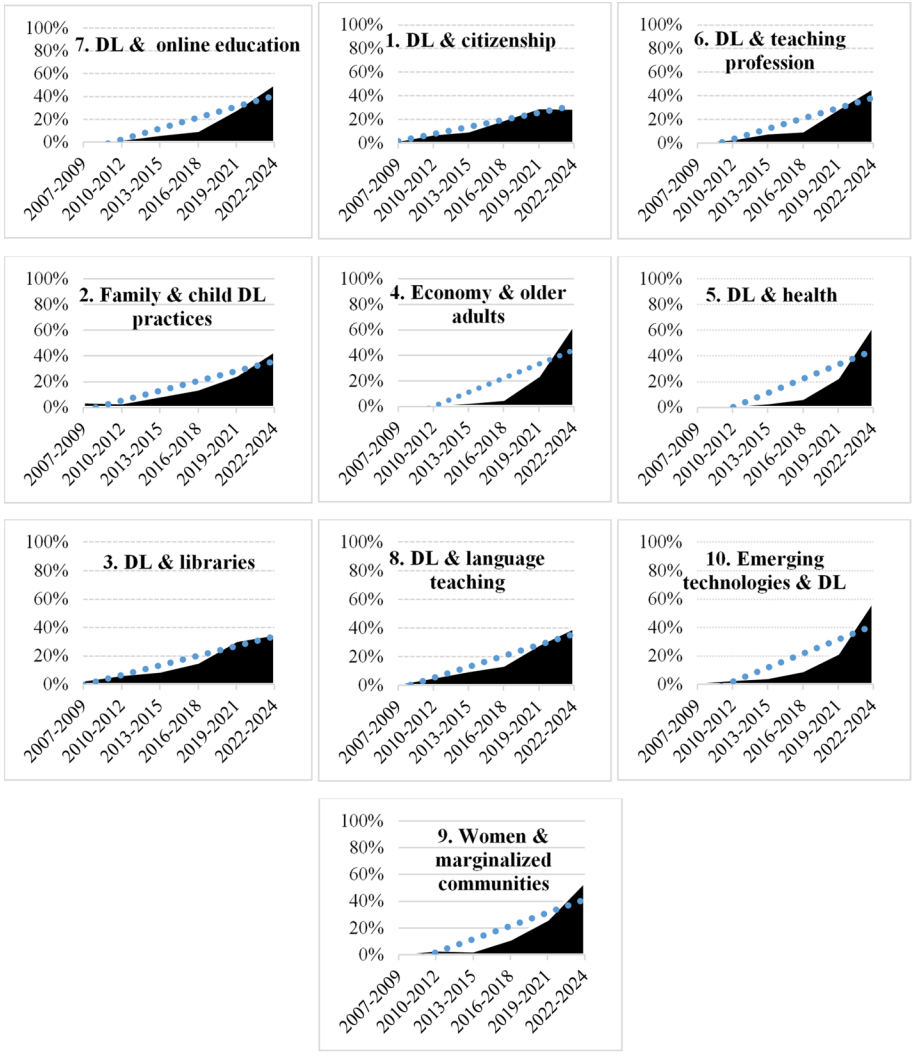


Fig. 3 The change of the volumes of topics within themselves according to time intervals

appearance over the last two time intervals. In terms of exhibiting a downward trend, it is observed that "Family & child DL practices" ranks second with an acceleration value of -2.2, while "DL & libraries" is in third place. Furthermore, the primary reason for "Family & child DL practices" achieving this ranking is the relatively high ratio it attained in the initial time period (9/37) compared to other time periods (Table 6). Lastly, the ratio of the volumes of other topics to the total volume of all studies is evaluated to exhibit a stable appearance across time periods.

The pronounced acceleration observed in the "DL & health" and "Economy & older adults" topics can be preliminarily interpreted as a direct consequence of the global COVID-19 pandemic. The pandemic significantly increased the reliance on digital platforms for healthcare delivery (e.g., telemedicine) and highlighted the digital vulnerabilities and needs of older populations, who faced social isolation. This context likely catalyzed research interest in these areas, a point which will be further elaborated in the discussion section.

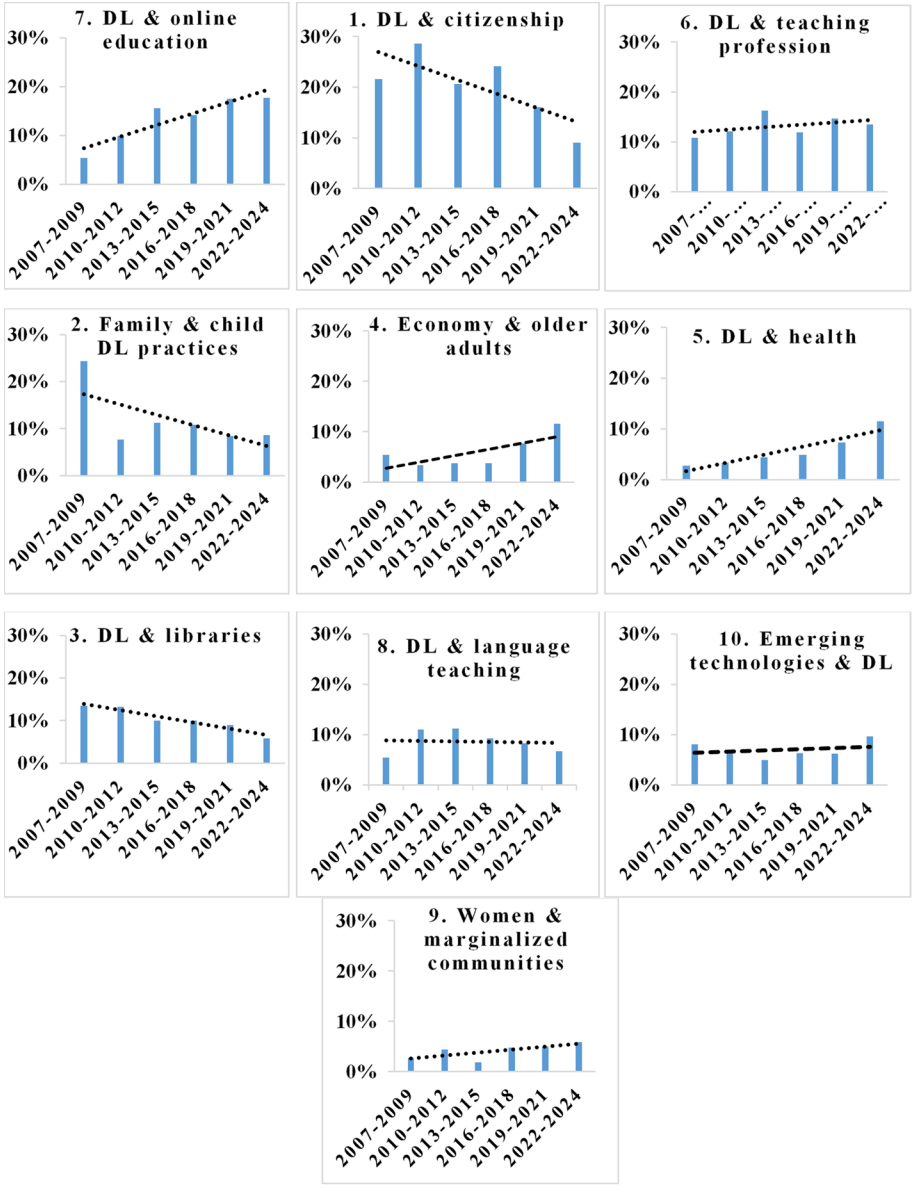


Fig. 4 The change of the volumes of topics relative to each other

4 Discussion

In the present investigation, an extensive examination utilizing descriptive statistics and topic modeling methodologies was conducted on articles related to DL covering the period from 2007 to the middle of 2024. The foremost aim of this investigation is to clarify the current terrain and identify the dominant research themes and trends prevalent within the field.

In analyzing the progression of published research on DL throughout the years (Fig. 1), it becomes evident that there was a consistent rise in the number of publications up until 2019. Following this period, the upward trajectory transformed into a parabolic shape. The emergence of this trend supports the conclusions of earlier research [7, 10, 55], indicating that DL plays a crucial role in academic pursuits and remains an area of interest for researchers from various academic disciplines. Moreover, the positive acceleration

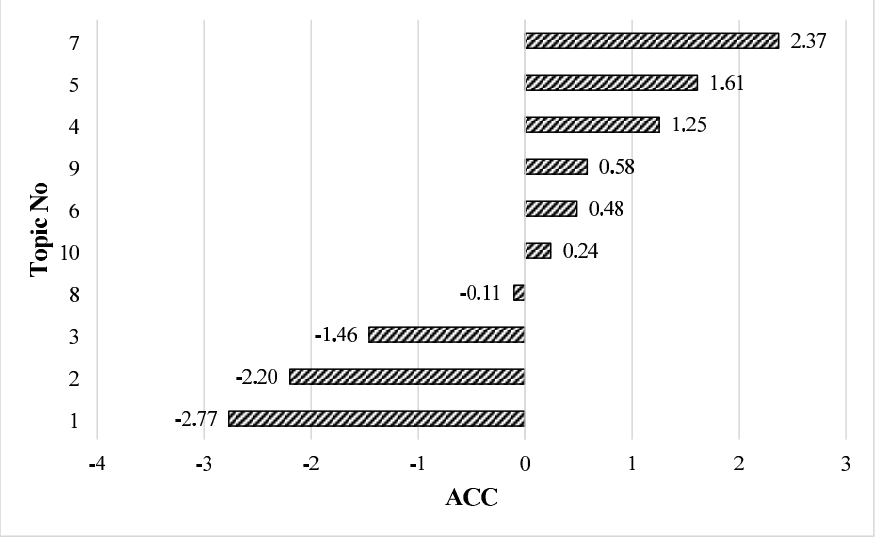


Fig. 5 The ACC values of topics

in the number of articles, particularly after 2019, supports previous studies [58, 62] that indicate the COVID-19 pandemic significantly impacted research related to DL.

The examination of topic modeling indicated that within the domain of DL, scholarly articles were classified into ten principal topics. Upon an evaluation and synthesis of the findings illustrated in Fig. 2; Table 4, and Table 5, the following conclusions can be articulated. Figure 2 reveals that topics Family & child DL practices (2), Economy & older adults (4), DL & health (5), Women & marginalized communities (9), and to a certain extent, Emerging technologies & DL (10) are characterized by their independence and maintain a consistent positioning relative to one another. Conversely, topics DL & teaching profession (6), DL & online education (7), and DL & language teaching (8) exhibit a high degree of interrelation, leading to a complex interplay. This intricate relationship is substantiated by the terminology associated with topics 6, 7, and 8, which is entirely justifiable given their focus on various aspects of the education field. Consequently, it is anticipated that the articles pertaining to these topics will share common lexicons or conceptual frameworks, thereby increasing the probability of their classification across multiple topics. Moreover, an analysis of the terminology representative of each topic indicates that the theme of distance learning is predominantly featured in topic 7, while the professional development of both in-service and pre-service educators is a central focus of topic 6. On the other hand, topic 8 is primarily concerned with foreign language acquisition, with a particular emphasis on the English language.

In addition, the topics of DL & citizenship (1) and DL & libraries (3) are interrelated to some extent. When considering the terms representing these topics (see Table 4), it becomes evident that the term "society" is present in both topics. However, Topic 1 primarily approaches the societal theme through the lenses of education and youth. More generally, relevant studies within Topic 1 focus on how DL can enhance educational outcomes and societal awareness among young individuals (e.g., [41]). On the other hand, Topic 3 frequently associates the societal theme with the role of libraries in fostering DL and promoting societal inclusion (e.g., [56]). Additionally, the term "information literacy" is another common concept shared between the two topics. This term has found its place in numerous articles categorized within both topics. In the articles belonging to

both groups, information literacy is collectively viewed as an important tool for enhancing social participation and achieving social justice [29]. Furthermore, it is regarded as a necessary skill for individuals to be effective in the digital age [37]. For this reason, it is believed that these relationships between the two topics play a significant role in their interconnectedness.

Among the identified topics, the three largest in volume are DL & online education, DL & citizenship, and DL & the teaching profession. Despite the substantial quantity of literature concerning DL & citizenship, an analysis of the number of studies conducted during the most recent two time periods reveals a notable decline in its popularity. This observation is further supported by the fact that both ACC values associated with this topic are the lowest when compared to others'. The notion of digital citizenship has evolved into a multifaceted and intricate discourse, as indicated by Jæger, who emphasizes the proliferation of various dimensions and aspects within the scholarly discussion on digital citizenship [28]. This complexity may inhibit researchers from thoroughly exploring the subject due to the need to traverse a diverse and often fragmented terrain. From another perspective, the increase in the focus on digital solutions in education and social interaction due to the prioritization of urgent needs during the pandemic may be another factor contributing to the decreased importance of this topic.

The prevalence of studies on DL within the contexts of online education and the teaching profession, along with their high acceleration values derived from their volume compared to the total volume of all studies, can be attributed to several factors. First, the rapid advancement and integration of digital technologies in educational contexts have made the emphasis on DL an essential component of contemporary pedagogy. Additionally, the digital transformation in education is closely linked to the competencies and preparedness of teachers in this area. This connection was particularly highlighted during the COVID-19 pandemic, which accelerated the shift towards remote learning options and underscored the importance of digital competence for both learners and educators [8, 38]. The pandemic prompted a swift transition to digital platforms, emphasizing the need for educational stakeholders to understand and enhance DL for effective pedagogical practices [22, 49, 58]. Notably, nearly half of all studies conducted on these topics (see Table 6) fall within the recent timeframe, indicating that the pandemic has been a significant factor in bringing these topics to the forefront in terms of volume. Furthermore, the comprehensive competencies of DL, as detailed in the introduction, are essential for effective engagement in online educational environments. The prioritization of DL in online education is also reinforced by the necessity to develop digital competencies among educators, which are crucial for facilitating successful online learning experiences [58]. In conclusion, it is thought that all these factors have collectively contributed to the substantial volume of research and the acceleration they have gained.

When analyzing the acceleration values of the topics both independently and in relation to one another, it becomes apparent that the topics of DL & health, as well as Economy & older adults, exhibit a significant distinction. These topics have achieved the highest two values in terms of acceleration when evaluated within themselves, while they rank as the second and third highest acceleration values when compared across different topics. Furthermore, approximately 64% of all studies conducted on these two topics have taken place within the last time interval. Consequently, these two domains

have surfaced as the most prominent topics in recent discourse. The outcomes we have derived align with the conclusions drawn in prior studies (e.g., [5, 32, 39]).

The growing interest in DL research within the health sector is driven by multiple factors identified in recent studies. DL enhances the implementation of telemedicine and digital health tools, which are vital for remote healthcare delivery [40]. Furthermore, digital health literacy, defined as the capacity to seek, comprehend, and utilize health information in a digital environment, is essential for effectively navigating contemporary healthcare systems and accessing digital health services and applications [27]. Additionally, evidence suggests a positive correlation between DL and healthy lifestyle behaviors across various generations, highlighting its potential to enhance health outcomes and foster health equity [13]. In summary, it seems the intersection of technological advancements, the necessity for effective health communication, and the commitment to health equity has catalyzed the rising prominence of DL research in the health domain.

The UNFPA [60] reports that the global proportion of individuals aged 65 and older nearly doubled from 1974 to 2024, with projections indicating a further doubling by 2074. As a result of this, the integration of older adults into a societal life that increasingly involves technology has become more important. It is argued that DL is crucial for older adults to maintain social connections, manage health, and access resources, which directly impacts their well-being and autonomy [59]. Moreover, the COVID-19 pandemic has further accelerated the need for digital skills among older populations, as digital platforms became essential for communication and accessing services, yet this rapid digitalization has also highlighted the digital divide, particularly among older adults [11]. From another point of view, the digital economy offers transformative opportunities for economic activities, where DL empowers older adults to participate in digital life and entrepreneurship, particularly in rural areas [64]. In conclusion, the increasing elderly population in both developed and developing countries, along with the significance of DL in the integration of older adults into areas such as the economy and health, is considered to have led to the emergence of the trends observed in this topic in the recent past.

Although not as pronounced as the topics of DL & health and Economy & older adults, there is a notable acceleration in the volume of studies within the topic of Women & marginalized communities relative to the overall volume of research. Adam, [1] emphasizes that the shift to online education, particularly in light of the COVID-19 pandemic, has exacerbated pre-existing educational inequalities. This scenario highlights the urgent need for research that specifically targets the DL needs of marginalized groups. Such a necessity may serve as a plausible rationale for the growing scholarly focus on this topic.

5 Conclusion

In conclusion, this comprehensive analysis of digital literacy literature from 2007 to mid-2024 reveals a dynamic and evolving research landscape characterized by a consistent rise in publications, particularly after 2019. The study identified ten principal themes and highlighted the significant acceleration in research related to "DL & health" and "Economy & older adults," reflecting pressing societal needs likely catalyzed by the COVID-19 pandemic. Furthermore, the growing focus on "Women & marginalized communities" signals an emerging awareness of disparities in digital access and literacy.

The unique contribution of this study, however, lies in the novel insights generated through the LDA method. Specifically, the novel contributions are threefold. First, the semantic analysis provides a granular map of the research field, distinguishing between subtly different but important sub-topics (e.g., "DL & Teaching Profession" vs. "DL & Online Education") that systematic and bibliometric reviews might overlook. Second, the trend analysis objectively quantifies the momentum of emerging research frontiers, providing concrete evidence for the field's future trajectory. Third, the study empirically identifies "Women & marginalized communities" as a comparatively under-researched area, thereby highlighting a critical gap for future inquiry.

The findings of this study also offer significant practical implications. For education policy, the prominence and continued growth of "DL & Online Education" underscore the need for national policies that support sustainable digital infrastructure in schools and continuous professional development for teachers. In terms of digital curriculum design, the research trends suggest a move beyond basic technical skills towards fostering critical digital literacy, including the ability to identify misinformation and engage in digital citizenship. Finally, the identification of older adults, women, and marginalized communities as key areas of recent research highlights the urgency of targeted interventions and inclusive digital literacy programs designed to empower these vulnerable groups, ensuring they are not left behind in an increasingly digital society.

Overall, by leveraging LDA, this study offers a nuanced, evidence-based overview that equips researchers, educators, and policymakers with a deeper understanding of the structure and evolution of digital literacy research, paving the way for more targeted and impactful future studies.

6 Limitations, and future research

This investigation aimed to elucidate the focal points of research and the trends in publication within the domain of DL, tracing the evolution from historical documentation to contemporary studies. However, it is imperative to recognize certain limitations. Primarily, the dataset is restricted to publications indexed in the Scopus database. This selection is predicated on the perception of the Scopus database as both a reliable and comprehensive resource within the academic community. Secondly, the review undertaken considered exclusively those studies published in the English language. The choice of these parameters may have led to the exclusion of pertinent research, potentially introducing a degree of bias. Thirdly, while this study effectively captures the significant impact of the COVID-19 pandemic on the digital literacy research landscape, the data collection timeframe (up to mid-2024) limits a comprehensive analysis of long-term, post-pandemic shifts.

A further limitation of this study is that the analysis was based on the combined titles, abstracts, and author keywords, rather than the full texts of the articles. While this combination provides a rich summary of a paper's focus and key terminology, it may still not fully capture the detailed discussions, methodological nuances, or the depth of the findings present in the full text. Therefore, the identified topics and trends should be interpreted as a reflection of the literature at this combined metadata level.

These limitations suggest avenues for future work. Expanding the dataset to include diverse databases and multiple languages could yield a more holistic understanding of the research themes prevalent in this field. Future research could also build upon this

study by analyzing data from 2025 onwards to explore how research trends evolve and potentially stabilize in the years following the pandemic's most acute phase. Furthermore, as articulated in the literature review section, future research may undertake systematic review studies that consider the topics extrapolated from this investigation.

Author contributions
Erdem Çekmez: Conceptualization, Literature Review, Writing – Original Draft, Writing – Review & Editing, Interpretation of Results, Discussion, Conclusion, Assisted with Data Analysis.Özcan Özyurt: Conceptualization, Literature Review, Data Collection, Writing, Data Analysis (LDA).

Funding
This work was not supported by any external funding.

Data availability
The dataset analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
This study utilizes published research obtained from the Scopus database on DL, which was accessible through Institutional subscription. Consequently, ethical approval for the execution of this study was deemed unnecessary.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 17 June 2025 / Accepted: 17 February 2026
Published online: 23 March 2026

References

1. Adam T. Digital literacy needs for online learning among peri-urban, marginalised youth in South Africa. Int J Mob Blended Learn. 2022;14(3):133–52. https://doi.org/10.4018/IJMBL.310940.
2. Aggarwal CC, Zhai C, editors. An introduction to text mining. In: Mining text data. Springer; 2012. p. 1–10. https://doi.org/10.1007/978-1-4614-3223-4_1.
3. Akil W, Adnan A. Promoting digital literacy in higher education: Case study of a medical laboratory program in Malaysia. Int J Acad Res Progressive Educ Dev. 2022;11(4). https://doi.org/10.6007/ijarped/v11-i4/15906.
4. Alowais M, Rudd G, Besa V, Nazar H, Shah T, Tolley C. Digital literacy in undergraduate pharmacy education: A scoping review. J Am Med Inform Assoc. 2024;31(3):732–45. https://doi.org/10.1093/jamia/ocad223.
5. Arias López MDP, Ong BA, Frigola B, Fernández X, Hicklent AL, Obeles RS, Celi AJ, L. A. Digital literacy as a new determinant of health: A scoping review. PLOS Digit Health. 2023;2(10):e0000279. https://doi.org/10.1371/journal.pdig.0000279.
6. Arici F, Yildirim P, Caliklar Ş, Yilmaz RM. Research trends in the use of augmented reality in science education: Content and bibliometric mapping analysis. Comput Educ. 2019;142:103647. https://doi.org/10.1016/j.compedu.2019.103647.
7. Arissaputra R, Sobandi A, Sentika S, Adib Sultan M, Wijaya PNP, N. Trend analysis using bibliometric study on digital literacy in education. Int J Humanit Educ Social Sci. 2023;3(3):1637–45. https://doi.org/10.55227/ijhess.v3i3.667.
8. Bacalja A, Beavis C, O'Brien A. Shifting landscapes of digital literacy. Australian J Lang Lit. 2022;45(2):253–63. https://doi.org/10.1007/s44020-022-00019-x.
9. Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. J Mach Learn Res. 2003;3(4–5):993–1022.
10. Caldevilla-Domínguez D, Martínez-Sala A, Barrientos–Báez A. Tourism and ICT. Bibliometric study on digital literacy in higher education. Educ Sci. 2021;11(4):172. https://doi.org/10.3390/educsci11040172.
11. Carlo S, Nanetti S. Covid-19 emergency: Boosting digitalization and a missed opportunity for acquiring digital literacy in the old population. Stepp. 2024;28:54–70. https://doi.org/10.15388/STEPP.2023.27.12.
12. Çekmez E. Geometry education and technology: a 25-year journey through topic modeling. Int J Technol Math Educ. 2025;32(3):115–31. https://doi.org/10.1564/tme_v32.3.02.
13. Çetin M, Gümüş R. Research into the relationship between digital health literacy and healthy lifestyle behaviors: An intergenerational comparison. Front Public Health. 2023;11:1259412. https://doi.org/10.3389/fpubh.2023.1259412.
14. Chen X, Zou D, Xie H, Cheng G, Liu C. Two decades of artificial intelligence in education: Contributors, collaborations, research topics, challenges, and future directions. Educational Technol Soc. 2022;25(1):28–47. https://www.jstor.org/stable/48647028.
15. Choudhary H, Bansal N. July 27, (2023). Emerging themes in digital literacy research: Trends and future directions. SSRN. https://ssrn.com/abstract=4477628
16. Chu K, Keikhosrokiani P, Asl M. A topic modeling and sentiment analysis model for detection and visualization of themes in literary texts. Pertanika J Sci Technol. 2022;30(4):2535–61. https://doi.org/10.47836/pjst.30.4.14.
17. Elsevier. (n.d.). How Scopus works: Content. Elsevier. Retrieved May 24. 2024, from https://www.elsevier.com/solutions/scopus/how-scopus-works/content
18. Farias-Gaytán S, Aguaded I, Ramírez-Montoya M. Transformation and digital literacy: Systematic literature mapping. Educ Inform Technol. 2021;27(2):1417–37. https://doi.org/10.1007/s10639-021-10624-x.

19. Farias-Gaytan S, Aguaded I, Ramírez-Montoya M. Digital transformation and digital literacy in the context of complexity within higher education institutions: A systematic literature review. Humanit Social Sci Commun. 2023;10:386. https://doi.org/10.1057/s41599-023-01875-9.
20. Gensim. (n.d.). Parallelized Latent Dirichlet Allocation. Retrieved March 15. 2024, from https://radimrehurek.com/gensim/models/ldamulticore.html
21. Grisales A, Robledo S, Zuluaga M. Topic modeling: Perspectives from a literature review. Ieee Access. 2023;11:4066–78. https://doi.org/10.1109/access.2022.3232939.
22. Gutiérrez-Ángel N, Sánchez-García J, Rubio I, Martín J, Brito-Costa S. Digital literacy in the university setting: A literature review of empirical studies between 2010 and 2021. Front Psychol. 2022;13:896800. https://doi.org/10.3389/fpsyg.2022.896800.
23. Hu Y, Boyd-Graber J, Satinoff B, Smith A. Interactive topic modeling. Mach Learn. 2014;95(3):423–69. https://doi.org/10.1007/s10994-013-5413-0.
24. Hwang S, Flavin E, Lee JE. Exploring research trends of technology use in mathematics education: A scoping review using topic modeling. Educ Inform Technol. 2023;28:10753–80. https://doi.org/10.1007/s10639-023-11603-0.
25. Ilomäki L, Paavola S, Lakkala M, Kantosalo A. Digital competence – an emergent boundary concept for policy and educational research. Educ Inform Technol. 2014;21(3):655–79. https://doi.org/10.1007/s10639-014-9346-4.
26. Ilomäki L, Lakkala M, Kallunki V, Mundy D, Romero M, Romeu T, Gouseti A. Critical digital literacies at school level: A systematic review. Rev Educ. 2023;11(3):e3425. https://doi.org/10.1002/rev3.3425.
27. Imeri A, Schorr S, Merkel S. Digital health applications and health literacy: an explorative analysis. Inform Res Int Electron J. 2024;29(2):378–85. https://doi.org/10.47989/ir292833.
28. Jæger B. Digital citizenship: A review of the academic literature. J Public Policy Law Manage. 2021;14(1):24–42. https://doi.org/10.3224/dms.v14i1.09.
29. Jaeger PT, Shilton K, Koepfler J. The rise of social justice as a guiding principle in library and information science research. Libr Q. 2016;86(1):1–9. https://doi.org/10.1086/684142.
30. Jia W, Huang X. Digital literacy and vocational education: Essential skills for the modern workforce. Int J Acad Res Bus Social Sci. 2023;13(5):2195–202. https://doi.org/10.6007/ijarbss/v13-i5/17080.
31. Kang J, Kim S, Roh S. A topic modeling analysis for online news article comments on nurses' workplace bullying. J Korean Acad Nurs. 2019;49(6):736–47. https://doi.org/10.4040/jkan.2019.49.6.736.
32. Kartanegara HT, Kamil M, Sudiapermana E, Saefudin A, Pramudia JR, Abdullah KH. Tracing the footprints of digital literacy and adult learners: A bibliometric perspective. Multidisciplinary Reviews. 2024;7(10):2024218. https://doi.org/10.31893/multirev.2024218.
33. Khanlou N, Khan A, Vázquez L, Zangeneh M. Digital literacy, access to technology and inclusion for young adults with developmental disabilities. J Dev Phys Disabil. 2020;33(1):1–25. https://doi.org/10.1007/s10882-020-09738-w.
34. Kherwa P, Bansal P. Semantic N-Gram Topic Modeling. EAI Endorsed Trans Scalable Inform Syst. 2020;7(26):e7. https://doi.org/10.4108/eai.13-7-2018.163131.
35. Kurniasih N. Digital literacy: Education for safe internet usage. Engagem Jurnal Pengabdian Kepada Masyarakat. 2023;7(1):139–50. https://doi.org/10.29062/engagement.v7i1.1534.
36. Lau JH, Baldwin T, Newman D. On collocations and topic models. ACM Trans Speech Lang Process. 2013;10(3):1–14. https://doi.org/10.1145/2483969.2483972.
37. Leaning M. An approach to digital literacy through the integration of media and information literacy. Media Communication. 2019;7(2):4–13. https://doi.org/10.17645/mac.v7i2.1931.
38. Li M, Yu Z. Teachers' satisfaction, role, and digital literacy during the COVID-19 pandemic. Sustainability. 2022;14(3):1121. https://doi.org/10.3390/su14031121.
39. Mainz A, Nitsche J, Weirauch V, Meister S. Measuring the Digital Competence of Health Professionals: Scoping Review. JMIR Med Educ. 2024;10(1):e55737. https://doi.org/10.2196/55737.
40. Meherali S, Rahim K, Campbell S, Lassi Z. Does digital literacy empower adolescent girls in low- and middle-income countries: A systematic review. Front Public Health. 2021;9:761394. https://doi.org/10.3389/fpubh.2021.761394.
41. Mihailidis P. The civic potential of memes and hashtags in the lives of young people. Discourse: Stud Cult Politics Educ. 2020;41(5):762–81. https://doi.org/10.1080/01596306.2020.1769938.
42. Muskhir M, Luthfi A, Watrianthos R, Usmeldi U, Fortuna A, Samala AD. Emerging research on virtual reality applications in vocational education: A bibliometric analysis. J Inform Technol Education: Innovations Pract. 2024;23:5284. https://doi.org/10.28945/5284.
43. Nichols T, Stornaiuolo A. Assembling digital literacies: Contingent pasts, possible futures. Media Communication. 2019;7(2):14–24. https://doi.org/10.17645/mac.v7i2.1946.
44. Njenga JK. Digital literacy: The quest of an inclusive definition. Read Writing-Journal Read Association South Afr. 2018;9(1):1–7. https://doi.org/10.4102/rw.v9i1.183.
45. Nokel M, Loukachevitch N. (2016). Accounting ngrams and multi-word terms can improve topic models. In *Proceedings of the 12th workshop on multiword expressions* (pp. 44–49). https://doi.org/10.18653/v1/W16-1806
46. Oh S, Kim K, Kim M, Oh J, Chu S, Choi J. Measurement of digital literacy among older adults: Systematic review. J Med Internet Res. 2021;23(2):e26145. https://doi.org/10.2196/26145.
47. Ozyurt O. Empirical research of emerging trends and patterns across the flipped classroom studies using topic modeling. Educ Inform Technol. 2023;28(4):4335–62. https://doi.org/10.1007/s10639-022-11396-8.
48. Park KC, Lee CH. A study on the research trends for smart city using topic modeling. J Internet Comput Serv. 2019;20(3):119–28. https://doi.org/10.7472/jksii.2019.20.3.119.
49. Peng D, Yu Z. A literature review of digital literacy over two decades. Educ Res Int. 2022;2022:2533413. https://doi.org/10.1155/2022/2533413.
50. Prinsloo M. Commentary: Critical digital literacies, practices and contexts. Tesol Q. 2022;56(3):1068–73. https://doi.org/10.1002/tesq.3152.
51. Reddy P, Sharma B, Chaudhary K. Digital literacy: A review of literature. Int J Technoethics. 2020;11(2):65–94. https://doi.org/10.4018/IJT.20200701.OA1.

52. Reichert F, Zhang D, Law N, Wong G, Torre J. Exploring the structure of digital literacy competence assessed using authentic software applications. Education Tech Research Dev. 2020;68(6):2991–3013. https://doi.org/10.1007/s11423-020-09825-x.
53. Samala AD, Mhlanga D, Bojic L, Howard N-J, Coelho DP. Blockchain technology in education: Opportunities, challenges, and beyond. Int J Interact Mob Technol. 2024;18(1):20–42. https://doi.org/10.3991/ijim.v18i01.46307.
54. Sik D, Németh R, Katona E. Topic modelling online depression forums: Beyond narratives of self-objectification and self-blaming. J Mental Health. 2023. https://doi.org/10.1080/09638237.2021.1979493.
55. Spante M, Hashemi SS, Lundin M, Algers A. Digital competence and digital literacy in higher education research: Systematic review of concept use. Cogent Educ. 2018;5(1):1519143. https://doi.org/10.1080/2331186X.2018.1519143.
56. Subaveerapandiyan A, Maurya A. Digital media and information literacy: A way to paperless society. Libr Philos Pract (e-journal). 2021;6079. https://doi.org/10.17613/ypa0-kh02.
57. Susanty L. Critical analysis of the research on digital literacy. Education. 2024;2(1):12–25. https://doi.org/10.61194/education.v2i1.149.
58. Tan Y. A bibliometric analysis of scholarly literature related to digital literacy in higher education during the pandemic period. Cogent Educ. 2024;11(1):2341587. https://doi.org/10.1080/2331186x.2024.2341587.
59. Tellado I, Girbés-Peco S, Joanpere-Foraster M, Burgués-Freitas A. Digital literacy of older women with smartphones: A dialogic approach to overcoming barriers. Res Ageing Social Policy. 2024;12(1):44–61. https://doi.org/10.17583/rasp.13864.
60. UNFPA. (2024). *Ageing*. United Nations Population Fund. https://www.unfpa.org/ageing
61. Vodă A, Căutișanu C, Gradinaru C, Tanasescu C, Moraes G. Exploring digital literacy skills in social sciences and humanities students. Sustainability. 2022;14(5):2483. https://doi.org/10.3390/su14052483.
62. Wang C, Si L. A bibliometric analysis of digital literacy research from 1990 to 2022 and research on emerging themes during the Covid-19 pandemic. Sustainability. 2023;15(7):5769. https://doi.org/10.3390/su15075769.
63. Weninger C. Skill versus social practice? Some challenges in teaching digital literacy in the university classroom. Tesol Q. 2022;56(3):1016–28. https://doi.org/10.1002/tesq.3134.
64. Xu F, Zhang X, Zhou D. Does digital financial inclusion reduce the risk of returning to poverty? Evidence from China. Int J Finance Econ. 2024;29(3):2927–49. https://doi.org/10.1002/ijfe.2812.
65. Yang C. Research on strategies for school-based training of digital literacy for university teachers. Adult High Educ. 2023;5(20):148–53. https://doi.org/10.23977/aduhe.2023.052021.
66. Yin B, Yuan CH. Detecting latent topics and trends in blended learning using LDA topic modeling. Educ Inform Technol. 2022;27(9):12689–712. https://doi.org/10.1007/s10639-022-11118-0.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.