Contemporary Educational Technology, 2025, **17**(4), ep609
ISSN: 1309-517X (Online)

OPEN ACCESS

Review Article

Sustainability education meets artificial intelligence: A post-2022 bibliometric map and thematic analysis

Alfiya R. Masalimova [1*]
0000-0003-3711-2527

Yuliya P. Kosheleva [2]
0000-0001-5653-2143

Aleksandr I. Burov [3]
0000-0002-7909-2993

Olga V. Payushina [3]
0000-0001-8467-0623

Natalia L. Sokolova [4]
0000-0002-0667-5098

Maria A. Khvatova [5]
0000-0002-2156-8805

[1] The Institute of Psychology and Education, Kazan Federal University, Kazan, RUSSIA
[2] Laboratory of Differential Psychology and Psychophysiology, Federal Scientific Center of Psychological and Multidisciplinary Research, Moscow, RUSSIA
[3] I. M. Sechenov First Moscow State Medical University, Moscow, RUSSIA
[4] Peoples' Friendship University of Russia (RUDN University), Moscow, RUSSIA
[5] Bauman Moscow State Technical University, Moscow, RUSSIA
[*] Corresponding author: alfkazan@mail.ru

Citation: Masalimova, A. R., Kosheleva, Y. P., Burov, A. I., Payushina, O. V., Sokolova, N. L., & Khvatova, M. A. (2025). Sustainability education meets artificial intelligence: A post-2022 bibliometric map and thematic analysis. *Contemporary Educational Technology, 17*(4), ep609. https://doi.org/10.30935/cedtech/17482

ARTICLE INFO

Received: 12 May 2025

Accepted: 17 Nov 2025

ABSTRACT

This study maps 417 peer reviewed publications (2022-2025) at the intersection of sustainability education (SE) and artificial intelligence (AI) using bibliometric methods. We chart venues, co authorship, keyword evolution, and technique usage. The results reveal that "ChatGPT" and "generative AI" are becoming the most popular terms after 2022. Outputs are still mostly from North America and Europe, although contributions from Saudi Arabia, India, and Malaysia are growing. Institutional networks are broken, which means that institutions don't cooperate together very often. Supervised learning predominates, and neural networks are the most used single technique. We synthesize scattered findings into three practical principles–personalization–protection, competence alignment, and multi-level synchronization–that link AI uses to core SE competencies and support course to institution coordination. The study also shows a dual sustainability lens: AI can help fight climate change, but it also has implications for privacy and the environment. This shows the need for energy reporting and bias safeguards. We suggest causal and longitudinal assessments, collaborative datasets and rubrics, and capacity enhancement for resource-limited environments. Some of the problems are a short citation window (2022-2025), a bias against English speakers, and the possibility of missing databases. Overall, the subject is growing swiftly, but it requires more proof, common standards, and more environmentally friendly ways of doing things to turn AI into lasting educational value.

Keywords: sustainability education, artificial intelligence, generative AI, bibliometric analysis

INTRODUCTION

Sustainability education (SE) is a normative and transformative pedagogy that aims to develop learners' critical reasoning, systemic understanding, foresight, strategic action, and collaboration competencies in sustainability issue areas (Corres et al., 2020; Sandri, 2020). On the other hand, education sustainability looks at how long an organization can last, from how it is run to how it works on campus. Even though the two subjects are typically talked about together in writing, they have different analytical focuses: We look at SE by looking at the competency outcomes of learning designs, and we look at education sustainability by looking at institutional performance metrics including energy, waste, carbon, equity, and stakeholder engagement (Rieckmann, 2012; Usak et al., 2021; Velazquez et al., 2006; Wiek et al., 2011). This study notes the institutional dimension as a contextual background and focuses its analysis on the pedagogical outcomes of SE.

The post-2022 generative artificial intelligence (AI) wave has triggered both a methodological and normative repositioning in SE literature. Intelligent teaching systems, generative AI-based writing and feedback tools, and learning analytics present innovative instruments for transformative learning via personalization and self-regulation. Nonetheless, these instruments present issues related to privacy, bias, explainability, and the environmental consequences associated with AI's energy and hardware consumption (Lin et al., 2023; Rillig et al., 2023). Initial meta-analytic results reveal medium-to-high effect sizes in overall learning outcomes for chatbots (Deng & Yu, 2023); however, evidence is still scarce regarding the mechanisms through which these effects translate into fundamental SE competencies (e.g., systems thinking and action competence). Concurrently, the "whole-institution" approach is found to improve outcomes by integrating organizational design with instruction (Holst et al., 2024). In short, the SE-AI intersection is a rapidly growing but fragmented field of knowledge that requires the simultaneous management of opportunities and risks.

This article aims to systematically map how AI has transformed SE into the post-2022 literature:

(1) which SE competencies are supported by AI and how they are measured,

(2) which pedagogical designs (ITS, generative AI feedback, XR/Metaverse, learning analytics) appear effective,

(3) how risks such as security-privacy-bias-environmental footprint are managed, and

(4) to assess the quality and clarity of evidence.

The contribution of this study is to present a framework that matches SE competencies with AI intervention types through the "competency-mechanism-evidence" triad and to make ethical and environmental dimensions visible together by not limiting the evaluation to pedagogical outputs alone.

There are notable gaps in the relevant literature:

1. Designs that test the causal and longitudinal impact of AI-based interventions on SE competencies are rare.
2. standardized measures (systems thinking scales, action competence, ethics/AI literacy) and shared datasets are insufficient.
3. AI's environmental footprint is rarely integrated into decision-making processes.
4. the equity and inclusivity dimension (access, language, cost, disability) is heterogeneous and often marginal.
5. When governance alignment cannot be established across class-program-institution scales, the policy-practice gap widens (Holst et al., 2024; Lin et al., 2023; Rillig et al., 2023; Rosen, 2025).

This study aims to systematically code and map these gaps to lay the groundwork for subsequent empirical designs and institutional strategies.

Ultimately, our claim is this: The use of AI in SE produces sustainable pedagogical value when the principles of personalization-protection (data/ethics), competency alignment (measurable contribution), and multi-level synchronization (course-program-institution purpose alignment) are collectively ensured.

The remaining questions are clearly visible:

(1) multi-stage/longitudinal designs that reveal the effects on SE competencies,

(2) cross-context comparability with shared criteria and datasets,

(3) integration of equity and inclusivity parameters into the design-evaluation chain,

(4) embedding environmental footprint and financial sustainability into decision-making processes, and

(5) adaptive governance and micro-crediting/professional development mechanisms that bridge the policy-practice gap.

Based on these findings, we propose three principles to guide SE-YZ applications: personalization + protection, competency alignment, and multi-level synchronization. The following sections interpret the methodology and findings in light of these principles.

METHOD

This study employed a mixed-methods approach. In the first phase, a bibliometric analysis using quantitative methods was conducted. In the second phase, the findings of the bibliometric analysis were examined and synthesized.

Data Collection Process

For the bibliometric phase, the Web of Science and Scopus databases were searched; the scope covers the period from 1 January 2022 to 15 September 2025. The search was conducted using the following query family in the title/abstract/keyword fields: ("sustainability education" OR "education for sustainable development" OR 'ESD' OR "education sustainability") AND ("artificial intelligence" OR "AI" OR "machine learning" OR "generative AI" OR 'chatbot' OR "large language model" OR "LLM").

The full search query is as follows: TITLE-ABS-KEY (("education for sustainable development" OR "sustainability education" OR "sustainable education" OR "educational sustainability" OR "education for sustainability" OR "environmental education" OR "green education" OR "SDG* education" OR "sustainab* pedagog*") AND ("artificial intelligence" OR "AI" OR "generative AI" OR "machine learning" OR "deep learning" OR "LLM" OR "large language model*" OR "ChatGPT" OR "GPT*" OR "intelligent tutoring system*" OR "adaptive learning" OR "personalized learning system*" OR "educational AI" OR "conversational AI").

Inclusion/exclusion criteria were established. Inclusion criteria were peer-reviewed articles and reviews; English language; studies addressing AI use in the SE context or the impact of AI on SE theoretically/empirically. Exclusion criteria include editorials/forewords/errata; texts focused solely on corporate sustainability operations; engineering/model development studies lacking an educational context; and records without full-text access. Author/institution names and keywords were standardized (AI/AI/LLM, ESD/SE synonym mappings). A title-abstract pre-screening was performed in a PRISMA-like flow. Publications were cleaned because the keyword "ESD" is used in abbreviations in different fields (e.g., endoscopic submucosal dissection [ESD] in medicine and electrostatic discharge [ESD] in physics). After cleaning, 335 publications were obtained in the Scopus database, while 235 publications were obtained in Web of Science. Duplicate records matching DOI/ISBN were removed. As a result of the removal, 417 publications remained.

Data Analysis

Scientific mapping and performance analyses were conducted using Bibliometrix v5 (R); bibliometrix's biblioshiny module was used for visualizations. In the preprocessing steps, author/institution/country names and keywords were lemmatized and standardized; the dataset was enriched with field-normalized citation metrics. Performance analysis reported annual publication trends, source and author productivity (Lotka), journal core (Bradford), most cited works, and collaboration indicators.

Co-authorship (author/institution/country), co-citation (author/source/document), bibliographic matching, and co-word networks were extracted. Association Strength normalization was used for edge weights; Louvain clustering (with Leiden sensitivity analysis for validation) was employed. Threshold values (e.g., keyword min. frequency = 3; document co-citation threshold = 5) were applied for low-frequency nodes to reduce noise. Thematic evolution and trend topic analyses were produced using time-segmented networks; the conceptual structure was mapped using correspondence analysis/MCA to reduce multiple writing overlaps.

Table 1. Descriptive statistics on publications

Description		Results
Main information about data	Timespan	2022:2025
	Sources (journals, books, etc.)	204
	Documents	417
	Annual growth rate (%)	67.77
	Document average age	0.906
	Average citations per document	4.650
Document contents	Keywords plus (ID)	1288
	Author's keywords (DE)	2,050
Authors	Authors	1,537
	Authors of single-authored docs	50
Authors collaboration	Single-authored docs	52
	Co-authors per document	3.94
	International co-authorships (%)	17.27
Document types	Article	243
	Book	9
	Book chapter	55
	Conference paper	59
	Proceedings paper	29
	Review	22

Table 2. Number of publications and citations

Year	MeanTCperArt	N	MeanTCperYear
2022	7.97	36	1.99
2023	15.27	59	5.09
2024	4.45	152	2.22
2025	0.44	170	0.44

In terms of qualitative content analysis, articles in the purposeful sample were examined using thematic coding based on SE core competencies, AI intervention types, measurement indicators, and ethical-environmental dimensions. The authors jointly decided on themes and quotations in the content analysis.

FINDINGS

The dataset is large and very recent. It covers 417 documents across 2022-2025 with a high annual growth rate (67.77%) (**Table 1**). The average document age is under one year (0.906), so citation counts are naturally modest (4.65 per doc). Dissemination is dispersed. Work is spread over 204 sources, suggesting no single outlet dominates yet. This fits a fast-emerging topic where authors publish opportunistically across education, sustainability, and computing venues.

Authorship shows collaborative but not highly global teams. There are 1,537 authors for 417 items, and about four co-authors per paper (3.94). Single-authored works are ~13% (52/417). International co-authorship is moderate at 17.27%, indicating room for stronger cross-country networks as the field consolidates.

Document types point to an implementation and methods phase. Articles are the majority (243; ~58%). Conference outputs (59; ~14%) and proceedings papers (29; ~7%) show rapid, time-sensitive diffusion typical after a technological shock (Generative AI). Reviews are still few (22; ~5%), implying the synthesis stage is only beginning. Book chapters (55) add breadth but will accrue citations slowly.

Keyword volume signals thematic fragmentation and experimentation. With 2,050 author keywords and 1,288 keywords plus for 417 items, topics are diverse and still coalescing. This aligns with your later maps where "education," "sustainability," "higher education," and "AI" act as hubs while many niche themes remain peripheral. Overall, the profile is of a young, fast-growing literature sparked after 2022: broad venue spread, collaborative teams, moderate internationalization, and many primary studies with few syntheses so far. As cohorts age, we should expect rising citations, more cross-border projects, and consolidation in a smaller set of core journals.

Output grows very fast after 2022 (**Table 2**). Publications rise from 36 (2022) to 59 (2023), then jump to 152 (2024) and 170 (2025). This confirms a post-ChatGPT expansion of research on AI in SE. Impact per paper

Table 3. Most influential source in AI and SE

Source	h_index	g_index	m_index	TC
Sustainability	12	19	3.00	464
International Journal of Sustainability in Higher Education	4	8	1.33	75
Cogent Education	3	5	1.50	31
Discover Sustainability	2	3	1.00	9
Smart Learning Environments	2	3	0.67	108
E3s Web of Conferences	2	3	0.67	10
Education Sciences	2	2	0.50	5
Education and Information Technologies	2	2	1.00	5
Journal of Cleaner Production	2	2	1.00	207
Administrative Sciences	2	2	0.67	96
Applied Sciences-Basel	2	2	0.50	22
Resources Policy	2	2	0.67	18
Scientific Reports	2	2	1.00	14
Radiography	2	2	1.00	9
Lecture Notes in Networks and Systems	1	4	0.50	18
Lecture Notes in Computer Science	1	1	0.50	2
Environmental Science & Technology	1	1	0.33	113
International Journal of Data and Network Science	1	1	0.33	105
ASEAN Journal of Science and Engineering	1	1	1.00	5

peaks in 2023. MeanTCperArt is 15.27 and MeanTCperYear is 5.09, about 2× and 2.6× higher than 2022, respectively. 2023 functions as the "early-mover" year when foundational and agenda-setting papers attracted the most attention. In 2024, the field scales up but averages dilute. N increases by 158%, while MeanTCperArt falls to 4.45 and MeanTCperYear to 2.22. This pattern fits a diffusion phase: many more papers enter the literature, often empirical or case-based, which need time to accumulate citations. The 2025 figures are not yet comparable. Output continues to climb (+12%), but MeanTCperArt and MeanTCperYear are 0.44 because the citation window is short and indexing may lag. Treat 2025 as provisional. A rough aggregation underscores the story: 2023 accounts for ~46% of all citations so far, 2024 ~35%, 2022 ~15%, and 2025 ~4%. Overall, the trajectory suggests a shift from novelty-driven, highly cited 2023 studies to broader implementation work in 2024-2025. As newer cohorts age, their averages should rise, indicating consolidation rather than waning interest.

Sustainability is the clear anchor journal (**Table 3**). It leads to volume (NP = 54) and cumulative impact (TC = 464), with the highest h-index (12) and g-index (19). Its m-index of 3.00 signals very fast, recent accumulation of influential papers–consistent with a surge after the diffusion of generative AI.

The International Journal of Sustainability in Higher Education plays a complementary, more specialized role. It publishes fewer papers (NP = 10) but maintains steady influence (TC = 75; h = 4). This suggests it is a go-to venue for institutional and curricular perspectives rather than technical AI advances.

Education-technology outlets are prominent signal carriers. Smart Learning Environments stands out: only 3 papers but 108 citations, indicating that design frameworks or early GenAI classroom studies in sustainability are being widely referenced. Education and Information Technologies and Education Sciences appear as steady, mid-tier hubs that disseminate practical implementations.

High-impact, low-count sources point to cross-disciplinary "flagship" articles. Journal of Cleaner Production (2 papers, 207 citations), Environmental Science & Technology (1, 113), and International Journal of Data and Network Science (1, 105) each host single contributions with outsized reach. These are likely methods or reviews that link AI capabilities to sustainability outcomes and are reused across education syllabus and research.

Contribution of the Authors

The author landscape is young and fragmented (**Table 4**). Most names have only 1-2 papers, which is typical for a post-2022 surge driven by generative AI. Influence is therefore concentrated on a few highly cited papers rather than long publication careers.

Two impact patterns stand out. First, very high TC with low h/g (e.g., Abulibdeh A., Abulibdeh R., and Zaidan E. at 170 citations and several co-authors with 113) signals single landmark papers attracting field-wide

Table 4. Most influential authors in AI and SE

Author	h_index	g_index	m_index	TC	NP
Lei C.	3	3	0.75	13	6
Wang Y.	2	2	0.50	8	6
Kim J.	1	1	0.25	3	5
Chen Y.	1	1	0.50	2	4
Sun Q.	1	3	0.25	12	4
Ronaghi M.	2	2	2.00	4	2
Leong W.	2	2	1.00	20	2
Al S. A.	2	2	0.67	109	2
Singh K.	2	2	0.67	55	2
Singh P.	2	2	0.67	55	2
Sharma S.	2	2	0.67	18	2
Nguyen M.	2	2	0.67	12	2
Deng X.	2	2	0.50	119	2
Yu Z.	2	2	0.50	119	2
Abulibdeh A.	1	2	0.50	170	2
Abulibdeh R.	1	1	0.50	170	1
Zaidan E.	1	1	0.50	170	1
Agerstrand M.	1	1	0.33	113	1
Bi M.	1	1	0.33	113	1
Gould	1	1	0.33	113	1
Kenneth A. K.	1	1	0.33	113	1
Rillig M.	1	1	0.33	113	1
Sauerland U.	1	1	0.33	113	1

attention. Because g-index cannot exceed the number of papers, these authors remain capped at g = 1-2 despite large citation counts.

Second, a small set couples output with steady citations (e.g., Lei C.: h = 3 across 6 papers and Wang Y.: h = 2 across 6). Their TC is modest but diversified, suggesting ongoing programs that may compound as the literature matures.

m-index highlights the speed of influence since first publication. Very high m (Ronaghi M. = 2 and Leong W. = 1) points to rapid early impact, consistent with a field where influential pieces can gain traction quickly after 2022. Authors with m ≈ 0.33-0.67 (many of the 2023 cohorts) show stable but slower trajectories.

Geography inferred from names and earlier institutional results aligns: Chinese and Asian affiliations are prominent (e.g., Lei C., Wang Y., Chen Y., and Sun Q.), matching the strong production from Hong Kong and mainland China institutions. This suggests Asia is shaping the research agenda in AI-for-SE.

Overall, "influence" currently reflects two engines: blockbuster, often multi-author pieces that define concepts, and emerging labs publishing several smaller studies. As citations accumulate, we should expect h and g to rise for the productive group, while single-paper leaders will need follow-ups to translate early visibility into durable author-level indices.

The co-authorship network is sparse and modular (**Figure 1**). Most nodes form dyads or small triads with limited links across groups. This fits a young field that expanded rapidly after 2022: teams publish within their own circles, and cross-cluster projects are still rare.

A single collaboration core sits around **Lei C.** and **Wang Y.** (light-blue cluster). Their nodes connect to multiple co-authors (e.g., Chen Y., Wang H., Wang X., Li X., and Chiu D.), giving them the highest centrality and PageRank in **Table 4**. This mirrors **Table 4** where both authors showed higher h-index with several papers. They function as early bridges who can diffuse methods and topics across subgroups.

Two additional hubs are visible but more insular. **Kim J.** works in a compact cluster with **Mahmud S.**, **Sun Q.** links to **Aljohani N.** and **Hsiao I**. Their closeness scores are moderate, yet betweenness is low–signals of cohesive teams that are not yet brokering between communities. Similar tight units appear around **Al-Emran M.–Al-Qaysi N.–Arpaci I.**, **Arnold O.–Jantke K.**, and **Kerkeb M.–Khoulji S.**

Several authors with high closeness = 1 and zero betweenness (e.g., **Hassan N.** and **Ismail A.**) are endpoints of tiny components. They are well connected locally (within their mini-cluster) but structurally

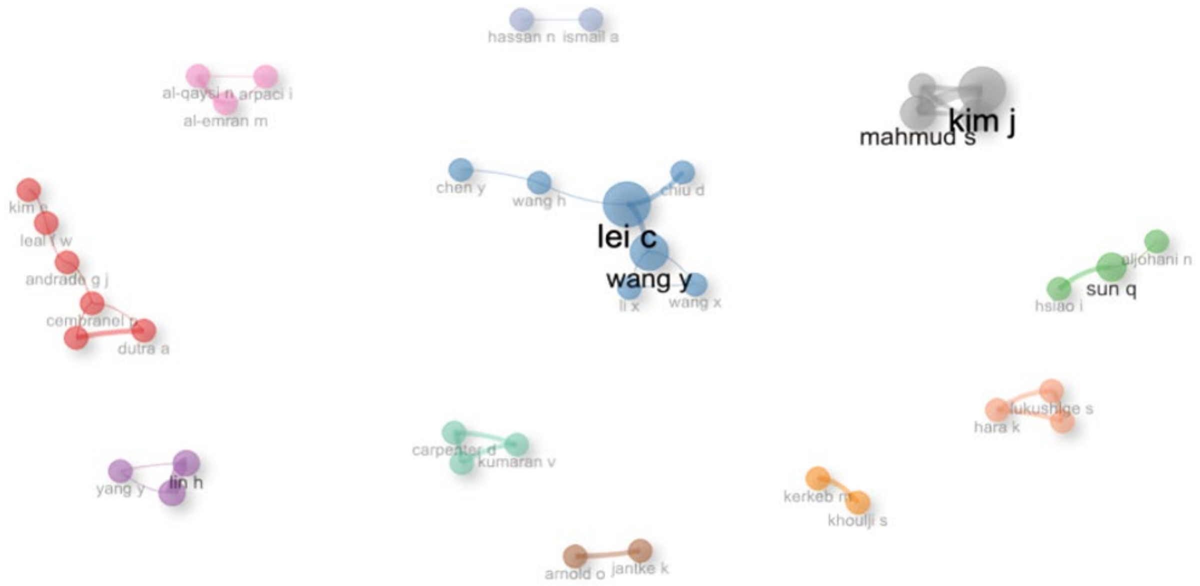

Figure 1. Network graph for authors (Created by the authors)

Table 5. Top 10 institutions in AI and SE

Affiliation	Articles
The University of Hong Kong	15
Central China Normal University	11
University of Aveiro	11
University of Helsinki	11
Monash University	10
East China Normal University	9
University of Connecticut	9
University of Ljubljana	8
University of Southern Santa Catarina	8
California Institute of Technology	7

isolated from the wider network. This pattern is common in fast-growing topics where new teams enter with one or two joint papers.

Regional signals are consistent with other results. East-Asian names dominate the central core (Lei, Wang, Sun), while Middle-East and European/Latin American clusters remain peripheral and separate. Cross-regional collaboration appears limited, which may slow theory convergence and tool standardization in “AI for SE.” The field is vulnerable to fragmentation: impact concentrates on a few bridging authors while many teams remain siloed. Coordinated special issues, shared datasets, and multi-institution grants that deliberately span clusters–especially connecting the Lei C./Wang Y. core with the Kim J. and Sun Q. clusters–would likely accelerate integration and raise overall betweenness across the network.

Contribution of the Institutions

The institutional landscape is broad but still emerging (**Table 5**). The University of Hong Kong leads with 15 papers, yet the gap to the rest is small (11-7 papers). This narrow range suggests a fragmented field without a single dominant hub. It fits a post-2022 surge where many groups are entering at once rather than a few long-established centers driving output. Geographically, the top 10 span five continents: Asia (Hong Kong; Central China Normal; East China Normal), Europe (Aveiro, Helsinki, Ljubljana), North America (Connecticut; California Institute of Technology), South America (UNISUL), and Oceania (Monash). Asia and Europe are most visible, with China appearing twice and the EU represented by three universities. This spread indicates that AI for SE is international from the outset, not confined to traditional Anglophone centers.

The mix of institutions is telling. Two “normal universities” highlight the role of teacher-education and pedagogy research, which aligns with classroom-level AI adoption after 2022. Monash and Helsinki bring strong learning-sciences traditions, while Caltech’s presence signals a complementary, engineering-driven

Table 6. Contribution of countries

Country	TC	Average article citations	SciProduct
China	373	6.0	173
Saudi Arabia	200	16.7	30
USA	153	4.9	112
India	149	3.3	68
Germany	119	14.9	31
Brazil	76	5.8	35
Mexico	42	6.0	26
United Arab Emirates	40	10.0	3
Italy	39	7.8	14
Lithuania	38	19.0	6
Spain	35	2.2	53
Serbia	24	12.0	4
United Kingdom	21	2.3	37
Czech Republic	21	7.0	10
Malaysia	17	1.7	31
Portugal	14	2.0	39
Australia	9	1.0	32

stream where AI and sustainability link through technical domains. Together, these profiles show that the topic sits at the intersection of education research and STEM.

The USA is present but not dominant (Connecticut; California Institute of Technology), and the UK does not appear in the top 10. The inclusion of UNISUL (Brazil) and Ljubljana (Slovenia) points to growing Global South and Central/Eastern European participation. This diversification may support localized curricula and context-sensitive sustainability challenges.

Overall, counts are modest and close to each other, reinforcing that this is a young, rapidly forming research space post-Generative-AI uptake. We would expect collaboration networks to consolidate next, likely around Asian and European hubs, with opportunities for cross-regional projects that connect pedagogy-focused groups with engineering-focused ones.

Contribution of the Countries

China is the clear volume leader (**Table 6**). It accounts for about one quarter of all author-affiliation appearances in this subset (173 of ~704) and about 27% of total citations (373 of ~1,370). The USA (112 appearances; ~16%) and India (68; ~10%) follow as major producers. Spain is also large in volume (53) but with low average citations. High average impact is concentrated in a different set. Lithuania (avg 19, from a small base of 6) and Saudi Arabia (avg 16.7, 30 appearances) sit at the top, with Germany close behind (avg 14.9, 31). Serbia (avg 12) and the United Arab Emirates (avg ~10, but only 3 appearances) also show strong per-article influence. The combination of mid-scale volume and high averages makes Saudi Arabia and Germany especially notable.

Several large producers show modest citation averages. The USA (avg 4.9), China (6), India (3.3), Spain (2.2), the UK (2.3), Malaysia (1.7), Portugal (2.0), Australia (1.0) and Brazil (5.8) indicate breadth and community engagement, but not yet high per-article impact. Given the 2022+ window, recency likely depresses averages for countries that published more in 2024-2025. Regional patterns emerge. East and South Asia lead on sheer output (China, India, and Malaysia). Gulf countries punch above their weight on impact (Saudi Arabia, UAE). Europe is mixed: Germany stands out on influence, while Spain and the UK contribute many papers with lower short-run citations. Latin America is present via Brazil and Mexico, both with moderate averages.

If the goal is reach, partnerships with Germany and Saudi Arabia look promising given their consistent influence at moderate scale. For scale and visibility, China, the USA, and India anchor the network. Countries with high participation but low averages (Spain, UK, Malaysia, Portugal, and Australia) may host many applied or emerging-area studies that have not yet accumulated citations–worth watching as generative-AI-related work diffuses. In short, the field shows a “quantity vs. quality” split across countries, shaped by collaboration patterns and the short citation window. **Table 6** makes both dimensions visible and supports targeted collaboration and monitoring strategies.

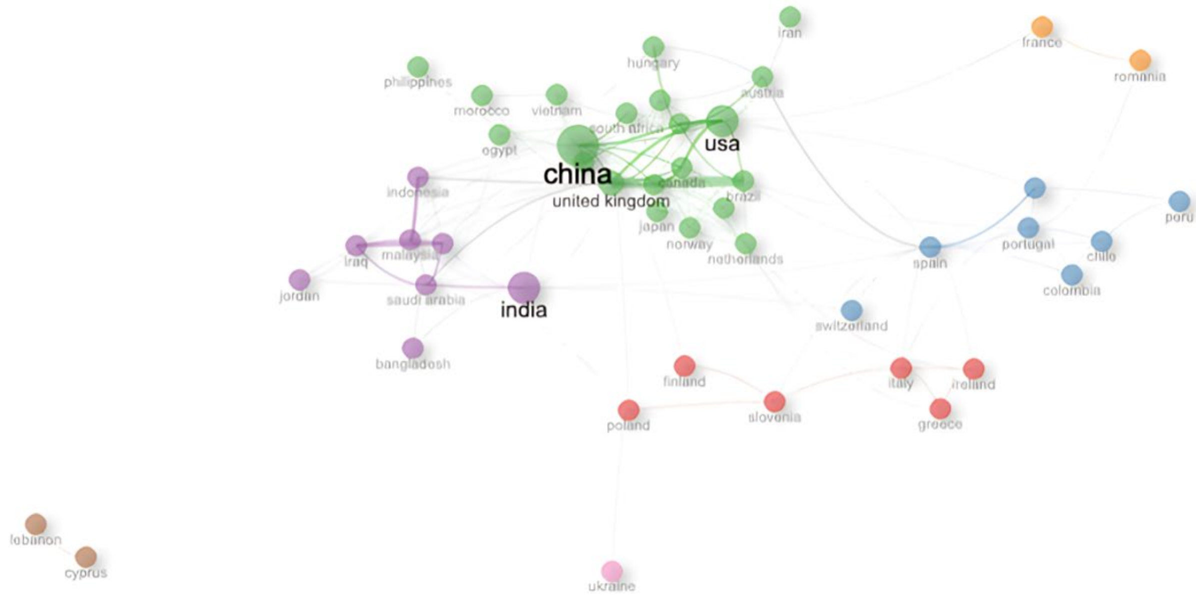

Figure 2. Network graph for countries (Created by the authors)

Figure 2 shows four main modules. A large "global" cluster (green) is organized around the USA, China, the UK, Germany, the Netherlands, and other OECD countries. A South/Southeast Asia-Gulf cluster (purple) is centered on India with Malaysia, Saudi Arabia, Indonesia, Bangladesh, and Jordan. An Ibero-American cluster (blue) links Spain and Portugal to Mexico, Chile, Colombia, and Peru. A smaller Central/Eastern Europe cluster (red) includes Italy, Ireland, Greece, Slovenia, Finland, and Poland. Two dyads (Cyprus-Lebanon) and a peripheral node (Ukraine) sit apart.

Brokerage is concentrated in a few hubs. The UK has the highest betweenness and PageRank. It is the key bridge across modules and the most "reachable" node by closeness. The USA and China are also central, but the UK appears to connect otherwise weakly linked regions more often. Spain is the standout broker to Latin America: high betweenness with strong ties to Mexico, Chile, Portugal, and Colombia. Within Europe, the Netherlands and Germany act as internal connectors; Poland is a notable bridge toward the Eastern subgroup.

In the Asia-Gulf module, India leads. Saudi Arabia and Malaysia show strong centrality given their size. Saudi Arabia's PageRank is comparable to Australia and Brazil, signaling visibility beyond its immediate neighbors. The United Arab Emirates is smaller but well connected, consistent with its high average citations in **Table 6**. Comparing with **Table 6**, influence and position are aligned in useful ways. Germany and Saudi Arabia combined high average citations with good network centrality–this makes them attractive partners for both impact and reach. Spain had low average citations but very high betweenness; it functions as a corridor to the Spanish-speaking Americas, which may pay off as citations mature in a 2022+ window. For global diffusion and cross-cluster projects, target the UK as a coordinating hub and pair with USA/China as volume anchors. To reach Latin America, work through Spain and Portugal. To deepen South/Southeast Asia, prioritize India and Saudi Arabia/Malaysia links. Within the EU, the Netherlands and Germany help span sub-regions; Poland can pull in Eastern partners.

Node size and thick edges reflect collaboration frequency, so countries with many co-authored papers look larger and sit nearer the center. The very high "closeness = 1" for the Cyprus-Lebanon dyad is an artifact of being a separate component; interpret centrality inside the giant component for fair comparison. Results would be even clearer with fractional counting of multi-country papers and with time-normalized edges (e.g., 2022-2025 only) to handle the post-GenAI surge. Overall, the field is modular but already well linked through a few brokers. Strengthening bridges among these modules–especially UK↔Spain↔Latin America and India/Saudi↔Europe/USA–should accelerate knowledge flow in AI-for-SE.

Figure 3 separates two core logics. In the upper-right "motor themes," **artificial intelligence**, **sustainable development**, and **sustainable education** appear together. They are both central and internally cohesive.

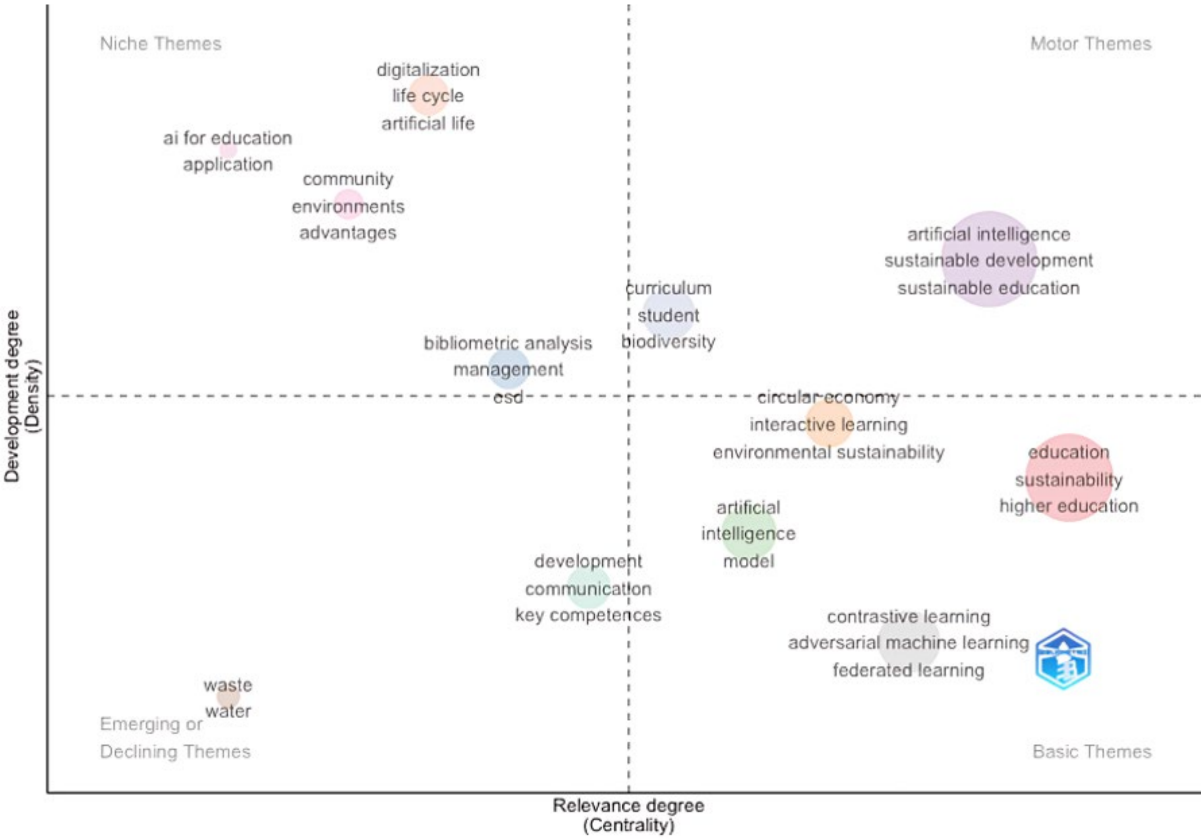

Figure 3. Thematic analysis of keywords (Created by the authors)

This indicates a mature research front where AI is explicitly framed as an enabler of ESD and sustainable development goals (SDGs). Studies in this area likely include course redesigns, assessment of AI-supported learning, and university pilots.

In the lower-right "basic themes," **education-sustainability-higher education** form a broad, central but low-density cluster. This is the backbone of the field. It ties many subtopics together but is conceptually diffuse. It confirms this pattern: *education* (59), *sustainability* (44), and *higher education* (30) have the highest occurrences and strong betweenness, yet moderate PageRank–typical for foundational, cross-cutting labels rather than tight specialties.

Also in the basic quadrant, a small technical cluster (contrastive learning, adversarial ML, and federated learning) sits low in density and centrality. These methods are present but not yet integrated with sustainability outcomes. They represent capability "plugs" that the community can mobilize for privacy, personalization, or robustness in ESD contexts.

The upper-left "niche themes" (digitalization, life cycle, artificial life; community and environments) show respectable density but weaker centrality. They are internally developed, yet peripheral to the main conversation. If you want to surface future specialties, track whether *life-cycle* and *community* topics move rightward (gaining links to the core) as ESD curricula adopt AI for systems thinking and local problem-solving.

Lower-left shows "emerging/declining" items around *waste* and *water*. In an education setting, these are surprisingly under linked. It suggests that domain-specific sustainability topics are still weakly connected to AI-for-learning research. This gap is actionable: project-based learning and data-driven labs on water/waste can become testbeds for GenAI tutors, AI-supported inquiry, and sensor/IoT data analysis.

Signals from the GenAI wave are clear but still consolidating. In **Table 6**, *ChatGPT* (19), *generative ai* (7), *large language models* (5), and *GenAI* (3) have non-trivial betweenness but lower PageRank than generic *artificial intelligence* (91). On the map, "artificial intelligence" appears both as a motor theme (when tied to SD/ESD) and as a basic enabler (when treated method-first). This split reflects a field that is rapidly adopting GenAI tools while still stabilizing concepts and outcomes.

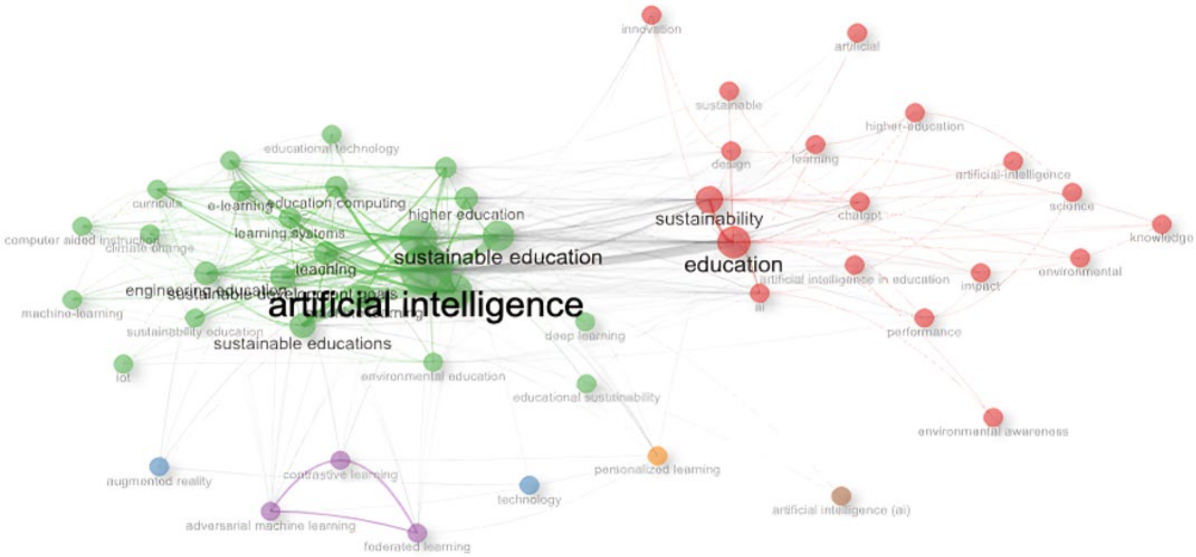

Figure 4. Network graph for keywords (Created by the authors)

Pedagogical levers such as *interactive learning*, *personalized learning*, and *learning systems* sit right of center but below the density cut-off. They are central enough to matter yet not fully developed for sustainability aims. This is a near-term opportunity: combine GenAI (feedback, scaffolding, code assistants) with ESD competencies (systems thinking, futures, values) to raise the density of these themes.

Figure 4 shows two macro-themes that meet in the middle. On the right, a red cluster centers on "education" and "sustainability," with terms like "impact," "performance," "higher-education," and "artificial intelligence in education." On the left, a green cluster centers on "artificial intelligence" and "sustainable education," surrounded by practice/discipline terms such as "environmental education," "engineering education," "curricula," "e-learning," "learning systems," and "educational technology." These two lobes are densely connected, which suggests an active convergence between AI-for-learning and education-for-sustainability.

The most influential hub in **Figure 4** is "artificial intelligence" (highest PageRank and closeness; very high betweenness). "Education" and "students" also have high betweenness. This means AI and the learner are the main bridges that link otherwise separate subtopics. In other words, research in this area is organized around applying AI to student-centered teaching and learning, more than around a single content strand.

"Sustainable education," "sustainable development," "SDGs," and "higher education" also score high on centrality. This confirms that the ESD frame–often in university settings–is the dominant educational context. "Engineering education" is well-connected, hinting that sustainability problems and AI tools are being explored through engineering curricula and projects.

Signals from the generative AI wave are visible but still consolidating. "ChatGPT" appears in the core red cluster, connected to "learning," "impact," and "performance," yet its centrality values are modest compared with the generic "artificial intelligence." This is consistent with a fast, recent uptake after 2022: GenAI is present and growing, but literature is still anchored in broader AI terms.

There are several technological side-clusters. "Augmented reality" (blue) forms a distinct pocket linked to education but weakly to sustainability. "Personalized learning" (orange) and a small methods cluster with "federated learning," "adversarial machine learning," and "contrastive learning" (purple) connect to AI more than to ESD. These look like methodological or pedagogical enablers that have not yet been deeply integrated with sustainability outcomes.

Topic variants and duplicates appear ("ai," "artificial-intelligence," "artificial intelligence (ai)"; "higher education" vs. "higher-education"; "machine learning" vs. "machine-learning"; "sustainable educations," "high educations"). These reduce metric precision and may slightly inflate fragmentation. A light thesaurus/cleanup would likely strengthen the core and raise the centrality of ESD-specific nodes.

Substantively, the picture points to three active lines:

Table 7. Most influential paper in

Paper	DOI	Total citations	TC per year	Normalized TC
Abulibdeh et al. (2024)	10.1016/j.jclepro.2023.140527	170	85.00	38.22
Rillig et al. (2023)	10.1021/acs.est.3c01106	113	37.67	7.40
Alshahrani (2023)	10.5267/j.ijdns.2023.6.010	105	35.00	6.88
Lin et al. (2023)	10.1186/s40561-023-00260-y	99	33.00	6.48
George and Wooden (2023)	10.3390/admsci13090196	94	31.33	6.16
Deng and Yu (2023)	10.3390/su15042940	89	29.67	5.83
Mokski et al. (2022)	10.1108/IJSHE-07-2021-0306	50	16.67	3.27
Arpaci and Bahari (2023)	10.1080/10494820.2022.2164313	48	24.00	10.79
Johri et al. (2024)	10.1016/j.jclepro.2024.140610	37	18.50	8.32
Shwedeh et al. (2024)	10.1007/978-3-031-52280-2_14	36	18.00	8.09
Damaševičius and Sidekerskienė (2024)	10.3390/su16052032	31	15.50	6.97
Ferk Savec and Jedrinović (2024)	10.3390/su17010183	6	6.00	13.60
Al Husaeni and Haristiani (2025)	10.17509/ajse.v5i2.87176	5	5.00	11.33
Iqbal et al. (2025)	10.1038/s41598-025-01676-x	4	4.00	9.07

(1) deploying AI to support ESD in universities, with students at the center,

(2) evaluating "impact" and "performance" of AI-supported learning, and

(3) embedding sustainability topics in computing/engineering and technology-enhanced learning (e-learning, learning systems, AR).

GenAI is an emerging connector that is beginning to tie these lines together.

Contribution of the Articles

The distribution of citations is significantly biased (**Table 7**). **Table 7** contains approximately 44% of all citations, which are attributed to the best three papers (388 of 887). The best five achieve approximately 66%. This focus suggests that a small number of early "magnet" studies are shaping the conversation. The paper (Abulibdeh et al., 2024) is an obvious mistake. It has the most citations overall (170), the most annualized impact (TC/year = 85), and the most normalized impact (38.22). The quick acceptance and influence that are inferred are supposed to go beyond the effects of publication year, as shown by the fact that all three measures are dominant. The next stratum contains many papers from 2023 (Rillig et al., 2023). Their TC/year values (approximately 30-38) suggest that the initial surge following the diffusion of generative AI in 2022 rapidly garnered attention. However, their normalized citations are negligible in comparison to those of 2024-2025 items, suggesting that 2023 is penalized by field/year normalization due to the significant number of papers from that year that received citations. Despite the low raw counts, the normalized metric emphasizes the early momentum among 2025 publications. Ferk Savec and Jedrinović (2024) (13.60), Al Husaeni and Haristiani (2025) (11.33), and Iqbal et al. (2025) (9.07) have already surpassed numerous 2023-2024 items on a per-year adjusted basis. This indicates a rapid initial adoption and implies that these papers could serve as new anchors if the current trajectory persists.

Arpaci and Bahari (2023), Johri et al. (2024), and Shwedeh et al. (2024) also perform well after normalization (≈8-11). They appear to bridge educational technology, sustainability management/production, and data/AI method domains–consistent with a cross-disciplinary field. Venue patterns are notable. Influential items are split between sustainability/environmental science journals (Journal of Cleaner Production; Environmental Science & Technology; Sustainability) and education/learning journals (Smart Learning Environments; Interactive Learning Environments; IJSHE). This confirms that impact is created both from the "sustainability science" side and the "education technology" side.

TC/year is especially helpful for recency-sensitive judgment in this fast-moving topic. It shows the speed of diffusion. Normalized TC adds a fairness layer across years and fields. In this table the two metrics together reveal a dual reality: 2023 papers built the first wave of citations, while selected 2024-2025 papers catch up quickly once year effects are removed.

The Focus of the Studies

The studies treat AI as both a means to advance sustainable education and as an object of sustainability inquiry. Conceptual framing papers position AI within long-term educational transformation agendas and quality education targets. One review stresses that AI systems are discussed for “personalized learning experiences” while also raising privacy, bias, and infrastructure challenges (Lin et al., 2023). One approach for policy and strategy initiatives to promote institutional redesign is the creation of a “smart and sustainable management system for universities” that enables the effective utilization of AI (George & Wooden, 2023). A regional case study views AI as a mechanism that connects higher education to SDG4 by enhancing social accountability and sustainability through collaborative initiatives between educational institutions and corporate entities (Ferk Savec & Jedrinović, 2024). A chapter in an academic guide asserts that “the time is ripe for a globally transformative shift within the educational framework,” suggesting that generative AI is ideally situated to facilitate this transformation (Shwedeh et al., 2024). A conceptual analysis of curricular architecture emphasizes that interdisciplinarity in ESD continues to revolve around the “three pillars” and need deliberate integration upon the introduction of AI tools (Mokski et al., 2022).

Numerous research at the learner level examine the educational impacts of AI tools. A meta-analysis of 32 experiments indicates that “chatbot technology exerted a medium-to-high effect on overall learning outcomes” across all fields and designs (Deng & Yu, 2023). A comprehensive survey and structural analysis including preservice teachers establishes a connection between the utilization of generative AI tools and academic achievements, specifically aligning with SDG4. The findings corroborate the direct and indirect connections of generative AI tool usage with academic achievement through metacognitive and cognitive offloading mechanisms (Iqbal et al., 2025). Studies in specific fields, notably ASEAN language education, deem technology-enhanced learning as “essential” for achieving inclusive and equitable outcomes for SDG4 (Al Husaeni & Haristiani, 2025). Authors caution that “students can adopt harmful behavior” if they are not instructed in critical thinking and ethical conduct in the use of AI (Alshahrani, 2023).

At the system level, the AI lens is based on tutoring and data analysis. A closer look at intelligent tutoring systems (ITS) shows that they can adapt to your needs, but they also have issues with fairness, privacy, and cost. It stresses that ITS “give students personalized learning experiences” and useful information for teachers, but they also have to deal with issues of bias and security breaches (Lin et al., 2023).

The second group of themes looks at technology that immerse people. A narrative review of virtual worlds asserts that “virtual worlds are effective in this model for hybrid education and meeting the needs, especially for disadvantaged or disabled students,” underscoring the importance of inclusive design for access and participation (Damaševičius & Sidekerskienė, 2024). A bibliometric-systematic review demonstrates that metaverse applications enhance education for sustainable development (ESD) and the SDGs through immersive learning, establishing immersive environments as venues for collaboration and the cultivation of sustainability competencies (Johri et al., 2024). A modeling study examines the determinants of the “educational sustainability of metaverse,” indicating that “hedonic motivation was the most crucial input factor,” with autonomy identified as a major predictor in both CB-SEM and deep ANN analyses (Arpaci & Bahari, 2023).

Governance, ethics, and curricular conformity are enduring focal themes. A scoping assessment at the intersection of AI, ESD, and Industry 4.0 posits that “ethics, curriculum design, continuous learning, and industry alignment come to the forefront,” highlighting deficiencies in practical techniques for integration and in evidence for classroom-level AI pedagogy (Abulibdeh et al., 2024). It also says that combining ESD and AI in higher education might “revolutionize the way we approach education” by making courses more flexible and allowing for transdisciplinary collaboration. However, it also points out the concerns of access, bias, and resistance to change (Abulibdeh et al., 2024). Institutional strategy documents align with this by suggesting university-level frameworks and change management for the proper integration of AI in sustainable education (George & Wooden, 2023).

The agenda also includes risk, responsibility, and environmental sustainability. One position paper on large language models says that the “‘energy footprint’ of LLMs must be considered” along with data-center loads and e-waste. This makes the environmental implications of AI a clear element of how decisions are made in schools (Rillig et al., 2023). Reviews also say that AI-enabled personalization creates “privacy and data security” and “bias” issues that need to be fixed to make sure that education stays fair and works overtime (Lin et al., 2023).

The set combines evidence synthesis with empirical modeling in terms of methodology. The chatbot paper uses PRISMA meta-analysis on 32 controlled experiments to measure the impact of multi-domain learning (Deng & Yu, 2023). The metaverse review combines bibliometrics and SLR to show where research on ESD and SDGs is going (Johri et al., 2024). The scoping study compiles 134 peer-reviewed articles to delineate potential, hazards, and policy strategies at the AI-ESD interface (Abulibdeh et al., 2024). Lastly, structural-equation and deep-learning models examine the psychological and behavioral precursors of sustainable learning in immersive environments (Arpaci & Bahari, 2023; Iqbal et al., 2025).

In sum, the topics coalesce around four foci. First, AI as an enabler of sustainable education through personalization, tutoring, and metacognition. Second, immersive and metaverse settings as locations for sustainable, inclusive learning design. Third, governance and curriculum alignment that bring together ethics, different fields of study, and connections with businesses. Fourth, the sustainability of AI itself, which includes its impact on the environment and the hazards of unfairness. The authors explicitly assert that generative AI tools "can enhance climate literacy" but necessitate stringent verification for errors and bias (Iqbal et al., 2025). Furthermore, the research is progressing towards institutional frameworks that facilitate the "smart and sustainable" application of AI at scale (George & Wooden, 2023).

Syntheses of the Findings

AI systems consistently improve personalization, engagement, and access–key levers for sustainable education. Across studies, AI and ITS enable "personalized learning experiences that cater to … unique learning styles and preferences" and give teachers "data-driven insights" for timely support. As the review (Lin et al., 2023) concludes, addressing privacy and bias while investing in infrastructure can "unlock the full potential of these technologies and support a more equitable and sustainable education system". In blended and hybrid formats, ChatGPT is reported to enhance motivation and self-direction through immediate feedback; the paper's conclusion states that integrating ChatGPT "has promise for promoting student engagement, motivation, and self-directed learning" (Alshahrani, 2023). The author also highlight environmental co-benefits (e.g., paper and travel reductions) as part of sustainable delivery.

Second one is related to trust (privacy/security). It is a strong predictor of educational sustainability; policy frameworks help–but may lag practice. A higher-education study in the UAE finds a significant positive path from AI adoption to educational sustainability ($\beta = 0.222$; $p < .05$) and an even stronger effect of trust (privacy/security) on sustainability ($\beta = 0.416$; $p < .05$). By contrast, "data policies and regulations exhibited an insignificant relationship" with sustainability, suggesting regulation trails adoption; however, policies moderate the effects when calibrated well (Shwedeh et al., 2024).

The third one is related to interdisciplinarity. It is essential to scale ESD; a three-pronged model is recommended, but staff capacity and leadership are barriers. The ESD literature synthesis proposes three simultaneous pillars:

(1) stand-alone, widely accessible ESD courses,

(2) embedded ESD content across existing programs, and

(3) interdisciplinary research initiatives–"all of which rely on the extensive involvement of technology and e-learning".

The paper urges this combinatory approach to reach non-STEM learners as well, who are frequently left out. Persistent barriers include limited staff expertise, low motivation, crowded curricula, and weak leadership; professors may "pay lip service" without engagement, and "structural inadequacies" impede integration (Mokski et al., 2022).

Fourt one is related to immersive and metaverse-style environments. They show potential contributions to sustainability, including education. A 2024 cleaner-production mapping of "metaverse for sustainable development" identifies a theme of "virtual technology for designing sustainable education," alongside social and economic sustainability benefits (e.g., reducing physical overheads) (Johri et al., 2024). While the field is nascent, the study calls for theory- and context-sensitive pathways so XR/VR can mature toward SDG-aligned use (Johri et al., 2024).

Design with customization in mind, but also safety. Evidence shows that AI/ITS makes things more engaging and easier to access. When these benefits are combined with strong privacy, security, and bias reduction, education becomes more sustainable (Alshahrani, 2023; Lin et al., 2023). Put governance and trust first. Trust is a big and steady thing that has an effect on sustainability results. Policies are useful when they can adapt and work together. When regulations are set in stone or not explicit, they don't work as well (Shwedeh et al., 2024). Colleges and universities should employ a three-pillar ESD plan. Include non-STEM areas in dedicated courses, curriculum integration, and research involvement, and support this with leadership, staff development, and e-learning infrastructure (Mokski et al., 2022). Consider immersive/XR as a cost-effective way to learn that fits with the SDGs. Preliminary mapping suggests that metaverse/XR could promote sustainable education and social inclusion if implemented with appropriate care (Johri et al., 2024).

The Syntheses of Recommendations For Researchers

Evidence and study design

Across the set, authors call for stronger designs, larger and more diverse samples, and longitudinal follow-ups. According to the findings of a study, it is recommended that future research should make use of "longitudinal and multi-stage designs across regions and disciplines in order to validate and extend the framework" (Iqbal et al., 2025). Another one of the calls for "Extensive studies … to examine the long-term impact of AI chatbots on student learning outcomes" is made (Alshahrani, 2023). According to Ferk Savec and Jedrinović (2024), a case study draws attention to the fact that "future studies should incorporate broader datasets to further minimize bias."

Analytics, measurement, and benchmarks

Authors ask for richer learning analytics and shared benchmarks. One paper advises: "Researchers should explore the use of learning analytics … to inform personalized instruction and adaptive learning approaches" (Alshahrani, 2023). A cleaner-production study adds that future work should "develop standardized datasets, benchmark methods, and cross-industry applications to support circular economy transitions" (Johri et al., 2024).

Pedagogy, teacher capacity, and curriculum

Several authors move beyond tools to teacher learning and curriculum standards. One paper recommends that "training programs should adopt pedagogical frameworks like the SAMR model … Micro-credential modules could equip educators" and that curricula "embed AI literacy into standards using TPACK" (Iqbal et al., 2025).

Personalization and learning components

Future studies should broaden what is evaluated beyond chat interfaces. A meta-analysis notes: "Future research could expand chatbot research by including different learning components" (Deng & Yu, 2023).

Interdisciplinarity and systems thinking

Authors stress joining technical with social and policy lenses. A review in cleaner production argues that future research should "integrate social and political sustainability … use longitudinal and multi-source designs… cross-cultural comparisons … [and] dynamic modeling of policy shifts" (Abulibdeh et al., 2024). An interdisciplinarity report suggests to "strengthen interdisciplinary approaches in educational curricula to link sustainability, digitalization, and ethics" and to test AI "with industry and civil society in real contexts" (Mokski et al., 2022). A bibliometric SDG-language study adds that "future research developments can be directed at uniting these three dimensions" of conceptual work, practice, and technical innovation (Al Husaeni & Haristiani, 2025).

Equity, access, and inclusion

Future agendas emphasize marginalized learners and basic education contexts. One synthesis identifies "research gaps that require strategic follow-up in future investigations" including the need to integrate digital inclusion “into basic language curricula and community-based programs” and to ensure “gender, cultural sensitivity, and global accessibility” (Al Husaeni & Haristiani, 2025).

Trust, explainability, and ethics

Explainable, trustworthy systems and governance are recurring needs. A review of AI-embedded tutoring systems notes that outcomes “need to be more explainable for learners and instructors to adopt AI technology” (Lin et al., 2023). An environmental viewpoint warns that misuse could flood the public sphere with misinformation: “We think this is the biggest concern of the more widespread use of LLMs for environmentally relevant topics” (Rillig et al., 2023). The blended-learning paper also calls for “further analysis of the ethical issues … [and] potential biases in AI-generated content” (Alshahrani, 2023).

Technical portability, data, and infrastructure

Several studies list technical gaps that guide future work. The ITS review’s limitation table highlights needs such as “more learner’s participation for sampling good quality data,” “features extraction … not fully representing the learning process,” and portability across systems, alongside “privacy” and “trustworthy” concerns (Lin et al., 2023). Authors studying virtual worlds likewise conclude there is a “need for further research and development to fully realize their potential in educational contexts” (Damaševičius & Sidekerskienė, 2024).

Institutional data ecosystems and knowledge graphs

One conceptual blueprint outlines concrete infrastructure tasks for future work: “cross-checking information and justifying the classification, building a knowledge graph for campus data … and teaching a robot how to handle classification decisions” (George & Wooden, 2023).

Metaverse and immersive environments

For immersive learning, authors point to stronger causal evidence and beyond self-reports. A metaverse study notes that “the use of self-report data was a critical limitation … [and] because of the cross-sectional nature … causal inferences cannot be drawn” (Arpaci & Bahari, 2023). This directly motivates experimental and longitudinal designs in future research.

Programmatic research agendas that link AI and SDGs

A university case study proposes a broad SDG-aware agenda: future research should “include more detailed evaluations in several faculties,” develop “quantitative instruments with statistical validation,” study “inclusive AI in education … to mitigate digital inequalities,” and analyze “cost-effectiveness … to provide guidelines for decision-makers” (Ferk Savec & Jedrinović, 2024).

DISCUSSION

This study analyzes 417 papers published between 2022 and 2025 at the intersection of SE and AI. We use a bibliometric approach to map rapid growth and a complex structure. Our findings match key patterns in recent reviews and also offer original contributions. We describe key clusters, leading venues, and recurring keywords in clear terms. We also point to links among topics across years so that readers can see how the field moves.

The growth in publications is clear. It aligns with recent work. Deng and Yu (2023) report a sharp rise in chatbot use in education after 2020. Delen et al. (2024) show exponential growth of AI in education since 2010, with a strong jump in 2020-2022. These patterns suggest two drivers: the digital transformation during COVID-19 and the spread of generative AI after 2022, including tools like ChatGPT. The trend appears in different regions and in different journal types. It also appears in both computer science and education venues, which suggests broad adoption.

Our focus on the post 2022 period adds a distinct contribution. Kavitha and Joshith (2024) identify a peak in 2023. This matches our data and supports the transformative effect of generative AI in education. The dominance of “ChatGPT” and “generative AI” as keywords signals a strong shift in the field. The shift shows up in titles, abstracts, and author chosen keywords. It also shows up in the kinds of tools that classrooms try first, such as conversational tutors and automated feedback. Additionally, studies examining the pedagogical aspects of SE indicate that there are still gaps in transferring technological innovations and content knowledge into classroom practice. A systematic review of STEAM teachers reports that both in-service and pre-service teachers struggle to connect theoretical knowledge with real-world applications (Álvarez Ariza & Olatunde-Aiyedun, 2024). This finding explains why the initial adoption areas of generative AI tools are often "low-threshold" uses such as guidance, explanation, and automated feedback.

Geographic patterns in our data are consistent with prior research. Bond et al. (2024) highlight a concentration of publications in North America and the value of team collaborations. We observe a similar Western weight. At the same time, contributions from Saudi Arabia, India, and Malaysia grow. This points to a global shift. It reflects the rising importance of SE in the Global South and ongoing efforts in digital transformation. It also suggests new research partnerships and context sensitive designs, because local needs differ. These regional differences also show that perceptions of SE are still tied to institutional and cultural boundaries in some contexts. For example, a study on school administrators in Greece reveals that the sustainability mission of modern schools is still largely determined by social expectations and institutional opportunities (Doulami & Dimitriou, 2025). This supports the notion that, despite the increase in global publications, practices vary greatly in local contexts.

Institutional collaboration is still limited. Sahar and Munawaroh (2025) show leadership by Symbiosis International Deemed University and the Bucharest University of Economic Studies. Our results also point to fragmentation in institutional networks. This indicates a need for stronger interdisciplinary work and broader partnerships. Shared datasets and open evaluation tasks could help. Joint labs that include educators, data scientists, and policy makers could also help.

The variety of AI applications in SE is large. Lin et al. (2023) underline the role of ITS in personalized learning and data driven decisions. Our study offers a holistic frame that fits this literature. We propose three principles: personalization-protection, competence alignment, and multi-level synchronization. These principles connect AI uses to SE goals in a simple and practical way. Below we give short explanations and examples for each.

Personalization-protection means that tailored learning should respect privacy and safety. For example, a tutor can adapt tasks to a student's level, but the system should minimize data collection and protect identity. Logs should be stored for short periods. Students should see why the tool suggests. Competence alignment means that tool use should match target SE competencies, such as systems thinking and action competence. For example, a simulation can build systems thinking if it makes feedback loops visible. Multi-level synchronization means that change should line up across course, program, and institution levels. For example, a course pilot should have program level support and basic rules for assessment and integrity.

A recent systematic review (Greif et al., 2024) reports the distribution of AI techniques in education as follows: 65% supervised learning, 18% unsupervised learning, and 17% reinforcement learning. It also notes that artificial neural networks are the most common single technique (23%). These figures support our technical landscape and show why method reporting matters. Shares may differ by subfield and year. Many studies also combine methods, so clear reporting helps readers compare results.

One of our original contributions is to address the environmental footprint of AI within SE. A comprehensive review (Leal Filho et al., 2025) shows a clear paradox. AI can support climate action and resource management. Yet AI systems also have their own ecological footprint and a risk of bias. This paradox strengthens the case for our "protection" principle. Energy demand comes from training, inference, data storage, and cooling. Bias risks come from data sources and from design choices. Running large AI models has notable environmental costs. These costs can conflict with sustainability goals unless systems use renewable energy. This supports our call to integrate environmental metrics into SE design and AI adoption. Courses can include simple carbon estimators for assignments. Institutions can prefer efficient models, schedule compute in greener grids, and audit energy use. Environmental awareness in SE encompasses not only the use of technology but also teachers' perceptions. The study examining the science capital approach emphasizes that developing sustainability awareness in students is closely linked to teachers' skills in modeling, contextualizing, and relating to everyday life (Fischer et al., 2024). This finding indicates that discussions about the ecological footprint of AI should also extend to the dimension of teacher education.

The evaluation of AI supported learning outcomes in SE is still developing. Abulibdeh et al. (2024) discuss challenges, opportunities, and ethics at the intersection of AI and ESD. Our findings are similar. We see few causal or longitudinal designs in the current work. This gap is important for future research. Multi-site trials and preregistered studies would help. Shared rubrics and public datasets would also improve comparability.

Insights from Yang et al. (2025) map ten years of AI literacy research. Core themes persist machine learning, AI literacy, accountability, and generative AI. These themes overlap with the competence areas we identify. They also suggest how AI tools can build systems thinking and action competence in SE. In practice, students can analyze model errors, read model cards, and debate tradeoffs. These tasks build both technical sense and ethical awareness.

Our "multi-level synchronization" principle aligns with the whole institution approach. Kulkov et al. (2023) show that AI supports sustainability through three domains: organizational, technical, and operational. The organizational domain focuses on integration in companies and industries, adoption barriers, and stakeholder relations. This supports our claim that the policy-practice gap must close. Clear rules, simple guidance for staff, and training plans can reduce that gap.

The literature also supports our directions for future research. Liang et al. (2025) review early impacts of AI in higher education. They see opportunities for personalization and instant feedback. They also warn about academic integrity and the spread of biased information. This dual pattern mirrors our balance of opportunities and risks. It calls for mixed methods designs that track learning outcomes and integrity outcomes at the same time.

Rohde et al. (2024) work on criteria and indicators for sustainable AI. They call for standardized measurement tools. Our results agree. We find a shortage of validated scales for SE, including systems thinking, action competence, and ethics/AI literacy. Checklists and reporting standards could help. They would also lower the cost of replication and meta-analysis.

This study contributes to SE theory. We provide empirical evidence that AI can reshape traditional pedagogies. AI can enhance learning processes in higher education (Caleb et al., 2023; Siminto et al., 2023). Our framework adds how these gains link to SE competencies in a structured way. The mapping from tool types to competencies can guide course design and assessment.

There are also clear practical implications. Educators and policy makers can use our three principles as guidance. The personalization-protection balance helps institutions adopt AI while guarding privacy and ethics. Competence alignment helps instructors connect AI tools to pedagogical goals. Multi-level synchronization supports integration at course, program, and institution levels. A simple path is to start with a pilot, measure learning and integrity, report energy use, and then scale.

In sum, our study offers a map of a critical moment in the SE-AI space. The field grows fast, but it still needs systematic approaches to create pedagogical value. Our three principles offer a clear path forward. The comparisons with the literature show both alignment and originality. The integration of environmental footprints and ethics is essential to address the paradox of AI in SE. Future work should test the three principles empirically and in diverse contexts. We also encourage cross cultural studies, equity focused outcomes, and cost studies that track benefits and tradeoffs over time.

CONCLUSION

This study maps 417 papers published between 2022 and 2025 on SE and AI. We find fast growth after 2022, driven by digital transformation and the spread of generative AI; "ChatGPT" and "generative AI" become dominant keywords. Output is still concentrated in Western countries, while contributions from Saudi Arabia, India, and Malaysia are rising. Institutional networks remain fragmented. Across the corpus, AI tools support personalization, timely feedback, and data driven instruction, yet concerns about privacy, academic integrity, and environmental cost also grow. Our analysis further shows that supervised learning is the most common approach and that neural networks are the most used single technique. To organize these results for practice, we offer three principles – personalization-protection, competence alignment, and multi-level synchronization – that link AI uses to SE goals and help coordinate change at the course, program, and institution levels.

Contribution to the Literature

This work contributes to the literature in five ways. First, it offers a focused, post 2022 view of the period when generative AI entered mainstream education, with a simple map of clusters, venues, and keyword shifts. Second, it distills scattered findings into three actionable principles that connect AI tools to SE competencies such as systems thinking and action competence, making design and evaluation easier. Third, it adopts a dual sustainability lens by integrating environmental footprints and ethics, treating AI both as a tool for sustainability and as an object of sustainability. Fourth, it identifies key gaps–a shortage of causal and longitudinal designs and a lack of validated SE measurement scales–and calls for shared rubrics, reporting checklists, and open datasets. Fifth, it documents signals of a global shift through increased participation from the Global South, opening new contexts and equity questions for future work.

Recommendations

For researchers, we recommend greater use of causal and longitudinal designs where feasible, along with preregistration and open analysis code. Methods should be reported transparently, including model classes, datasets, prompts, baselines, and at least minimal energy or carbon estimates. Validated SE scales for systems thinking, action competence, and AI ethics or literacy should be developed and shared, and open datasets and common tasks for SE-AI should be created to support replication and meta-analysis. Research should also address equity, access, and cost through partnerships that include Global South institutions and context specific constraints.

For educators and institutions, we suggest starting with small pilots, measuring learning, integrity, and student experience, and then scaling successful practices. Personalization-protection should guide adoption by minimizing data retention, building privacy by design, and explaining model outputs to students. Competence alignment should ensure that each tool maps explicit SE outcomes–for example, using simulations that make feedback loops visible to build systems thinking. Multi-level synchronization should align course policies, program outcomes, and institutional rules and be supported through staff training. When possible, institutions should prefer efficient tools, audit energy use, and schedule compute in greener grids.

For policy makers and funders, we advise issuing clear guidance on privacy, integrity, transparency, and energy reporting in education; investing in standardized measures and shared evaluation tasks for SE-AI; and supporting capacity building in teacher education and in under resourced regions so that adoption is both responsible and inclusive.

Limitations

This analysis covers 2022-2025 and draws on a bibliometric sample with a strong English language focus; database coverage, indexing biases, and keyword choices may exclude relevant work, and some misclassification may remain despite data cleaning. Reported technique shares are aggregated across subfields and years and should be read as indicative, not definitive. Because the evidence reflects published literature, it may overrepresent successful or well-resourced projects. As citation windows mature, specific patterns may shift, but the main directions we report are likely to hold.

Author contributions: ARM: conceptualization, formal analysis, methodology, writing – original draft; **YPK:** conceptualization, data curation, methodology, writing – review & editing; **AIB:** data curation, methodology, writing – review & editing; **OVP:** data curation, formal analysis, writing – original draft; **NLS:** conceptualization, writing – original draft, writing – review & editing; **MAK:** formal analysis, methodology, writing – review & editing. All authors approved the final version of the article.

Funding: The authors received no financial support for the research and/or authorship of this article.

Ethics declaration: This study is based on the analysis of scientific studies already published in the literature using bibliometric methods. The data used in the study were obtained from secondary sources (Web of Science and Scopus) and do not involve any application, survey, or experiment on humans or animals. Therefore, due to the nature of the study, it does not require ethics committee approval/permission. Full compliance with the 'Higher Education Institutions Scientific Research and Publication Ethics Guidelines' was ensured throughout the study process.

Declaration of interest: The authors declared no competing interest.
Data availability: Data generated or analyzed during this study are available from the authors on request.

REFERENCES

Abulibdeh, A., Zaidan, E., & Abulibdeh, R. (2024). Navigating the confluence of artificial intelligence and education for sustainable development in the era of industry 4.0: Challenges, opportunities, and ethical dimensions. *Journal of Cleaner Production, 437*, Article 140527. https://doi.org/10.1016/j.jclepro.2023.140527

Al Husaeni, D. N., & Haristiani, N. (2025). What evidence supports the advancement of language learning through digital innovation? Toward achieving sustainable development goals (SDGs) in the 21st century completed with bibliometric analysis. *ASEAN Journal of Science and Engineering, 5*(2), 327-356. https://doi.org/10.17509/ajse.v5i2.87176

Alshahrani, A. (2023). The impact of ChatGPT on blended learning: Current trends and future research directions. *International Journal of Data and Network Science, 7*(4), 2029-2040. https://doi.org/10.5267/j.ijdns.2023.6.010

Álvarez Ariza, J., & Olatunde-Aiyedun, T. G. (2024). A systematic literature review on STEAM pre- and in-service teacher education for sustainability: Are teachers ready? *Eurasia Journal of Mathematics, Science and Technology Education, 20*(9), Article em2498. https://doi.org/10.29333/ejmste/14982

Arpaci, I., & Bahari, M. (2023). Investigating the role of psychological needs in predicting the educational sustainability of Metaverse using a deep learning-based hybrid SEM-ANN technique. *Interactive Learning Environments, 32*(4), 2957-2969. https://doi.org/10.1080/10494820.2022.2164313

Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education, 21*, Article 4. https://doi.org/10.1186/s41239-023-00436-z

Caleb, K., Lenny, W., Dazzy, I., & Azra, E. (2023). The impact of A.I on teaching and learning. *London Journal of Social Sciences*, (6), 124-129. https://doi.org/10.31039/ljss.2023.6.111

Corres, A., Rieckmann, M., Espasa, A., & Ruiz-Mallén, I. (2020). Educator competences in sustainability education: A systematic review of frameworks. *Sustainability, 12*(23), Article 9858. https://doi.org/10.3390/su12239858

Damaševičius, R., & Sidekerskienė, T. (2024). Virtual worlds for learning in metaverse: A narrative review. *Sustainability, 16*(5), Article 2032. https://doi.org/10.3390/su16052032

Delen, I., Sen, N., Ozudogru, F., & Biasutti, M. (2024). Understanding the growth of artificial intelligence in educational research through bibliometric analysis. *Sustainability, 16*(16), Article 6724. https://doi.org/10.3390/su16166724

Deng, X., & Yu, Z. (2023). A meta-analysis and systematic review of the effect of chatbot technology use in sustainable education. *Sustainability, 15*(4), Article 2940. https://doi.org/10.3390/su15042940

Doulami, E., & Dimitriou, A. (2025). Exploring the perceptions of primary education executives in Greece on the role of the modern school in promoting environmental and sustainability education. *Interdisciplinary Journal of Environmental and Science Education, 21*(1), Article e2504. https://doi.org/10.29333/ijese/15812

Ferk Savec, V., & Jedrinović, S. (2024). The role of AI implementation in higher education in achieving the sustainable development goals: A case study from Slovenia. *Sustainability, 17*(1), Article 183. https://doi.org/10.3390/su17010183

Fischer, A., Havu-Nuutinen, S., Kontkanen, S., & Suortti, E. (2024). Investigating sustainability education in the science capital teaching approach: Competence development and pillar considerations. *Interdisciplinary Journal of Environmental and Science Education, 20*(4), Article e2418. https://doi.org/10.29333/ijese/15038

George, B., & Wooden, O. (2023). Managing the strategic transformation of higher education through artificial intelligence. *Administrative Sciences, 13*(9), Article 196. https://doi.org/10.3390/admsci13090196

Greif, L., Röckel, F., Kimmig, A., & Ovtcharova, J. (2024). A systematic review of current AI techniques used in the context of the SDGs. *International Journal of Environmental Research, 19*, Article 1. https://doi.org/10.1007/s41742-024-00668-5

Holst, J., Grund, J., & Brock, A. (2024). Whole institution approach: Measurable and highly effective in empowering learners and educators for sustainability. *Sustainability Science, 19*(4), 1359-1376. https://doi.org/10.1007/s11625-024-01506-5

Iqbal, J., Hashmi, Z. F., Asghar, M. Z., & Abid, M. N. (2025). Generative AI tool use enhances academic achievement in sustainable education through shared metacognition and cognitive offloading among preservice teachers. *Scientific Reports, 15*, Article 16610. https://doi.org/10.1038/s41598-025-01676-x

Johri, A., Joshi, P., Kumar, S., & Joshi, G. (2024). Metaverse for sustainable development in a bibliometric analysis and systematic literature review. *Journal of Cleaner Production, 435*, Article 140610. https://doi.org/10.1016/j.jclepro.2024.140610

Kavitha, K., & Joshith, V. P. (2024). The transformative trajectory of artificial intelligence in education: The two decades of bibliometric retrospect. *Journal of Educational Technology Systems, 52*(3), 376-405. https://doi.org/10.1177/00472395241231815

Kulkov, I., Kulkova, J., Rohrbeck, R., Menvielle, L., Kaartemo, V., & Makkonen, H. (2023). Artificial intelligence-driven sustainable development: Examining organizational, technical, and processing approaches to achieving global goals. *Sustainable Development, 32*(3), 2253-2267. https://doi.org/10.1002/sd.2773

Leal Filho, W., Kim, E., Borsatto, J. M. L. S., & Marcolin, C. B. (2025). Using artificial intelligence in sustainability teaching and learning. *Environmental Sciences Europe, 37*, Article 124. https://doi.org/10.1186/s12302-025-01159-w

Liang, J., Stephens, J. M., & Brown, G. T. L. (2025). A systematic review of the early impact of artificial intelligence on higher education curriculum, instruction, and assessment. *Frontiers in Education, 10*. https://doi.org/10.3389/feduc.2025.1522841

Lin, C.-C., Huang, A. Y. Q., & Lu, O. H. T. (2023). Artificial intelligence in intelligent tutoring systems toward sustainable education: A systematic review. *Smart Learning Environments, 10*, Article 41. https://doi.org/10.1186/s40561-023-00260-y

Mokski, E., Leal Filho, W., Sehnem, S., & Andrade Guerra, J. B. S. O. d. (2022). Education for sustainable development in higher education institutions: An approach for effective interdisciplinarity. *International Journal of Sustainability in Higher Education, 24*(1), 96-117. https://doi.org/10.1108/ijshe-07-2021-0306

Rieckmann, M. (2012). Future-oriented higher education: Which key competencies should be fostered through university teaching and learning? *Futures, 44*(2), 127-135. https://doi.org/10.1016/j.futures.2011.09.005

Rillig, M. C., Agerstrand, M., Bi, M., Gould, K. A., & Sauerland, U. (2023). Risks and benefits of large language models for the environment. *Environmental Science & Technology, 57*(9), 3464-3466. https://doi.org/10.1021/acs.est.3c01106

Rohde, F., Wagner, J., Meyer, A., Reinhard, P., Voss, M., Petschow, U., & Mollen, A. (2024). Broadening the perspective for sustainable artificial intelligence: Sustainability criteria and indicators for Artificial Intelligence systems. *Current Opinion in Environmental Sustainability, 66*, Article 101411. https://doi.org/10.1016/j.cosust.2023.101411

Rosen, M. A. (2025). Artificial intelligence and sustainable development. *European Journal of Sustainable Development Research, 9*(1), Article em0275. https://doi.org/10.29333/ejosdr/15819

Sahar, R., & Munawaroh, M. (2025). Artificial intelligence in higher education with bibliometric and content analysis for future research agenda. *Discover Sustainability, 6*, Article 401. https://doi.org/10.1007/s43621-025-01086-z

Sandri, O. (2020). What do we mean by 'pedagogy' in sustainability education? *Teaching in Higher Education, 27*(1), 114-129. https://doi.org/10.1080/13562517.2019.1699528

Shwedeh, F., Salloum, S. A., Aburayya, A., Fatin, B., Elbadawi, M. A., Al Ghurabli, Z., & Al Dabbagh, T. (2024). AI adoption and educational sustainability in higher education in the UAE. In A. Al-Marzouqi, S. A. Salloum, M. Al-Saidat, A. Aburayya, & B. Gupta (Eds.), *Artificial intelligence in education: The power and dangers of ChatGPT in the classroom* (pp. 201-229). Springer. https://doi.org/10.1007/978-3-031-52280-2_14

Siminto, S., Akib, A., Hasmirati, H., & Widianto, D. S. (2023). Educational management innovation by utilizing artificial intelligence in higher education. *Al-Fikrah: Jurnal Manajemen Pendidikan, 11*(2), 284-296. https://doi.org/10.31958/jaf.v11i2.11860

Usak, M., Hsieh, M. Y., & Chan, Y.-K. (2021). A concretizing research on making higher-education sustainability count. *Sustainability, 13*(5), Article 2724. https://doi.org/10.3390/su13052724

Velazquez, L., Munguia, N., Platt, A., & Taddei, J. (2006). Sustainable university: What can be the matter? *Journal of Cleaner Production, 14*(9-11), 810-819. https://doi.org/10.1016/j.jclepro.2005.12.008

Wiek, A., Withycombe, L., & Redman, C. L. (2011). Key competencies in sustainability: A reference framework for academic program development. *Sustainability Science, 6*(2), 203-218. https://doi.org/10.1007/s11625-011-0132-6

Yang, Y., Zhang, Y., Sun, D., He, W., & Wei, Y. (2025). Navigating the landscape of AI literacy education: Insights from a decade of research (2014-2024). *Humanities and Social Sciences Communications, 12*, Article 374. https://doi.org/10.1057/s41599-025-04583-8