A Human-Centered Explainable AI Framework for Enhancing Trust and Decision Support in Learning Analytics Systems

Dr. Ravi Soni
Lincoln University College
47301 Petaling Jaya, Selangor Darul Ehsan, Malaysia
Dr. Sharad Salunke
Lincoln University College
47301 Petaling Jaya, Selangor Darul Ehsan, Malaysia

ABSTRACT
Artificial Intelligence has become widely adopted in learning analytics systems for predicting student performance, enabling adaptive learning, and supporting institutional decision-making. Despite improvements in predictive accuracy, many systems remain opaque, difficult to interpret, and insufficiently aligned with human-centered design principles. Insufficient transparency often weakens user trust and can limit effective adoption by both educators and learners. This study presents a systematic review of Explainable Artificial Intelligence and Human-Centered AI research in educational contexts published between 2015 and 2025. Relevant peer-reviewed studies were identified through Web of Science, Scopus, and IEEE Xplore using a PRISMA-guided screening approach. The review identifies critical gaps in the integration of explainability, trust modeling, and decision-support mechanisms within learning analytics systems. Based on thematic synthesis, a conceptual Human-Centered Explainable AI framework is proposed to enhance transparency, usability, and trust in intelligent educational systems. The framework establishes a structured foundation for future empirical validation and system implementation in subsequent stages of the research.

Index Terms — *Explainable Artificial Intelligence, Human-Centered AI, Learning Analytics, Trust, Educational Technology, Decision Support*

ABSTRACT
Artificial intelligence has undergone common usage in learning analytics systems, which have enabled student performance prediction, adaptive pedagogy and promulgate institutional decisions. Although significant improvements can be observed in the field of predictive power, a significant part of the systems is still oblique and uneasy to comprehend, not highly developed based on the principles of human-centered design. Lack of openness often undermines user trust and may hinder successful adoption even by both educators and learners. The study is a systematic review of the Explainable Artificial Intelligence (XAI) and Human-Centred AI research in the field of education since 2015. Sources of peer-reviewed publication were gathered by using Web of science, Scopus, IEEE Xplore using a PRISMA-based screening process. Within the context of learning analytics systems, the lack of significant gaps in explainability, trust modelling, and decision support is outlined in the review. Thematic synthesis provides a proposed Human-Centred Explainable AI architecture, which will increase the levels of transparency, usability, and trust in intelligent educational systems. This framework provides the empirical basis to the subsequent validation and implementation in a further research phase.

Index Terms Indexable artificial intelligence, human-centred artificial intelligence, Learning AI, Trust, Learning technology, Decision support.

I. INTRODUCTION

The modern learning analytics products actively apply artificial intelligence methods to predict the student path, identify academic risk areas, and support the instruction. Intelligent tutoring and adaptive learning environments widely include machine-learning models which include classification, clustering and recommendation. All of the operational deployments have limited interpretability and transparency, even though predictive fidelity has been becoming significantly better.

The recent research predicts the demands of Explainable Artificial Intelligence (XAI) in education, where it is apparent that the clarity of model reasoning as well as accessible outputs to users need to be made required [1], [4]. However, explainability of technology alone does not recommend effective human interaction. The paradigms of Human-Centred AI focus on usability and meaningful user agency, and consonance with cognitive processing [2]. In the teaching environment, it is evident that the lack of transparency undermines trust and restrains the uptake by the teacher and student [3].

Although the literature has assisted in studying explainability and human-centered design individually, their combination in terms of learning analytics is scanty and fragmented. Current research mostly has common research objectives of optimising the performance of their algorithms, usually to cost of interpretability, little is proposed that puts the strands of explainability, trust, and decision-support into one cohesive scaffold.

In order to fill this gap, the current study discusses such research questions:
1) What is the effect of explainability on the trust inlearning analytics systems?
2) How can the interpretability and usability be enhanced with the help of human-centred design?
3) How can we use transparency, trust and decision-support in smart systems of learning through a conceptual framework?

This study uses a systematic review as a research approach to comprise available literature and propose a systematic Human-Centred Explainable version of AI in learning analytics.

II. RELATED WORK

The past research studies have questioned the place of explainability in AI systems in education. Zhai et al. [1] have conducted a systematic review of XAI in the educational data analytics, focusing on the role of transparency in enhancing user understanding and model transparency. According to Chen and Xie [4], interactive adaptive learning systems increase perceived fairness and acceptance in learners and it is consequent to note that interpretable decision-making mechanisms are required within the education settings.

The study of Human-Centred AI pays more emphasis to usability, interaction with the user, and its correspondence with the way of cognition. Yang and Lee [2] suggested principles of designing intelligent tutoring system by coming up with personnel designs which anticipate interpretability, agency, and substantive feedbacks to the user. They conclude their analysis that transparency should be accompanied with proper interaction design to have productive human-AI and collaboration. Patel and Singh [6] also explored the concept of usability and trust related to the AI-led educational tools, and interpretability, reliability, and clarity are listed as the key factors shaping the user trust.

The topic of trust modelling becomes an increasing point of focus in explainable systems. Wang and Ho [3] examined structural relationships between explainability and user trust and showed that the use of the lucid reasoning leads to increase in the perceived system reliability and decision confidence. Kim et al. [9] examined explainable decision-support mechanism in the education setting and this study reveals the importance of accountability and comprehension of the intelligent system by the user.

Though these contributions have been made, there exist three major limitations. To begin with, explainable AI and human-centred design are often studied separately, which leads to the disjunction between technical transparency and user-centred design. Second, there is limited research that suggests frameworks that are cohesive explaining why explainability, trust, and decision-support processes are interlinked in the context of learning analytics. Third, the existing literature lays emphasis on predictive performance supreme to human interpretability and practical decision assistance.

Such weaknesses serve to fuel the current research, which aims at generalising the available findings and suggest a formal Human-Centred Explainable AI framework of learning analytics systems.

III. RESEARCH METHODOLOGY

The current investigation is rooted in a literature review with the purpose of questioning the intersection of Explainable artificial intelligence and Human-Centred artificial intelligence in the context of learning analytics. The aim is to reveal the gaps encountered by the research and condense thematic information on transparency, trust, as well as decision support in educational AI.

A. Search Strategy

Three main academic databases (Web of Science, Scopus, and IEEE Xplore) were selected as the literature search. The time frame was between 2015 and 2025. Combination of key words used:
"Explainable AI in education"
"Human-centred AI"
Learning analytics clarification easy.
"Trust in AI systems"
The interpretable machine learning in education.
Results were refined using the Boolean operators and database filters that narrowed by the publication year, type of documents and relevance of the subject.

B. Inclusion Criteria

Qualified articles were peer reviewed journal and conference articles indexed in Web of science, Scopus, or IEEE Xplore, which were published in the last 5 years (2015-2025) and which covered the topic of explainability, trust, usability or learning analytics in educational settings.

C. Exclusion Criteria

The excluded studies were not peer-reviewed articles, editors, reports, and literature only examining the effectiveness of algorithms without making reference to humans, as well as articles that did not focus on the educational AI or learning analytics.

D. Screening and Selection Process.

To guarantee curation, a PRISMA-guided workflow was adopted in a light manner. This involved determination of the records of the sampled databases, removal of duplicates, screening of abstracts, and full-text screening. Thereafter, future-related studies were classified and analysed through thematic synthesis to shed light on the prevailing research directions and gaps. In the first place, 148 records were discovered in Web of Science, Scopus, and IEEE Xplore. After the window of duplicates and eligibility has been screened, 32 studies have been left to be included in the final thematic synthesis.

In order to ensure the analytical rigour, the included studies were assessed based on relevance to explainable AI, integration of human-centred design, and methodology transparency.

IV. Result analysis and thematic analyses FINDINGS and thematic analyses.

Thematic analysis of thirty-two peer-reviewed articles provides three main clusters of themes that help to identify the existing gap in the current literature on explainability, human-centered design, and trust in learning analytics systems. These topics define the current academic outlook and reveal the gaps in the organization of technical transparency and human oriented system engineering.

The first involves explaining learning analytics.
Numerous studies anticipate the importance of explainable artificial intelligence as ways of increasing the degree of transparency and interpretability in predictive modelling. Modern methodological directions are focused on feature attribution, visualisation modalities, and post-hoc interpretability modalities that provide the user with intelligible presentations of the model outputs [1], [4]. Even with the significant improvements, several systems used in deployment are still lacking in an exemplific introduction to users such as easy-to-understand interfaces thus limiting real use in educational institutions.

B. Human -Centered Design and Usability.
Human-centered AI studies also anticipate usability, clarity of interaction, and meaningful feedback as the necessary elements of the intelligent platforms. It has been shown through empirical studies that user-centered design improves the quality of the interaction and reduces the cognitive load of pedagogical systems [2], [6]. However, these contributions often separate out the considerations of usability and the considerations of model transparency, and have produced artefacts that are interpretable, but not helpful to user experience, or convivial, but not substantively explainable.

C. Trust and Decision Support
Trust is a crucial factor to the adoption of AI-driven learning analytics. The research indicates that the perceived transparency, reliability, and fairness are directly related to the user confidence and hence the effectiveness of making decisions [3], [9]. Trust is, however, less tangible and quite difficult to define, which makes it highly elusive and difficult to operationalised in learning analytics systems. Most systems are mainly designed to emphasise predictive accurateness as opposed to supportiveness on decision-making and learner trust. Table I summarizes the exemplar studies that were found in the systematic review, pre-empting their part in clarifying explainable AI and human-centered learning analytics studies.

Table I. Summary of Key Studies on Explainable AI and Human-Centered Learning Analytics

Ref	Study Focus	AI/XAI Approach	Human-Centered Aspect	Key Contribution
[1] Zhai et al., 2022	Systematic review of Explainable AI in education	Review of interpretable ML and XAI methods	Discusses transparency and educational decision support	Identifies importance of explainability for educational AI adoption
[2] Yang & Lee, 2022	Human-centered design in intelligent tutoring systems	AI-driven tutoring models	Usability and user interaction design	Emphasizes integrating user-centered design in AI educational tools
[3] Wang & Ho, 2023	Trust in explainable AI systems	Model interpretability frameworks	Trust modeling and perception	Shows transparency increases user confidence in AI decisions

[4] Chen & Xie, 2023	Transparent adaptive learning systems	Adaptive learning algorithms	Student perception and fairness	Demonstrates impact of explainability on learning acceptance
[5] Natale et al., 2023	Conceptual framework for explainable AI adoption	Framework-based analysis	Ethical and transparency considerations	Proposes adoption model for explainable AI systems
[6] Patel & Singh, 2024	Usability in AI-driven educational systems	Learning analytics tools	User trust and usability evaluation	Identifies key usability factors influencing trust
[7] Aziz & Rahman, 2024	Explainable dashboards for learning analytics	Visualization-based explanation	Decision-support interface design	Improves interpretability through visual analytics
[8] Singh & Gupta, 2025	Interpretability in intelligent educational systems	Interpretable ML techniques	User trust and understanding	Demonstrates relationship between interpretability and trust
[9] Kim et al., 2025	Explainable AI for decision support	Explainable decision-support models	Accountability and transparency	Highlights role of XAI in responsible AI adoption

D. Overview of the Important Consumption Research Gaps.
A thorough look into the data through thematic synthesis shows three gaps:

1. Lack of explainability AI applied to human-centered design.
2. Absence of coherent models which interconnect accountability, trust, and decision support.
3. Minor consideration of human interpretability and pragmatic decision-support systems in learning analytics systems.

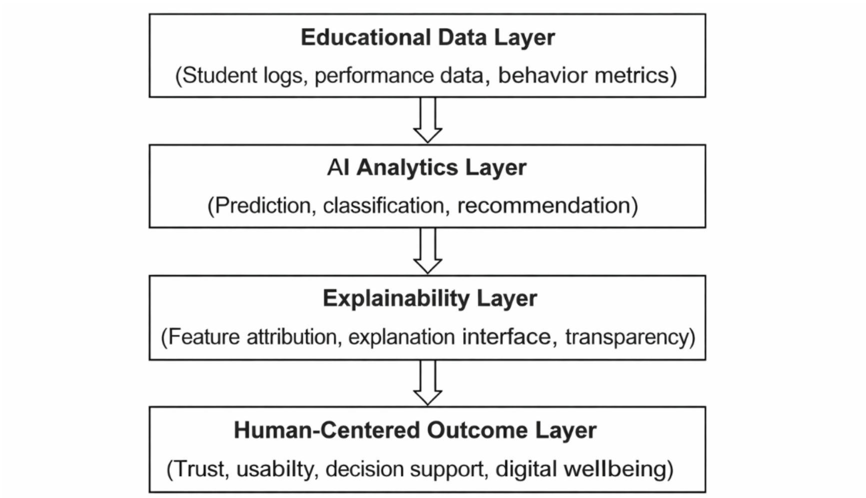
The lacunae converge in an effort to encourage the expression of an organized Human-Attributes Explainable AI system, discussed in the following section.

V. Suggested framework of explainable AI FMC.

Based on the thematic synthesis and perceived gaps in research, this paper fulfills a structured Human-Centered Explainable AI (HC-XAI) model of learning analytics systems. The paradigm overlaps technical explicability and humanist/overall design principles to supplement openness, belief, and decision-providing in clever educative settings.

The architecture has four layers, which are interdependent:

Fig. 1. Explainable Artificial Intelligence Framework of Human-Centered Learning Analytics Systems.



A. Educational Data Layer
Stratigraphic layer will contain the structured and unstructured data collected in the learning environments, including student performance measures, interaction records, behavioural cues, and assessment records. Most critical in this layer is data quality and contextual relevance so that the further analyses can result in semantically rich and pedagogically valid results.

B. AI Analytics Layer
The layer contains machine-learning models used to make a prediction, a classification, and a recommendation. Common applications include student risk prediction, adaptive learning paths and performance analytics. Although and especially the accuracy of the model is forever irreplaceable, this layer is increasingly focused on models that can be subject to explainability systems hence establishing the foundation of intelligible results.

C. Explainability Layer
Explainability Layer releases explainable results and agentic logic to the final user. Some of its repertoire consists of feature attribution methods, explanation interfaces, and transparent reporting systems. This layer reduces the impact of opaqueness by refining algorithmic decision-making to readable narrative to educators and learners to enhance quality decision-making. In order to prevent cognitive overload explanations are presented through the method of progressive disclosure: brief overviews are given first followed by optional more in-depth presentations where appropriate depending on user wish.

Models: A representation of the collective behavioral pattern of human society.<|human|>D. Human-Centered Outcome Layer
The layer that determines system impact is the terminal layer that evaluates the system impact using the lenses of the user. It incorporates the level of trust, usability testing, efficacy of decision support, and the issue of digital well-being. The HC-XAI model ensures the involvement of transparency in the architecture by integrating an analysis system that is user-based and offers high-quality interaction.

The HC-XAI framework of weaving four layers enables the shift in paradigm of performance-based learning analytics systems into the realm of transparent, trust-building, educational AI systems. In opposition to the extant fragmented strategies, the given architecture clearly links the explainability mechanisms with human-oriented results, hence setting a consistent root which may be integrated in further empirical proof.

Flow of Interaction of Framework.
The workflow of system elements and individuals interested in a particular issue is schematized in Fig.�rieben and carried out in the form of a chronological but recursive process. Education data,

obtained by learning management systems, tests, and interactions with students, are first stored in the Educational Data Layer, where learning indicators are filtered out to the salient data, then pre-processed. The processed data is then passed through the AI Analytics Layer on which predictive or classification models are created that produce judgements in the form of judgements like performance estimates or risk indicators. These results are then interpreted using the Explainability Layer that marks the key contributors to every prediction and makes them as coherent, intelligible stories. Lastly, such explanations are made available to educators and learners through easy to use interfaces that comply with the principles of the human-centered design.

Finally, the Human-Centered Outcome Layer conducts a stringent evaluation of the effect of these explanatory mechanisms on the trust of the users, their usability as well as decision making effectiveness. The responses to the end users can be used to drive the series of iterative changes in both the model design and the explanatory methods hence, creating a self-reinforcing process of constant improvement.

VI. Discussion and Future Research direction.

The results of our empirical study highlight the necessity to go beyond a single emphasis on algorithmic performance in learning analytics and point to the need to focus on the provision of both transparent and highly human-centered systems. Although explainability facilitated by explainable artificial intelligence increases comprehension, its actual practical feature can be achieved only through an aidless combination with design principles, in which usability and building trust in the experimental subjects are key considerations. The Human-Centered Explainable AI framework proposed closes this gap by linking the technical understanding of goals to the human goals, as well not the number of things depended upon, supported, or meaningfully engaged with the system.

This framework provides a systematic base towards the operationalization of transparency in the context of learning-analytics. The framework cultivates a deeper awareness among the users and promotes responsible use of AI in education through integrating explainability and providing systematic usability and trust assessments. As a result, it corresponds to the emerging trend in the direction of the human-AI partnerships and the overall trend to the trustful intelligent systems.

In future research, we will strive to put the idea scaffold into scientifically based and action-oriented items. The following step will be outlining evaluation constructs that will include trust, usability, and the effectiveness of decision support. Subsequently, a prototype learning-analytics dashboard will be developed by us incorporating explanatory outputs. The empirical validation will be done with the help of the user-based assessment, where educators and learners will participate, and interpretability, confidence, and the quality of decisions will be evaluated. Such planned programs will enable an African system of refinement and extensive validation of the suggested model.

VII. CONCLUSION

It is clear that artificial intelligence has provided learning-analytics systems with a new boost; however, some acute issues related to transparency, interpretability and maintenance of the trust of users are left unanswered. This systematic review appreciated the interaction of Explainable Artificial Intelligence and Human-Centered AI in the education processes. Our review indicated a divided research field where technical explainability and human-centered design are extensively discussed as single concepts, so the overall effect of the two on trust and decision-support is diluted.

The paper contributes to the growing body of research on reliable artificial intelligence in the educational field by illustrating a Human-Centered Explainable AI model that directly links the transparency of the model to the trust towards it and decipherment of findings. This framework presents a systematic framework on how to achieve the empirical implementation and appraisal of this framework in the future, supporting the creation of accountable, explainable, and truly human-focused AI-based systems of learning analytics.

REFERENCES

[1] H. Khosravi, S. Sadiq, and D. Gasevic, "Explainable Artificial Intelligence in Education," *Computers and Education: Artificial Intelligence*, vol. 3, 100074, 2022.

[2] R. Farrow, "The possibilities and limits of explicable artificial intelligence in education," *Learning, Media and Technology*, vol. 48, no. 3, pp. 324–338, 2023.

[3] A. Holzinger, "Explainable AI and Multi-Modal Causability in Medicine," *Artificial Intelligence in Medicine*, vol. 117, 102134, 2021.

[4] J. Longo, D. Dondio, and S. Barrett, "Explainable Artificial Intelligence: A Survey," *IEEE Access*, vol. 8, pp. 115–130, 2020.

[5] B. Shneiderman, "Human-Centered Artificial Intelligence: Reliable, Safe and Trustworthy," *International Journal of Human-Computer Interaction*, vol. 36, no. 6, pp. 495–504, 2020.

[6] D. Gunning and D. Aha, "DARPA's Explainable Artificial Intelligence (XAI) Program," *AI Magazine*, vol. 40, no. 2, pp. 44–58, 2019.

[7] A. Abdul, J. Vermeulen, D. Wang, B. Lim, and M. Kankanhalli, "Trends and trajectories for explainable, accountable and intelligible systems," Proc. ACM CHI Conf. Human Factors Comput. Syst., 2018.

[8] A. Khosravi, S. Sadiq, and D. Gasevic, "Learning Analytics and Artificial Intelligence in Education: A Systematic Review," *Computers and Education*, vol. 174, 104309, 2021.

[9] Z. Altukhi and S. Pradhan, "Explainable AI in Education: A Systematic Review of Concepts and Methods,"2025. (Preprint research review widely cited in current XAI discussions.)