Vol. 7, No. 1, 2026, pp. 859-870
DOI: https://doi.org/10.29210/08jces696600

Contents lists available at Journal ELORA

Journal of Counseling, Education and Society
ISSN: 2716-4896 (Print) , ISSN 2716-4888 (Electronic)

Journal homepage: https://jces.eloracenter.org/jces

Human-centered leadership for artificial intelligence integration in k-12 and higher education: a systematic review

Bambang Styawan*), Ilham Ilham, Imran Imran, Azainil Azainil, Usfandi Haryaka
University of Mulawarman, Samarinda, Indonesia

Article Info

Article history:

Received Sept 12th, 2026
Revised Oct 20th, 2026
Accepted Nov 26th, 2026

Keyword:

Artificial intelligence,
Educational leadership,
Human-centered leadership,
School principals,
Systematic review

ABSTRACT

This systematic review examined how school and higher education leaders integrated artificial intelligence (AI) into their work and which human-centered principles guided that integration. We followed the PRISMA 2020 statement and searched Scopus in October 2026. Eligible reports were peer-reviewed empirical studies published in English between 2020 and 2026, in which formal education leaders were the main participants or a core variable and AI was examined in relation to leadership roles, practices, competencies, or governance. An AI model proposed title-abstract decisions, and the authors verified every proposal and made all final decisions. We appraised quality with the Mixed Methods Appraisal Tool and used thematic synthesis with hybrid coding: human-centered AI principles served as deductive codes, and leadership practices emerged inductively. Of 214 records, 25 reports from 24 studies were included, and 14 were rated high quality. Most studies examined leaders' perceptions of generative AI, and only eight reported actual implementation. AI use shifted leadership toward four practices: reallocating routine work, critically evaluating AI output, enabling teachers' AI use, and governing ethical and legal risks. Equity, human agency, and ethics were the principles cited most often, yet few studies described how leaders enacted them. Competence gaps, unclear policy, and weak infrastructure were the main constraints. The resulting framework places educational purpose and human judgment at its core, links each practice to a principle, and rests on four enabling conditions: leader-specific training, infrastructure, clear policy, and an innovation culture. Because most evidence is perception-based, the framework offers testable propositions rather than established effects.

© 2026 The Authors.
This is an open access article under the CC BY-NC-SA license (https://creativecommons.org/licenses/by-nc-sa/4.0)

Corresponding Author:

Bambang Styawan,
University of Mulawarman, Samarinda, Indonesia
Email : bambangstyawan3017@gmail.com

Introduction

Artificial intelligence (AI), and generative AI in particular, is now part of everyday work in schools and universities. Most research on AI in education still concentrates on learning, for example intelligent tutoring systems, automated assessment, and personalized learning. An early review of AI in higher education noted that educators' perspectives were almost absent from this literature (Zawacki-Richter et al., 2019). International guidance has since called for a human-centered approach in which institutions, not the technology, set the purposes and limits of AI use (UNESCO, 2023). Turning this principle into practice in schools and universities is, at its core, a leadership task.

Educational leaders decide whether and how AI enters their institutions. They allocate resources, set rules, model use, support staff, and judge whether a new technology is ethically acceptable. Conceptual work describes the relationship between leaders and AI as human–machine collaboration: AI can speed up data-based decisions, but it can also carry biases that leaders must recognize and correct (Y. Wang, 2021a, 2021b). Dai, Thomas, and Rawolle (2025) separate the informational role of AI from the leader's role in setting a vision, managing conflict, and distributing authority. Other authors stress the moral side of leadership and propose ethical principles and leadership domains for responsible AI integration (Ali, 2026; DeMatthews et al., 2026). These ideas are consistent with human-centered AI, which calls for high levels of automation to be matched by equally high levels of human control (Shneiderman, 2020).

Two constructs guide this review. Educational leadership refers to the work of people who hold a formal leadership position in a school or higher education institution, such as principals, deputy principals, middle leaders, deans, and heads of department, and of teacher leaders when leadership is the focus of the study. Their core practices are setting direction, developing people, and managing the organization (Leithwood et al., 2008). AI integration refers to the use of an AI tool or system in that leadership work, either directly by the leader or through the decisions a leader makes about AI use by staff. We traced both constructs into the eligibility criteria (Table 1) and into the extraction form, where AI integration was coded by type of AI, domain of use, and adoption stage.

Several reviews have mapped the field, but each leaves a specific question open. Bibliometric studies describe publication trends and thematic clusters (Arar et al., 2025, 2026) but do not report what leaders actually do with AI. Bellibaş and Karaferye (2026) synthesized ethical issues, implementation challenges, and practical use, and Mifsud et al. (2026) mapped policy enactment, but both reviews covered K-12 school leadership only. Other reviews took a single lens: ethical leadership (Ho, 2026), capacity building (Psyrras & Kafa, 2026), or professional capital (Ho & MacGregor, 2026). As far as we could determine, no review has (a) synthesized empirical evidence across both schools and higher education, (b) used human-centered AI principles as an explicit coding frame, and (c) separated evidence on implemented AI from evidence on perceptions. Without this synthesis, we cannot tell whether the leadership changes reported in the literature rest on observed practice, or which human-centered principles leaders turn into procedures. The evidence also grows quickly: 171 of the 214 records retrieved for this review were published in 2025 or 2026.

We chose a systematic review because the gap concerns what the existing empirical record shows across countries, educational levels, and adoption stages. A single survey or case study cannot make that comparison. The review aims to describe how AI is integrated into educational leadership, to explain how this integration changes leadership practice, and to identify which human-centered principles are enacted rather than only stated. It addresses four research questions (RQs): (1) How is AI integrated into educational leadership across contexts (RQ1)? (2) How does AI integration change the roles, practices, and competencies of educational leaders (RQ2)? (3) Which human-centered principles, opportunities, and challenges shape this integration (RQ3)? (4) How can the evidence be synthesized into a framework for human-centered educational leadership in AI integration (RQ4)? For research, the review turns a fragmented literature into a framework with testable propositions. For practice, it gives school principals, academic leaders, and policymakers a structured account of how to integrate AI while keeping human judgment at the center of leadership.

Method

This study is a systematic literature review reported in line with the PRISMA 2020 statement (Page et al., 2021). The review question was framed with the PICo framework, which suits syntheses of qualitative and mixed evidence (Lockwood et al., 2015). The population was educational leaders at all levels; the phenomenon of interest was the integration of AI into leadership roles, practices, competencies, decision-making, or governance; and the context was formal education from K-12 schools to higher education.

The search was run in Scopus on 20 October 2026. Scopus was chosen because it indexes the main journals in educational leadership and educational technology and provides complete metadata for screening. Two blocks of terms were combined with the AND operator: AI terms (e.g., "artificial intelligence", "generative AI", "ChatGPT") and educational leadership phrases (e.g., "school leader*", "school principal*", "academic leader*", "educational management"). Both blocks were restricted to titles and keywords so that the retrieved studies treated AI and leadership as their main topics. The search was limited to English-language journal articles published between 2020 and 2026. The string was refined over five iterations, from 1,124 to 214 records, to balance sensitivity and precision; the final string is given in Appendix A.

The eligibility criteria are listed in Table 1. Selection took place in two stages: title-abstract screening and full-text assessment. At the title-abstract stage, an AI model (Claude, Anthropic) received the criteria in Table 1 and the title, abstract, and keywords of each record, and returned a proposed decision with a short reason. The model made no final decision. The authors checked every proposal against Table 1 and made all final decisions. This procedure reflects evidence that large language models can support screening but still need human verification (Khraisha et al., 2024). We did not use a second independent screener, so inter-rater reliability statistics are not available; we discuss this limitation at the end of the Results and Discussions section. The decision rule for unclear records was explicit: when a record met the context and design criteria but its population or phenomenon could not be judged from the abstract, it moved to full-text assessment rather than being excluded. At the full-text stage, a report was included only if leaders' data were reported separately and AI was analyzed in relation to at least one leadership role, practice, competency, decision, or governance task. Studies that measured leadership through teacher ratings, or that examined teacher leaders without a formal position, were retained because leadership was a core variable or focus of the study, in line with the population criterion in Table 1. Two reports, Berkovich (2026) and Berkovich (2025), drew on the same survey of 302 school leaders in Israel and were therefore counted as one study in the synthesis.

Table 1. Inclusion and exclusion criteria

Criterion	Inclusion	Exclusion
Population	Educational leaders are the main participants, the unit of analysis, or a core variable	Teachers, lecturers, or students only; leadership is a peripheral factor
Phenomenon	AI is examined in relation to leadership roles, practices, competencies, decisions, or governance	AI only as a learning tool or analytic method; system development without analysis of leadership
Context	Formal K-12, vocational, and higher education	Corporate, workplace, or non-educational settings
Design	Quantitative, qualitative, or mixed-methods empirical studies	Conceptual articles, commentaries, essays, and policy briefs
Publication	Peer-reviewed journal articles	Reviews, meta-syntheses, and bibliometric analyses
Language and year	English; published 2020–2026	Other languages; published before 2020

Data were extracted into a 28-field form mapped to the four RQs: study identification, context and design, type and domain of AI use, adoption stage, changes in roles, practices, and competencies, human-centered principles, opportunities and challenges, and key findings. The same AI model assisted with extraction, and the authors checked every entry against the full text. Methodological quality was appraised with the Mixed Methods Appraisal Tool (MMAT), version 2018 (Hong et al., 2018). Results are reported as the number of criteria met (0–5): 14 reports were rated high (4–5), 4 moderate (3), and 7 low (≤2). No study was excluded on quality grounds. Instead, each theme was checked to see whether it was still supported once the low-quality reports were set aside.

The synthesis had three steps. First, descriptive statistics summarized the distribution of studies by year, region, educational level, type of AI, domain of use, and adoption stage (RQ1). Second, thematic synthesis followed the three stages described by Thomas and Harden (2008): line-by-line coding of findings, development of descriptive themes, and generation of analytical themes (RQ2 and RQ3). Coding was hybrid. The principles of human-centered AI, namely human control, transparency, fairness, and well-being, served as deductive codes (Shneiderman, 2020; UNESCO, 2023), while leadership practices, competencies, opportunities, and challenges were coded inductively. Every retained theme was supported by at least three reports. We rated confidence in a theme as stronger when at least five high-quality reports supported it and as moderate when fewer did. Third, the analytical themes were combined into a framework (RQ4), and each component was traced back to the reports that support it.

Results and Discussions

Of the 214 records, 162 were excluded at the title-abstract stage, mainly because AI was not linked to leadership work (56), leaders were not the focus of the study (44), or the study was not empirical (40). The remaining 22 were reviews or bibliometric studies (20) or were not set in an educational context (2). Of the 52 reports sought in full text, 25 could not be retrieved. Of the 27 reports assessed, one was excluded because leader and teacher data were pooled without separation, and one because it was written in Spanish. The final corpus comprised 25 reports from 24 studies (Figure 1). Each report was assigned a code from S25 to S52, detailed in Appendix B; these codes are used in Table 2, Table 3, and Figure 3.

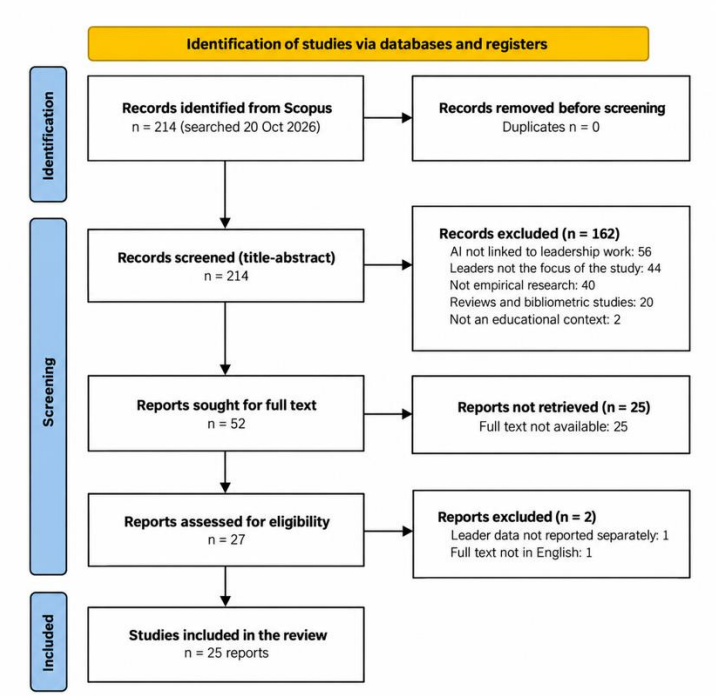

Figure 1. PRISMA 2020 flow diagram of study selection

Regarding RQ1, the evidence on AI in educational leadership is very recent and concentrated. Of the 25 reports, 24 were published in 2025–2026, 18 came from K-12 settings, and 16 examined generative AI (Figure 2). The studies came from the Middle East (9 reports), Asia (8), North America (3), Europe (3), and Africa (1), plus one survey spanning 12 countries (Eryilmaz et al., 2025). The designs comprised 11 qualitative, 9 quantitative, and 5 mixed-methods studies. Most important for interpreting the results, only 8 reports examined AI that had actually been implemented; 13 were at the exploration stage and 4 at the awareness stage. In other words, most of the evidence describes leaders' perceptions, readiness, and intentions rather than observed practice.

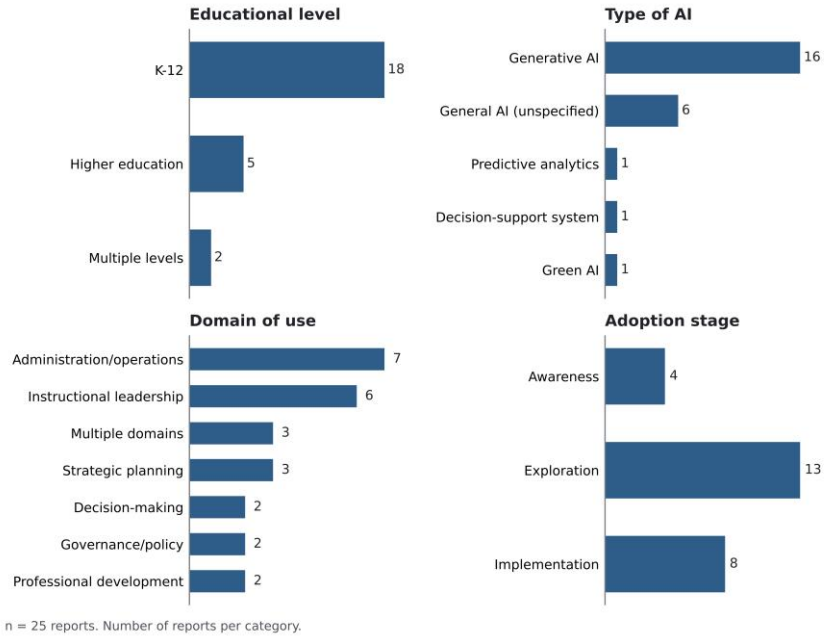

Figure 2. Characteristics of the 25 included reports by educational level, type of AI, domain of use, and adoption stage

AI integration took three forms. The first and most common was generative AI used as a personal work tool for drafting documents, reports, and communications (Flogie et al., 2026; Salvatierra, 2026; Sterrett et al., 2026). The

second was institutional AI systems, such as an AI-based teaching management system in China (Tahir et al., 2026), green AI for university sustainability governance (Salem & Khalil, 2026), and predictive analytics for identifying leadership talent in Oman (AL-Housni et al., 2026). The third was AI in teachers' practice under the direction of leaders (Chassida, 2026; Ghamrawi, Shal, et al., 2026; Hsieh et al., 2026). The adoption pattern was similar across contexts: AI entered low-risk administrative tasks first, while integration into teaching and research moved more slowly (Flogie et al., 2026; X. Wang et al., 2025). The exception was a survey of school leaders in Israel, which found that the instructional domain made the most use of generative AI, for example in preparing educational programs (60.3%) and lesson plans for teams (57.3%) (Berkovich, 2025).

Context also shaped how leaders made sense of AI. In higher education, AI adoption was treated as a matter of institutional governance and strategy (Mohasseb et al., 2026; Salem & Khalil, 2026; Selmi Arrooqi & Miqad Alruqi, 2025; X. Wang et al., 2025), whereas in K-12 the emphasis was on individual use and support for teachers. Differences in infrastructure created a clear divide. Principals in Nigeria faced unstable electricity and high data costs (Awodiji & Adeoye, 2026), while middle leaders in Korea worked with mature digital infrastructure and national policy support (Lee & Kim, 2026). Cyprus's centralized system left school leaders waiting for training that had so far been offered only to teachers (Kafa, 2025), and leaders in Slovenia interpreted AI through European Union data protection rules (Flogie et al., 2026). These findings extend earlier K-12 reviews (Bellibaş & Karaferye, 2026; Mifsud et al., 2026) by showing that adoption in higher education follows a more institutional logic.

For RQ2, the synthesis identified four changes in leadership roles (Table 2). The first was the reallocation of routine work. AI took over the drafting of compliance documents, minutes, letters, and reports (Berkovich, 2025; Flogie et al., 2026; Salvatierra, 2026; Sterrett et al., 2026), which freed time for instructional priorities and relationships with staff (Lee & Kim, 2026). Efficiency, however, did not automatically lead to transformation. Some principals felt that AI only helped them finish less meaningful tasks faster (Sterrett et al., 2026), and the time saved could be absorbed by new demands (Lee & Kim, 2026). A survey of 636 school leaders in China even found that the intention to use generative AI did not predict actual use (Adams & Feng, 2026).

Table 2. Changes in the roles, practices, and competencies of educational leaders (RQ2)

Role change	Practices identified	Competencies required	Supporting reports
Reallocating routine work	Drafting compliance documents, reports, and communications with AI; redirecting time to instructional priorities	AI literacy; time and workload management	S25, S26, S30, S41, S42, S47
Critical evaluator of AI output	Verifying AI output, upholding standards of evidence, and keeping the final decision	Critical thinking; metacognition and epistemic judgment	S25, S28, S32, S40, S42, S52
Enabler of teachers and an innovation culture	Modeling use, creating safe spaces for experimentation, and sharing authority with teacher leaders	Transformational and digital leadership; change leadership	S26, S29, S31, S32, S34, S35, S36, S39, S48, S49
Governance and strategy	Writing guidelines, aligning with regulations, and setting the strategic direction for AI	Ethical and legal judgment; strategic planning	S27, S33, S40, S41, S42, S43, S49

Note: Report codes refer to Appendix B.

The second change positioned leaders as critical evaluators of AI output. Rather than reducing the role of professional judgment, AI made it more important: leaders had to verify accuracy, assess fit with the local context, and remain responsible for the final decision (Lee & Kim, 2026; Salvatierra, 2026). Teacher leaders in six Arab countries expected principals to uphold standards of evidence when AI output looked convincing (Ghamrawi, Al-Fadala, et al., 2026). A comparative study found that ChatGPT's solutions to problems of leadership practice resembled those of leaders but were less sensitive to context and drew on fewer resources (Y. Kim et al., 2024). A scale validation with 712 principals in Jordan identified six barriers in this role, including epistemic ambiguity and technological identity conflict (Khasawneh et al., 2026).

The third change was the role of enabling teachers and building an innovation culture. Digital instructional leadership predicted teachers' AI integration ($\beta = 0.264$), but its effect on pedagogy was only indirect (Hsieh et al., 2026). Transformational leadership strengthened the relationship between teachers' perceptions and their AI use (Chassida, 2026), and leaders' openness to experience was related to AI integration only in schools with a strong innovation culture (Berkovich, 2026). In practice, this role involved modeling use and encouraging hesitant teachers (Sterrett et al., 2026), creating safe spaces for experimentation and giving teacher leaders a say in AI rules (Ghamrawi, Al-Fadala, et al., 2026), and relying on teacher leaders to draft school AI guidelines (Ghamrawi, Shal, et al., 2026). These findings suggest that leaders influence AI use in teaching through organizational conditions rather than through direct instruction.

The fourth change was a governance and strategy role. In Saudi Arabia, university leaders' intention to adopt green AI was a governance decision framed by institutional social responsibility, with facilitating conditions as the strongest predictor (β = 0.311) (Salem & Khalil, 2026). Leaders in Slovenia aligned AI use with the GDPR and the EU AI Act (Flogie et al., 2026), and 62.5% of higher education leaders in the Philippines treated data protection rules as a precondition for adoption (Salvatierra, 2026). Principals in China displayed four leadership styles, ranging from "gatekeepers" who filtered AI against educational values to strategy-driven leaders (Yuan et al., 2026). By contrast, in a survey of rural school leaders in the United States, the lack of a strategic purpose for AI adoption appeared in 20.5% of responses about barriers (J. Kim & Wargo, 2025).

These four changes call for competencies that leaders do not yet have. AI literacy was mentioned most often (17 reports), but leaders' understanding was often shallow and equated with general digital technology (Awodiji & Adeoye, 2026). Readiness lagged behind awareness: higher education leaders in the Philippines reported high awareness (M = 3.82) but only moderate readiness (M = 3.36) (Salvatierra, 2026). In the United Kingdom, 82.7% of academic leaders needed professional development, and 86.5% cited lack of time as a barrier (Mohasseb et al., 2026). The training on offer was also mostly designed for teachers rather than leaders (Kafa, 2025; Salvatierra, 2026). Beyond AI literacy, the studies called for metacognitive and epistemic skills (Khasawneh et al., 2026), ethical and legal judgment (Flogie et al., 2026; Selmi Arrooqi & Miqad Alruqi, 2025), and change leadership (Chassida, 2026; X. Wang et al., 2025).

The RQ2 findings support the idea of human–AI collaboration (Dai et al., 2025; Y. Wang, 2021a), with two qualifications. First, the shift in roles is uneven and remains largely aspirational. Second, middle leaders emerged as key actors: subject heads, coordinators, and department heads mediated between AI policy and classroom practice (Berkovich, 2025; Lee & Kim, 2026; Mohasseb et al., 2026; Tahir et al., 2026). This role is less visible in earlier reviews centered on principals (Bellibaş & Karaferye, 2026; Mifsud et al., 2026). An alternative explanation should also be considered. Several studies recruited participants who were already active AI users, such as principals with digital leadership awards (Sterrett et al., 2026) and schools already using AI (Ghamrawi, Al-Fadala, et al., 2026; Ghamrawi, Shal, et al., 2026), so the picture of new roles may be more positive than the general situation.

For RQ3, the human-centered principles that appeared most often were equity and inclusion (10 reports), human agency and control (9), and ethics (8), followed by accountability, transparency, privacy, and well-being (6 each) and trust (5) (Table 3). Human agency was enacted by treating AI as an assistant that follows professional judgment, not as the decision-maker (Ghamrawi, Shal, et al., 2026; Salvatierra, 2026). Equity appeared in concerns that minority voices would be lost in aggregated data and that staff skills were uneven (Lee & Kim, 2026), in support for students with special needs and from low socioeconomic backgrounds (Wu et al., 2026), and in equal access for rural schools (J. Kim & Wargo, 2025). Staff well-being was visible in the role of workload as the link between principals' positive and negative views of AI across 12 countries (Eryilmaz et al., 2025) and in concerns that AI experimentation would add to teachers' workload (Ghamrawi, Al-Fadala, et al., 2026).

Table 3. Human-centered principles, opportunities, and challenges (RQ3)

Category	Element	n	Supporting reports
Principles	Equity and inclusion	10	S25, S27, S36, S37, S40, S41, S42, S46, S48, S49
	Human agency and control	9	S25, S26, S28, S32, S33, S35, S40, S42, S52
	Ethics	8	S32, S35, S36, S37, S43, S46, S48, S49
	Accountability	6	S25, S27, S32, S41, S42, S43
	Transparency	6	S31, S40, S41, S42, S43, S46
	Privacy	6	S32, S36, S41, S42, S48, S49
	Staff well-being	6	S25, S31, S32, S33, S45, S47
	Trust	5	S25, S28, S31, S39, S46
Opportunities	Efficiency, data-informed decisions, personalization, wider participation	—	S25, S36, S40, S41, S42, S48
Challenges	Competence gaps	12	S25, S28, S29, S36, S37, S41, S42, S43, S44, S46, S48, S49
	Unclear policy	9	S27, S31, S33, S35, S37, S41, S42, S44, S49
	Ethical and privacy risks	9	S31, S32, S36, S40, S41, S42, S46, S48, S49
	Resistance to change	9	S25, S26, S28, S29, S34, S42, S43, S46, S48
	Infrastructure	8	S27, S29, S33, S36, S37, S41, S44, S49
	Cost	7	S29, S33, S37, S41, S42, S46, S48

Note: n = number of reports containing the element; report codes refer to Appendix B.

The opportunities mentioned most often were administrative efficiency, data-informed decisions, and personalized learning (Flogie et al., 2026; X. Wang et al., 2025; Wu et al., 2026). AI could also widen participation in leadership by lowering the barriers for teachers to take on new roles (Lee & Kim, 2026) and help identify and develop future leaders (AL-Housni et al., 2026). The challenges were more consistent: competence gaps (12 reports), absent or unclear policy (9), ethical and privacy risks (9), resistance to change (9), infrastructure (8), and cost (7). In Israel, for example, the absence of a national policy left principals to interpret AI use on their own (Chassida, 2026).

Quantitative studies showed that institutional conditions operated differently depending on where the decision was made. Facilitating conditions were the strongest predictor when AI adoption was an institutional governance decision (Salem & Khalil, 2026) but had no significant effect when school leaders decided whether to use generative AI in their own work (Adams & Feng, 2026). Ethical considerations were the strongest predictor of academic leaders' readiness in Saudi Arabia (Selmi Arrooqi & Miqad Alruqi, 2025). The most important RQ3 finding is a gap between principle and practice. Human-centered principles appeared more often as concerns than as procedures in use. Only a few studies described concrete mechanisms, and these were mostly proposals: rules for disclosing AI use and human review of official documents (Salvatierra, 2026), the ethical boundaries around data and care that teacher leaders expected from principals (Ghamrawi, Al-Fadala, et al., 2026), and institutional guidelines aligned with regulation (Flogie et al., 2026). This confirms the criticism in an earlier review that ethical guidance for leaders remains limited (Ho, 2026).

For RQ4, the themes from RQ2 and RQ3 were combined into a framework for human-centered educational leadership in AI integration (Figure 3). At its core are educational purpose and human judgment, supported by studies that place AI below professional judgment (Ghamrawi, Shal, et al., 2026; Y. Kim et al., 2024; Lee & Kim, 2026). Around the core sit the four leadership practices from RQ2, each tied to principles from RQ3: reallocating routine work to staff well-being, critical evaluation of AI output to accountability and calibrated trust, enabling teachers to equity and participation, and governance to privacy, transparency, and ethics. The foundation consists of four enabling conditions derived from the RQ3 challenges: leader-specific AI professional development (Awodiji & Adeoye, 2026; Kafa, 2025; Mohasseb et al., 2026), infrastructure and resources (Awodiji & Adeoye, 2026; J. Kim & Wargo, 2025), clear policy and regulation (Chassida, 2026; Flogie et al., 2026), and a culture of innovation and trust (Berkovich, 2026; Hsieh et al., 2026).

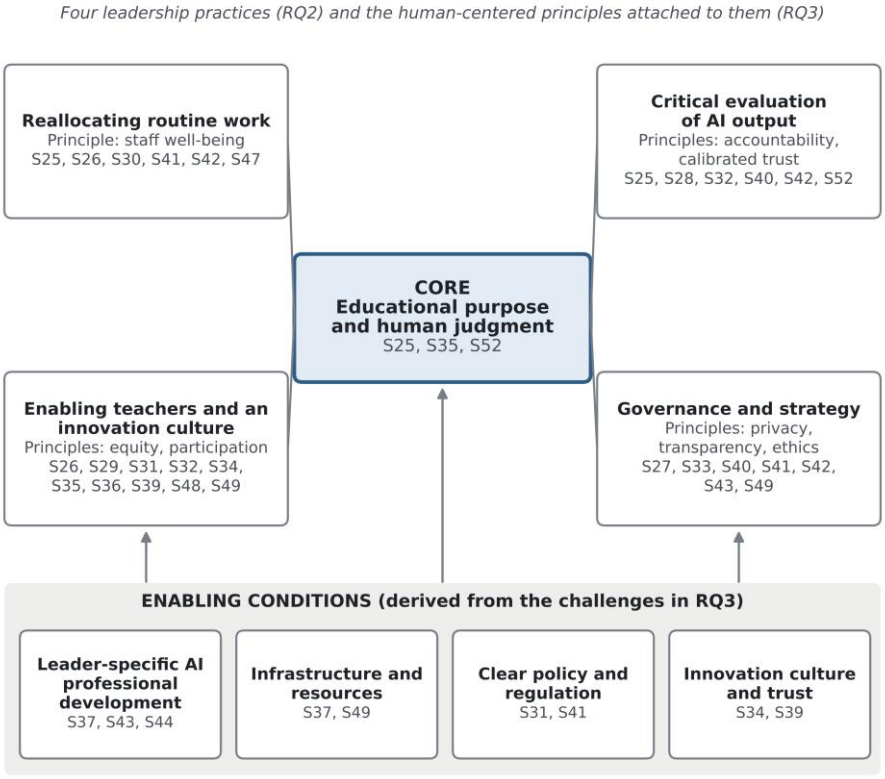

Figure 3. Framework for human-centered educational leadership in AI integration

The framework offers three propositions for future research to test. First, efficiency gains from AI will improve the quality of leadership only if the time saved is deliberately reallocated to instructional and relational work (Lee & Kim, 2026; Sterrett et al., 2026). Second, leaders influence teachers' AI use through culture and legitimacy rather than through direct instruction (Berkovich, 2026; Chassida, 2026; Hsieh et al., 2026). Third, human-centered principles are more likely to be put into practice when they are written into explicit governance procedures (Ghamrawi, Al-Fadala, et al., 2026; Salvatierra, 2026). Compared with earlier frameworks that focus on domains of school leadership (DeMatthews et al., 2026) or on capacity building (Psyrras & Kafa, 2026), this framework is built from empirical evidence across educational levels and explicitly links practices to the conditions that enable them.

Two theories help explain why perception studies dominate the corpus. Diffusion of innovations theory describes adoption as a sequence from knowledge and persuasion to decision and implementation (Rogers, 2003). Most included studies captured leaders at the first two stages. This explains why the intention to use generative AI did not predict actual use among school leaders in China (Adams & Feng, 2026) and why awareness exceeded readiness among higher education leaders in the Philippines (Salvatierra, 2026). Institutional theory offers a second explanation. Under uncertainty, organizations respond to regulation and imitate their peers (DiMaggio & Powell, 1983). Leaders in Slovenia and the Philippines tied AI use to data protection rules (Flogie et al., 2026; Salvatierra, 2026). Leaders in Israel and Cyprus, without national guidance, interpreted AI use on their own or waited for central training (Chassida, 2026; Kafa, 2025). This contrast suggests that regulatory pressure, more than individual readiness, moves leaders from exploration to implementation. Our findings extend the K-12 reviews by Bellibaş and Karaferye (2026) and Mifsud et al. (2026) in two ways. First, higher education studies framed adoption as institutional governance rather than individual use. Second, the quality appraisal shows that the evidence for the governance practice is thinner than the evidence for the other three practices.

We checked the robustness of the findings by quality and adoption stage. Reallocating routine work was supported by six reports, five of them high quality. Critical evaluation of AI output was supported by six reports, five high quality. Enabling teachers was supported by ten reports, five high quality and four low quality. Governance was supported by seven reports, three high quality and three moderate. We therefore hold stronger confidence in the first three practices and moderate confidence in governance. Most of the low-quality reports were quantitative studies of adoption predictors (Berkovich, 2025, 2026; Chassida, 2026; Hsieh et al., 2026; Salem & Khalil, 2026; Selmi Arrooqi & Miqad Alruqi, 2025), so their effect sizes should be read with caution.

Adoption stage limits the claims further. Only 8 of the 25 reports examined implemented AI, and 17 described perceptions or intentions. The four practices therefore describe what leaders do or expect to do in early adoption, not established routines, and the principle-practice gap in RQ3 may partly reflect this early stage.

This review has limitations. The search covered only Scopus and was restricted to titles and keywords, so studies that mention leaders only in the abstract may have been missed. The full texts of 25 of the 52 reports could not be obtained, including one study of rural primary schools in Indonesia, which skews the corpus toward open-access journals. Selection and extraction were carried out by the authors with the help of an AI model and without an independent second reviewer, so inter-rater reliability cannot be reported and the risk of selection bias cannot be fully ruled out. We reduced this risk in three ways: the authors verified every AI proposal and every extracted entry against the source, unclear records went to full-text assessment instead of being excluded, and each theme was tested for robustness against study quality. Future research should examine actual implementation longitudinally, measure effects on teachers and students, and add evidence from Southeast Asia, including Indonesia.

Conclusions

AI integration in educational leadership is at an early stage. Of the 25 reports, 17 described leaders' perceptions or intentions, and only eight examined AI in use. Within this limited evidence, AI use was associated with four leadership practices: reallocating routine work (6 reports), critically evaluating AI output (6), enabling teachers through an innovation culture (10), and governing ethical and legal risks (7). The first three practices rest on at least five high-quality reports each, and governance rests on three. In none of these practices did AI replace the leader's decision. Instead, leaders took on more checking work: they verified the accuracy of AI output, judged its fit with the local context, and kept responsibility for the final decision. Human-centered principles were often cited but rarely written into procedures. The framework therefore places educational purpose and human judgment at its core and names four enabling conditions: leader-specific professional development, clear policy, infrastructure, and an innovation culture. These findings are preliminary. They suggest that AI training should be designed for leaders, not only for teachers, and should come with explicit governance guidelines. Education systems not represented in the corpus, including Indonesia, need local studies before these implications are applied there.

Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The authors used Claude (Anthropic) to assist with pre-screening records, extracting data, and drafting the manuscript. All such output was reviewed, verified against the original sources, and edited by the authors, who take full responsibility for the content of this article.

References

Adams, D., & Feng, X. (2026). School leaders' acceptance and use of artificial intelligence in China. *Educational Studies*. https://doi.org/10.1080/03055698.2026.2708320

AL-Housni, H. A., Abunaser, F., Bani-Oraba, A. M. N., & Al Harthy, R. A. H. (2026). To what extent can artificial intelligence sustain leadership talents in education? Voices of educational leaders and experts. *Education Sciences*, *16*(4), 601. https://doi.org/10.3390/educsci16040601

Ali, N. (2026). Educational leadership in the age of artificial intelligence: The moral compass in the machine. *Management in Education*. https://doi.org/10.1177/08920206261473366

Arar, K., Tlili, A., & Salha, S. (2026). Human-machine symbiosis in educational leadership in the era of artificial intelligence (AI): Where are we heading? *Educational Management Administration & Leadership*, *54*(3), 669–692. https://doi.org/10.1177/17411432241292295

Arar, K., Tlili, A., Schunka, L., Salha, S., & Saiti, A. (2025). Reimagining educational leadership and management through artificial intelligence: An integrative systematic review. *Leadership and Policy in Schools*, *24*(1), 4–26. https://doi.org/10.1080/15700763.2025.2451982

Awodiji, O. A., & Adeoye, M. A. (2026). Exploring basic school leaders' AI readiness: The role of professional development. *Computers and Education Open*, *11*, 100409. https://doi.org/10.1016/j.caeo.2026.100409

Bellibaş, M. Ş., & Karaferye, F. (2026). Artificial intelligence (AI) in educational leadership: A systematic review of ethical issues, implementation challenges, and practical use. *Improving Schools*. https://doi.org/10.1177/13654802261469999

Berkovich, I. (2025). The rise of AI-assisted instructional leadership: Empirical survey of generative AI integration in school leadership and management work. *Frontiers in Education*, *10*, 1643023. https://doi.org/10.3389/feduc.2025.1643023

Berkovich, I. (2026). Innovation culture in schools, school leaders' openness to experience, and the integration of generative AI in school leadership work: A UTAUT based model. *Improving Schools*, *28*(2), 443–460. https://doi.org/10.1177/13654802261434283

Chassida, Y. (2026). Transformational leadership as a contextual enabler of teachers' AI use. *Education Sciences*, *16*(4), 572. https://doi.org/10.3390/educsci16040572

Dai, R., Thomas, M. K. E., & Rawolle, S. (2025). The roles of AI and educational leaders in AI-assisted administrative decision-making: A proposed framework for symbiotic collaboration. *The Australian Educational Researcher*, *52*(2), 1471–1487. https://doi.org/10.1007/s13384-024-00771-8

DeMatthews, D., Reyes, P., Hart, T. D., & James, L. , I. (2026). Leadership for artificial intelligence use in schools: A six-domain framework for ethical, equitable, and effective integration. *Educational Management Administration & Leadership*. https://doi.org/10.1177/17411432261418940

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, *48*(2), 147–160. https://doi.org/10.2307/2095101

Eryilmaz, N., Mah, D.-K., Bellibas, M. S., & Pietsch, M. (2025). School leaders' expectations of AI's impact on students and teachers: Insights from ICILS 2023. *Educational Management Administration & Leadership*. https://doi.org/10.1177/17411432251401487

Flogie, A., Hari, D., Skrbinjek, V., & Brejc, M. (2026). Use and aspects of generative artificial intelligence in the educational system among leadership personnel. *Education and Information Technologies*, *31*, 4929–4963. https://doi.org/10.1007/s10639-026-13986-2

Ghamrawi, N., Al-Fadala, A., Shal, T., & Ghamrawi, N. A. R. (2026). Architecting the artificial intelligence-era principal: Teacher leaders reimagine school leadership. *Management in Education*. https://doi.org/10.1177/08920206261473364

Ghamrawi, N., Shal, T., & Ghamrawi, N. A. R. (2026). Irreplaceable teachers in the AI era are teacher leaders. *Education and Information Technologies*, *31*, 5243–5261. https://doi.org/10.1007/s10639-026-14006-z

Ho, C. S. M. (2026). Principals' ethical leadership in the AI era: A narrative literature review of emerging challenges, strategies, and outcomes. *Computers & Education*, *243*, 105517. https://doi.org/10.1016/j.compedu.2025.105517

Ho, C. S. M., & MacGregor, S. (2026). Reconceptualizing professional capital: From human, social, and decisional to digital-adaptive in principals' leadership in the AI era. *Journal of Professional Capital and Community*. https://doi.org/10.1108/JPCC-10-2025-0119

Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., Vedel, I., & Pluye, P. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for Information*, *34*(4), 285–291. https://doi.org/10.3233/EFI-180221

Hsieh, C.-C., Bali, S., & Li, H.-C. (2026). Untangling pathways from digital instructional leadership to AI pedagogy: A diffusion of innovation theory perspective. *BMC Psychology*, *14*, 63. https://doi.org/10.1186/s40359-025-03792-9

Kafa, A. (2025). Exploring integration aspects of school leadership in the context of digitalization and artificial intelligence. *International Journal of Educational Management*, *39*(8), 98–115. https://doi.org/10.1108/IJEM-11-2024-0703

Khasawneh, Y. J. A., Khasawneh, N. A. S., Khasawneh, M. A. S., Sufyan, N. S., & Khasawneh, A. J. (2026). Development and validation of the AI Learning, Metacognitive, and Epistemic Barriers Scale (AIME-B) for secondary school principals in Jordan. *Journal of Computer Assisted Learning*, *42*(5), e70301. https://doi.org/10.1002/jcal.70301

Khraisha, Q., Put, S., Kappenberg, J., Warraitch, A., & Hadfield, K. (2024). Can large language models replace humans in systematic reviews? Evaluating GPT-4's efficacy in screening and extracting data from peer-reviewed and grey literature in multiple languages. *Research Synthesis Methods*, *15*(4), 616–626. https://doi.org/10.1002/jrsm.1715

Kim, J., & Wargo, E. (2025). Empowering educational leaders for AI integration in rural STEM education: Challenges and strategies. *Frontiers in Education*, *10*, 1567698. https://doi.org/10.3389/feduc.2025.1567698

Kim, Y., Curry, K. A., & Fiegener, A. M. (2024). Exploring suggestions from educational leaders and ChatGPT for addressing problems of practice in TeleED: A qualitative case study. *Journal of School Administration Research and Development*, *9*(2), 56–65.

Lee, Y., & Kim, M. (2026). The use of artificial intelligence in educational leadership practice: A qualitative study of middle leaders in Korean secondary schools. *Management in Education*. https://doi.org/10.1177/08920206261487370

Leithwood, K., Harris, A., & Hopkins, D. (2008). Seven strong claims about successful school leadership. *School Leadership & Management*, *28*(1), 27–42. https://doi.org/10.1080/13632430701800060

Lockwood, C., Munn, Z., & Porritt, K. (2015). Qualitative research synthesis: Methodological guidance for systematic reviewers utilizing meta-aggregation. *International Journal of Evidence-Based Healthcare*, *13*(3), 179–187. https://doi.org/10.1097/XEB.0000000000000062

Mifsud, D., Campbell, P., & Niesche, R. (2026). The impact of AI on school leadership practices and policy enactment: A scoping literature review. *Leadership and Policy in Schools*. https://doi.org/10.1080/15700763.2026.2679510

Mohasseb, A., Beukman, R., & Kanavos, A. (2026). The AI Literacy Leadership Framework (AILLF): A framework for AI-enabled leadership in higher education. *AI*, *7*(9), 328. https://doi.org/10.3390/ai7090328

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, n71. https://doi.org/10.1136/bmj.n71

Psyrras, N., & Kafa, A. (2026). Artificial intelligence in school leadership: A multi-level framework for leadership capacity building. *International Journal of Educational Management*, *40*(8), 124–138. https://doi.org/10.1108/IJEM-04-2026-0573

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.

Salem, M. A., & Khalil, Z. A. (2026). Unlocking GAI in universities: Leadership-driven corporate social responsibility for digital sustainability. *Administrative Sciences*, *16*(2), 58. https://doi.org/10.3390/admsci16020058

Salvatierra, E. G. (2026). Generative AI in higher education management: A mixed-methods study of school leaders' awareness, readiness, and adoption intentions. *International Journal of Learning, Teaching and Educational Research*, *25*(6), 413–435. https://doi.org/10.26803/ijlter.25.6.17

Selmi Arrooqi, A., & Miqad Alruqi, M. (2025). Academic leadership attitudes toward employing artificial intelligence applications in developing administrative processes. *Humanities and Social Sciences Communications*, *12*, 1342. https://doi.org/10.1057/s41599-025-05598-x

Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human–Computer Interaction*, *36*(6), 495–504. https://doi.org/10.1080/10447318.2020.1741118

Sterrett, W. L., Richardson, J. W., & Frey, M. A. (2026). Perceptions of how AI has changed teaching, learning, and leading in K-12 schools: Insights from award-winning principals. *Frontiers in Education*, *11*, 1854863. https://doi.org/10.3389/feduc.2026.1854863

Tahir, L. M., Liu, H., Ahmad, J., Abdul Aziz, N., Abd Rani, N. S., & Mohamad, N. (2026). Middle school leaders' acceptance of artificial intelligence: A mixed-methods study from Shandong, China. *Asian Journal of University Education*, *22*(2), 431–455. https://doi.org/10.24191/ajue.v22i2.027

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, *8*, 45. https://doi.org/10.1186/1471-2288-8-45

UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO.

Wang, X., Zhao, S., Xu, X., Zhang, H., & Lei, V. N.-L. (2025). AI adoption in Chinese universities: Insights, challenges, and opportunities from academic leaders. *Acta Psychologica*, *258*, 105160. https://doi.org/10.1016/j.actpsy.2025.105160

Wang, Y. (2021a). Artificial intelligence in educational leadership: A symbiotic role of human-artificial intelligence decision-making. *Journal of Educational Administration*, *59*(3), 256–270. https://doi.org/10.1108/JEA-10-2020-0216

Wang, Y. (2021b). When artificial intelligence meets educational leaders' data-informed decision-making: A cautionary tale. *Studies in Educational Evaluation*, *69*, 100872. https://doi.org/10.1016/j.stueduc.2020.100872

Wu, S., Xu, J., & Feng, S. (2026). Opportunities and challenges of generative AI in STEAM education: A thematic co-occurrence network analysis of primary and secondary school leaders' perspectives. *Technology, Knowledge and Learning*. https://doi.org/10.1007/s10758-026-09976-5

Yuan, P., Chen, X., & Song, H. (2026). School principals' perspectives and leadership styles for digital transformation: A Q-methodology study. *Behavioral Sciences*, *16*(2), 165. https://doi.org/10.3390/bs16020165

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education*, *16*, 39. https://doi.org/10.1186/s41239-019-0171-0