https://jsras.rcc.edu.ly/
Vol.3 No.2 (2026), 210-219
Article history:
Received: 15 Aug .2026
Accepted: 4.Sep . 2026
Published: 20.Oct . 2026

Journal of Sustainable Research in Applied Sciences

JSRAS

A Systematic Review of AI-Powered Learning Assistants in Higher Education: Personalization Strategies, Adaptive Frameworks, and Deployment Challenges

Laila Esmeda [1], Anwar Alhenshiri [2]

[1] Faculty of Information Technology, Misurata University, Misurata, Libya, l.esmeda@it.misuratau.edu.ly, ORCID:0000-0002-9245-7635

[2] Faculty of Information Technology, Misurata University, Misurata, Libya, alhenshiri@it.misuratau.edu.ly, ORCID: 0009-0007-8825-7685

keywords

Adaptive learning; AI learning assistant; Systematic review; Higher Education Resource-constrained; Cold start; Libya

Abstract

This review examines personalization strategies, adaptivity frameworks, implementation challenges, and transferability to resource-constrained higher education, with emphasis on settings that have no learning management system (LMS). Following PRISMA 2020, six search sources were searched in English and Arabic between March and May 2026 for studies published from January 2018 to May 2026. Of 357 records identified, 45 studies were included. One researcher coded the studies and re-coded a 20% sample in a second pass (initial agreement 62.5%); the review is therefore exploratory, and its percentages are directional indicators. Eighteen of the 45 studies (40.0%) used assessment or LMS-generated data as the primary learner-modeling input, and four (8.9%) used interaction-timing signals. Cold start was discussed in eight studies and addressed directly in two. Only four studies reported an operational system deployed without LMS integration; most other studies coded as LMS-free were surveys or reviews that describe no deployed system. Studies from Libya most often reported infrastructure, training, and policy barriers. Evidence on deployed adaptive assistants outside LMS-based settings rests on a small number of studies, and none was found to combine platform independence, infrastructure-aware design, and handling of Arabic dialect input. Controlled evaluations in Arabic-language settings without an LMS are needed.

مراجعة منهجية للمساعدات التعليمية المدعومة بالذكاء الاصطناعي في التعليم العالي: استراتيجيات التخصيص، وأطر التكيف، وتحديات التطبيق

ليلى صميدة[1,*]، أنور الهنشيري[2]

[1] جامعة مصراتة، كلية تقنية المعلومات، قسم علوم الحاسوب، ليبيا، La.esmeda@asmarya.edu.ly، 0000-0002-9245-7635

[2] جامعة مصراتة، كلية تقنية المعلومات، قسم علوم الحاسوب، ليبيا، alhenshiri@it.misuratau.edu.ly، 0009-0007-8825-7685

الكلمات المفتاحية

التعلم التكيفي؛ مساعد التعلم الذكي؛ مراجعة منهجية استكشافية؛ تعليم عالٍ محدود الموارد؛ البداية الباردة؛ ليبيا.

الملخص

تفحص هذه المراجعة المنهجية استراتيجيات التخصيص وأطر التكيف وتحديات التنفيذ وقابلية النقل إلى مؤسسات التعليم العالي محدودة الموارد، مع التركيز على البيئات التي لا يتوفر فيها نظام لإدارة التعلم. اتُّبعت إرشادات PRISMA 2020، وجرى البحث في ست منصات بحثية باللغتين الإنجليزية والعربية خلال الفترة من مارس إلى مايو 2026 عن دراسات نُشرت بين 2018 ومايو 2026. وبعد إزالة التكرارات من بين 357 سجلاً، أُدرجت 45 دراسة، كما أجرى باحث واحد الترميز مع إعادة ترميز عينة نسبتها 20% في جولة ثانية (الاتفاق الأولي 62.5%)؛ ولذلك تُعدّ المراجعة استكشافية، وتمثل النسب المئوية مؤشرات اتجاهية. اعتمدت 18 دراسة من 45 (40.0%) على بيانات التقييم أو بيانات نظام إدارة التعلم مدخلاً رئيساً لنمذجة المتعلم، واستخدمت أربع دراسات (8.9%) إشارات توقيت التفاعل. نوقشت مشكلة البداية الباردة في ثماني دراسات، وعولجت مباشرة في دراستين. وأفادت أربع دراسات فقط بنظام تشغيلي منشور دون تكامل مع نظام إدارة التعلم، في حين كانت معظم الدراسات الأخرى المصنفة خالية من نظام إدارة التعلم مسوحاً أو مراجعات لا تصف نظاماً منشوراً. وأشارت الدراسات الليبية في أغلبها إلى عوائق البنية التحتية والتدريب والسياسات. تقتصر الأدلة على المساعدات التكيفية المنشورة خارج البيئات القائمة على أنظمة إدارة التعلم على عدد قليل من الدراسات، ولم تُوجد دراسة تجمع بين استقلالية المنصة والتصميم المراعي للبنية التحتية ومعالجة المدخلات المختلفة للهجة العربية. وتبرز الحاجة إلى تقييمات ضابطة في البيئات العربية الخالية من أنظمة إدارة التعلم..

1. INTRODUCTION

1.1 Problem and setting

AI-powered learning assistants, including intelligent tutoring systems, adaptive recommenders, and chatbot-based tutors, are a growing research topic in higher education (Chen et al., 2020; Crompton & Burke, 2023; Zawacki-Richter et al., 2019). Research on educational chatbots has expanded in particular (Ait Baha et al., 2024; Huang et al., 2022), and intelligent tutoring systems have been reported to approach the effectiveness of human tutoring under some conditions (VanLehn, 2011). Broader accounts of AI in education also exist (Chaudhry & Kazim, 2022). Many reported systems assume an institutional digital layer: a learning management system (LMS) that records learner activity, dependable connectivity, and technical support staff. In some Arab and African universities this layer is partial or absent, and course communication runs through messaging applications and email. At the Faculty of Information Technology of Al-Asmarya Islamic University (Zliten, Libya), which motivated this review, no LMS is in use; consequently, no structured record exists of late submissions, repeated questions, or declining engagement. Where interaction produces no structured data, personalization that depends on such data cannot operate.

1.2 Concentration of the evidence base

The published evidence is concentrated in particular settings. In a scoping review of 69 studies of personalized adaptive learning in undergraduate education, 29% of the studies originated in the United States, Moodle and McGraw-Hill Connect LearnSmart were the most frequently used platforms, and pre-knowledge quizzes were the most common trigger for adaptation (58%); the authors also noted a clear underrepresentation of Africa du Plooy et al. (2024). This review asks whether findings obtained in LMS-based settings transfer to settings without an LMS. In the present corpus, most studies from Libya are attitude or adoption surveys and scoping reviews rather than evaluations of deployed systems (Section 3.4). One faculty survey conducted at Al-Asmarya Islamic University is included (Gratim & Elatrash, 2025); no included study reports the design or evaluation of a learning assistant at that institution.

1.3 Aim, review questions, and scope

The review addresses four questions (Table I). RQ1-RQ3 correspond to three analytical dimensions: personalization strategies (what systems know about learners and how they use it), adaptivity frameworks (the computational mechanisms of adaptation), and implementation challenges. RQ4 is a cross-cutting question on transferability to resource-constrained Arab higher education. The questions were fixed before searching because the review is embedded in a doctoral project at Al-Asmarya Islamic University; they were not derived inductively from the search results. Resource-constrained is used here for institutions in which three conditions co-occur: no deployed LMS generating structured interaction data, intermittent or unreliable connectivity, and limited technical support capacity.

TABLE 1. REVIEW QUESTIONS (RQ1–RQ4)

RQ	Dimension	Research Question
RQ1	Personalization	What personalization strategies have been applied in AI learning assistants, and what learner data do they require?
RQ2	Adaptivity	What adaptivity frameworks underlie these systems, and what are their infrastructure requirements?
RQ3	Challenges	What implementation challenges are documented, particularly where digital infrastructure is absent?
RQ4	Transferability	To what extent do findings transfer to resource-constrained Arab HE?

2. METHODOLOGY

2.1 Design and reporting

The review followed the PRISMA 2020 reporting guideline (Page et al., 2021). No protocol was registered in PROSPERO because the review was designed as a component of a doctoral program; this is an accepted approach for embedded reviews (Booth et al., 2016), but it means that the analysis plan was not publicly fixed in advance. Because coding was performed by one researcher (Section 2.5), the review is classified as an exploratory systematic review rather than a confirmatory one: its percentages describe the corpus and are not tests of hypotheses.

2.2 Search sources and strategy

Six sources were searched between March and May 2026 (Table II). English-language queries were run in Google Scholar, Consensus, ResearchGate, Semantic Scholar, and BASE; Arabic-language queries were run in Google Scholar, Consensus, and ResearchGate, and Dar Al Mandumah was searched as an Arabic-language database. The Arabic searches therefore reused platforms already listed and did not add sources. An institutional subscription to Dar Al Mandumah was unavailable, so papers identified through its index were retrieved from publisher sites or open-access repositories. Table II summarizes the databases searched, the query languages used for each source.

TABLE 2. DATABASES SEARCHED

Source	Type and reason for inclusion	Searched directly; query language
Google Scholar	Academic search engine; broad multidisciplinary coverage of English and Arabic publications	Yes; English, Arabic
Consensus	AI-assisted academic search engine that retrieves findings from peer-reviewed papers	Yes; English, Arabic
Research Gate	Academic social network; authors share full texts, including from regional journals	Yes; English, Arabic
Semantic Scholar	AI-assisted academic search engine with strong computer-	Yes; English

	science coverage	
BASE	Open-access aggregator of repositories and journals	Yes; English
Dar Al Mandumah	Arabic-language research database; index of Arabic-language journals	Index searched; full texts retrieved from publisher or open-access sources; Arabic

The English Boolean string was: ("AI learning assistant" OR "intelligent tutoring system" OR "chatbot" OR "virtual assistant") AND ("personalized learning" OR "adaptive learning" OR "learner model") AND ("higher education" OR "university" OR "undergraduate") AND ("e-learning" OR "online learning" OR "LMS"). Arabic queries used semantically equivalent phrases rather than literal translations, because Arabic morphology does not map cleanly onto English Boolean logic: (مساعد ذكاء اصطناعي + تعليم جامعي), (تعلم تكيفي + تعليم عالٍ + عربي), and (روبوت محادثة + تعليم إلكتروني). This follows cross-lingual review practice (Gough et al., 2017).

Three of the six sources (Google Scholar, Consensus, and ResearchGate) are search engines or networks rather than curated bibliographic indexes, and their rankings can vary with time and personalization; exact reproduction of the retrieved set therefore cannot be guaranteed. Curated indexes such as Scopus, Web of Science, ERIC, and IEEE Xplore were not searched (Section 4.5).

2.3 Eligibility criteria

Eligible publications appeared between January 2018 and May 2026; the searches themselves were run between March and May 2026. The window begins in January 2018 so that the corpus covers the period leading up to and following the introduction of transformer-based language models (Devlin et al., 2019). Earlier studies were excluded because their architectural assumptions differ from current deployments, not because of their quality. Table III summarizes the inclusion and exclusion criteria applied consistently across databases.

TABLE 3. Inclusion And Exclusion Criteria

Criterion	Inclusion	Exclusion
Publication window	January 2018 – May 2026	Before January 2018
Language	English or Arabic	Other languages
Publication type	Peer-reviewed journal articles or conference papers	Theses, book chapters, editorials, technical reports, video transcripts
Context	Higher education (undergraduate or postgraduate)	K–12, corporate training, language-only learning
AI component	At least one personalization or adaptive AI component	AI tools without an adaptive component; general e-learning adoption surveys
Evidence type	Empirical studies or systematic reviews	Opinion pieces

2.4 Study selection

The searches returned 357 records. After removal of 19 duplicates by title-similarity matching, performed manually by the lead author with cross-checking against DOI and author lists, 338 records were screened by title and abstract; 144 were excluded as off-topic, outside the publication window, or of ineligible publication type, leaving 194 reports for full-text assessment. Of these, 149 were excluded because they lacked an adaptive or personalization component, were not conducted in higher education, were not peer-reviewed journal articles or conference papers, or were unavailable in full text, and 45 studies were included (Fig. 1). Screening was performed by the lead author; no independent second screener was available, which is acknowledged as a limitation; all discrepancies were resolved through full-text re-reading

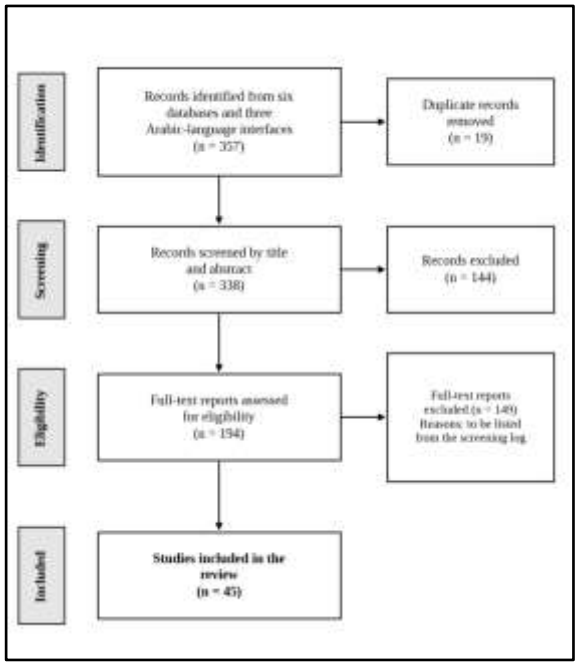


Figure 1. PRISMA 2020 Flow Diagram

2.5 Data coding and quality appraisal

2.5.1 Coding framework

One researcher coded each included study with a structured extraction instrument covering five dimensions (Table IV): AI system type, primary learner-modeling input, adaptivity framework, LMS dependency, and relevance to each review question (High, Medium, or Low). Table V gives worked examples of coding decisions.

TABLE 4. Coding Framework And Decision Rules

Dimension	Categories	Decision rule
AI system type	Intelligent tutoring system; recommender; chatbot or conversational agent; adaptive assessment; hybrid; survey or review	Coded from the system the study builds or evaluates; studies that describe no system are coded as survey or review.
Primary learner-modeling input	Assessment performance or LMS logs; interaction timing; self-report; none explicit	Coded from the input the study reports as primary. If no single dominant input is reported, the study is coded none explicit (the value "Multiple" was removed).

Adaptivity framework	Rule-based; machine learning (ML)-based; large language model (LLM)-based; hybrid; not applicable	Hybrid combines two or more adaptation approaches in one system (e.g., Rule+ML, Rule+LLM, or ML+LLM). All Twelve hybrid studies were re-verified against this definition.
LMS dependency	LMS required; LMS optional; LMS free; not applicable	LMS required: the system cannot function without LMS data or integration. LMS optional: integration is possible but not required. LMS free: described without an LMS. Studies that describe no deployed system are reported separately (Table XI).
Relevance to RQ1–RQ4	High; Medium; Low	High: empirical evaluation or deployment evidence that directly addresses the question (for RQ4, in a resource-constrained setting). Medium: partial or theoretical evidence. Low: tangential or contextual.

TABLE 5. Worked Examples Of Coding Decisions

Study	Initial or alternative coding	Final coding	Rule applied
Essel et al. (2022)	—	Chatbot; LMS free; RQ4 High	A deployed system with an outcome evaluation is rated High for RQ4.
Alenezi (2023)	Recommender (first pass)	Survey	Coded from the method reported in the full text, not from the title: a survey of students with follow-up interviews.
Gratim and Elatrash (2025)	—	Survey; no deployed system; LMS free by default	Studies that describe no system are reported separately from operational systems.
A study reporting several inputs without a dominant one	Multiple	None explicit	The value "Multiple" was removed from the coding scheme.

2.5.2 Self-review agreement

A random 20% sample (9 of 45 studies) was re-coded by the same researcher in a second pass. Initial agreement was 62.5% (45 of 72 decisions across eight dimensions; (Table VI); for the five core dimensions alone it was 62.2% (28 of 45 decisions). Agreement was highest for system type, which is usually explicit in titles and abstracts, and lowest for the RQ1 and RQ4 relevance ratings, which require interpretive judgment. All 27 disagreements were resolved by re-reading the full text and applying the rules in (Table VII) . Agreement after resolution is complete by construction, so initial agreement is the informative figure. No second coder was available, and no inter-rater statistic was computed.

TABLE 6 Initial Self-Review Agreement By Dimension

Dimension	Agreeing decisions	Agreement (%)
AI system type	8 of 9	88.9
Relevance to RQ2	7 of 9	77.8
Adaptivity framework	6 of 9	66.7
Relevance to RQ3	6 of 9	66.7
Primary learner-modeling input	5 of 9	55.6
LMS dependency	5 of 9	55.6
Relevance to RQ1	4 of 9	44.4
Relevance to RQ4	4 of 9	44.4
All dimensions	45 of 72	62.5

TABLE 7 Main Sources Of Coding Disagreement And How They Were Resolved

Source	Example	Resolution
Ambiguous input value	Three studies were first coded "Multiple" because the source text reported several inputs without a dominant one.	Recoded as none explicit; "Multiple" removed from the scheme.
Title–content mismatch	Alenezi (2023) was coded as a recommender in the first pass and as a survey in the second.	Full-text reading confirmed a survey design.
Undefined High/Medium boundary for RQ4	The boundary between High and Medium for RQ4 was not defined before coding.	High requires an actual empirical evaluation in a resource-constrained setting; Medium covers theoretical discussion only.

2.5.3 Quality appraisal

Each study was rated High, Medium, or Low on four dimensions (design appropriateness, sampling adequacy, measurement validity, and transferability to resource-constrained deployment) using a purpose-built checklist adapted from the Mixed Methods Appraisal Tool (Hong et al., 2018) and the CASP systematic review checklist (Critical Appraisal Skills Programme [CASP], 2018). Studies rated Low on design appropriateness or measurement validity were retained for completeness of the literature map but are not used as primary evidence for empirical claims.

3. RESULTS

Percentages are descriptive, are computed on the 45 included studies unless stated otherwise, and are directional indicators (Section 2.5.2). Figures taken from other reviews are labeled as such and are not directly comparable with figures from this corpus.

3.1 Study characteristics and appraisal

Forty-five studies were included, 12 of them from Libya. (Table VIII) summarizes the appraisal. Transferability to resource-constrained deployment was the weakest dimension: 17 studies (37.8%) were rated Low.

TABLE 8 Quality appraisal of the 45 included studies

Dimension	High	Medium	Low
Design appropriateness	18 (40.0)	20 (44.4)	7 (15.6)
Sampling adequacy	22 (48.9)	17 (37.8)	6 (13.3)
Measurement validity	16 (35.6)	21 (46.7)	8 (17.8)
Transferability to low-resource higher education	9 (20.0)	19 (42.2)	17 (37.8)

3.2 RQ1: Personalization strategies

Assessment performance or LMS-generated data was the primary learner-modeling input in 18 studies (40.0%), interaction-timing signals (response latency, session frequency, retry patterns) in 4 (8.9%), and self-report instruments in 8 (17.8%); 15 studies (33.3%) reported no explicit primary input or described no system (Table IX). The Felder-Silverman learning styles model (Felder & Silverman, 1988) was the most frequently applied self-report framework.

TABLE 9 Primary learner-modeling input (n = 45)

Input	n	%
Assessment performance or LMS logs	18	40.0
Interaction timing	4	8.9
Self-report	8	17.8
No explicit primary input, or no system described	15	33.3

For context, du Plooy et al. (2024) reported that pre-knowledge quizzes triggered adaptation in 40 of 69 studies (58%) and learning-analytics activity logs in 15 (22%). The corpora, inclusion criteria, and coding schemes differ, so these figures cannot be compared directly with (Table IX); both sets indicate that assessment and platform-log data are the dominant triggers.

The same review reported improved academic performance in 41 studies (59%), no impact in 28 (41%), and increased engagement in 25 (36%), and noted that publication in peer-reviewed journals may favour positive results.

Interaction-timing signals appear in four studies, mostly in systems designed without LMS dependence; these form the subset relevant to messaging-only deployments. Cold start, the inability to personalize for a user with no interaction history, was discussed in eight studies and addressed directly in two: Sajja et al. (2023), which uses explicit goal elicitation and early formative tasks to build a minimal learner profile from the first five interactions, and Salim et al. (2024), which uses message timing and question repetition as proxy difficulty signals.

Neither approach has been independently replicated. In settings without an LMS, no cohort history exists to resolve cold start, so it persists across semesters. Essel et al. (2022) report that students who interacted with a chatbot teaching assistant performed better academically than students who interacted with the course instructor (post-test means 81.1 vs. 65.2, n = 68)

3.3 RQ2: Adaptivity frameworks

Machine-learning-based frameworks were the most common (17 studies, 37.8%), followed by LLM-based (12, 26.7%), hybrid (9, 20.0%), and rule-based (5, 11.1%) frameworks; 2 studies (4.4%) were not applicable (Table X).

These percentages use all 45 studies as the denominator, including studies that describe no deployed system (Table XI), so they describe the corpus rather than deployed systems. Conceptually, rule-based adaptation has the lowest infrastructure demand and runs without connectivity but is brittle when learner behavior departs from the anticipated rules; ML-based adaptation requires accumulated interaction records and offers no advantage over rules in a first semester with no history; and LLM-based adaptation offers the greatest conversational flexibility at the highest infrastructure cost.

These are design considerations rather than coded findings, and the corpus contains no controlled comparison of the three approaches. The ordering is a map of deployment feasibility, not a ranking of sophistication.

Salim et al. (2025) describe an LLM-based question-answering chatbot builder and report local deployment of a quantized open-source model on 8 GB RAM without a GPU - a configuration that matches the hardware profile of resource-constrained institutions such as Al-Asmarya. This is the most practically relevant technical result in the corpus for zero-infrastructure deployment.

El Hajji et al. (2026) describe an open-source LLM-based learning platform with offline capabilities; this work is a preprint and has not been peer reviewed, and its findings should be interpreted accordingly.

TABLE 10 Adaptivity framework (n = 45)

Framework	n	%
ML-based	17	37.8
LLM-based	12	26.7
Hybrid	9	20.0
Rule-based	5	11.1
Not applicable	2	4.4

3.4 RQ3: Implementation challenges

The Libyan sub-corpus comprises **twelve** studies. Most are surveys, scoping reviews, or acceptance studies rather than evaluations of deployed systems. Infrastructure or connectivity was named as the primary barrier in **six of the twelve**, absent institutional policy in **five**, training gaps in four, an ethics or disclosure gap in two, and inter-regional inequity in one. A single study could report more than one barrier, so these counts are not mutually exclusive. All twelve studies are discussed below.

Gratim and Elatrash (2025) applied an extended unified theory of acceptance and use of technology model to 101 academics at Al-Asmarya Islamic University and found that trust in AI and self-efficacy were the main

determinants of adoption intention. Baroud et al. (2024) found that 79.6% of 119 surveyed staff at the University of Zawia reported no AI-related professional development. Alennabi et al. (2026) surveyed 130 Libyan researchers using structural equation modelling, Random Forest, and social network analysis: AI adoption was 96.2%, consistent disclosure of AI assistance was 11.5%, and institutional policy clarity was the strongest predictor of responsible AI behaviour ($\beta = 0.45$, $p < .001$).

Aboseta and Jamah (2025) propose a staged integration roadmap and highlight inter-regional equity — the only Libyan study in the corpus to foreground geographic disparities alongside the infrastructure–training–policy sequence identified across the other studies. Atbiqa (2025) conducted a field study of Libyan media students and found broadly positive attitudes toward AI in scientific research, extending the demand-side readiness signal from faculty to students. Abdallah (2026) examined awareness and use of AI-based conversational agents among Libyan PhD students; the study is by N. S. Abdallah (earlier drafts cited the study under a different spelling, which has been corrected), and it confirms high awareness alongside limited structured use.

Abolkasim and Hasan (2024) documented widespread consent for ChatGPT integration in Libyan universities, but noted concerns about overreliance and academic integrity. Emran and Elhony (2025) identified weak educational policies, inadequate infrastructure, and insufficient training as the principal obstacles to AI implementation in Libyan higher education. Alsayd et al. (2025) examined postgraduate AI usage patterns at the University of Zawia and found that integration remains limited and underexplored despite positive dispositions. Alnagrat et al. (2025) — published under the AlQalam Journal of Medical and Applied Sciences — reported strong positive perceptions of AI integration in Libyan education, particularly for personalized learning, while flagging infrastructure and policy gaps. A 2023 study of an AI assistant in Libyan higher education documented early adoption experiences but noted the absence of institutional support structures. A 2024 study of ChatGPT use at a Libyan university reported widespread familiarity but uneven pedagogical integration. A 2025 study on the future of AI in Libyan education weighed the potential benefits of personalization against the risks of inequitable deployment.

Across these twelve studies, infrastructure limits were the barrier reported most often; training gaps and missing policy were also reported repeatedly. These counts describe how often each barrier was reported. The included studies did not test whether the barriers are causally related or occur in sequence. What they consistently document is a conditional rather than attitudinal barrier pattern: students and faculty report willingness to use AI, but limited training, unclear policy, and weak infrastructure prevent that willingness from translating into sustained, responsible adoption.

3.5 RQ4: Transferability

The LMS-dependence code "LMS free" can mean either that a deployed system runs without an LMS or merely that a study describes no system. (Table XI) therefore reclassifies the studies into four groups: (1) operational zero-LMS systems; (2) LMS-independent systems, which work with or without an LMS; (3) studies that do not specify LMS use; and (4) surveys and reviews that describe no deployed system. Four studies (8.9%) form Group 1 (El Hajji et al., 2026; Essel et al., 2022; Sajja et al., 2023; Salim et al., 2025).

Two further chatbot studies (P131 and P291) were initially tagged as LMS-free but on full-text review describe no deployed operational system and were reclassified to Group 4.

The other 20 studies coded LMS free are surveys or scoping reviews and belong to Group 4, together with the seven studies coded not applicable. Of the seven studies coded LMS optional, three describe a deployed system that integrates with an institutional platform (Group 2), and four describe surveys or conceptual work only (Group 4). Seven studies require an LMS and fall outside the four groups. Seven studies require an LMS and fall outside the four groups.

Conclusions about transferability are based on Group 1 and, once classified, Group 2; the broad "LMS free" code is not used as evidence of transferability.

TABLE 11 LMS-dependence codes and reclassified groups (n = 45)

Original code	n	Reclassified group
LMS free	24	4 to Group 1 (operational zero-LMS systems); 20 to Group 4 (surveys and reviews)
LMS optional	7	The seven studies coded LMS optional were reviewed individually. None describes a deployed LMS-independent system; all seven are surveys or scoping reviews that describe no deployed system, and they have therefore been assigned to Group 4 together with the other surveys and reviews. As a result, Group 2 (LMS-independent systems) and Group 3 (LMS use not specified) contain no studies in this corpus
Not applicable	7	Group 4 (surveys and reviews outside the LMS classification)
LMS required	7	Outside Groups 1–4: the system depends on an LMS
Total	45	

LMS independence is not the same as infrastructure independence. For example, Sajja et al. (2023) describe a framework based on GPT-3 that generates course-specific assistants from course materials; it is platform-independent but relies on an external language-model service, so it does not by itself tolerate intermittent connectivity.

Transferability to resource-constrained Arab higher education requires three conditions: platform independence, infrastructure-aware design, and cultural-linguistic adaptation. No included study was found to combine all three, and none was found to describe normalization of Arabic dialect or Latin-script Arabic (Arabizi) input before processing. These features were identified by reading the studies; they are not coded dimensions (Table IV).

A keyword check of the full texts was performed for "dialect", "Arabizi", and "MSA"; no included study reported normalization of Arabic dialect or Latin-script Arabic input before processing.

(Table XII) lists the studies that carry most of the evidence for the four questions.

TABLE XII SELECTED KEY STUDIES AND RELEVANCE RATINGS

Study	Context	Design	n	RQ1	RQ2	RQ3	RQ4	Key finding (as coded)
Gratim & Elatrash (2025)	Libya	Survey (UTAUT)	101	L	L	H	H	Trust and self-efficacy drive adoption intention.
Baroud et al. (2024)	Libya	Faculty survey	119	L	L	H	H	79.6% reported no AI-related training.
Alennabi et al. (2026)	Libya	Mixed methods (SEM+RF+SNA)	130	L	L	H	H	Policy clarity was the strongest predictor of responsible use.
Aboseta & Jamah (2025)	Libya	Scoping review	—	L	L	H	H	Staged roadmap; inter-regional equity.
Atbiqa (2025)	Libya	Field study	—	L	L	H	H	Positive attitudes toward AI in research.
Essel et al. (2022)	Ghana	Chatbot; achievement test	68	M	M	H	H	Chatbot group outperformed the instructor group (81.1 vs 65.2).
Salim et al. (2025)	Bangladesh	Chatbot builder (RAG+LLM)	—	M	H	H	H	Local deployment on 8 GB RAM, no GPU.
El Hajji et al. (2026)	Morocco	Platform (preprint)	—	H	H	M	H	Open-source LLM platform; not peer reviewed.
Sajja et al. (2023)	USA	GPT-3 framework	100 docs	H	H	M	H	Course-specific assistants; depends on external service.
du Plooy et al. (2024)	Multi-country	Scoping review	69 studies	H	H	M	M	Performance improved in 59.4%; Africa underrepresented.
Abdulraheem et al. (2025)	Iraq	Survey	—	L	L	H	H	Infrastructure + connectivity = primary barriers.
Keyhan (2025)	Iran	Qualitative	—	H	H	M	M	25 design categories for AI assistant.
Alshammari (2026)	Multi	SLR	65 studies	H	H	M	M	Digital divide + hyper-personalization frontier.

Note. RQ columns show relevance to each review question: H = High, M = Medium, L = Low, as defined in Table IV. Ratings were assigned by the single coder and should be read as exploratory classifications. This table presents a selection of the most directly relevant studies; the full corpus of 12 Libyan studies is discussed in Section 3.4.

4. DISCUSSION

4.1 Data dependence of personalization

In this corpus, 40.0% of studies used assessment or LMS-generated data as the primary input, and the scoping review of du Plooy et al. (2024) found pre-knowledge quizzes and platform logs to be the dominant triggers for adaptation. These inputs are produced by LMS or courseware platforms. Where such platforms are absent, personalization must rely on signals that exist without them, such as message timestamps, question repetition, session frequency, and stated goals; four studies in the corpus used interaction-timing signals. Whether these signals suffice for learner-state inference is not established by the corpus. Theory offers design

hypotheses rather than findings: self-regulated learning theory (Pintrich, 2004; Zimmerman, 2000) suggests that the timing and targeting of support may matter as much as its content, and cognitive load theory (Sweller, 2011) suggests limiting the information presented in each interaction. In Saudi Arabia, a mixed-methods study identified adequate resources, training and support, and clear goals among the conditions for effective personalization (Alenezi, 2023).

4.2 Cold start as a persistent condition

In recommender-system research, cold start is usually treated as a transitional problem that resolves as interaction history accumulates. In a first semester without an LMS, every user is a cold-start user, and the only source of history is the interactions themselves. A design therefore has to build a learner model from the first few exchanges or deliver uniform content. The corpus contains few studies that address this directly: Sajja et al. (2023) and Salim et al. (2024) propose cold-start strategies (explicit goal elicitation and interaction-timing signals, respectively), but neither has been replicated, and no included study describes a phased design that starts with rules and progressively adds learned models as data accumulate.

4.3 The barrier pattern in the Libyan studies

The Libyan studies report infrastructure, training, and policy barriers repeatedly. The picture is of conditional rather than attitudinal barriers: the faculty, student, and researcher samples report willingness to use AI, but limited training, unclear policy, and weak infrastructure. At Al-Asmarya Islamic University, trust and self-efficacy were the main determinants of adoption intention (Gratim & Elatrash, 2025). The ordering infrastructure, training, policy is a description of how often each barrier was reported and of how an initiative could fail at a given level; it is not an intervention roadmap and not a causal model, because none of the included studies tested the sequence.

4.4 LMS independence and transferability

The classification in (Table XI) shows that the evidence for deployed assistants outside LMS-based settings is thin: four operational systems, one of which is a non-peer-reviewed preprint (El Hajji et al., 2026), and one of which depends on a cloud language model (Sajja et al., 2023). Surveys and scoping reviews dominate the remainder of the LMS-free group. These documents are useful for describing barriers, but they cannot show that an adaptive assistant works. The absence of Arabic dialect and Arabizi handling among the included studies matters for Arabic-language deployment, where informal written input is common; this gap was identified by reading the studies, not demonstrated by an evaluation.

4.5 Limitations

Eight limitations apply. (1) Coding and screening were performed by one researcher, and initial self-agreement was 62.5% (not 62.2%; the 62.2% figure refers to the five core dimensions only); all percentages are therefore directional. (2) Three of the six search sources are search engines or networks, curated indexes were not searched, and no protocol was pre-registered, so retrieval may be incomplete and is not exactly reproducible. (3) Publication bias could not be assessed: the corpus is dominated by surveys and reviews rather than outcome evaluations, and du Plooy et al. (2024) noted that publication in peer-reviewed journals may favour positive results, so effects reported for the few evaluated systems should be read cautiously. (4) The effect of excluding the seven studies rated Low on design appropriateness has not been quantified in this version. (5) The Arabic-language sub-corpus is thinner in technical depth than the English-language corpus. (6) Evidence on cold start rests on two studies that have not been replicated. (7) Transferability evidence relies heavily on a single study from one African institution (Essel et al., 2022). (8) Some included works, such as a preprint (El Hajji et al., 2026) and recent 2026 publications, have not been through a long period of scholarly scrutiny.

5. CONCLUSION

The central constraint identified by this review is that no validated adaptive AI learning assistant has been designed from the ground up for zero-LMS, low-connectivity higher education. The findings are organized below by review question.

RQ1. Assessment performance or LMS-generated data was the most common primary input (40.0%); interaction-timing signals, which do not require an LMS, were used in four studies (8.9%). RQ2. ML-based (37.8%) and LLM-based (26.7%) frameworks predominate, while rule-based adaptation (11.1%) has the lowest infrastructure demand. RQ3. The Libyan studies most often report infrastructure, training, and policy barriers, and describe willingness to use AI; none evaluates a deployed adaptive assistant. RQ4. Only four of 45 studies report an operational system without LMS integration, and none was found to combine platform independence, infrastructure-aware design, and handling of Arabic dialect input. Limitations. Coding was done by one researcher with an initial self-agreement of 62.5%; three of the six search sources are search engines or networks and no protocol was registered; and the percentages for input type and adaptivity framework use all 45 studies, including studies that describe no system (Section 4.5). The findings are therefore best read as a map of where evidence exists and where it does not. Controlled evaluations of adaptive assistants in Arabic-language higher education without an LMS remain to be done.

6. RECOMMENDATIONS

Based on the findings of this review, the following recommendations are proposed for three stakeholder groups. For universities in resource-constrained settings, it is recommended to assess connectivity and device access prior to design selection, prioritise LMS-free and

locally hosted assistants, plan for cold start from the first interaction, provide staff training alongside a clear AI use and disclosure policy, and pilot new assistants with controlled designs while reporting null results. For researchers, future work should explicitly report LMS dependence, hosting, hardware, and connectivity assumptions, conduct controlled evaluations in Arabic-language settings without an LMS while assessing dialect and Latin-script Arabic input handling, replicate cold-start strategies, and employ a second coder and a registered protocol in subsequent reviews.

7 References.

Abdallah, N. S. (2026). Exploring Libyan PhD students' awareness and use of (AI) based conversational agents and speech technologies to improve academic writing. *International Science and Technology Journal, 38*(1), 1–17. https://doi.org/10.62341/anas951 4

Abolkasim, E., & Hasan, M. (2024). Integrating ChatGPT in education and learning: A case study on Libyan universities. *Journal of Pure and Applied Sciences, 23*(2), 19–24. https://doi.org/10.51984/jopas.v23i2.308 2

Aboseta, A. K. A., Alhaj Salem, A. A. M., & Alnagrat, A. J. A. (2025). Harnessing artificial intelligence for education reform in Libya: Opportunities and challenges. *PERFECT: Journal of Smart Algorithms, 2*(2), 78–83. https://doi.org/10.62671/perfect.v2i2.11 7

Ait Baha, T., El Hajji, M., Es-Saady, Y., & Fadili, H. (2024). The impact of educational chatbot on student learning experience. *Education and Information Technologies, 29*(8), 10153–10176. https://doi.org/10.1007/s10639-023-12166- w.

Alam, A. (2021). *Employing adaptive learning and intelligent tutoring robots*. Springer.

Alenezi, A. (2023). Personalized learning strategies in higher education in Saudi Arabia. *TEM Journal, 12*(4), 2023–2037. https://doi.org/10.18421/TEM124-1 3

Aleven, V., McLaren, B. M., Sewall, J., & Koedinger, K. R. (2016). Help helps, but only so much. *User Modeling and User-Adapted Interaction, 26*(1), 13–65.

Al-Hafdi, F. S., & AlNajdi, S. M. (2024). The effectiveness of using chatbot-based environment on learning process. *Education and Information Technologies, 29*(15), 20633–20664.

Alennabi, K., Shaaban, O., Ejmaa, A., Elbahri, E., & Mukhtar, M. (2026). Artificial intelligence in Libyan academia: Adoption, ethics, and institutional responses. *AlQalam Journal of Medical and Applied Sciences, 9*(5), 1421–1436. https://doi.org/10.54361/ajmas.26954 3

Alshammari, S. H. (2026). Integrating artificial intelligence in educational technology: A systematic review. *Journal of Educational Technology*.

Atbiqa, A. M. A. (2025). Libyan media students' attitudes toward AI technologies in scientific research. *Sirte University Journal of Humanities, 15*(2), 74–89. https://doi.org/10.37375/sujh.v15i2.367 6

Baroud, N., Alouzi, K., Elfzzani, Z., Ayad, N., & Albshkar, H. (2024). Educators' perspectives on using AI as a content creation tool in Libyan HE. *JERIT, 1*(2), 61–70.

Bilquise, G., Ibrahim, S., & Salhieh, S. M. (2023). Investigating student acceptance of an academic advising chatbot. *Education and Information Technologies, 28*(5), 5813–5836.

Booth, A., Sutton, A., & Papaioannou, D. (2016). *Systematic approaches to a successful literature review* (2nd ed.). SAGE.

Chaudhry, M. A., & Kazim, E. (2021). Artificial intelligence in education (AIEd). *AI and Ethics, 1*(4), 584–595.

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access, 8*, 75264–75278.

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education, 20*, 22. https://doi.org/10.1186/s41239-023-00392- 8

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers. In *Proceedings of NAACL-HLT 2019* (pp. 4171–4186). https://doi.org/10.18653/v1/N19-142 3

du Plooy, E., Casteleijn, D., & Franzsen, D. (2024). Personalized adaptive learning in higher education: A scoping review of key characteristics and impact on academic performance and engagement. *Heliyon, 10*(21), e39630. https://doi.org/10.1016/j.heliyon.2024.e3963 0

El Hajji, M., Ait Baha, T., Dakir, A., Fadili, H., & Es-Saady, Y. (2026). Open TutorAI: An open-source platform for personalized and immersive learning. *arXiv*. https://doi.org/10.48550/arXiv.2602.0717 6

Emran, H. A. A., & Elhony, F. M. (2025). Challenges in implementing artificial intelligence in Libyan higher education. *East Asian Journal of Multidisciplinary Research, 4*(10), 5137–5146.

Essel, H. B., Vlachopoulos, D., Tachie-Menson, A., Johnson, E. E., & Essel, P. K. (2022). The impact of a virtual teaching assistant (chatbot) on students' learning in Ghanaian higher education. *International Journal of Educational Technology in Higher Education, 19*(1), 57. https://doi.org/10.1186/s41239-022-00362- 6

Felder, R. M., & Silverman, L. K. (1988). Learning and teaching styles in engineering education. *Engineering Education, 78*(7), 674–681.

Gligorea, I., Cioca, M., Oancea, R., Gorski, A.-T., Gorski, H., & Tudorache, P. (2023). Adaptive learning using artificial intelligence in e-learning. *Education Sciences, 13*(12), 1216.

Gough, D., Oliver, S., & Thomas, J. (2017). *An introduction to systematic reviews* (2nd ed.). SAGE.

Gratim, H. M., & Elatrash, E. M. (2025). Examining faculty members' adoption of ChatGPT in Libyan higher education. *HNSJ, 6*(1), 145–167. https://doi.org/10.53796/hnsj61/1 0

Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning -Are they really useful? *Journal of Computer Assisted Learning, 38*(1), 237–257.

Kabudi, T., Pappas, I., & Olsen, D. H. (2021). AI-enabled adaptive learning systems: A systematic mapping. *Computers and Education: Artificial Intelligence, 2*, 100017.

Keyhan, J. (2025). Designing a model for an AI-based intelligent assistant. *International Journal of Technology in Education and Science, 9*(2). https://doi.org/10.46328/ijtes.61 6

Lin, C. C., Huang, A. Y. Q., & Lu, O. H. T. (2023). Artificial intelligence in intelligent tutoring systems toward sustainable education. *Smart Learning Environments, 10*, 41.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … Moher, D. (2021). The PRISMA 2020 statement. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n7 1

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning. *Educational Psychology Review, 16*(4), 385–407.

Sajja, R., Sermet, R., Cikmaz, M., Cwiertny, A., & Demir, I. (2023). Platform-independent and curriculum-oriented intelligent assistant for higher education. *Intelligent Systems with Applications, 17*, 200192. https://doi.org/10.1016/j.iswa.2022.20019 2

Salim, M. S., Hossain, S. I., Jalal, T., Bose, D. K., & Basher, M. J. I. (2025). LLM based QA chatbot builder: A generative AI-based chatbot builder for question answering. *SoftwareX, 29*, Article 102029. https://doi.org/10.1016/j.softx.2024.10202 9

Sweller, J. (2011). Cognitive load theory. *Psychology of Learning and Motivation, 55*, 37–76.

UNESCO. (2022). *Technology in education: A tool on whose terms?* UNESCO Publishing.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems. *Educational Psychologist, 46*(4), 197–221.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education, 16*, 39.

Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., … Li, Y. (2021). A review of artificial intelligence (AI) in education from 2010 to 2020. *Complexity, 2021*, 8812542.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press.