Volume 4 Number 2 (2026)
July-December 2026
Page: 525-553

Journal of Creative Power and Ambition (JCPA)
https://edujavare.com/index.php/jcpa
E-ISSN: 3031-4054
DOI: 10.70610/jcpa.1648

Transforming Islamic Religious Education Assessment through Explainable Artificial Intelligence (XAI): Developing a Transparent and Human-Centered Evaluation Model

Wulandari Retnaningrum[1], Mariana[2]

[1] Institut Agama Islam Suan Giri Ponorogo, Indonesia
[2] Universitas Nahdlatul Ulama Al Ghazali Cilacap, Indonesia
* Correspondence e-mail; mar14na1212@gmail.com, retnaningrum44@gmail.com

Article history Submitted: 2026/05/01; Revised: 2026/06/11; Accepted: 2026/07/15

Abstract

Artificial Intelligence (AI) has transformed educational assessment by enabling automated evaluation, learning analytics, and predictive decision-making. However, AI implementation in Islamic Religious Education (IRE) assessment remains limited by conventional assessment practices, algorithmic opacity, and inadequate integration of teachers' professional judgment. This study aims to develop a Human-Centered Explainable Artificial Intelligence (XAI)-Based Assessment Model that promotes transparent, accountable, and evidence-based assessment in IRE. A qualitative library research design with a conceptual model development approach was employed. Data were collected from peer-reviewed journal articles, scholarly books, and international policy documents on AI, Explainable AI, educational assessment, learning analytics, and Islamic educational philosophy. The literature was analyzed through systematic document analysis and interactive qualitative synthesis. The findings reveal that conventional assessment relies on fragmented evidence and subjective judgment, while existing AI models prioritize predictive accuracy over interpretability and accountability. The proposed framework integrates seven components: Learning Data, Machine Learning, Learning Analytics, Explainable AI, Teacher Validation, Professional Judgment, and Continuous Assessment Improvement. The model integrates Islamic values of *amanah*, *'adl*, *shūrā*, and *maslahah*, positioning AI as a transparent decision-support system that strengthens teacher-centered assessment rather than replacing professional judgment.

Keywords Explainable Artificial Intelligence; Educational Assessment; Islamic Religious Education; Human-Centered AI; Learning Analytics.

INTRODUCTION

Artificial Intelligence (AI) has become one of the most influential technological innovations driving educational transformation in the twenty-first century. The integration of AI into educational systems has shifted the paradigm from conventional instructional practices toward intelligent, data-driven educational

environments capable of supporting personalized learning, automated assessment, predictive analytics, and evidence-based decision making. Consequently, AI is no longer perceived merely as a technological innovation but as a strategic instrument for improving educational quality, institutional effectiveness, and learning outcomes (Holmes et al., 2019; UNESCO, 2023).

Recognizing this transformation, UNESCO identifies Artificial Intelligence as a strategic technology capable of accelerating educational innovation while emphasizing that its implementation must remain human-centered, transparent, accountable, and ethically responsible. Educational institutions are therefore encouraged to establish governance mechanisms that ensure algorithmic transparency, data privacy, human oversight, and responsible AI implementation rather than allowing automated systems to replace professional educational judgment (UNESCO, 2023). Similarly, the Organisation for Economic Co-operation and Development (OECD) highlights that AI significantly improves educational assessment, adaptive learning, and institutional decision-making but simultaneously introduces challenges related to algorithmic bias, fairness, accountability, and public trust (OECD, 2024).

The application of AI has expanded rapidly within educational assessment. Machine learning, learning analytics, and predictive analytics enable automated scoring, learner profiling, formative feedback generation, and early identification of learning difficulties. Unlike conventional assessment systems that primarily measure final learning outcomes, AI-supported assessment integrates multidimensional learning evidence derived from digital learning environments, thereby providing more comprehensive information for instructional improvement and educational planning (Holmes et al., 2019; OECD, 2024). These developments indicate that educational assessment is evolving from retrospective measurement toward predictive and evidence-based evaluation.

Despite these advantages, recent studies reveal that AI-based assessment continues to face substantial methodological and ethical challenges. The increasing complexity of machine learning algorithms has resulted in the emergence of so-called **black-box AI**, where highly accurate recommendations are generated without providing understandable explanations regarding how those recommendations are produced. Such opacity raises concerns about algorithmic bias, interpretability, transparency, fairness, and accountability, particularly when AI-generated decisions influence high-stakes educational outcomes (Bulut et al., 2024; Gunning & Aha, 2019). Consequently, educational institutions require governance approaches capable

of ensuring that AI recommendations remain transparent, interpretable, and professionally justifiable.

These challenges become even more significant in the context of Islamic Religious Education (Pendidikan Agama Islam—PAI). Unlike many academic subjects that primarily evaluate cognitive achievement, PAI aims to develop learners' religious understanding, moral character, spiritual awareness, worship practices, and ethical behaviour. Accordingly, assessment within PAI should evaluate not only academic performance but also affective development, behavioural change, religious practice, and character formation. However, current assessment practices in many Islamic educational institutions remain dominated by written examinations, teacher observation, and subjective judgement, while the utilization of digital learning evidence and educational analytics remains limited. As a result, assessment frequently provides only partial representations of students' holistic development.

Artificial Intelligence offers considerable opportunities to transform PAI assessment into a more adaptive and comprehensive system. Learning analytics enables educators to integrate multiple sources of educational evidence, including learning management system activities, formative assessment results, classroom participation, reflective learning journals, and longitudinal learning progress. These multidimensional datasets provide educators with richer insights into learners' development than conventional assessment methods. Nevertheless, AI-generated recommendations cannot be fully trusted if educators are unable to understand the analytical processes underlying algorithmic decisions.

To address these limitations, Explainable Artificial Intelligence (XAI) has emerged as a promising approach for enhancing transparency and interpretability within AI-supported decision-making. Explainable AI provides understandable explanations regarding the variables, reasoning processes, and confidence levels underlying algorithmic recommendations, thereby enabling educators to evaluate, validate, and justify AI-assisted assessment results. Rather than replacing teachers, XAI strengthens teacher agency by positioning Artificial Intelligence as an intelligent decision-support system operating under human supervision (Gunning & Aha, 2019; UNESCO, 2023).

Although AI in education has received considerable scholarly attention during the past five years, existing studies primarily investigate intelligent tutoring systems, educational chatbots, automated scoring, adaptive learning, and learning analytics. Comparatively fewer studies focus on AI-supported assessment, while research integrating Explainable Artificial Intelligence with Islamic Religious Education

assessment remains extremely limited. Existing literature rarely combines learning analytics, explainability, teacher professional judgement, and human-centered AI principles into a comprehensive assessment framework specifically designed for PAI.

Based on this review, three major research gaps can be identified. First, previous studies predominantly emphasize AI applications for instructional delivery rather than the transformation of assessment systems within Islamic Religious Education. Second, limited research has incorporated Explainable Artificial Intelligence to improve transparency, interpretability, and accountability in AI-supported PAI assessment. Third, no comprehensive conceptual framework has been developed that integrates learning analytics, Explainable AI, human-centered assessment, and teachers' professional judgement into a unified evaluation model for Islamic Religious Education.

Therefore, this study aims to develop a Human-Centered Explainable Artificial Intelligence (XAI)-Based Assessment Model for Islamic Religious Education that integrates learning analytics, algorithmic explainability, teacher validation, and ethical AI governance. The proposed framework extends existing AI-based educational assessment by positioning teachers as the final decision-makers while utilizing Artificial Intelligence as a transparent and accountable decision-support system. The findings are expected to contribute theoretically to the advancement of AI-assisted educational assessment and practically to the development of transparent, ethical, and evidence-based assessment practices in Islamic Religious Education.

METHODS

This study employed a qualitative library research design using a conceptual model development approach to formulate an Explainable Artificial Intelligence (XAI)-based assessment model for Islamic Religious Education (PAI). Rather than collecting primary empirical data, the study synthesized theories, empirical findings, and international policy documents to construct a comprehensive conceptual framework integrating Artificial Intelligence, Explainable AI, learning analytics, educational assessment, and Islamic educational values. This approach was selected because conceptual model development enables researchers to integrate multidisciplinary knowledge into a coherent governance framework that addresses emerging educational challenges (Creswell & Poth, 2018; Snyder, 2019).

The sources of data consisted of peer-reviewed journal articles indexed in Scopus and Web of Science, scholarly books published by internationally recognized academic publishers, conference proceedings, and official reports from UNESCO, the Organisation for Economic Co-operation and Development (OECD), the National

Institute of Standards and Technology (NIST), and the European Parliament. To ensure the relevance and quality of the reviewed literature, publications were selected based on four inclusion criteria: (1) discussing Artificial Intelligence in education, educational assessment, or educational management; (2) focusing on Explainable Artificial Intelligence (XAI), learning analytics, educational data governance, or Human-Centered AI; (3) published primarily between 2019 and 2025; and (4) directly supporting the development of an AI-based assessment framework for Islamic Religious Education. Publications that lacked scientific rigor or were not relevant to the research objectives were excluded from the analysis (Boell & Cecez-Kecmanovic, 2015).

Data were collected using systematic document analysis. Relevant publications were identified, reviewed, classified, and coded according to major analytical themes, including educational assessment, AI governance, Explainable Artificial Intelligence, learning analytics, educational data governance, human-centered AI, and Islamic educational assessment. The coded information was subsequently synthesized to identify conceptual relationships among the principal components of the proposed assessment model (Bowen, 2009).

The collected data were analyzed using the interactive qualitative analysis model proposed by Miles, Huberman, and Saldaña (2020), consisting of four interconnected stages: data collection, data condensation, data display, and conclusion drawing/verification. During data condensation, concepts obtained from multiple literature sources were critically compared, interpreted, and synthesized to identify recurring patterns, theoretical convergence, and conceptual relationships. The synthesized findings were then integrated into a Human-Centered Explainable Artificial Intelligence framework for Islamic Religious Education assessment. To enhance the credibility of the proposed model, theoretical triangulation was conducted by comparing perspectives from educational assessment, Artificial Intelligence, Explainable AI, educational management, and Islamic educational philosophy, while source triangulation was applied through the integration of international policy reports and peer-reviewed empirical studies.

Figure 1 illustrates the conceptual research framework guiding the development of the proposed Explainable Artificial Intelligence-based assessment model for Islamic Religious Education.

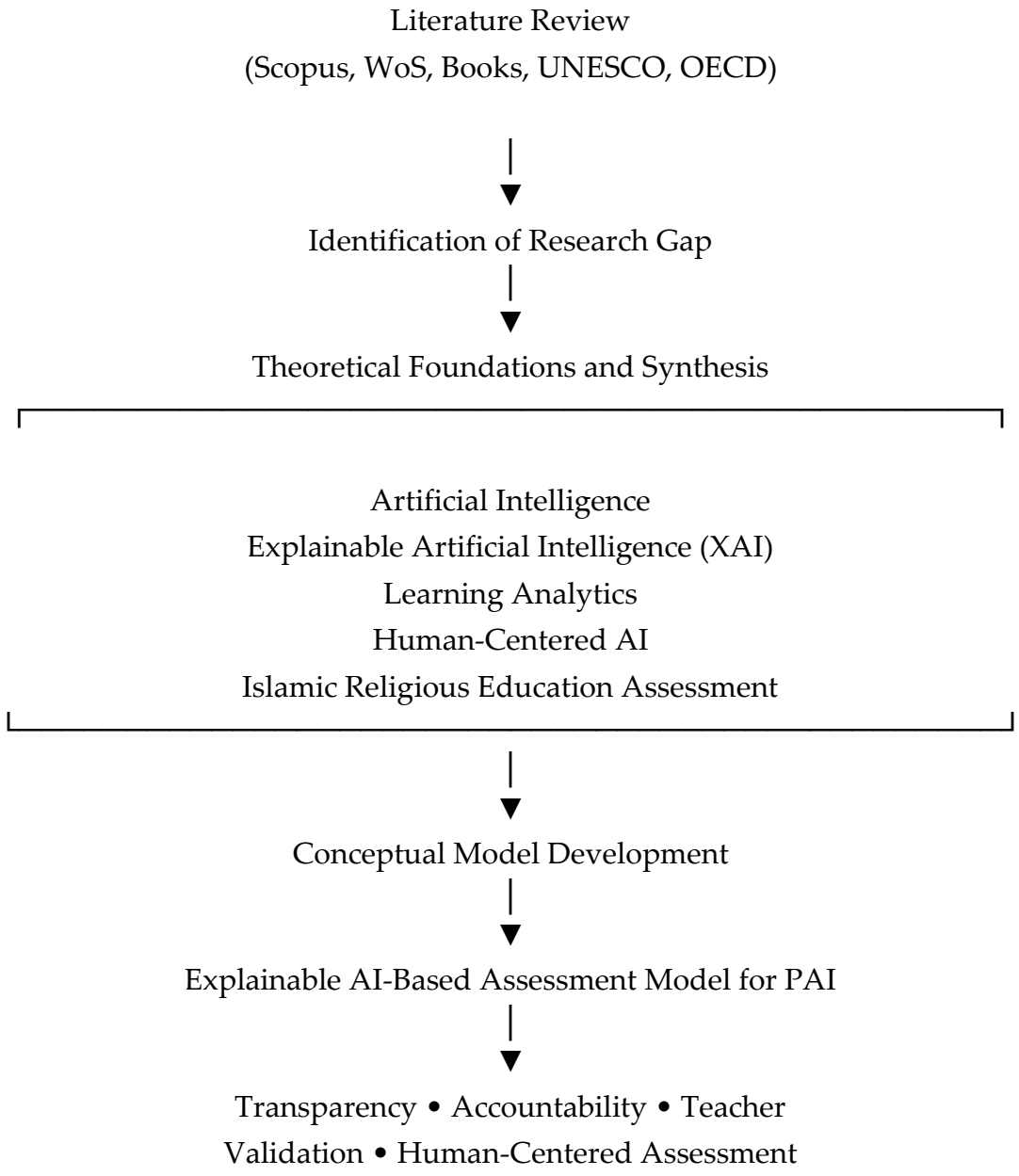


Figure 1. Research framework for developing an Explainable Artificial Intelligence-based assessment model for Islamic Religious Education.

FINDINGS AND DISCUSSION

Limitations of Conventional Assessment in Islamic Religious Education

Assessment constitutes one of the most fundamental components of the educational process because it provides evidence regarding students' learning progress, instructional effectiveness, and curriculum achievement. Beyond measuring learning outcomes, assessment functions as a strategic mechanism that guides instructional planning, supports educational decision-making, and promotes continuous improvement in teaching and learning. Contemporary educational assessment has gradually shifted from emphasizing the measurement of learning

outcomes (*assessment of learning*) toward supporting learning processes (*assessment for learning*) and fostering students' self-regulation (*assessment as learning*). This paradigm shift requires assessment systems capable of collecting multidimensional evidence that accurately reflects learners' cognitive, affective, and behavioral development rather than relying exclusively on standardized testing (Black & Wiliam, 2018; OECD, 2024).

Within Islamic Religious Education (Pendidikan Agama Islam—PAI), assessment performs an even broader educational function because the objectives of PAI extend beyond academic achievement to include the development of faith (*aqīdah*), worship practices (*ibādah*), moral character (*akhlāq*), religious attitudes, and social responsibility. National educational standards in Indonesia emphasize that PAI should develop students holistically by integrating knowledge, values, attitudes, and religious practices into everyday life. Consequently, assessment in PAI should evaluate not only what students know about Islam but also how Islamic values are internalized and manifested through behavior, habits, ethical decision-making, and spiritual development. Such multidimensional educational objectives require assessment approaches capable of capturing complex learning processes rather than merely recording examination scores (Ministry of Religious Affairs of the Republic of Indonesia, 2023).

Despite these comprehensive objectives, assessment practices in many Islamic educational institutions continue to rely predominantly on conventional assessment methods. Written examinations, oral questioning, teacher observation, assignments, and portfolio assessment remain the primary instruments used to evaluate student achievement. These approaches undoubtedly provide valuable information regarding learners' cognitive understanding and classroom participation; however, they frequently depend upon periodic evaluation rather than continuous evidence collection. Consequently, assessment often emphasizes measurable academic performance while providing only limited information concerning students' long-term moral, spiritual, and behavioral development (Darling-Hammond & Wentworth, 2020).

One of the principal limitations of conventional assessment concerns its dependence on fragmented and isolated sources of evidence. Teachers generally assess students based on classroom observations, written tests, homework, and limited behavioral records accumulated during specific instructional sessions. Such data rarely capture students' learning trajectories over extended periods or across multiple educational contexts. Moreover, many important indicators of religious

learning, including consistency in worship practices, ethical interactions, reflective thinking, and spiritual growth, are inherently dynamic and difficult to evaluate using traditional assessment instruments. As a result, assessment outcomes frequently represent isolated snapshots of student performance rather than comprehensive representations of learners' educational development (Brookhart, 2020).

Another significant limitation involves the subjectivity associated with teacher judgment. Although professional judgment remains indispensable within educational assessment, observational evaluation is frequently influenced by individual experience, personal interpretation, instructional workload, and contextual constraints. Teachers often encounter difficulties maintaining consistency when assessing affective and behavioral dimensions because evaluation criteria may be interpreted differently across classrooms or educational institutions. Variations in teacher perceptions, combined with the absence of systematic evidence integration, may reduce assessment reliability and generate inconsistencies in educational decision-making. Consequently, ensuring fairness and objectivity remains a persistent challenge within Islamic Religious Education assessment (Brown & Harris, 2019).

The increasing complexity of educational environments further amplifies these challenges. Digital learning platforms, blended learning environments, online discussions, electronic portfolios, and learning management systems continuously generate substantial volumes of educational data describing students' learning behaviors, participation patterns, formative assessment results, and collaborative activities. Nevertheless, much of this information remains underutilized within conventional assessment practices because educators often lack appropriate analytical tools capable of transforming raw educational data into meaningful pedagogical insights. Consequently, valuable learning evidence that could support more comprehensive and personalized assessment frequently remains disconnected from institutional decision-making processes (Ifenthaler & Yau, 2020).

These limitations become particularly evident when educational assessment is expected to support evidence-based instructional improvement. Contemporary educational policies increasingly encourage teachers to formulate instructional decisions using empirical learning evidence rather than intuition alone. However, conventional assessment systems generally provide descriptive information regarding students' current achievement while offering limited capacity to identify learning patterns, predict future difficulties, or recommend targeted interventions. This reactive orientation restricts teachers' ability to implement timely instructional

adjustments and personalized educational support, especially within classrooms characterized by diverse learner needs (OECD, 2024).

From the perspective of Islamic educational philosophy, assessment should function not merely as an administrative procedure but as an integral component of educational guidance (*tarbiyah*). Evaluation should encourage continuous self-improvement, strengthen learners' moral responsibility, and facilitate holistic human development consistent with the objectives (*maqāṣid*) of Islamic education. Therefore, assessment practices that emphasize numerical achievement while overlooking spiritual growth, ethical conduct, and character formation only partially fulfill the broader mission of Islamic education. The discrepancy between the multidimensional objectives of PAI and the relatively limited capacity of conventional assessment systems indicates the necessity for more comprehensive, adaptive, and evidence-informed assessment approaches capable of integrating diverse forms of educational evidence into meaningful professional judgment.

Taken together, these limitations demonstrate that conventional assessment in Islamic Religious Education faces substantial methodological and pedagogical constraints. Fragmented educational data, dependence on periodic evaluation, subjectivity in teacher judgment, limited utilization of digital learning evidence, and insufficient support for evidence-based decision-making reduce the capacity of assessment to represent learners' holistic development. These challenges create an important opportunity for emerging technologies, particularly Artificial Intelligence, learning analytics, and Explainable Artificial Intelligence, to complement rather than replace teachers' professional expertise by providing more comprehensive, transparent, and human-centered assessment systems. The following section therefore discusses how Artificial Intelligence has emerged as a transformative force capable of addressing these limitations and redefining educational assessment within Islamic Religious Education.

The limitations of conventional assessment identified in the previous section indicate that the challenge is not merely methodological but also conceptual. Contemporary educational systems increasingly demand assessment practices capable of generating valid, comprehensive, and actionable evidence to support instructional improvement. However, conventional assessment in Islamic Religious Education (PAI) remains predominantly teacher-centered, episodic, and retrospective, providing information about what students have achieved rather than explaining how learning develops or identifying factors that influence educational progress. Consequently, assessment frequently functions as a mechanism for

reporting achievement rather than as an instrument for improving learning quality and educational decision-making (Schildkamp, 2019).

Recent educational assessment literature consistently argues that assessment should be integrated with evidence-based instructional decision-making. Rather than functioning as an isolated activity conducted at the end of instruction, assessment should continuously collect, interpret, and utilize multiple forms of learning evidence to support adaptive teaching strategies. Black and Wiliam (2018) emphasize that effective assessment promotes learning when educational evidence is systematically interpreted and transformed into instructional actions. Similarly, OECD (2024) argues that assessment systems should support continuous educational improvement by integrating diverse learning indicators capable of describing learners' cognitive, behavioral, and affective development. These perspectives indicate that educational assessment should evolve from a product-oriented evaluation toward a continuous learning intelligence system.

Nevertheless, the implementation of such assessment principles remains challenging within Islamic Religious Education because many dimensions of religious learning are difficult to quantify using conventional instruments. Indicators such as sincerity in worship, moral responsibility, honesty, empathy, social interaction, and spiritual awareness develop gradually through continuous educational experiences and therefore require longitudinal observation rather than one-time measurement. Existing classroom assessment instruments generally provide only fragmented evidence that is insufficient to represent the complexity of learners' religious development. Consequently, educational decisions based solely on written examinations or periodic teacher observations may fail to capture the holistic objectives of Islamic education.

Another important issue concerns the increasing diversity of educational data generated during contemporary learning processes. Digital transformation has significantly expanded the amount of information available to educators through Learning Management Systems (LMS), online assessments, electronic portfolios, discussion forums, classroom interaction records, attendance systems, and digital learning analytics. These educational datasets potentially provide comprehensive insights regarding learning engagement, participation patterns, collaborative behavior, formative achievement, and learning progression. However, conventional assessment practices rarely integrate these diverse data sources into coherent evaluation systems because teachers generally lack technological support capable of processing complex educational information efficiently (Ifenthaler & Yau, 2020). As a

result, substantial educational evidence remains unused despite its potential contribution to improving assessment quality.

This situation demonstrates a significant discrepancy between the availability of educational data and the capacity of conventional assessment systems to utilize those data effectively. Educational institutions are increasingly rich in digital information, yet many assessment decisions continue to depend primarily on teachers' subjective interpretation of limited evidence. This discrepancy illustrates that the principal challenge is no longer the absence of educational data but rather the inability to transform abundant educational information into meaningful knowledge that supports professional judgment and instructional improvement. Therefore, educational assessment requires analytical approaches capable of integrating multiple learning indicators into comprehensive evidence while preserving teachers' pedagogical authority.

The findings of previous studies also reveal that the integration of Artificial Intelligence into educational assessment has largely focused on improving technical efficiency rather than addressing broader pedagogical and ethical considerations. Existing AI applications commonly emphasize automated scoring, intelligent tutoring systems, adaptive learning environments, plagiarism detection, and predictive learning analytics (Holmes et al., 2019). Although these innovations significantly improve assessment efficiency and analytical capacity, they frequently prioritize algorithmic performance while paying comparatively less attention to transparency, interpretability, teacher validation, and ethical accountability. Consequently, AI-supported assessment often functions as an automated prediction system rather than as an educational decision-support framework.

Recent systematic reviews further indicate that the majority of AI research in education concentrates on learning personalization, student performance prediction, and educational analytics, whereas relatively few studies investigate how Artificial Intelligence can strengthen teachers' professional assessment practices through explainable and human-centered approaches (Zawacki-Richter et al., 2019). Even fewer studies address AI implementation within Islamic Religious Education, where educational assessment encompasses cognitive, affective, spiritual, and moral dimensions simultaneously. This imbalance demonstrates that current research has not fully addressed the distinctive assessment requirements of value-based educational systems.

From a governance perspective, another critical limitation concerns the absence of transparency in AI-assisted educational assessment. Highly sophisticated

machine learning algorithms frequently operate as black-box systems, generating recommendations without clearly explaining the relationships among input variables, prediction mechanisms, and final assessment outcomes. While such algorithms may demonstrate high predictive accuracy, educators often experience difficulty understanding why specific recommendations are generated or how algorithmic decisions should be interpreted within particular educational contexts (Gunning & Aha, 2019). This lack of interpretability potentially reduces teachers' confidence in AI-generated recommendations and weakens institutional accountability because assessment decisions cannot easily be justified to students, parents, or school administrators.

These observations suggest that the future development of educational assessment should not focus exclusively on technological sophistication but also on strengthening educational transparency, explainability, and human agency. Artificial Intelligence should function as an intelligent assistant that enhances teachers' analytical capacity rather than replacing professional judgment. Educational technologies therefore require governance mechanisms ensuring that algorithmic recommendations remain understandable, verifiable, and pedagogically meaningful. Such principles are particularly important in Islamic Religious Education, where educational decisions should always reflect ethical responsibility, fairness, and the broader objectives of human development.

Based on this critical analysis, it can be argued that the limitations of conventional assessment create an important opportunity for the integration of Artificial Intelligence into Islamic Religious Education. However, the adoption of AI should not simply automate existing assessment procedures; instead, it should transform assessment into a transparent, evidence-based, and human-centered process that complements teachers' expertise. This transformation requires the integration of learning analytics, machine learning, educational data, teacher validation, and explainable decision-making within a unified conceptual framework. Accordingly, the following section discusses how Artificial Intelligence can function as a transformative technology capable of overcoming the methodological limitations of conventional assessment while supporting more comprehensive, transparent, and accountable evaluation practices in Islamic Religious Education.

Artificial Intelligence as a Driver of Assessment Transformation

The limitations of conventional assessment have encouraged educational institutions worldwide to explore innovative approaches capable of generating more comprehensive, accurate, and evidence-based evaluation systems. Among the most

transformative innovations, Artificial Intelligence (AI) has emerged as a catalyst for reshaping educational assessment by enabling institutions to collect, analyze, and interpret large-scale educational data beyond the capacity of conventional human observation. Rather than functioning merely as an automation tool, AI represents a paradigm shift from retrospective assessment toward intelligent, adaptive, and predictive evaluation systems capable of supporting continuous educational improvement (Holmes et al., 2019; UNESCO, 2023).

The transformation introduced by AI is closely associated with the rapid expansion of digital learning ecosystems. Learning Management Systems (LMS), virtual classrooms, mobile learning applications, online assessment platforms, and educational information systems continuously generate extensive educational data reflecting students' learning behaviors, participation patterns, assessment performance, collaborative activities, and learning trajectories. Unlike traditional assessment approaches that primarily utilize isolated examination scores, Artificial Intelligence processes multidimensional educational evidence to identify meaningful learning patterns and generate actionable insights for educators and institutional leaders (OECD, 2024). Consequently, assessment evolves from measuring isolated learning outcomes toward understanding the entire learning process.

One of the principal technologies driving this transformation is machine learning, a branch of Artificial Intelligence that enables computational systems to recognize patterns, learn from historical educational data, and improve predictive performance without explicit programming. Within educational assessment, machine learning algorithms analyze relationships among multiple learner characteristics, including assessment scores, learning engagement, attendance, online participation, assignment completion, collaborative interaction, and behavioral indicators. Through iterative learning processes, these algorithms generate predictive models capable of identifying students at risk of academic difficulties, forecasting future learning achievement, and recommending personalized instructional interventions (Luckin et al., 2016; Holmes et al., 2019).

Machine learning substantially expands the analytical capacity of educational assessment by integrating variables that would be difficult for teachers to evaluate simultaneously. Conventional assessment generally considers only a limited number of observable indicators during classroom instruction, whereas machine learning synthesizes hundreds or even thousands of interconnected variables within large educational datasets. Such computational capability enables educational institutions to move beyond descriptive evaluation toward predictive educational intelligence,

where learning difficulties can be identified before they become significant barriers to student achievement (Baker & Inventado, 2014). This predictive capacity represents one of the most significant contributions of Artificial Intelligence to educational transformation because it supports proactive rather than reactive instructional decision-making.

Complementing machine learning, learning analytics has become another major driver of assessment transformation. Learning analytics refers to the systematic measurement, collection, analysis, and reporting of educational data for the purpose of understanding and optimizing learning and the environments in which learning occurs (Siemens & Long, 2011). Unlike conventional assessment, which primarily evaluates final achievement, learning analytics continuously monitors learners' educational experiences throughout the learning process. Information derived from online learning activities, digital assessments, classroom interactions, discussion forums, and electronic portfolios provides educators with comprehensive evidence regarding learning engagement, motivation, participation, persistence, and conceptual understanding.

The integration of learning analytics significantly enhances teachers' ability to understand learner diversity. Educational data generated through digital platforms reveal not only whether students answer questions correctly but also how they approach learning tasks, how frequently they participate in discussions, how much time they spend completing assignments, and how learning behaviors evolve throughout an academic semester. Such multidimensional evidence enables teachers to identify individual learning needs, detect early signs of disengagement, and provide personalized instructional support that would be difficult to achieve through conventional assessment alone (Ifenthaler & Yau, 2020). Consequently, assessment becomes a continuous source of pedagogical intelligence rather than merely an administrative reporting mechanism.

Another important development concerns the implementation of automated assessment**,** which applies Artificial Intelligence to evaluate educational performance efficiently while maintaining consistency and scalability. AI-supported automated assessment has been widely implemented for objective testing, essay scoring, formative feedback generation, language assessment, and adaptive testing. Automated scoring systems significantly reduce teachers' administrative workload by providing immediate evaluation and individualized feedback, allowing educators to allocate greater attention to instructional planning and learner support (Zawacki-Richter et al., 2019). Moreover, AI-assisted feedback systems facilitate continuous

formative assessment by delivering personalized recommendations immediately after learning activities, thereby strengthening students' self-regulated learning and instructional responsiveness.

Despite these operational advantages, automated assessment should not be interpreted as replacing teachers' professional judgment. Instead, AI functions as an intelligent decision-support system that augments educators' analytical capacity while preserving human responsibility for educational evaluation. UNESCO (2023) emphasizes that Artificial Intelligence should strengthen teacher agency, enabling educators to make better-informed decisions through evidence generated by intelligent technologies rather than delegating educational authority entirely to computational systems. Human-centered AI therefore remains a fundamental principle of responsible educational transformation.

Beyond descriptive and automated assessment, Artificial Intelligence has also introduced predictive assessment, representing one of the most advanced developments in educational evaluation. Predictive assessment utilizes historical educational data together with machine learning algorithms to estimate future learning performance, identify students at risk of academic failure, recommend targeted interventions, and support institutional planning. Rather than waiting until students experience learning difficulties, predictive assessment enables educators to intervene proactively based on probabilistic forecasts generated through educational analytics (OECD, 2024). This capability significantly enhances educational responsiveness by allowing instructional strategies to be adjusted before learning problems become irreversible.

The emergence of predictive assessment fundamentally changes the role of educational assessment within institutional decision-making. Assessment is no longer limited to documenting past achievement but becomes an intelligent forecasting mechanism supporting evidence-based educational planning. School leaders and teachers may utilize predictive insights to optimize curriculum implementation, allocate learning resources more effectively, identify vulnerable learners, and improve institutional quality assurance through continuous monitoring of educational performance. Consequently, Artificial Intelligence transforms assessment into a strategic component of educational governance rather than merely an instrument for measuring academic achievement.

Nevertheless, while machine learning, learning analytics, automated assessment, and predictive assessment substantially improve educational intelligence, they simultaneously introduce new ethical and pedagogical challenges.

AI-generated recommendations often rely on highly complex computational models whose reasoning processes remain inaccessible to teachers and educational stakeholders. Although predictive accuracy may improve significantly, educators frequently encounter difficulty understanding how algorithmic recommendations are produced or determining whether such recommendations are pedagogically appropriate within specific educational contexts. Therefore, technological sophistication alone is insufficient to establish trustworthy educational assessment. Artificial Intelligence must also ensure transparency, interpretability, and meaningful human oversight so that teachers remain capable of validating, questioning, and contextualizing AI-generated recommendations before translating them into educational decisions.

These emerging challenges indicate that the next stage of assessment transformation extends beyond improving algorithmic performance toward developing Artificial Intelligence systems that are transparent, understandable, and accountable. Such requirements have stimulated the rapid development of Explainable Artificial Intelligence (XAI), which seeks to make AI-generated decisions interpretable for educators and other stakeholders. Accordingly, the following section discusses how Explainable Artificial Intelligence addresses the limitations of conventional AI by strengthening transparency, professional judgment, and human-centered assessment within Islamic Religious Education.

Artificial Intelligence as a Driver of Assessment Transformation

The transformation of educational assessment through Artificial Intelligence has attracted substantial scholarly attention over the past decade. Existing research consistently demonstrates that AI significantly enhances educational assessment by improving efficiency, scalability, personalization, and predictive capability. However, a closer examination of recent literature indicates that the majority of AI-based assessment studies continue to prioritize computational performance rather than educational transparency and pedagogical accountability. Consequently, although AI has become increasingly sophisticated in processing educational data, its contribution to educational decision-making remains constrained by limited explainability and insufficient integration with teachers' professional judgment.

Holmes et al. (2019) argue that Artificial Intelligence should be understood as an educational partner rather than an autonomous decision-maker. Their work emphasizes that AI possesses considerable potential to support teachers through automated feedback, adaptive learning environments, and intelligent assessment systems. Nevertheless, they also caution that educational decisions involve

pedagogical, ethical, and contextual considerations that cannot be fully replicated through computational algorithms alone. Similarly, UNESCO (2023) proposes a human-centered approach to educational AI in which intelligent technologies augment teachers' professional capacity while preserving educators' authority over instructional planning, assessment, and learner development. These perspectives collectively reinforce the principle that AI should function as an evidence-generation mechanism rather than as a replacement for professional educational judgment.

Recent developments in learning analytics further strengthen this argument by demonstrating that educational data can provide valuable insights into learners' cognitive progress, engagement patterns, and instructional effectiveness. Through continuous analysis of digital learning activities, AI enables teachers to monitor learning progression, identify students requiring additional support, and personalize instructional interventions based on empirical evidence (Ifenthaler & Yau, 2020). Unlike conventional assessment, which frequently provides static information regarding learning outcomes, AI-supported analytics facilitate dynamic monitoring throughout the learning process. Consequently, educational assessment evolves from measuring achievement toward supporting continuous learning improvement.

Despite these advantages, existing empirical studies also reveal several persistent limitations within AI-supported assessment. First, most AI applications focus primarily on cognitive assessment because objective academic performance is comparatively easier to quantify using machine learning algorithms. In contrast, affective development, moral reasoning, ethical behavior, and spiritual growth remain considerably more difficult to represent through computational models. This imbalance is particularly problematic within Islamic Religious Education (PAI), where educational objectives extend beyond knowledge acquisition toward the formation of character, religious commitment, and moral responsibility. Accordingly, AI systems designed exclusively for cognitive prediction may fail to represent the holistic educational outcomes expected in Islamic education.

Second, previous studies frequently evaluate AI systems based on predictive accuracy rather than educational usefulness. High-performing algorithms are generally assessed according to statistical indicators such as classification accuracy, precision, recall, or prediction error. Although these indicators are important from a computational perspective, they do not necessarily indicate whether AI-generated recommendations are pedagogically meaningful, ethically acceptable, or professionally interpretable. Educational assessment differs fundamentally from many other AI applications because assessment outcomes influence instructional

planning, student evaluation, promotion decisions, and educational opportunities. Consequently, algorithmic performance alone cannot become the sole criterion for evaluating educational AI systems (Floridi & Cowls, 2022).

Third, the literature indicates that transparency remains one of the least developed dimensions of AI-supported educational assessment. Many machine learning models—including deep neural networks and ensemble learning algorithms—operate through highly complex computational processes that are difficult for educators to interpret. Although these systems frequently generate accurate predictions, teachers often receive only the final recommendation without understanding the reasoning process that produced it. This situation creates significant challenges for educational accountability because teachers cannot adequately explain or justify AI-supported decisions to students, parents, school leaders, or policymakers (Gunning & Aha, 2019). Consequently, algorithmic opacity reduces institutional trust and limits the practical adoption of AI within educational settings.

These limitations become even more significant within the context of Islamic Religious Education. Unlike standardized academic assessment, PAI requires evaluation systems capable of interpreting students' cognitive understanding together with behavioral consistency, ethical conduct, religious practice, social interaction, and spiritual development. Such multidimensional educational objectives require professional interpretation informed by pedagogical experience, contextual understanding, and ethical reasoning. Artificial Intelligence may successfully identify behavioral patterns and learning trends; however, determining whether those patterns genuinely reflect learners' moral development continues to require human interpretation grounded in educational values. Therefore, AI should strengthen rather than replace teachers' professional judgment within Islamic educational assessment.

From the perspective of Islamic educational philosophy, teachers occupy a central position as murabbī**,** mu'allim, and mu'addib, whose responsibilities extend beyond transmitting knowledge toward nurturing character, guiding moral development, and cultivating spiritual awareness. Assessment therefore constitutes not merely a technical procedure but an educational responsibility closely associated with the principles of amanah (trustworthiness), 'adl (justice), ihsan (excellence), and maslahah (public benefit). Delegating educational decisions entirely to opaque computational systems would therefore contradict the humanistic orientation of Islamic education, which emphasizes wisdom (*hikmah*), consultation (*shūrā*), and

ethical responsibility as essential components of educational leadership (Al-Attas, 1999; Kamali, 2008).

These observations indicate that future educational assessment should not pursue automation as an end in itself but should seek an appropriate balance between computational intelligence and professional educational expertise. Artificial Intelligence possesses remarkable capacity to process educational data, identify hidden patterns, and generate predictive insights that exceed human analytical capability. Nevertheless, educational decisions require interpretability, contextual understanding, ethical reflection, and professional responsibility that remain fundamentally human competencies. Consequently, trustworthy educational assessment requires AI systems capable of supporting—not replacing—teachers through transparent, interpretable, and accountable analytical processes.

The synthesis of previous studies therefore reveals an important conceptual gap. While current research has successfully demonstrated the analytical capabilities of Artificial Intelligence, comparatively little attention has been devoted to ensuring that AI-generated recommendations remain understandable and professionally justifiable for educators. Existing AI applications largely emphasize prediction accuracy, automation, and efficiency, whereas issues concerning explainability, algorithmic transparency, human oversight, and teacher validation remain insufficiently explored, particularly within Islamic Religious Education. Addressing this gap requires moving beyond conventional Artificial Intelligence toward systems specifically designed to make algorithmic reasoning transparent and interpretable.

Accordingly, the transformation of educational assessment enters a new phase through the emergence of Explainable Artificial Intelligence (XAI). Rather than focusing solely on prediction, XAI seeks to reveal how AI reaches particular conclusions, which variables influence decision-making, and why specific recommendations are generated. By making algorithmic reasoning understandable, XAI enables teachers to evaluate, verify, and contextualize AI-generated evidence before translating it into educational decisions. This capability is particularly relevant for Islamic Religious Education because it preserves teachers' pedagogical authority while benefiting from the analytical power of Artificial Intelligence. Therefore, the following section discusses Explainable Artificial Intelligence as the conceptual foundation for developing transparent, accountable, and human-centered assessment within Islamic Religious Education.

Development of the Proposed XAI-Based Assessment Model

The findings of this study resulted in the development of a Human-Centered Explainable Artificial Intelligence (XAI)-Based Assessment Model for Islamic Religious Education (PAI). The proposed framework was developed through the synthesis of educational assessment theory, Artificial Intelligence, Explainable Artificial Intelligence (XAI), learning analytics, and Islamic educational philosophy. Unlike existing AI-supported assessment models that primarily emphasize automation and predictive accuracy, the proposed model positions transparency, teacher professional judgment, and ethical accountability as integral components of the assessment process. Consequently, Artificial Intelligence is conceptualized not as an autonomous evaluator but as an intelligent decision-support system that strengthens teachers' capacity to conduct transparent, evidence-based, and human-centered assessment (Holmes et al., 2019; UNESCO, 2023).

The model consists of seven interrelated components that collectively transform conventional assessment into an intelligent assessment ecosystem. These components include Learning Data**,** Machine Learning**,** Learning Analytics**,** Explainable Artificial Intelligence**,** Teacher Validation**,** Professional Judgment**,** and Continuous Assessment Improvement**.** Rather than operating independently, each component functions sequentially while maintaining continuous interaction through iterative feedback, enabling assessment to become adaptive, transparent, and continuously refined.

The first component, Learning Data, represents the foundation of the proposed assessment model. Educational assessment should no longer rely exclusively on examination scores but instead integrate multiple forms of learning evidence generated throughout the educational process. These include formative assessment results, classroom participation, Learning Management System (LMS) activities, online discussions, digital portfolios, reflective journals, attendance records, collaborative learning behavior, and students' longitudinal learning progress. Such multidimensional educational data provide a comprehensive representation of learners' cognitive, affective, behavioral, and spiritual development, thereby reflecting the holistic objectives of Islamic Religious Education (Ifenthaler & Yau, 2020). Accordingly, assessment shifts from episodic measurement toward continuous evidence collection.

The second component, Machine Learning, functions as the analytical engine that processes educational data into meaningful patterns. Machine learning algorithms identify relationships among learner characteristics, classify educational

behaviors, detect learning difficulties, and generate predictive information regarding students' future learning performance. Unlike conventional statistical analysis, machine learning continuously improves its predictive capability through iterative learning from historical educational data. This enables educators to recognize learning patterns that may not be observable through conventional classroom assessment alone (Luckin et al., 2016). Nevertheless, within the proposed model, machine learning serves exclusively as an analytical instrument rather than an autonomous decision-maker.

Following computational analysis, educational evidence enters the third component, Learning Analytics. Whereas machine learning focuses on prediction, learning analytics emphasizes educational interpretation. Learning analytics transforms algorithmic outputs into pedagogically meaningful information by describing learners' engagement, participation, conceptual understanding, collaborative interaction, and learning progression. Through visual dashboards, learning indicators, and longitudinal reports, learning analytics provides teachers with comprehensive insights into students' educational development. Consequently, assessment extends beyond measuring learning outcomes toward understanding the processes through which learning occurs (Siemens & Long, 2011).

Although machine learning and learning analytics substantially enhance educational intelligence, they remain insufficient if teachers cannot understand the reasoning underlying AI-generated recommendations. Therefore, the fourth component, Explainable Artificial Intelligence (XAI), constitutes the principal innovation of the proposed framework. Explainable AI provides interpretable explanations concerning how algorithmic recommendations are generated, which educational variables influence predictions, the relative importance of each indicator, and the confidence level associated with algorithmic outputs. Rather than presenting teachers with unexplained recommendations, XAI enables educators to examine the logical relationships connecting educational evidence with assessment outcomes (Gunning & Aha, 2019). Explainability therefore transforms Artificial Intelligence from a computational "black box" into a transparent educational partner.

The fifth component, Teacher Validation, reinforces the principle that educational assessment should remain under professional human supervision. AI-generated recommendations are not accepted automatically but undergo validation by teachers, who evaluate their pedagogical relevance, contextual appropriateness, ethical implications, and consistency with students' actual learning conditions. This validation process ensures that algorithmic outputs are interpreted within the

broader educational context, recognizing that many aspects of Islamic Religious Education—including sincerity, moral responsibility, religious commitment, and spiritual development—cannot be fully represented through computational data alone. Teacher validation therefore strengthens educational accountability while preserving professional autonomy (UNESCO, 2023).

Building upon validated evidence, the sixth component, Professional Judgment, positions teachers as the final decision-makers within the assessment process. Professional judgment integrates AI-generated evidence with pedagogical experience, classroom observation, ethical reflection, institutional policy, and Islamic educational values. Unlike fully automated assessment systems, the proposed framework explicitly maintains that final educational decisions should remain the responsibility of educators. This principle reflects UNESCO's concept of human-centered Artificial Intelligence, which emphasizes that AI should augment rather than replace human expertise (UNESCO, 2023). Within Islamic educational philosophy, professional judgment also embodies the principles of amanah (trustworthiness), *'adl* (justice), *shūrā* (consultative deliberation), and *maslahah* (public benefit), ensuring that educational decisions remain ethically grounded and socially responsible (Al-Attas, 1999; Kamali, 2008).

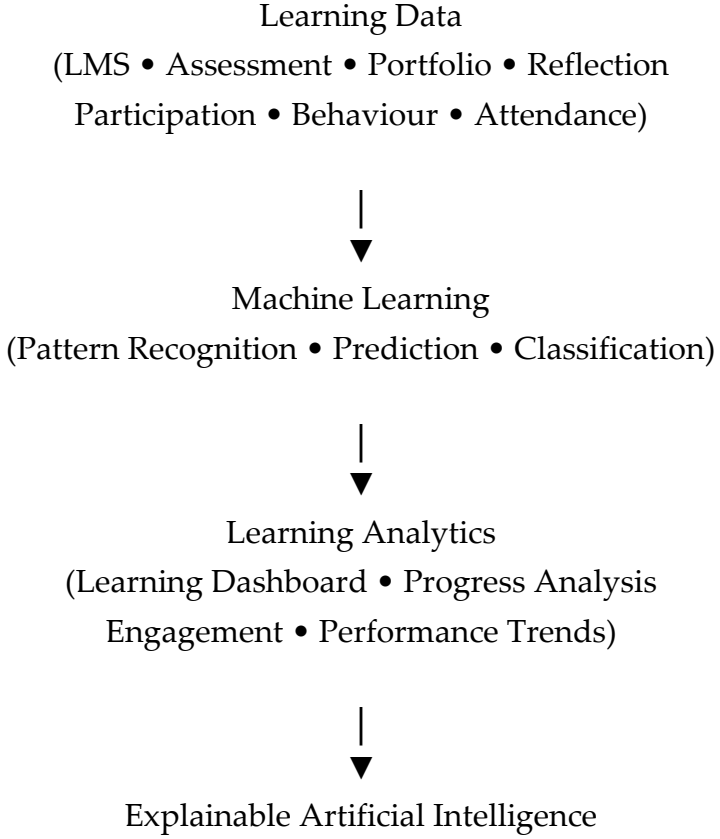
The final component, Continuous Assessment Improvement, transforms assessment into an ongoing cycle of institutional learning. Assessment outcomes, teacher feedback, learner responses, and AI performance evaluations continuously generate new educational evidence that feeds back into the learning data repository. This iterative process enables machine learning algorithms to improve predictive performance while simultaneously allowing teachers to refine instructional strategies, assessment criteria, and pedagogical interventions. Consequently, assessment becomes a dynamic process characterized by continuous organizational learning rather than periodic evaluation.

Figure 2 illustrates the conceptual relationships among these seven components. The framework operates through a cyclical process beginning with multidimensional learning data, followed by computational analysis through machine learning, pedagogical interpretation using learning analytics, transparent explanation through Explainable Artificial Intelligence, professional validation by teachers, educational decision-making based on professional judgment, and continuous improvement through institutional feedback. Unlike previous AI-supported assessment models that conclude with algorithmic recommendations, the proposed framework introduces teacher validation and professional judgment as

mandatory stages before educational decisions are implemented. This human-centered architecture constitutes the principal novelty of the study because it reconciles the analytical power of Artificial Intelligence with the pedagogical authority and ethical responsibility of educators.

The proposed model also contributes conceptually to Islamic Religious Education by integrating advanced Artificial Intelligence technologies with the philosophical objectives of Islamic education. While AI provides computational intelligence capable of processing extensive educational data, Islamic educational philosophy emphasizes the formation of holistic human beings whose intellectual, moral, spiritual, and social development cannot be reduced to algorithmic prediction alone. Accordingly, the framework positions Artificial Intelligence as a supportive educational instrument operating under human guidance, ethical governance, and Islamic educational values. Through this integration, the proposed model extends current AI-based assessment literature by demonstrating that transparency, explainability, teacher agency, and value-oriented educational leadership are indispensable requirements for responsible Artificial Intelligence implementation in Islamic Religious Education.

Figure 2. Proposed Human-Centered Explainable AI Assessment Model



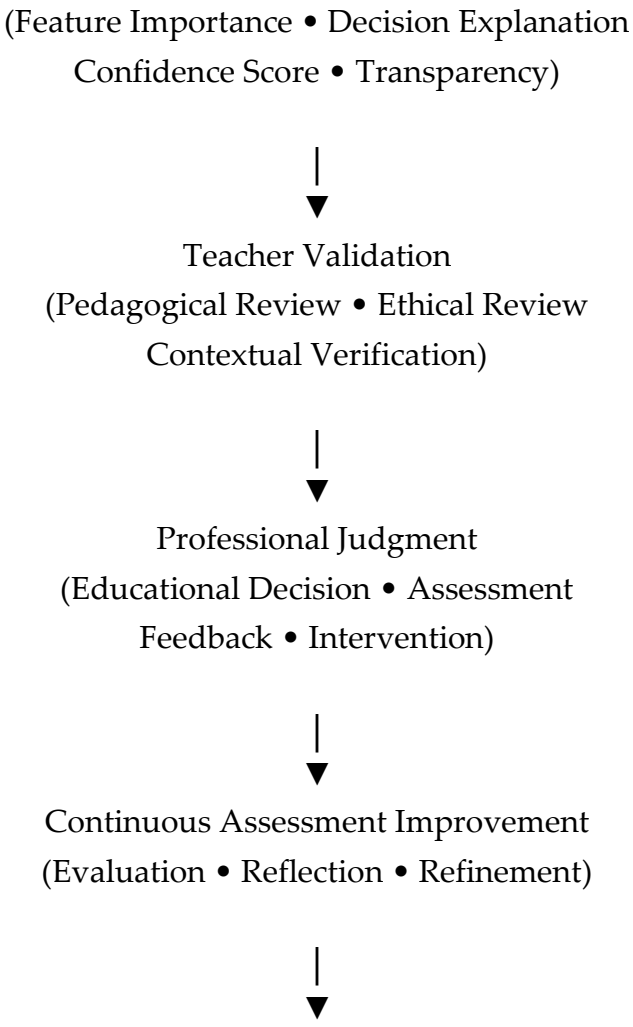


Figure 2. Proposed Human-Centered Explainable Artificial Intelligence (XAI)-Based Assessment Model for Islamic Religious Education.

Theoretical and Practical Implications

The development of the proposed Human-Centered Explainable Artificial Intelligence (XAI)-Based Assessment Model provides significant theoretical and practical contributions to the advancement of educational assessment, Artificial Intelligence in education, and Islamic Religious Education (PAI). Unlike previous studies that primarily emphasize technological innovation or predictive performance, this study demonstrates that responsible AI implementation requires the integration of computational intelligence, pedagogical expertise, ethical governance, and human-centered educational values. Consequently, the proposed framework contributes not only to the technical development of AI-supported assessment but also to the conceptual understanding of how Artificial Intelligence should function within educational systems.

Theoretical Implications

The first theoretical implication concerns the advancement of educational assessment theory. Conventional assessment theories generally conceptualize assessment as a process of measuring learning outcomes, monitoring instructional effectiveness, and supporting educational improvement (Black & Wiliam, 2018). While these perspectives remain highly relevant, the increasing availability of educational data and Artificial Intelligence technologies requires a broader conceptualization of assessment that extends beyond measurement toward intelligent educational decision-making. The proposed model contributes to this theoretical development by positioning assessment as a dynamic system integrating multidimensional learning evidence, predictive analytics, pedagogical interpretation, and continuous improvement. Assessment therefore evolves from a retrospective evaluation mechanism into an intelligent educational ecosystem capable of supporting adaptive learning and evidence-based instructional planning.

Second, this study extends the theoretical discourse on Artificial Intelligence in Education (AIED). Existing AI research has predominantly concentrated on intelligent tutoring systems, automated scoring, adaptive learning, and learning analytics (Holmes et al., 2019; Zawacki-Richter et al., 2019). Although these studies demonstrate the effectiveness of AI in enhancing educational efficiency, they frequently conceptualize Artificial Intelligence as an autonomous computational technology whose primary objective is to optimize learning processes. The framework proposed in this study offers a different perspective by emphasizing that AI should function as an educational decision-support system operating under human supervision. This perspective reinforces UNESCO's concept of Human-Centered Artificial Intelligence**,** where intelligent technologies strengthen rather than replace educators' professional expertise (UNESCO, 2023). Accordingly, the study contributes to AIED theory by integrating algorithmic intelligence with pedagogical responsibility and ethical accountability.

A third theoretical contribution relates to the growing field of Explainable Artificial Intelligence (XAI)**.** Existing XAI literature primarily focuses on improving the interpretability of machine learning models within domains such as healthcare, finance, cybersecurity, and industrial automation (Gunning & Aha, 2019). Comparatively fewer studies investigate the educational implications of explainability, particularly within assessment practices. The proposed model extends XAI theory by demonstrating that explainability should not be regarded solely as a computational characteristic but also as a pedagogical requirement. Transparent

explanations enable teachers to understand algorithmic reasoning, evaluate AI-generated recommendations, and integrate computational evidence with professional educational judgment. Consequently, explainability becomes an essential prerequisite for trustworthy educational assessment rather than merely a technical feature of AI systems.

The fourth theoretical implication concerns Islamic Religious Education (PAI). Traditional theories of Islamic educational assessment emphasize the comprehensive evaluation of learners' cognitive, affective, moral, and spiritual development while recognizing teachers as the primary educational authorities (Al-Attas, 1999; Kamali, 2008). Nevertheless, these theoretical perspectives were developed prior to the widespread implementation of Artificial Intelligence and therefore provide limited guidance regarding AI-supported educational assessment. The proposed framework bridges this gap by integrating Explainable Artificial Intelligence with the Islamic educational principles of amanah (trustworthiness), 'adl (justice), shūrā (consultative decision-making), and maslahah (public benefit). Through this integration, the study contributes a novel theoretical perspective in which technological innovation and Islamic educational philosophy function synergistically to support holistic learner development.

Finally, this study introduces a human-centered conceptual model that integrates educational data, machine learning, learning analytics, explainability, teacher validation, professional judgment, and continuous assessment improvement into a unified framework. Unlike linear AI assessment models that conclude with algorithmic prediction, the proposed model establishes a cyclical assessment process in which teachers continuously evaluate AI recommendations and generate feedback for subsequent learning improvement. This recursive structure contributes to educational assessment theory by conceptualizing Artificial Intelligence as a collaborative educational partner embedded within an ongoing cycle of professional reflection and institutional learning.

CONCLUSION

This study develops a Human-Centered Explainable Artificial Intelligence (XAI)-Based Assessment Model for Islamic Religious Education (PAI) by synthesizing theories of educational assessment, Artificial Intelligence, Explainable Artificial Intelligence (XAI), learning analytics, and Islamic educational philosophy. The findings demonstrate that conventional assessment practices in PAI remain constrained by fragmented learning evidence, subjective teacher judgment, limited utilization of digital educational data, and insufficient support for evidence-based

instructional decision-making. While Artificial Intelligence has significantly expanded educational assessment through machine learning, learning analytics, automated assessment, and predictive analytics, existing AI applications continue to face substantial challenges related to algorithmic opacity, limited interpretability, and insufficient teacher involvement in educational decision-making.

To address these challenges, this study proposes a conceptual assessment framework consisting of seven interconnected components: Learning Data, Machine Learning, Learning Analytics, Explainable Artificial Intelligence (XAI), Teacher Validation, Professional Judgment, and Continuous Assessment Improvement. Unlike conventional AI-supported assessment systems that primarily emphasize automation and predictive accuracy, the proposed model positions teachers as the final educational decision-makers while utilizing Explainable Artificial Intelligence to provide transparent, interpretable, and accountable analytical evidence. Consequently, Artificial Intelligence functions as a collaborative decision-support system that strengthens teachers' professional expertise rather than replacing pedagogical authority.

The principal novelty of this study lies in the integration of Explainable Artificial Intelligence with teacher validation, professional judgment, and the Islamic educational principles of *amanah* (trustworthiness), *ʼadl* (justice), *shūrā* (consultative decision-making), and *maslahah* (public benefit). This integration extends current AI-based educational assessment literature by introducing a human-centered conceptual framework that balances computational intelligence with educational ethics and professional responsibility. The proposed model therefore contributes theoretically to the advancement of educational assessment, Artificial Intelligence in Education, and Islamic Religious Education, while practically providing strategic guidance for teachers, school leaders, curriculum developers, policymakers, and educational technology developers seeking to implement trustworthy Artificial Intelligence in educational assessment.

Since this study is conceptual and based on qualitative library research, empirical validation of the proposed framework remains necessary. Future research is therefore recommended to implement and evaluate the model within diverse educational contexts, including madrasahs, Islamic boarding schools (*pesantren*), and Islamic higher education institutions. Quantitative, mixed-method, and design-based research approaches may further examine the effectiveness of the proposed model in improving assessment quality, teacher decision-making, student learning outcomes, and educational accountability. Such empirical investigations will strengthen the

applicability of Explainable Artificial Intelligence as a transparent, accountable, and human-centered assessment framework for the future development of Islamic Religious Education.

REFERENCES

Al-Attas, S. M. N. (1999). *The concept of education in Islam: A framework for an Islamic philosophy of education*. International Institute of Islamic Thought and Civilization (ISTAC).

Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 253–272). Cambridge University Press.

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice, 25*(6), 551–575. https://doi.org/10.1080/0969594X.2018.1441807

Boell, S. K., & Cecez-Kecmanovic, D. (2015). On being "systematic" in literature reviews. *European Journal of Information Systems, 24*(3), 255–271. https://doi.org/10.1057/ejis.2014.26

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal, 9*(2), 27–40. https://doi.org/10.3316/QRJ0902027

Brookhart, S. M. (2020). *Classroom assessment and educational measurement*. Routledge.

Brown, G. T. L., & Harris, L. R. (2019). *Handbook of human and social conditions in assessment*. Routledge.

Bulut, O., Kunnan, A. J., & Gierl, M. J. (2024). The rise of artificial intelligence in educational measurement: Opportunities and ethical challenges. *Educational Measurement: Issues and Practice*. Advance online publication. https://doi.org/10.1111/emip.12620

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.

Darling-Hammond, L., & Wentworth, L. (2020). *Assessment for deeper learning*. Harvard Education Press.

Floridi, L., & Cowls, J. (2022). A unified framework of five principles for AI in society. *Harvard Data Science Review, 4*(1), 1–32. https://doi.org/10.1162/99608f92.8cd550d1

Gunning, D., & Aha, D. W. (2019). DARPA's explainable artificial intelligence (XAI) program. *AI Magazine, 40*(2), 44–58. https://doi.org/10.1609/aimag.v40i2.2850

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

Ifenthaler, D., & Yau, J. Y.-K. (2020). *Utilising learning analytics for study success*. Springer.

Kamali, M. H. (2008). *Shari'ah law: An introduction*. Oneworld Publications.

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2020). *Qualitative data analysis: A methods sourcebook* (4th ed.). SAGE Publications.

Ministry of Religious Affairs of the Republic of Indonesia. (2023). *Kurikulum Pendidikan Agama Islam dan Bahasa Arab pada Madrasah*. Ministry of Religious Affairs.

OECD. (2024). *Artificial intelligence and education and skills*. OECD Publishing.

OECD. (2024). *The potential impact of artificial intelligence on equity and inclusion in education*. OECD Publishing.

Schildkamp, K. (2019). *Data-based decision making for school improvement*. Springer.

Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review, 46*(5), 30–40.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

UNESCO. (2023). *Global education monitoring report 2023: Technology in education—A tool on whose terms?* UNESCO.

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

UNESCO. (2024). *AI competency frameworks for teachers*. UNESCO.

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*, Article 39. https://doi.org/10.1186/s41239-019-0171-0