

Mapping the Landscape of Social Media Literacy Research: A Bibliometric Analysis of Scopus-Indexed Publications (2010–2026)

Nurhafizah Azizan^{1*}, Aflah Isa²

¹Faculty of Information Science, Universiti Teknologi MARA (UiTM) Johor, Malaysia

²Faculty of Business and Management, Universiti Teknologi MARA (UiTM) Johor, Malaysia

*Corresponding Author

DOI: <https://doi.org/10.47772/IJRISS.2026.100800504>

Received: 22 August 2026; Accepted: 27 August 2026; Published: 10 September 2026

ABSTRACT

Social media literacy has emerged as a distinct construct in understanding how individuals navigate misinformation, algorithmic content curation, and platform-mediated social life. Studies on this topic have grown rapidly. However, existing bibliometric work has addressed the broader construct of digital media literacy rather than social media literacy specifically, leaving the latter's intellectual structure unmapped. This study addresses that gap through a bibliometric analysis of 215 Scopus-indexed documents published between 2010 and 2026. Data were retrieved through exact-phrase matching on the Title, Abstract, and Keywords fields and refined using explicit inclusion and exclusion criteria and a documented screening procedure. VOSviewer was used to construct keyword co-occurrence and country co-authorship networks with temporal overlays. The field records a compound annual growth rate of 28.9% between 2010 and 2025, with 72.1% of all output published from 2020 onward. Twenty-five countries met the three-document threshold. The United States leads output with 61 documents, yet records an average normalised citation score of 0.99, at the world average, while Belgium (3.14), Nigeria (2.74), and China (2.31) achieve higher normalised impact from smaller outputs. Malaysia ranks third with 16 documents and collaborates almost exclusively with Global South partners. Keyword co-occurrence analysis identified six thematic clusters, of which the two largest concern conceptual foundations and adolescent well-being. Contrary to the prominence of misinformation in public discourse, the intellectual centre of the field is adolescent psychosocial outcomes, appearance and well-being themes together account for 152 of 365 keyword occurrences. Temporal analysis shows movement from pedagogy-focused research before 2020 toward population- and outcome-focused research from 2023, with algorithmic literacy recording the most recent average publication year (2025.0) despite appearing in only three documents.

Keywords: Social Media Literacy, Bibliometric Analysis, VOSviewer, Adolescent Well-being, Scopus

INTRODUCTION

Social media platforms build on the technological and ideological foundations of Web 2.0, allowing users to create and exchange content [9]. They now shape how people acquire information, form attitudes, and adopt behaviours. Yet these platforms carry real costs. Health misinformation circulates widely on them, and its effect on adolescent well-being remains under-examined [21]. Much of what users see is determined by algorithmic systems they are often unaware of [6], and heavy use has been linked to poorer mental health outcomes among adolescents and young adults [18]. Researchers across education, psychology, communication, and information science have responded by studying what it takes to use social media competently, safely, and critically. This capacity is termed social media literacy [14, 19].

Social media literacy is conceptually related to, but distinct from, the broader constructs of media literacy and digital literacy. Media literacy traditionally emphasises the critical evaluation of professionally produced mass-

media content; digital literacy emphasises general technological competence. Social media literacy addresses platform-specific capacities including algorithm awareness, source verification within social feeds, privacy management, and critical engagement with user-generated content [5, 10]. Despite this distinction, the literature remains fragmented across disciplines, and no bibliometric study to date has mapped it using social media literacy as its primary unit of analysis. The closest existing work examines the adjacent construct of digital media literacy [1], or related phenomena such as social media addiction [13] and social media use in education [24]. To verify this claim, a supplementary search was conducted across Scopus, Web of Science, and Google Scholar using combinations of “social media literacy” with “bibliometric” or “scientometric”. No existing study was found that adopts social media literacy as its primary unit of analysis. The closely related work maps the broader construct of digital literacy at substantially larger scale [23], underscoring the narrower and more recent nature of the social-media-specific construct.

Defining Social Media Literacy and its Conceptual Boundaries

This study adopts the conceptualisation advanced by Cho and colleagues [5], who define social media literacy as a multidimensional construct comprising the knowledge, skills, and dispositions required to access, analyse, evaluate, produce, and act upon content within algorithmically mediated social platforms. Following Livingstone [10], the construct is understood to operate at the level of the individual user and to be specific to the affordances of social media environments, namely user-generated content, networked audiences, and algorithmic curation of visibility.

Four adjacent constructs are treated as outside the boundary of this analysis. Media literacy concerns the critical evaluation of professionally produced mass-media content and predates networked platforms. Digital literacy concerns general technological and information-handling competence across digital environments and is not platform-specific. News literacy concerns the evaluation of journalistic content irrespective of the channel through which it is encountered. Algorithmic literacy is treated not as a separate field but as an emerging sub-dimension of social media literacy [6], retained where situated within social media contexts.

Two further categories are excluded because they do not describe individual user literacy. The first is organisational or firm-level social media capability, which concerns institutional competence and marketing performance rather than user skill. The second is the study of social media use, adoption, or effects where no literacy, competence-development, or critical-evaluation component is present. This boundary specification was operationalised directly as the inclusion and exclusion criteria reported in the Methodology.

This study therefore maps the growth and geographic distribution of the field, identifies its intellectual structure through thematic clustering and citation analysis, and traces how its dominant themes have shifted over time.

Problem Statement

Social media use is near-universal among adolescents in the United States, where 95% of 13- to 17-year-olds report active use. The scale of associated mental health concerns prompted the U.S. Surgeon General to issue a national advisory [18]. Compounding this, health misinformation circulates widely on these platforms and represents an under-examined threat to adolescent well-being [21].

This urgency is equally salient in the Malaysian context, in which this study is situated. Malaysia is directly represented within the present dataset, contributing 16 documents including a highly cited study modelling how social media literacy skills moderate fake news sharing behaviour [22] and work on socio-demographic predictors of misinformation sharing among Malaysian young adults [3].

Despite the rapid growth of scholarship on social media literacy, the field lacks a comprehensive quantitative synthesis of its intellectual structure. Existing bibliometric studies in adjacent domains, most notably on digital media literacy [1], employ a broader construct that does not isolate the platform-specific competencies unique to social media environments. Among these, algorithm awareness is defined as the combination of users'

knowledge, evaluative skill, and capacity to cope with or influence the algorithmic curation of content [6]. Empirical evidence indicates that university students demonstrate limited understanding of algorithmic curation, and that this awareness varies substantially across platforms [4]. Algorithmic awareness therefore constitutes an emerging competence that existing social media literacy frameworks have yet to incorporate systematically.

The field consequently lacks a systematic mapping. Fundamental questions concerning its growth trajectory, geographic distribution, influential publications, thematic structure, and emerging directions remain unaddressed. In the absence of such a synthesis, research effort risks duplication, and emerging competences such as algorithmic literacy risk being overlooked in framework development.

Research Questions

This study is guided by three research questions:

RQ1: How has social media literacy research developed in terms of publication volume and geographic distribution between 2010 and 2026?

RQ2: What is the intellectual structure of the field, as reflected in its thematic clusters and its most influential publications?

RQ3: How have dominant themes shifted from the earlier to the more recent publication years, and what emerging directions do these shifts indicate?

METHODOLOGY

Research design

This study employs a bibliometric research design, a quantitative approach to identifying patterns across a body of literature through publication counts, citation analysis, co-authorship analysis, and keyword co-occurrence mapping. The procedure follows the methodological guidelines set out by Donthu et al. [7].

Data source and search strategy

Data were retrieved from the Scopus database, selected for its coverage of social science and multidisciplinary journals and its established use in prior bibliometric work on related constructs [1, 13]. Retrieval was based on exact-phrase matching rather than proximity operators, in order to maximise precision with respect to the construct boundary specified above. The following query was applied to the Title, Abstract, and Keywords field:

TITLE-ABS-KEY("social media literacy" OR "social media literacies" OR "social networking literacy" OR "SNS literacy")

The search was restricted to English-language publications and to the document types article, conference paper, and review. No subject-area restriction was applied. The search was executed on 27 August 2026 and returned 304 raw records. An earlier version of this query used a proximity operator combined with the truncated term *competenc** to capture organisational as well as individual competence framings; this broader formulation was found to admit a substantial volume of records concerning firm-level social media capability and marketing performance rather than individual user literacy. The *competenc** term and the associated proximity operator were therefore removed in favour of the exact-phrase query reported above, and the construct boundary is instead enforced explicitly through the inclusion and exclusion criteria in Table 1.

Inclusion and exclusion criteria

The construct boundaries specified in the Introduction were operationalised as the following criteria, applied uniformly to every retrieved record on the basis of its title, author keywords, and abstract (Table 1).

Table 1: Inclusion and exclusion criteria

Code	Criterion
I1	The record addresses the knowledge, skills, or dispositions of individual users in social media environments.
I2	The record measures, theorises, or develops an intervention targeting such literacy.
I3	The social media context is central to the study rather than incidental.
E1	The record addresses organisational, firm-level, or brand social media capability.
E2	The record addresses social media marketing, sales, or business performance outcomes.
E3	The record addresses professional journalistic or institutional verification workflows rather than user literacy.
E4	The record addresses social media use, adoption, or effects with no literacy or competence component.
E5	The record uses “social network” in a non-platform sense, including offline social network analysis.

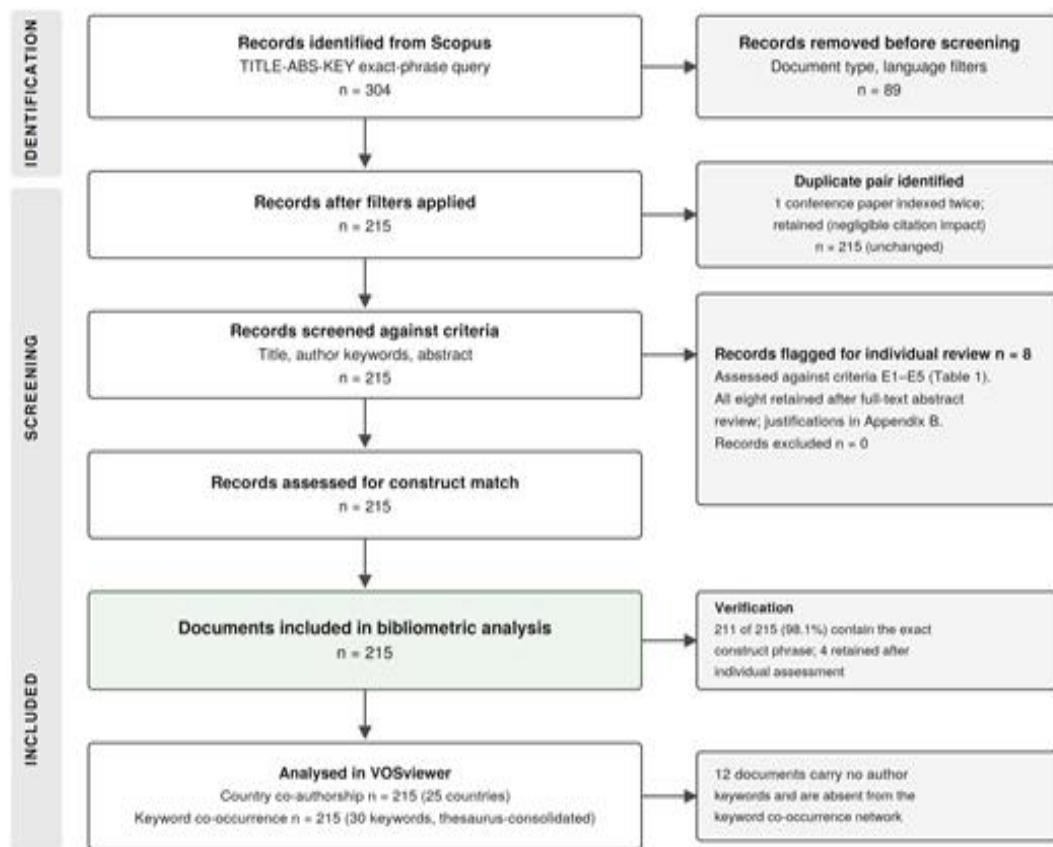
Criteria I1 to I3 show a record needs to include in the corpus. Its focus on individual literacy, including a literacy component that can be measured or theorised, and treat social media as a central topic, not just a side mention. Criteria E1 to E5 list five ways a record can fail this test. The two most common are addressing organisational rather than individual capability (E1, E2), and using “social media” in a non-platform sense, such as offline networks (E5)

Screening and data cleaning

Screening followed the sequence represented in Figure 1. The initial search returned 304 records. Application of document type and language filters removed 89 records, yielding 215. One duplicate pair was identified within the filtered set, corresponding to a single conference paper indexed separately under two Scopus proceedings volumes. Both entries were retained in the analytic corpus, as the duplicate contributes at most 14 citations to a corpus total of 3,401 and does not appear among the top-cited publications in Table 5 or Table 6. The retention does not materially affect any reported indicator. Each record was then assessed against the criteria in Table 1 on the basis of its title, author keywords, and abstract. Eight records were flagged for individual assessment. All eight were retained following review of the full abstract, and the per-record justifications are provided in Appendix B. The final verified corpus comprises 215 documents.

Verification confirmed that 211 of 215 records (98.1%) contain the exact construct phrase in the title, abstract, or author keywords. Of the remaining four, three employ the parenthetical form “(social) media literacy” within the abstract or author keywords. The fourth does not contain the construct phrase in any form but was retained on the basis that it addresses literacy development within a social-media-mediated higher education context, consistent with inclusion criterion I3.

Figure 1: PRISMA-style screening flow diagram



Analysis tools

VOSviewer version 1.6.20 [20] was used for three analyses: keyword co-occurrence mapping based on author keywords, co-authorship network analysis at country level, and temporal overlay visualisation displaying the average publication year of each node. A thesaurus file was applied to consolidate variant terminology, merging forms such as “social networking sites” into “social media” and “fake news” into “misinformation”; the complete thesaurus is provided in Appendix C.

Minimum thresholds were set at three occurrences for keywords and three documents for countries. The keyword threshold was lowered from the more conventional value of five because the restricted corpus yielded only 13 keywords at that level, insufficient to support cluster analysis. At a threshold of three, 30 keywords were retained. Full counting was applied throughout, meaning that a document co-authored across multiple countries is credited in full to each contributing country rather than divided fractionally among them; consequently, country-level document totals reported later sum to more than the corpus size. Normalised citation scores, which adjust raw citation counts for publication year, were used alongside raw counts to avoid privileging older publications.

Each research question is addressed by a distinct set of analyses. RQ1 draws on annual publication counts, compound annual growth rate, and country-level co-authorship analysis. RQ2 draws on keyword co-occurrence clustering and citation analysis. RQ3 draws on the temporal overlay visualisation of the keyword co-occurrence network.

RESULTS

Publication growth and geographic distribution (RQ1)

The corpus comprises 215 documents. 173 articles (80.5%), 31 conference papers (14.4%), and 11 reviews

(5.1%). Together, these accumulated 3,401 citations, a mean of 15.8 and a median of 5.0 per document. Fifty documents (23.3%) remain uncited. The earliest document in the corpus dates to 2010. No publication using social media literacy as an explicit construct appears before that year (Table 2).

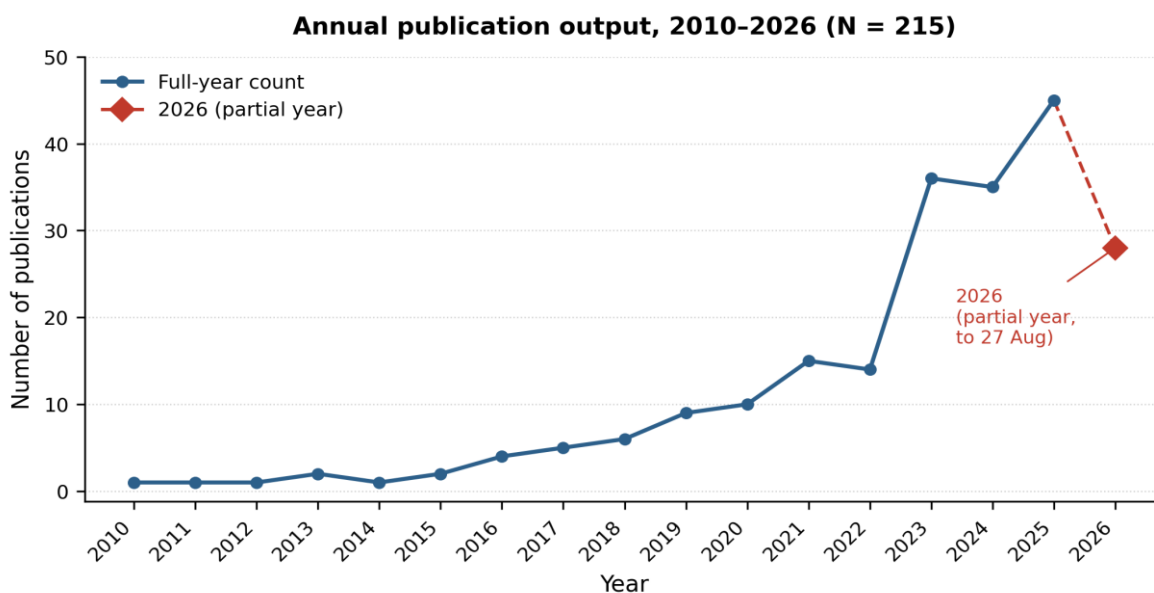
Table 2: Annual publication output, 2010–2026 (N = 215)

Year	N	Year	N
2010	1	2019	9
2011	1	2020	10
2012	1	2021	15
2013	2	2022	14
2014	1	2023	36
2015	2	2024	35
2016	4	2025	45
2017	5	2026*	28
2018	6		

* Partial year; data collection concluded 27 August 2026.

Based on Figure 2, output was trifling until 2015, grew steadily to 2022, then rose sharply from 2023. The compound annual growth rate between 2010 and 2025 is 28.9%. Phase-wise compound annual growth rates show clear acceleration over time, with 14.9% between 2010 and 2015, 38.0% between 2015 and 2020, and 35.1% between 2020 and 2025. Only 32 documents (14.9%) were published before 2020, against 155 (72.1%) between 2020 and 2025.

Figure 2: Annual publication output, 2010–2026, with the partial 2026 year marked separately (N = 215)



Twenty-five countries met the three-document threshold, drawn from 51 countries represented in the corpus overall (Table 3).

Table 3: All 25 countries meeting the three-document publication threshold

Rank	Country	Docs	Citations	Cites/Doc	Avg. Norm. Cit.	TLS	Avg. Yr.
1	United States	61	937	15.4	0.99	22	2022.9
2	Australia	21	847	40.3	1.43	12	2022.2
3	Malaysia	16	222	13.9	0.92	12	2022.6
4	United Kingdom	15	558	37.2	0.98	14	2021.7
5	Italy	15	204	13.6	0.86	15	2023.3
6	China	12	232	19.3	2.31	15	2023.6
7	Indonesia	12	78	6.5	0.41	5	2024.1
8	Germany	11	206	18.7	1.56	14	2023.3
9	France	9	231	25.7	1.27	13	2022.4
10	Spain	9	78	8.7	0.67	11	2023.6
11	Belgium	8	365	45.6	3.14	1	2020.1
12	Austria	8	62	7.8	1.23	4	2020.2
13	Canada	7	175	25.0	0.98	3	2022.6
14	Nigeria	7	171	24.4	2.74	8	2024.0
15	Turkey	7	71	10.1	1.07	3	2021.7
16	India	5	38	7.6	0.27	2	2022.2
17	Pakistan	5	19	3.8	0.50	5	2023.8
18	Japan	5	18	3.6	0.22	0	2020.2
19	Thailand	4	5	1.2	0.10	2	2024.8
20	Philippines	4	3	0.8	0.18	1	2025.2
21	Singapore	3	59	19.7	0.65	2	2022.7
22	Bangladesh	3	56	18.7	1.01	4	2024.0
23	Estonia	3	53	17.7	0.58	0	2018.3
24	Romania	3	21	7.0	1.02	0	2023.3
25	Finland	3	14	4.7	0.74	0	2025.0

A co-authored document is counted once for each country involved. In Table 3, this adds up to 256 across the 25 countries, 41 more than the actual corpus of 215. For example, a document co-authored by researchers based in Malaysia and China contributes one count to Malaysia and one count to China, and is therefore counted twice in Table 3 despite constituting a single document in the corpus. This is a standard convention in bibliometric co-authorship analysis. It is consistent with the counting method reported in the Methodology, and does not indicate an error in the underlying dataset.

Raw and normalised indicators diverge sharply. The United States accounts for 28.4% of corpus output yet

records an average normalised citation score of 0.99, precisely at the world average. Belgium (8 documents, 3.14), Nigeria (7 documents, 2.74), and China (12 documents, 2.31) achieve considerably higher normalised impact from smaller outputs.

The co-authorship network contains 41 links across the 25 countries, a density of 13.7%. Four countries meeting the publication threshold recorded no international co-authorship at all. Estonia, Finland, Japan, and Romania. Malaysia records a total link strength of 12, equal to Australia despite producing a quarter of its output. Its six collaborative partners are China (4), Canada (2), Nigeria (2), Indonesia (2), Bangladesh (1), and Pakistan (1); Malaysia holds no direct co-authorship link to the United States, the United Kingdom, or Australia (Figure 3, Figure 3b).

Figure 3: Country co-authorship network (VOSviewer; minimum three documents per country; node colour indicates cluster membership)

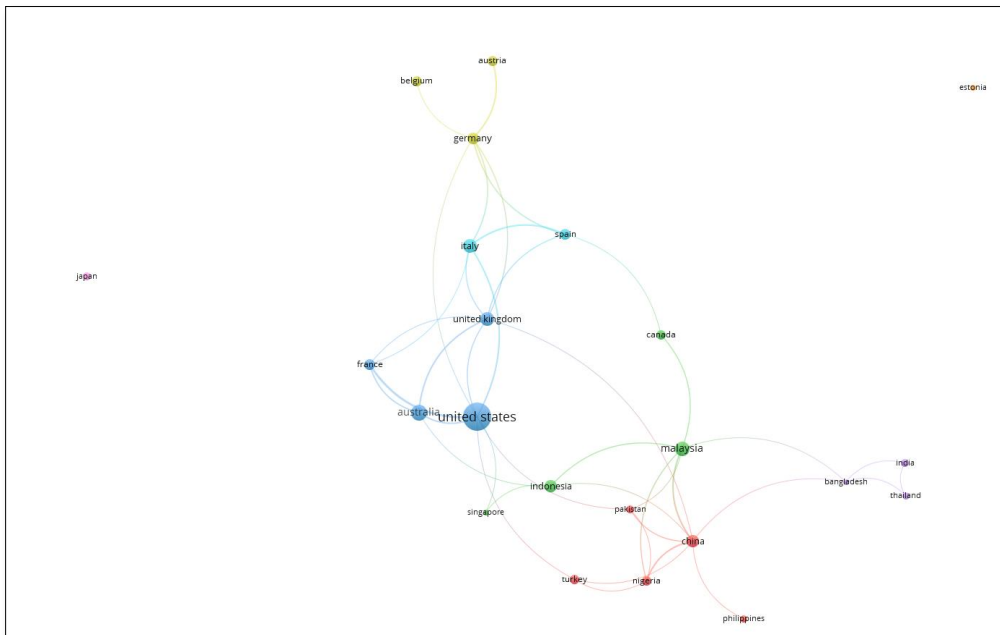
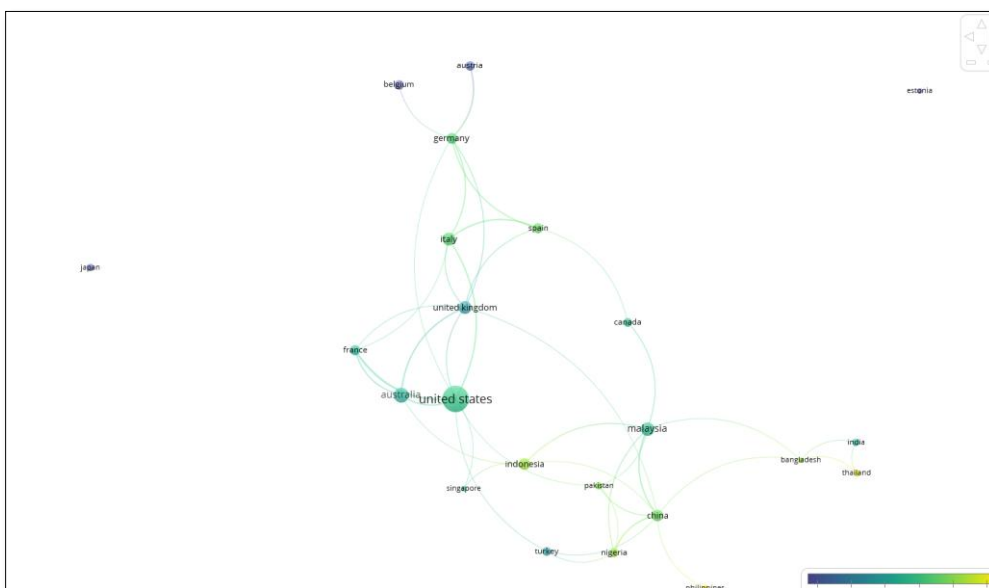


Figure 3b: Country co-authorship network, temporal overlay (node colour indicates average publication year, 2018.3–2025.3)



Intellectual structure: thematic clusters and influential publications (RQ2)

Keyword co-occurrence analysis of author keywords produced a network of 30 keywords and 111 links, a

density of 25.5%, resolving into six thematic clusters (Table 4, Figure 4).

Table 4: Thematic clusters from author keyword co-occurrence analysis

Theme	Kws	Occ.	TLS	Avg. Yr.	Representative Keywords	Representative Publications
Media literacy foundations	5	126	200	2022.1	social media (91), media literacy (25), digital media (4)	Livingstone (2014); Cho et al. (2024); Hanckel et al. (2019)
Adolescent SML & well-being	6	120	141	2023.2	social media literacy (75), adolescents (22), youth (11)	Schreurs & Vandenbosch (2021); Festl (2021); Schreurs et al. (2023)
Digital literacy in education	5	37	66	2021.0	digital literacy (20), higher education (5), teacher education (5)	Nagle (2018); Ungerer (2016); Pangrazio & Gaibisso (2020)
Body image & appearance	5	32	74	2022.5	body image (17), instagram (5), fitpiration (4)	Tamplin et al. (2018); Paxton et al. (2022); Slater et al. (2019)
Misinformation & emerging tech	3	31	48	2023.4	misinformation (20), covid-19 (8), artificial intelligence (3)	Wei et al. (2023); Raj & Goswami (2020); Balakrishnan (2024)
Critical & participatory pedagogy	6	19	29	2021.2	social justice (4), critical literacy (3), pedagogy (3)	Silvestri (2016); Valle et al. (2025); Shahzad & Khan (2024)

The largest cluster, Media literacy foundations, anchors the field's terminology and accounts for 126 of the 365 total keyword occurrences. The second largest, Adolescent SML & well-being, follows closely at 120 occurrences and carries the most recent average publication year (2023.2) of the six clusters, reflecting growing attention to youth psychosocial outcomes. The remaining four clusters are considerably smaller which are Digital literacy in education (37 occurrences) and Body image & appearance (32) form a second tier, while Misinformation & emerging technologies (31) and Critical & participatory pedagogy (19) are the smallest. Representative publications for each cluster are listed in the final column of Table 4.

Figure 4: Keyword co-occurrence network (VOSviewer; author keywords, thesaurus-consolidated, minimum three occurrences; node colour indicates cluster membership)

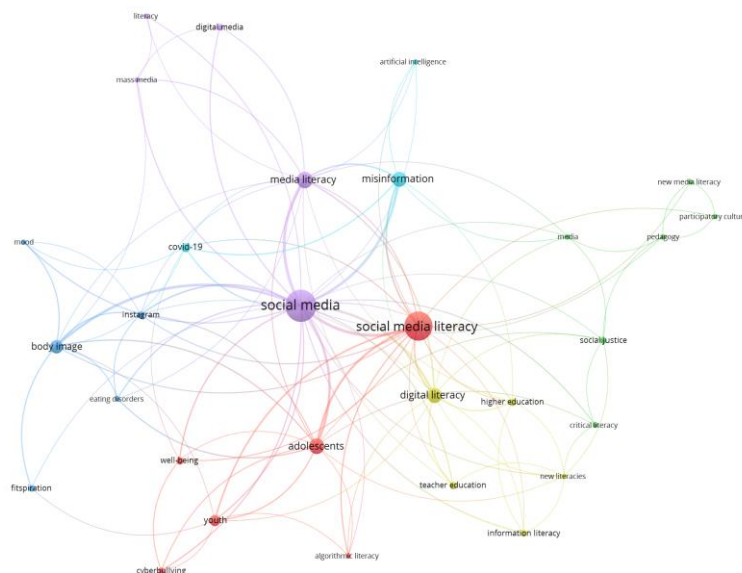
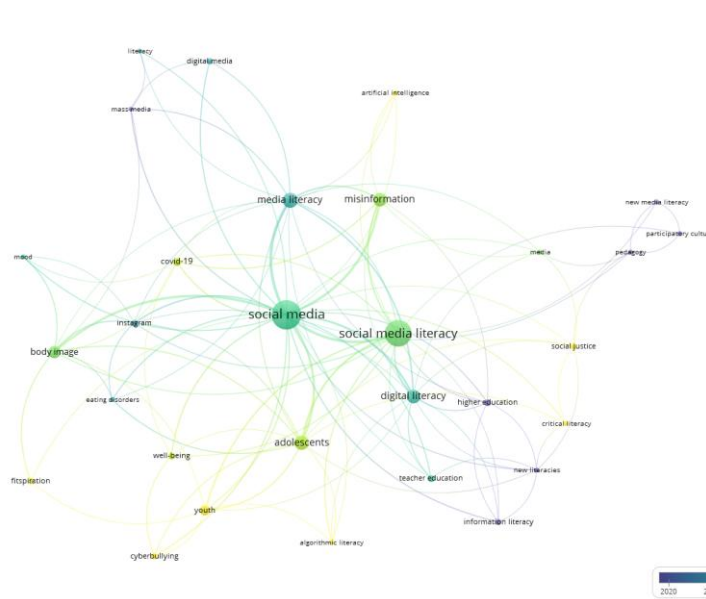


Figure 4b: Keyword co-occurrence network, temporal overlay (node colour indicates average publication year, 2016.3–2025.0)



The two largest clusters account for 246 of the 365 total keyword occurrences. The ten most-cited publications in the corpus are presented in Table 5.

Table 5: Ten most-cited publications in the corpus

Rk	Author (Year)	Source	Cit.
1	Livingstone (2014)	Communications	183
2	Cho et al. (2024)	New Media and Society	182
3	Hanckel et al. (2019)	Media, Culture and Society	161
4	Schreurs & Vandenbosch (2021)	Journal of Children and Media	146
5	Tamplin et al. (2018)	Body Image	143
6	Nagle (2018)	Teaching and Teacher Education	110
7	Wei et al. (2023)	Telematics and Informatics	98
8	Manca et al. (2021)	Computers and Education	98
9	Paxton et al. (2022)	Body Image	94
10	Festl (2021)	Journal of Children and Media	89

Livingstone's (2014) foundational study of how children interpret risk on social network sites is the most-cited work in the corpus, followed closely by Cho et al.'s (2024) conceptual framework. The ten most-cited publications span multiple clusters rather than concentrating in one: three concern conceptual foundations (Livingstone, Cho et al., Hanckel et al.), three concern body image and well-being (Tamplin et al., Paxton et al., Festl), and the remainder address education (Nagle, Manca et al.) and misinformation (Wei et al., Schreurs & Vandenbosch). This distribution indicates that citation impact in the field is shared across its major thematic clusters rather than dominated by a single research strand.

Eighty-six documents (40.0%) carry the construct explicitly in the title. Table 6 restricts the citation ranking to this core subset.

Table 6: Ten most-cited core social media literacy publications

Rank	Author (Year)	Source	Cit.
1	Livingstone (2014)	Communications	183
2	Cho et al. (2024)	New Media and Society	182
3	Schreurs & Vandenbosch (2021)	Journal of Children and Media	146
4	Tamplin et al. (2018)	Body Image	143
5	Nagle (2018)	Teaching and Teacher Education	110
6	Wei et al. (2023)	Telematics and Informatics	98
7	Manca et al. (2021)	Computers and Education	98
8	Paxton et al. (2022)	Body Image	94
9	Festl (2021)	Journal of Children and Media	89
10	Daneels & Vanwynsberghe (2017)	Cyberpsychology	82

Restricting the ranking to documents carrying the construct explicitly in the title removes one publication from the original top ten. Hanckel et al. (2019), whose title addresses platform affordances and identity curation without naming social media literacy directly. Daneels and Vanwynsberghe (2017) enter the core ranking in its place. The remaining nine publications are unchanged between Table 5 and Table 6, indicating that most of the field's most influential work already treats social media literacy as its explicit subject rather than an incidental theme.

Thematic evolution and emerging directions (RQ3)

The temporal overlay reveals a three-phase progression across the keyword network. Earliest keywords cluster around formal education and critical pedagogy consist of participatory culture (2016.3), new literacies (2018.0), pedagogy (2018.0), information literacy (2019.5), and higher education (2019.6). A consolidation phase spanning 2021 to 2022 corresponds to the maturing of the construct itself, with media literacy (2021.5), digital literacy (2021.8), social media (2022.4), and social media literacy (2022.8) all reaching their average publication year in this window, alongside the emergence of Instagram (2021.2) and body image (2022.9). The most recent phase, from 2023 onward, is dominated by population- and outcome-specific themes: misinformation (2023.2), adolescents (2023.4), well-being (2023.6), fitspiration (2023.8), cyberbullying (2024.3), youth (2024.5), and algorithmic literacy (2025.0).

Algorithmic literacy records the most recent average publication year in the network (2025.0), tied with critical literacy, yet appears in only three documents and does not form an independent cluster (Figure 4b).

DISCUSSION

Growth and geographic distribution (RQ1)

The geographic findings require careful interpretation (Table 3). The United States dominates output with 61 documents, representing 28.4% of the corpus, yet records an average normalised citation score of 0.99, precisely at the world average. Volume and impact are not equivalent in this field. Belgium (3.14), Nigeria (2.74), and China (2.31) all achieve substantially higher normalised impact from smaller outputs. Belgian scholarship in particular contributed early conceptual work, with an average publication year of 2020.1 and the highest normalised impact of any country in the corpus, reflecting the influence of the SMILE model and related studies [15, 16].

Malaysia ranks third in publication output with 16 documents and records a total link strength of 12, equal to that of Australia despite producing a quarter of its output. Its six co-authorship partners are China, Canada, Nigeria, Indonesia, Bangladesh, and Pakistan. It holds no direct link to the United States, the United Kingdom, or Australia. Except for Canada, Malaysian collaboration in this field is conducted entirely within the Global South. This is consistent with wider evidence that Global South scholarship collaborates less frequently across regional boundaries and that intra-regional collaboration typically shows lower citation impact than cross-regional collaboration [11], suggesting that Malaysian social media literacy research is developing along a regional rather than a transatlantic axis.

The network remains sparse, with 41 links connect 25 countries, a density of 13.7%, and four countries meeting the publication threshold record no international co-authorship at all. Notably, the keyword network is considerably denser (25.5%) than the collaboration network. The field possesses a coherent conceptual core but has not yet consolidated into an integrated international research community.

Intellectual structure (RQ2)

The thematic structure identified here reframes the field in a way that departs from how social media literacy is commonly introduced. The largest cluster is conceptual, anchoring the construct within the older media literacy tradition. The second largest is organised around adolescent populations and psychological well-being and carries the most recent average publication year of any substantial cluster. Taken together with the appearance-focused cluster, these two clusters account for 152 of the 365 total keyword occurrences, and two of the ten most-cited publications in Table 5 (Tamplin et al.; Paxton et al.) concern appearance-related outcomes specifically. Misinformation, by contrast, occupies a smaller cluster of three keywords totalling 31 occurrences. The intellectual centre of gravity of social media literacy research is therefore adolescent psychosocial outcomes rather than information verification, notwithstanding the prominence of misinformation in public and policy discourse on the subject.

This has consequences for how the construct is operationalised. The most-cited publications reflect the same orientation: the two most-cited works establish conceptual frameworks [5, 10], while subsequent high-impact work develops and tests protective-factor models linking social media literacy to body image and well-being outcomes [15, 17, 12], including at least one randomised controlled trial of a school-based programme [8]. For policy, this suggests that social media literacy programmes framed solely around misinformation detection address the smaller of the field's two evidence bases; the larger body of intervention evidence concerns adolescent body image and psychological well-being.

Thematic evolution (RQ3)

The field has moved from asking how social media literacy should be taught to asking whom it protects and from what. Algorithmic literacy warrants specific comment. It records the most recent average publication year in the network yet appears in only three documents, and is best characterised as an emergent rather than an established theme. Its position within the adolescent well-being cluster, rather than as an independent cluster, suggests that algorithmic concerns are currently being absorbed into existing research agendas rather than forming a distinct line of inquiry. This is consistent with evidence that algorithm awareness among users remains limited and platform-dependent [4], and that recommendation systems shape informational exposure in ways users do not readily perceive [2]. The gap between the recency of this theme and its volume represents the clearest opportunity for future contribution identified by this analysis.

CONCLUSION

This study presents the first bibliometric synthesis focused specifically on the construct of social media literacy, analysing 215 Scopus-indexed documents published between 2010 and 2026. The field is young and expanding rapidly, with a compound annual growth rate of 28.9% between 2010 and 2025 and 72.1% of its output published from 2020 onward. Geographically, the United States leads in volume but not in normalised impact, and Malaysia ranks third in output while collaborating almost exclusively within the Global South. Structurally, the field resolves into six thematic clusters whose combined weight lies with conceptual

foundations and adolescent well-being rather than with misinformation. Temporally, research has shifted from pedagogical questions toward population- and outcome-specific questions, with algorithmic literacy emerging as the newest but least developed theme.

Three directions follow. First, algorithmic literacy is recognised as important but remains empirically thin, representing the clearest opportunity for new contribution. Second, the concentration of intervention evidence in the body image and well-being domain suggests that social media literacy programmes should address psychosocial outcomes alongside critical evaluation, rather than treating misinformation detection as the sole objective. Third, the sparse and regionally segmented collaboration network indicates room for cross-regional partnership, particularly for Malaysian and other Southeast Asian scholarship currently developing along a regional axis.

Limitations

Several limitations qualify these findings. First, exact-phrase retrieval maximises precision at the cost of recall. Relevant work not using the specified terminology in the title, abstract, or keywords is excluded. Second, the boundary between social media literacy and adjacent constructs including media literacy, digital literacy, news literacy, and algorithmic literacy is contested in the primary literature. Therefore, the operational boundary applied here is one defensible position rather than a settled consensus. Third, the analysis was restricted to English-language publications, which may under-represent non-Anglophone research, including from several Asian countries identified as growing contributors.

Fourth, Scopus was the sole data source. Scopus indexing favours established international journals and under-represents regional and non-indexed outlets, which may understate output from developing research systems. Fifth, the 2026 publication count is a partial-year figure and should not be interpreted as evidence of decline. Sixth, citation counts are subject to citation-age bias; normalised citation scores mitigate but do not remove this, and the algorithmic literacy theme identified as a key emerging direction is precisely the theme most disadvantaged by this bias.

Seventh, keyword co-occurrence clustering relies on author-assigned keywords, which vary in consistency across disciplines, journals, and time periods; twelve documents in the corpus carry no author keywords and are therefore absent from the keyword network. Thesaurus consolidation, listed in full in Appendix C, mitigates but does not eliminate this inconsistency, and a small number of consolidation decisions. Most of it notably merging “fake news” into “misinformation” and “algorithm awareness” into “algorithmic literacy”, involve a degree of interpretive judgement that alternative coding schemes might resolve differently. Eighth, average normalised citation scores for individual keywords appearing three or four times are unstable and driven by single publications; for this reason normalised scores are reported at cluster level rather than keyword level. Ninth, the corpus was screened by the authors, and no second-rater reliability assessment was conducted. Finally, bibliometric indicators measure attention and visibility, not quality or causal importance; the thematic structure reported here describes what the field has studied rather than what it should study.

Ethical Considerations

This study analysed publicly available bibliographic metadata retrieved from the Scopus database. No human participants, personal data, or animal subjects were involved, and ethical approval was therefore not required.

The authors declare no conflict of interest.

Data Availability

The bibliographic dataset analysed in this study was retrieved from Scopus (Elsevier) under institutional subscription access and cannot be redistributed due to licensing restrictions. The search string, screening criteria, VOSviewer thesaurus file, and screening decision log are provided in Appendices A–D and are available from the corresponding author upon reasonable request.

REFERENCES

1. Bahlamar, A. R. U. (2024). Digital media literacy in scholarly discourse: A bibliometric analysis of Scopus-indexed publications. *Khizanah al-Hikmah*, 12(2), 253–266. <https://doi.org/10.24252/kah.v12i2a3>
2. Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239), 1130–1132. <https://doi.org/10.1126/science.aaa1160>
3. Balakrishnan, V. (2024). Socio-demographic predictors for misinformation sharing and authenticating amidst the COVID-19 pandemic among Malaysian young adults. *Information Development*, 40(2), 319–331. <https://doi.org/10.1177/02666669221118922>
4. Brodsky, J. E., Zomberg, D., Powers, K. L., & Brooks, P. J. (2020). Assessing and fostering college students' algorithm awareness across online contexts. *Journal of Media Literacy Education*, 12(3), 43–57. <https://doi.org/10.23860/JMLE-2020-12-3-5>
5. Cho, H., Cannon, J., Lopez, R., & Li, W. (2024). Social media literacy: A conceptual framework. *New Media & Society*, 26(2), 941–960. <https://doi.org/10.1177/14614448211068530>
6. Dogruel, L., Facciorusso, D., & Stark, B. (2022). 'I'm still the master of the machine.' Internet users' awareness of algorithmic decision-making and their perception of its effect on their autonomy. *Information, Communication & Society*, 25(9), 1311–1332. <https://doi.org/10.1080/1369118X.2020.1863999>
7. Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285–296. <https://doi.org/10.1016/j.jbusres.2021.04.070>
8. Gordon, C. S., Rodgers, R. F., Slater, A. E., McLean, S. A., Jarman, H. K., & Paxton, S. J. (2020). A cluster randomized controlled trial of the SoMe social media literacy body image and wellbeing program for adolescent boys and girls: Study protocol. *Body Image*, 33, 27–37. <https://doi.org/10.1016/j.bodyim.2020.02.003>
9. Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons*, 53(1), 59–68. <https://doi.org/10.1016/j.bushor.2009.09.003>
10. Livingstone, S. (2014). Developing social media literacy: How children learn to interpret risky opportunities on social network sites. *Communications*, 39(3), 283–303. <https://doi.org/10.1515/commun-2014-0113>
11. McManus, C., Baeta Neves, A. A., Finan, T. J., Pimentel, F., Pimentel, D., & Schleicher, R. T. (2024). The South-South dimension in international research collaboration. *Anais da Academia Brasileira de Ciências*, 96(3). <https://doi.org/10.1590/0001-3765202420230942>
12. Paxton, S. J., McLean, S. A., & Rodgers, R. F. (2022). My critical filter buffers your app filter: Social media literacy as a protective factor for body image. *Body Image*, 40, 158–164. <https://doi.org/10.1016/j.bodyim.2021.12.009>
13. Pellegrino, A., Stasi, A., & Bhatiasavi, V. (2022). Research trends in social media addiction and problematic social media use: A bibliometric analysis. *Frontiers in Psychiatry*, 13, 1017506. <https://doi.org/10.3389/fpsyt.2022.1017506>
14. Polanco-Levicán, K., & Salvo-Garrido, S. (2022). Understanding social media literacy: A systematic review of the concept and its competences. *International Journal of Environmental Research and Public Health*, 19(14), 8807. <https://doi.org/10.3390/ijerph19148807>
15. Schreurs, L., & Vandenbosch, L. (2021). Introducing the Social Media Literacy (SMILE) model with the case of the positivity bias on social media. *Journal of Children and Media*, 15(3), 320–337. <https://doi.org/10.1080/17482798.2020.1809481>
16. Schreurs, L., Meier, A., & Vandenbosch, L. (2023). Exposure to the positivity bias and adolescents' differential longitudinal links with social comparison, inspiration and envy depending on social media literacy. *Current Psychology*, 42(32), 28221–28241. <https://doi.org/10.1007/s12144-022-03893-3>
17. Tamplin, N. C., McLean, S. A., & Paxton, S. J. (2018). Social media literacy protects against the negative impact of exposure to appearance ideal social media images in young adult women but not men. *Body Image*, 26, 29–37. <https://doi.org/10.1016/j.bodyim.2018.05.003>
18. U.S. Department of Health and Human Services, Office of the Surgeon General. (2023). Social media and youth mental health: The U.S. Surgeon General's advisory.

<https://www.hhs.gov/surgeongeneral/reports-and-publications/youth-mental-health/social-media/index.html>

19. Valle, N., Zhao, P., Freed, D., Gorton, K., Chapman, A. B., Shea, A. L., & Bazarova, N. N. (2025). Towards a critical framework of social media literacy: A systematic literature review. *Review of Educational Research*, 95(4), 701–746. <https://doi.org/10.3102/00346543241247224>
20. van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538. <https://doi.org/10.1007/s11192-009-0146-3>
21. Wang, M. L., Togher, K., Griffiths, R., & Lin, T.-K. (2024). Health misinformation on social media and adolescent health. *JAMA Pediatrics*, 178(2), 109–110. <https://doi.org/10.1001/jamapediatrics.2023.5282>
22. Wei, L., Gong, J., Xu, J., Zainal Abidin, N. E., & Apuke, O. D. (2023). Do social media literacy skills help in combating fake news spread? Modelling the moderating role of social media literacy skills in the relationship between rational choice factors and fake news sharing behaviour. *Telematics and Informatics*, 76, 101910. <https://doi.org/10.1016/j.tele.2022.101910>
23. Wu, D., Sukumaran, S., Zhi, X., Zhou, W., Li, L., & You, H. (2025). Categories, themes and research evolution of the study of digital literacy: A bibliometric analysis. *Education and Information Technologies*, 30(4), 4907–4931. <https://doi.org/10.1007/s10639-024-12955-x>
24. Yu, Z., Sukjairungwattana, P., & Xu, W. (2023). Bibliometric analyses of social media for educational purposes over four decades. *Frontiers in Psychology*, 13, 1061989. <https://doi.org/10.3389/fpsyg.2022.1061989>

APPENDICES

Appendix A: Complete search strategy

Full query as executed on 27 August 2026, filters applied (English language; article, conference paper, review), and record counts at each stage (304 raw → 215 after filters → 215 final).

Appendix B: Screening decision log

Eight records could not be resolved by title-based screening alone and were assessed individually against the full abstract. All eight were retained; the criterion initially raised and the retention justification are listed below.

Year	Title (abbreviated)	Criterion raised	Retention justification
2025	Female youth and social media: How do I see myself through social media?	E4	Measures self-perceived media literacy among 581 university students; I1 and I2 satisfied.
2024	Building Resilience Against Hostile Information Influence Activities (Estonian Defense Forces)	E4	Literacy training platform developed for individual military personnel; I2 satisfied.
2023	Corporate training for developing social media literacy skills	E1	Targets individual employee literacy development, not firm-level capability.
2012	Participatory action research in the age of social media: Literacies, affinity spaces and learning	E4	Addresses literacy development within a social-media-mediated higher education context; I3 satisfied.
2018	The emergence of socio-material assemblages in a university, company, and municipality collaboration	E4	Examines literacy development within a platform-mediated collaborative learning setting.

2026	“How Do I Draw Something That Doesn’t Exist?” (Pre)Adolescents’ Constructions of Personal and Ideal Bodies	E4	Contains an explicit “(social) media literacy” reference in the Conclusions; relevant to the body-image cluster.
2025	Developing Tangkahan Tourism Village in Indonesia by Strengthening Social Media Literacy for SME	E1	Intervention targets individual participant literacy development, not organisational capability.
2022	Role of Social Media Marketing in SME Sector Performance	E2	Models social media literacy as an individual-level predictor variable within the study design.

Appendix C: VOSviewer thesaurus file

Complete mapping of source term to replacement term applied during keyword co-occurrence analysis (45 pairs). Source terms on the left of each pair map to the replacement term on the right.

Source term	Replacement	Source term	Replacement
social networking sites	social media	fake news	misinformation
social networking site	social media	disinformation	misinformation
social network sites	social media	body dissatisfaction	body image
social network site	social media	appearance	body image
social networking	social media	mental health	well-being
sns	social media	wellbeing	well-being
social media use	social media	psychological well-being	well-being
social media usage	social media	algorithm awareness	algorithmic literacy
social network	social media	algorithmic awareness	algorithmic literacy
social platforms	social media	algorithms	algorithmic literacy
social media literacy skills	social media literacy	artificial intelligence (ai)	artificial intelligence
critical social media literacy	social media literacy	ai	artificial intelligence
sml	social media literacy	university students	higher education
digital literacies	digital literacy	college students	higher education
digital media literacy	digital literacy	higher education students	higher education
media literacies	media literacy	pre-service teachers	teacher education
media education	media literacy	preservice teachers	teacher education
adolescence	adolescents	teachers	teacher education
adolescent	adolescents	covid-19 pandemic	covid-19

adolescent girls	adolescents	covid 19	covid-19
teenagers	youth	media literacy intervention	intervention
young people	youth	social media literacy intervention	intervention
young adults	youth		

Appendix D: Data-cleaning rules

Duplicate identification. Titles were normalised to lowercase with punctuation stripped, then compared across all 215 records. One duplicate pair was identified: a single conference paper indexed separately under two Scopus proceedings volumes (EIDs 2-s2.0-105046939267 and 2-s2.0-85181535664, cited 7 and 14 times respectively). Both entries were retained in the analytic corpus, as their combined citation contribution is negligible relative to the corpus total and neither appears among the top-cited publications in Table 5 or Table 6.

Author keyword normalisation. Author keywords were lowercased and stripped of leading or trailing whitespace before matching. Variant forms differing only in singular or plural (e.g., “adolescent” versus “adolescents”), British and American spelling, hyphenation, or common abbreviation were consolidated to a single preferred term using the thesaurus file listed in Appendix C. Terms not covered by the thesaurus were retained in their original form. Twelve documents carried no author keywords and were therefore excluded from the keyword co-occurrence network, though they remain in all other analyses.

Country attribution. Country was attributed at the level of individual author affiliations rather than the corresponding author alone. A document co-authored by researchers based in different countries was credited in full to each country represented, consistent with the full counting convention reported in the Methodology. No fractional or first-author-only counting was applied. This is why the Documents totals in Table 3 sum to more than the corpus size of 215.