

Explainable Artificial Intelligence Framework for Trustworthy Decision Support Systems

Ali Ahmad

Universitas Darunnajah

*Corresponding Author: Ali Ahmad, ali.ahmad@darunnajah.ac.id

ARTICLE INFO

Keywords: Explainable Artificial Intelligence, Decision Support System, User Trust, Transparency, Thematic analysis.

Received : 27, June

Revised : 29, July

Accepted: 25, August

©2026 Ahmad: This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/).



ABSTRACT

The application of artificial intelligence in decision support systems faces transparency challenges that can reduce user confidence in the results of recommendations. This research aims to develop an Explainable Artificial Intelligence framework that supports a transparent, understandable, and accountable decision support system. The research used an exploratory qualitative approach with semi-structured interviews with 20 informants consisting of AI developers, information systems experts, and professional users. The data was analyzed using thematic analysis to identify the need for explanation, confidence-building factors, and the characteristics of information that were considered relevant. The results show that clarity of decision reasons, consistency of explanations, traceability of processes, and suitability of context are the main elements of forming trust. This research produces an XAI conceptual framework that can be the basis for the development of a more transparent, reliable, and user-oriented decision support system

INTRODUCTION

The development of artificial intelligence (AI) has driven the transformation of decision support systems in various high-risk fields, such as health, finance, public administration, education, and industry. Machine learning-based systems are capable of processing large amounts of data, finding complex patterns, and generating recommendations faster than conventional approaches. However, these predictive improvements are often achieved through black box models, making the reasons behind a recommendation difficult for users to understand. This condition raises issues of transparency, accountability, fairness, and traceability of decisions, especially when the results of the system affect the safety, rights, or welfare of individuals. Therefore, Explainable Artificial Intelligence (XAI) was developed to bridge computational performance and human needs for understandable, relevant, and accountable explanations in decision support systems (Kostopoulos et al., 2024; Vilone & Longo, 2021).

The urgency of developing an explainable decision support system is also seen in the Indonesian context. Susanto et al. (2025) involved 102 doctors from seven geographical regions of Indonesia to evaluate AI-based decision support systems in cardiovascular risk assessment. The use of the system increased the accuracy of risk classification by 27% and the accuracy of statin administration by 29% compared to decision-making without the assistance of the system. As many as 79% of participants stated that they were willing to follow the AI recommendations and 82% were willing to use the system if it is available. The findings show the great potential of the application of AI in supporting professional decisions in an environment with limited resources. However, the high willingness to follow AI recommendations also underscores the importance of explanations that allow users to assess the basis, limitations of ability, uncertainty, and relevance of recommendations before using them in their final decision. Thus, the application of AI in Indonesia needs to be accompanied by a transparency mechanism that can prevent automatic acceptance of recommendations that are not necessarily in accordance with the user's context (Susanto et al., 2025).

Theoretically, XAI is not only concerned with the ability of the system to display features or variables that affect predictions, but also with the formation of a mental model of the user regarding the working, competence, limitations, and uncertainties of the system. Mohseni et al. (2021) position XAI as a multidisciplinary field that needs to integrate technical objectives, user characteristics, task context, and evaluation methods. In XAI's human-centered perspective, the quality of explanations is not only judged based on fidelity to the model, but also on comprehension, usability, relevance, and explanatory ability to support user actions. The concept of trust in this context needs to be understood as calibrated trust, which is a level of trust that is proportional to the actual capabilities of the system, not just an increase in trust in general. Scharowski et al. (2023) found that human-oriented explanations can increase reliance on AI recommendations, but they do not necessarily increase user trust consistently. In fact, the characteristics of the decision task can have a greater

influence than the form of explanation. The findings suggest that explainable systems don't necessarily automatically become trustworthy systems.

Although XAI's research is progressing rapidly, its evaluation is still dominated by technical and experimental approaches that are less representative of the user's needs in real-world situations. A systematic review of Kim et al. (2024) identified 73 articles containing 77 user studies, but only 36 studies were conducted in the context of real applications. A total of 33 studies or around 43% actually involved lay participants who were not target users of the system. In addition, 70% of the research focused on a local explanation of a single prediction, while only 10% specifically evaluated a global explanation of the overall behavior of the system. Kim et al. also found that there is no consensus on terminology, meaningful explanatory components, and XAI evaluation standards. This condition shows a gap between the development of computational explanatory methods and the understanding of how explanations are received, interpreted, and used by stakeholders in the actual decision-making process (Kim et al., 2024; Vilone & Longo, 2021).

A number of previous studies have attempted to address this gap, but have still resulted in findings that are fragmented by domain and user group. Naiseh et al. (2021), through the think-aloud and co-design methods with health professionals, propose design principles in the form of user involvement, attention direction, friction, habitual action challenge, and learning support to improve trust calibration. Follow-up research by Naiseh et al. (2023) shows that the contribution of an explanation to trust depends on the class of explanation and its suitability with the scope of the user's work. Bienefeld et al. (2023) found that there are differences in mental models between developers and clinicians, especially regarding the purpose of explanation, the source of truth, and the relevance of information to clinical decisions. Meanwhile, interviews by Habiba et al. (2025) with 14 machine learning practitioners revealed communication barriers with non-technical stakeholders and the lack of a standard framework to translate the need for explanations into system requirements. These studies have not fully integrated the perspectives of developers, information systems experts, and professional users in a single conceptual framework that connects clarity, consistency, traceability, contextual suitability, and user trust, especially in the context of implementing decision support systems in Indonesia.

Based on these gaps, this study aims to develop a conceptual framework of Explainable Artificial Intelligence to support a transparent, understandable, reliable, and accountable decision support system. The research uses an exploratory qualitative approach through semi-structured interviews with 20 informants consisting of AI developers, information systems experts, and professional users. Thematic analysis is used to identify explanatory needs, confidence-forming factors, characteristics of relevant information, as well as the relationship between clarity of decision reasons, clarity of explanations, traceability of processes, and suitability of context of use. Theoretically, this study is expected to expand the human-centered study of XAI by integrating the concepts of explanatory quality, user mental models, and trust calibration in a single decision support system framework. In practical terms, the resulting

framework can be used as a guideline for developers and organizations in formulating system requirements, designing explanatory interfaces, evaluating the comprehension of recommendations, and reducing the risk of overconfidence and improper rejection of AI-based decisions.

LITERATURE REVIEW

Explainable Artificial Intelligence in Decision Support Systems

Explainable Artificial Intelligence (XAI) is an approach designed so that the process, reasoning, and basis of AI model predictions can be understood by humans. In a decision support system, XAI not only serves to open a black box mechanism, but also helps users assess the relevance of recommendations before making them the basis for decisions. Markus et al. (2021) distinguish the XAI method into intrinsically explainable models and post-hoc explanations, while the form of explanation can be global or local. The selection of these methods must be adjusted to the purpose of explanation, the characteristics of the decision, and the needs of the user. Antoniadi et al. (2021) also emphasized that the application of XAI in decision support systems still faces limitations in the form of a lack of user studies and the lack of effective integration of explanations into professional workflows.

Quality and Clarity of AI Explanations

The quality of the explanation is not sufficiently assessed based on the ability of the XAI method to show influential variables. The explanation must have clarity, precision, completeness, consistency, and conformity to the user's cognitive abilities. Shin (2021) explained that explainability and causability contribute to user understanding, trust, and acceptance. Causability shows the extent to which explanations can help users understand the cause-and-effect relationships that underlie system outcomes. Wang and Yin (2022) further formulated three characteristics of effective explanations, namely improving understanding of models, helping users recognize prediction uncertainties, and supporting confidence calibration. However, their experiments show that not all explanatory methods are able to meet all three characteristics, especially on complex models and tasks that users don't understand very well.

User Trust and Trust Calibration

Trust in AI systems needs to be distinguished from calibrated trust. XAI's goal is not just to increase user confidence, but to help users know when AI recommendations can be followed and when they need to be checked or rejected. Bayer et al. (2022) found that domain expertise influences the relationship between explanations, trust, and behavior following AI recommendations. Experienced users tend to be more critical of the system and need more in-depth explanations, while users with limited expertise may receive recommendations without being able to evaluate the quality of the explanations. Ferrario (2024) reinforces this view by distinguishing the user's perceived trust from the system's actual level of trust. Therefore, the XAI framework needs to support an assessment process that can align user perceptions with the system's capabilities, limitations, and real risks.

Evaluation of the Effectiveness of the Explanation

The XAI evaluation should include both the technical dimension and the human-oriented dimension. Nauta et al. (2023) identified 12 characteristics of explanatory evaluation, including truthfulness, completeness, consistency, stability, comprehension, and cohesiveness of information. The review found that about one-third of XAI's studies used only anecdotal evidence and only one-fifth involved users. Meanwhile, Jung et al. (2023) categorized the effectiveness of explanations based on the formation of mental models, user satisfaction, trust, task performance, and the user's ability to correct system recommendations. These findings suggest that the success of XAI cannot be determined solely by the fidelity of the explanation of the model, but must also be judged by its usefulness in helping users make decisions.

Traceability, Transparency, and Accountability

Transparency in the decision support system has a wider scope than the presentation of predictive explanations. Transparency also includes data documentation, process logging, traceability of model changes, communication of limitations, audit mechanisms, and division of responsibility. Kiseleva et al. (2022) view transparency as a multi-layered accountability system that involves developers, organizations, professional users, and parties affected by decisions. In the framework, traceability allows users to know the data source, processing stage, model version, and reason for the recommendation change. Eke and Shuib (2025) also show that explainability and transparency need to be combined with bias evaluation, robustness, validation, and governance in order to produce a system that is truly trustworthy.

Suitability of Research Context and Conceptual Framework

The literature shows that there is no one form of explanation that is suitable for all users and decision situations. The explanation should take into account the user's objectives, knowledge level, decision consequences, model complexity, information needs, and organizational environment. Therefore, this study places clarity of decision reasons, clarity of explanations, traceability of processes, and suitability of context as key elements of the XAI framework. A qualitative approach through interviews with AI developers, information systems experts, and professional users allowed the research to identify differences in needs as well as mental models between stakeholders. The synthesis is used to generate a conceptual framework that connects explanatory quality, process transparency, trust calibration, and responsible use of AI recommendations.

METHODOLOGY

Types of Research

This research uses a qualitative approach with an exploratory design. The approach was chosen to gain an in-depth understanding of the need for explanation, confidence-forming factors, and information characteristics that are considered relevant in artificial intelligence-based decision support systems. This design allows the research to develop an Explainable Artificial Intelligence (XAI) conceptual framework based on the perspective of stakeholders, rather than

testing the relationship between variables statistically. Thus, the research analysis unit is the views, experiences, and interpretations of informants regarding the transparency and reliability of the decision support system.

Population and Sample

The research population includes those who have knowledge or experience related to the development and use of AI-based decision support systems. The sample consisted of 20 informants that included AI developers, information systems experts, and professional users. Informants were selected using the non-probability purposive sampling technique based on the relevance of their knowledge and involvement to the phenomenon being studied. This technique is suitable for selecting participants who are able to provide in-depth and contextual information according to the research objectives (Ahmad & Wilkins, 2025). The number of 20 informants was maintained according to the research design and was assessed as adequate to accommodate the perspectives of several stakeholder groups, with an emphasis on depth and richness of data, rather than statistical representation (Hennink & Kaiser, 2022).

Research Instruments

The research instrument is in the form of semi-structured interview guidelines that are prepared based on the research objectives. The interview questions were directed at three main focuses, namely the need for AI explanations, the factors that shape user trust, and the characteristics of information relevant to decision-making. Core questions are used consistently to all informants, while in-depth questions are adjusted to the informant's answers and background. Semi-structured interviews were chosen because they allow the discussion to remain directional while providing space to explore the views that emerge during the interview process (Adeoye-Olatunde & Olenik, 2021). Because the instrument generates narrative data and not numerical scores, testing of construct validity and statistical reliability is not applied. The quality of the instrument and data was assessed through the suitability of the questions with the research objectives, the consistency of the application of the core questions, and the traceability of the data interpretation process.

Data Collection Procedure

The research is carried out through several stages. The initial stage includes the formulation of the research focus, the determination of the informant category, and the preparation of interview guidelines. Furthermore, prospective informants are selected purposively based on the suitability of their competencies or experience with the research topic. Before the interview, the informant received an explanation of the purpose of the research and the use of the data. The interviews were then conducted in a semi-structured manner and documented for transcription purposes. After all the data is collected, the interview results are transcribed and compiled based on the informant category. The identity of the informant is not used in the reporting process. The final stage includes transcript examination, coding, theme development, and preparation of XAI conceptual framework based on interthematic relationships.

Data Analysis Techniques

The data was analyzed using thematic analysis through the stages of data recognition, initial code preparation, code grouping, theme development, theme review, theme naming, and interpretation preparation. The analysis is focused on patterns of meaning that explain the need for explanation, confidence-forming factors, and the characteristics of relevant information. The themes that were formed were then synthesized into four main components of the XAI framework, namely clarity of decision reasons, consistency of explanation, traceability of processes, and suitability of context. Thematic analysis is appropriate to identify and interpret patterns of meaning in interview data systematically and still recognize the active role of researchers in the interpretation process (Byrne, 2022). The reliability of the analysis is maintained through checking the conformity of codes and themes with transcripts, documentation of analytical decisions, and the presentation of interpretations in a consistent and transparent manner (Adler, 2022). The name of the analysis software is not listed because it is not set in the research abstract; Therefore, this script does not claim the use of any particular software.

Development of the Conceptual XAI Framework

The conceptual XAI framework was developed through thematic synthesis of the interview data. Transcripts were reviewed repeatedly, coded, and grouped into broader categories based on recurring patterns across AI developers, information systems experts, and professional users. This process produced four main themes: clarity of decision rationale, consistency of explanation, process traceability, and contextual relevance.

These themes were integrated as core components of the proposed framework for supporting trust in AI-based decision support systems. The framework is conceptual and does not represent statistically or experimentally tested causal relationships. The arrows in the figure indicate conceptual integration from AI data and processing toward a more transparent and trustworthy decision support system, rather than a fixed causal sequence.

RESEARCH RESULT

Characteristics of Research Informants

This research involved 20 informants consisting of seven AI developers, five information systems experts, and eight professional users. The selection of informants is carried out purposively based on experience and involvement in the development, evaluation, or use of artificial intelligence-based decision support systems. The informant has a minimum undergraduate educational background and at least three years of professional experience. The variation of these characteristics is intended to obtain views from the technical, system governance, and practical use sides.

Table 1. Characteristics of Research Informants

Features	Category	Quantity	Percentage
Informant groups	AI Developer	7	35%
	Information systems expert	5	25%
	Professional users	8	40%
Gender	Male	12	60%
	Women	8	40%
Age	25-34 years old	8	40%
	35-44 years old	9	45%
	45-54 years old	3	15%
Final education	Bachelor	7	35%
	Magister	10	50%
	Doctor	3	15%
Professional experience	3-5 years	5	25%
	6-10 years	9	45%
	More than 10 years	6	30%
Total report		20	100%

The AI developer group consists of informants who have experience in developing, testing, or implementing artificial intelligence models. Information systems experts include informants who have expertise in the analysis, design, evaluation, or governance of information systems. Meanwhile, professional users are individuals who have used or evaluated decision support systems in work activities. The composition allows the research to gain diverse perspectives on the clarity of the reason for the decision, the consistency of explanation, the traceability of the process, and the suitability of the context.

The identity of the informant is anonymized using AI code for AI developers, SI for information systems experts, and IP for professional users. The number on each code is an anonymous identity that is randomly assigned and does not indicate the number of informants or the order in which the interviews were conducted.

Overview of Analysis Results

The research data was obtained through semi-structured interviews with 20 informants consisting of AI developers, information systems experts, and professional users. Informants are selected based on their knowledge or involvement in the development and use of artificial intelligence-based decision support systems. Interview data was transcribed, coded, grouped, and analyzed using thematic analysis.

The results of the analysis produced four main themes that build trust in *Explainable Artificial Intelligence* (XAI), namely: (1) clarity of decision reasons; (2) consistency of explanation; (3) process traceability; and (4) contextuality. The four themes show that trust is not only determined by the existence of explanations, but also by its quality and relevance.

Clarity of Reasons for Decision

Clarity of decision reasons emerges as the main need of the user for the system to provide not only results, but also an easy-to-understand reason. Users emphasize the importance of knowing the factors that affect recommendations.

"I need to know why the system chose that option, not just the end result." – IP-03, Professional User, March 12, 2024

"Often the explanation of the model is too technical, so non-technical users have a hard time understanding." – AI-07, AI Developer, March 18, 2024

"If I don't know the determining factor, I find it hard to believe the system's recommendations." – IP-11, Professional User, March 21, 2024

Clarity serves as the basis for evaluating decisions, not just additional information.

Consistency of Explanation

Explanatory consistency has to do with the stability of the reasons that the system gives under similar conditions. Inconsistencies can lower user trust.

"Sometimes the inputs are the same, but the reason for the system is different. It's confusing." – IP-05, Professional User, March 14, 2024

"Changes in results can indeed occur if the data changes, but they must be clearly explained." – AI-02, AI Developers, March 10, 2024

"The results and explanations must be aligned, otherwise the user will be hesitant." – SI-04, Information Systems Expert, March 20, 2024

Consistency helps users understand the system's working patterns more stably.

Process Traceability

Process traceability refers to the ability to trace data and the stages that lead to decisions. This is important for evaluation and audit of decisions.

"I want to know what data is being used until the system makes that recommendation." – IP-09, Professional User, March 16, 2024

"Process history is very important to check if there are any model errors." – SI-02, Information Systems Expert, March 11, 2024

"If there is an error, the trace process must be clear so that it can be corrected quickly." – AI-10, AI Developer, March 25, 2024

Traceability supports accountability and evaluation of the system.

Context Fit

Contextual suitability indicates that the explanation should be relevant to the user's needs and usage situation.

"The explanation should be in line with my work, not too technical." – IP-14, Professional User, March 22, 2024

"Technical and non-technical users need different levels of detail." – AI-06, AI Developer, March 17, 2024

"The information provided should not be excessive, just what is relevant." – SI-01, Information Systems Expert, March 09, 2024

Contextuality ensures that explanations can be used effectively in decision-making.

Interthematic Linkages

The four themes complement each other in forming trust in the system. Clarity helps to understand reason, consistency maintains the stability of understanding, traceability allows verification, and contextuality ensures the relevance of information.

Table 2. Interthematic Linkages

Theme	Main Functions
Clarity	Understand the reasons for the decision
Consistency	Stability of system understanding
Traceability	Verification and audit of results
Contextual fit	Relevance of information to users

The Resulting XAI Conceptual Framework

The results of the research resulted in an XAI framework that places four main elements as the basis of a trustworthy decision support system.

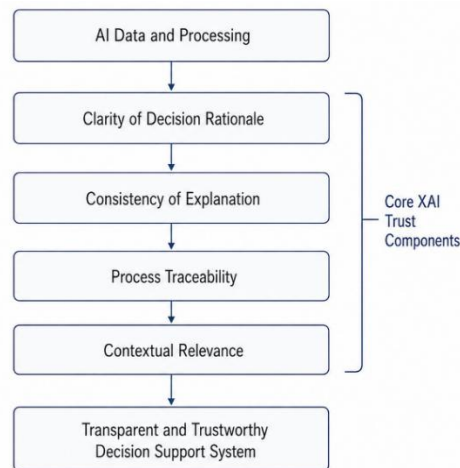


Figure 1. XAI conceptual framework

The image shows that a transparent and trustworthy decision support system is built through four main components of XAI, namely clarity of decision reasons, clarity of explanation, traceability of processes, and suitability of context. Clarity helps users understand the basis of the recommendation, while consistency ensures that explanations remain aligned under similar conditions. Traceability allows users to drill down into the data and processes that lead to decisions, while contextual suitability ensures that the information provided is relevant to the user's needs and characteristics. The integration of these four components forms a more transparent, accountable, and user-oriented XAI framework.

DISCUSSION

The results of the study show that trust in Explainable Artificial Intelligence-based decision support systems is not only determined by the existence of explanations, but by four interrelated components, namely clarity of decision reasons, consistency of explanations, traceability of processes, and suitability of context. These findings show that AI systems cannot be called trustworthy simply because they are able to display predictive reasons. Explanations should help users understand the recommendations, examine the processes that underlie them, and determine if the results are relevant to use. Thus, the trust formed is not an acceptance without criticism, but a calibrated trust, that is, a trust that is adjusted to the capabilities, limitations, and level of uncertainty of the system.

Clarity of the reason for the decision was found to be a fundamental need of the user. The informant expects the system to explain the factors that cause an alternative to be chosen, not just present the final result. These findings are supported by Rismala et al. (2025), who show that explanations at the criteria level help users understand the basis of a recommendation more comprehensively. Aprianto and Anasanti (2025) also proved that the SHAP and permutation importance methods can show the contribution of each feature to the prediction of heart disease, while Hendryani et al. (2025) used Grad-CAM to show the imaging areas underlying the prediction of pneumonia. Nonetheless, these studies mainly emphasize the technical interpretability of the model. This study fills this gap by showing that technical information is not necessarily considered clear if it is not translated into a language and format that suits the needs of professional users.

Consistency of explanations is also an important factor because users need to see the alignment between the input data, the results of recommendations, and the reasons given by the system. Changes in results are still acceptable if the change is accompanied by an explanation of the data, conditions, or models that affect it. These findings help explain the results of Scharowski et al. (2023), who show that human-oriented explanations do not necessarily increase user confidence consistently. These differences can occur because the explanations available are not necessarily stable, relevant, or appropriate to the characteristics of the task. Therefore, this study does not view explanation as a factor that automatically increases trust, but as a mechanism that must have internal and temporal consistency in order for users to build a stable understanding of system behavior.

Process traceability shows that users need the ability to drill down into the data, processing stages, and changes that result in a recommendation. Traceability allows errors to be checked and responsibility determined when the system produces inappropriate decisions. These results are in line with Busroh et al. (2025), who emphasized that the implementation of AI in government decision-making must be accompanied by transparency and accountability. Wibowo et al. (2025) also found that the lack of transparency and human oversight can pose risks in AI-based administrative dispute resolution. However, both studies discuss AI risks at the legal and governance levels. This research

clarifies its operational mechanism by placing the documentation of the data source, process history, model version, and reason for decision changes as concrete parts of the XAI framework.

Contextual suitability indicates that effective explanations should be tailored to the user's level of knowledge, tasks, and needs. AI developers, information systems experts, and professional users don't always need the same type and level of depth of explanation. This finding is supported by Amelia and Sartika (2025), who show that ease of use affects user trust in AI-based technology. Andriati et al. (2024) also emphasized that AI needs to be placed as a decision support tool because it is not yet able to fully consider moral context, empathy, and social conditions. Thus, a highly detailed explanation is not necessarily more effective when burdening users with irrelevant information. The explanation should be complete enough to support the evaluation, but still be consistent with the scope of the work and the consequences of the decision.

Overall, previous research has largely discussed technical interpretability, the quality of explanatory visualizations, user acceptance, and the legal risks of using AI. However, these studies tend to address each aspect separately, use a single group of users, or focus on a specific domain. The research gap found is that the perspectives of AI developers, information systems experts, and professional users have not been integrated into a single conceptual framework that connects clarity, consistency, traceability, and contextuality. This study fills this gap by showing that system reliability is the result of the interaction of the four components, not the result of one particular method of explanation.

These findings also expand on the research of Nauta et al. (2023), who identified consistency, stability, completeness, and comprehensibility as important characteristics in XAI evaluation. This research provides practical context by showing how these characteristics are understood by different stakeholder groups. Theoretically, the resulting framework links technical interpretability to user mental models, organizational accountability, and contextual relevance. In practical terms, this framework can serve as a guideline for developers to not only display predictive factors, but also provide stable explanations, auditable decision trails, and user-customizable levels of detail.

This study has limitations because it uses 20 purposively selected informants, so the results are not intended to be statistically generalized. The distribution of informants in each group has also not been explained in detail, while the resulting framework is still conceptual and has not been tested through prototypes. Follow-up research needs to develop a system based on these four components and test them through a mixed method approach. Testing can measure user understanding, decision accuracy, subjective trust, ability to detect errors, and behavior of accepting or rejecting recommendations. Cross-disciplinary research is also needed to find out whether the need for explanations differs in the health, education, finance, government, and industry sectors.

CONCLUSIONS AND RECOMMENDATIONS

This study concludes that the reliability of Explainable Artificial Intelligence-based decision support systems is not enough to be built just by providing an explanation of the model output. The results of the analysis show four components that need to be integrated, namely clarity of decision reasons, consistency of explanations, traceability of processes, and suitability of user context. Clarity helps users understand the factors underlying the recommendations, consistency forms a stable understanding of system behavior, traceability allows verification and audit of decisions, while contextuality ensures that explanations are relevant to the user's needs as well as knowledge levels. The integration of these four components results in an XAI conceptual framework that connects technical interpretability, system accountability, and user orientation to support more transparent and accountable decision-making.

Based on these findings, developers and organizations are advised not only to add technical explanatory features, but also to provide information that is easy to understand, stable, searchable, and adaptable to user characteristics. The system needs to be equipped with documentation of the data source, processing history, model version, level of uncertainty, and reasons for the change in recommendations. However, because this study uses 20 informants purposively and produces a framework that is still conceptual, its application needs to be further tested through prototypes and in various fields of use before it can be used as a standard for the development of decision support systems.

ADVANCED RESEARCH

Further research is recommended using a mixed method approach by developing a prototype based on the four components of the XAI framework. The effectiveness of the framework can be tested through comparison of systems with and without explanations to measure comprehension, calibrated trust, decision accuracy, error-detecting ability, and user behavior in accepting or rejecting recommendations. Cross-sectoral and longitudinal research is also needed to evaluate differences in the need for explanation as well as the stability of user trust over time.

ACKNOWLEDGMENT

The author expressed his gratitude to all informants who have provided experience, views, and information during the interview process. Appreciation was also conveyed to the parties who provided academic support and input in the preparation and completion of this research.

REFERENCES

- Adeoye-Olatunde, O. A., & Olenik, N. L. (2021). Research and scholarly methods: Semi-structured interviews. *JACCP: Journal of the American College of Clinical Pharmacy*, 4(10), 1358–1367. <https://doi.org/10.1002/jac5.1441>
- Adler, R. H. (2022). Trustworthiness in qualitative research. *Journal of Human Lactation*, 38(4), 598–602. <https://doi.org/10.1177/08903344221116620>
- Ahmad, M., & Wilkins, S. (2025). Purposive sampling in qualitative research: A framework for the entire journey. *Quality & Quantity*, 59, 1461–1479. <https://doi.org/10.1007/s11135-024-02022-5>
- Amelia, A., & Sartika, F. (2025). Factors affecting customer trust in chatbot usage: Evidence from Indonesia. *Journal of Indonesian Economy and Business*, 40(3), 327–342. <https://doi.org/10.22146/jieb.v40i3.11047>
- Andriati, S. L., Rizki, I. K., & Malian, A. N. B. M. (2024). Justice on trial: How artificial intelligence is reshaping judicial decision-making. *Journal of Indonesian Legal Studies*, 9(2), 909–942. <https://doi.org/10.15294/jils.v9i2.13683>
- Antoniadi, A. M., Du, Y., Guendouz, Y., Wei, L., Mazo, C., Becker, B. A., & Mooney, C. (2021). Current challenges and future opportunities for XAI in machine learning-based clinical decision support systems: A systematic review. *Applied Sciences*, 11(11), Article 5088. <https://doi.org/10.3390/app11115088>
- Aprianto, K., & Anasanti, M. D. (2025). Classifying heart disease through fusion of multi-source datasets: Integration of feature selection and explainable machine learning techniques. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 19(3), 247–258. <https://doi.org/10.22146/ijccs.92395>
- Bayer, S., Gimpel, H., & Markgraf, M. (2022). The role of domain expertise in trusting and following explainable AI decision support systems. *Journal of Decision Systems*, 32(1), 110–138. <https://doi.org/10.1080/12460125.2021.1958505>
- Bienefeld, N., Boss, J. M., Lüthy, R., Brodbeck, D., Azzati, J., Blaser, M., Willms, J., & Keller, E. (2023). Solving the explainable AI conundrum by bridging clinicians' needs and developers' goals. *npj Digital Medicine*, 6, Article 94. <https://doi.org/10.1038/s41746-023-00837-4>
- Busroh, F. F., Khairo, F., Niravita, A., & Nunna, B. P. (2025). Digital transformation with the impact of AI in government decision making. *Journal of Law and Legal Reform*, 6(4), 1839–1874. <https://doi.org/10.15294/jllr.v6i4.30339>
- Byrne, D. (2022). A worked example of Braun and Clarke's approach to reflexive thematic analysis. *Quality & Quantity*, 56(3), 1391–1412. <https://doi.org/10.1007/s11135-021-01182-y>
- Eke, C. I., & Shuib, L. (2025). The role of explainability and transparency in fostering trust in AI healthcare systems: A systematic literature review, open issues and potential solutions. *Neural Computing and Applications*, 37, 1999–2034. <https://doi.org/10.1007/s00521-024-10868-x>

- Ferrario, A. (2024). Justifying our credences in the trustworthiness of AI systems: A reliabilistic approach. *Science and Engineering Ethics*, 30, Article 55. <https://doi.org/10.1007/s11948-024-00522-z>
- Habiba, U.-e., Habib, M. K., Bogner, J., Fritzsche, J., & Wagner, S. (2025). How do ML practitioners perceive explainability? An interview study of practices and challenges. *Empirical Software Engineering*, 30, Article 18. <https://doi.org/10.1007/s10664-024-10565-2>
- Hendryani, A., Nurdinawati, V., & Komarudin, A. (2025). Explainable artificial intelligence model for pneumonia detection: A hybrid CNN-ViT and Grad-CAM. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, 14(4), 235–244. <https://doi.org/10.22146/jnteti.v14i4.19822>
- Hennink, M., & Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social Science & Medicine*, 292, Article 114523. <https://doi.org/10.1016/j.socscimed.2021.114523>
- Jung, J., Lee, H., Jung, H., & Kim, H. (2023). Essential properties and explanation effectiveness of explainable artificial intelligence in healthcare: A systematic review. *Heliyon*, 9(5), Article e16110. <https://doi.org/10.1016/j.heliyon.2023.e16110>
- Kim, J., Maathuis, H., & Sent, D. (2024). Human-centered evaluation of explainable AI applications: A systematic review. *Frontiers in Artificial Intelligence*, 7, Article 1456486. <https://doi.org/10.3389/frai.2024.1456486>
- Kiseleva, A., Kotzinos, D., & De Hert, P. (2022). Transparency of AI in healthcare as a multilayered system of accountabilities: Between legal requirements and technical limitations. *Frontiers in Artificial Intelligence*, 5, Article 879603. <https://doi.org/10.3389/frai.2022.879603>
- Kostopoulos, G., Davrazos, G., & Kotsiantis, S. (2024). Explainable artificial intelligence-based decision support systems: A recent review. *Electronics*, 13(14), Article 2842. <https://doi.org/10.3390/electronics13142842>
- Markus, A. F., Kors, J. A., & Rijnbeek, P. R. (2021). The role of explainability in creating trustworthy artificial intelligence for health care: A comprehensive survey of the terminology, design choices, and evaluation strategies. *Journal of Biomedical Informatics*, 113, Article 103655. <https://doi.org/10.1016/j.jbi.2020.103655>
- Mohseni, S., Zarei, N., & Ragan, E. D. (2021). A multidisciplinary survey and framework for design and evaluation of explainable AI systems. *ACM Transactions on Interactive Intelligent Systems*, 11(3–4), Article 24, 1–45. <https://doi.org/10.1145/3387166>
- Naiseh, M., Al-Thani, D., Jiang, N., & Ali, R. (2021). Explainable recommendation: When design meets trust calibration. *World Wide Web*, 24(5), 1857–1884. <https://doi.org/10.1007/s11280-021-00916-0>
- Naiseh, M., Al-Thani, D., Jiang, N., & Ali, R. (2023). How the different explanation

- classes impact trust calibration: The case of clinical decision support systems. *International Journal of Human–Computer Studies*, 169, Article 102941. <https://doi.org/10.1016/j.ijhcs.2022.102941>
- Nauta, M., Trienes, J., Pathak, S., Nguyen, E., Peters, M., Schmitt, Y., Schlötterer, J., van Keulen, M., & Seifert, C. (2023). From anecdotal evidence to quantitative evaluation methods: A systematic review on evaluating explainable AI. *ACM Computing Surveys*, 55(13s), Article 295. <https://doi.org/10.1145/3583558>
- Rismala, R., Maulidevi, N. U., & Surendro, K. (2025). Enhancing the comprehensiveness of criteria-level explanation in multi-criteria recommender system. *Journal of Information Systems Engineering and Business Intelligence*, 11(2), 160–172. <https://doi.org/10.20473/jisebi.11.2.160-172>
- Scharowski, N., Perrig, S. A. C., Svab, M., Opwis, K., & Brühlmann, F. (2023). Exploring the effects of human-centered AI explanations on trust and reliance. *Frontiers in Computer Science*, 5, Article 1151150. <https://doi.org/10.3389/fcomp.2023.1151150>
- Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human–Computer Studies*, 146, Article 102551. <https://doi.org/10.1016/j.ijhcs.2020.102551>
- Susanto, A. P., Lyell, D., Widyanoro, B., Juzar, D. A., Santoso, A., Berkovsky, S., & Magrabi, F. (2025). AI-assisted cardiovascular risk assessment by general practitioners in resource-constrained Indonesian settings using a conceptual prototype: Randomized controlled study. *Journal of Medical Internet Research*, 27, Article e73131. <https://doi.org/10.2196/73131>
- Vilone, G., & Longo, L. (2021). Notions of explainability and evaluation approaches for explainable artificial intelligence. *Information Fusion*, 76, 89–106. <https://doi.org/10.1016/j.inffus.2021.05.009>
- Wang, X., & Yin, M. (2022). Effects of explanations in AI-assisted decision making: Principles and comparisons. *ACM Transactions on Interactive Intelligent Systems*, 12(4), Article 27. <https://doi.org/10.1145/3519266>
- Wibowo, E. M., Nte, N. D., & Kusuma, B. H. (2025). AI in administrative dispute resolution: Exploring opportunities and risks. *Indonesian State Law Review*, 7(2), 137–146. <https://doi.org/10.15294/islrev.v7i2.18877>