Global Journal of Informatics
Vol. 1 No.1 2026 Page 31-40
E-ISSN: XXXX-XXXX, DOI: 10.XXXXX/XXXX

Development of an Explainable AI Framework in Information Systems to Enhance User Trust and Effectiveness

Maya Pratama[1]*, Nadia Wibowo[2], Rizki Lestari[3]
[1,2] Department of Information Systems, Faculty of Computer Science, Universitas Indonesia, Depok, Indonesia
[3] Department of Informatics, School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Bandung, Indonesia

Article Info

Received : 17 January 2026
Revised : 17 March 2026
Accepted : 11 April 2026

Keyword:

Explainable AI;
Information Systems;
User Trust;
Human-Centered AI;
System Effectiveness.

ABSTRACT

This study aims to develop an Explainable Artificial Intelligence framework in information systems to enhance user trust and system effectiveness. The study addresses the growing problem of limited user understanding of AI-based recommendations, especially when users must make decisions based on outputs that are difficult to interpret. A qualitative exploratory case study design was used to examine users' experiences, meanings, and social processes when interacting with AI-based information systems in organizational settings. Data were collected through semi-structured interviews, non-participant observation, and documentation involving 15 to 20 informants, including end users, system administrators, developers or data analysts, unit managers, and decision-makers who interact with AI-supported systems. The data were analyzed using reflexive thematic analysis supported by the Miles and Huberman interactive model. The findings identified five main themes: limited understanding of AI decision logic, trust shaped by explanation clarity, contextual and role-based explanation needs, the importance of human control, and the contribution of explainability to work effectiveness. These findings show that XAI strengthens user trust when explanations are simple, relevant, transparent, and connected to real work practices. The study contributes to human-centered XAI and trust calibration literature by emphasizing the importance of explanation quality, user role, and organizational context. Practically, the study suggests that AI-based information systems should include layered explanations, human oversight, and feedback mechanisms to support responsible and effective system use.

***) Corresponding Author:**

Maya Pratama,
Department of Information Systems, Faculty of Computer Science, Universitas Indonesia, Depok, Indonesia
Email: maya.pratama084@gmail.com

1. Introduction

The rapid development of artificial intelligence has changed how organizations manage information, make decisions, and deliver services to users. In information systems, AI no longer functions only as a technical support tool. It has become a core component in recommendation processes, prediction, classification, risk detection, and service automation. This transformation can be seen in education, healthcare, finance, public services, digital business, and organizational administration. However, the growing use of AI also creates a critical problem when users cannot understand how the system produces its decisions. Arrieta et al. (2020) explain that Explainable

Artificial Intelligence is needed to make AI processes and outputs more understandable, auditable, and accountable. Ali et al. (2023) state that XAI is a key element in developing trustworthy AI because users need not only accurate results, but also clear reasons behind those results. In the field of information systems, Brasse et al. (2023) show that XAI requires deeper investigation because information systems are directly connected to work processes, user experience, and organizational decision-making.

In both global and national contexts, AI adoption still faces problems related to user trust. Some users accept AI outputs because the system is perceived as advanced. Other users reject or question AI outputs because the process behind them remains unclear. Bach et al. (2024) found that user trust in AI-enabled systems is shaped by technical factors, system design, user characteristics, and socio-ethical considerations. Choung et al. (2023) also show that trust plays an important role in AI technology acceptance because users are more likely to use technology when they perceive it as useful, reliable, and understandable. In Indonesia, Judijanto et al. (2025) found that XAI can strengthen user trust and ethical awareness in AI adoption. However, its implementation still faces challenges related to technical complexity, limited resources, and resistance to change. These findings indicate that the main problem of AI in information systems is not only algorithmic capability. The problem also concerns how users understand, evaluate, and trust the outputs produced by the system.

Field problems become more evident when users must make decisions based on AI recommendations without knowing the reasons behind those recommendations. In many organizations, end users only see scores, risk labels, recommended actions, or prediction results without adequate explanation. Kaur et al. (2020) found that even data scientists can face difficulties in understanding and using AI interpretability tools properly. Liao et al. (2020) emphasize that XAI design should begin with users' questions, not only with the technical capacity of algorithms. This means that an explainable system must be able to answer users' real needs, such as why a decision appears, which factors have the strongest influence, when the system can be trusted, and what limitations the system has. Ehsan et al. (2021) expand this idea through the concept of social transparency, which means that AI explanation should not only explain the model. It should also explain the social context, actors, processes, and organizational goals that shape AI-based decisions.

This issue is important because user trust in AI has a direct impact on the effectiveness of information systems. If users do not trust the system, they may ignore its recommendations even when the outputs are accurate. In contrast, if users trust the system too much, they may accept AI recommendations without critical evaluation, even when the system makes errors. Shin (2021) shows that explainability and causability influence users' perception, trust, and acceptance of AI. Zhang et al. (2020) found that model confidence information can help calibrate trust in AI-assisted decision-making, but explanation does not automatically improve decision quality. Bansal et al. (2021) also show that AI explanations can increase users' tendency to accept AI recommendations, including when the recommendations are incorrect. Buçinca et al. (2021) argue that XAI design should encourage users to think critically so they do not become overly dependent on AI. Scharowski et al. (2023) further show that different types of explanations can produce different effects on user trust and reliance.

Previous studies on XAI have focused mainly on algorithmic techniques, model visualization, and quantitative measurement of user trust. These studies are important, but they do not fully explain the experiences, meanings, and social processes that users encounter when interacting with AI-based information systems. Langer et al. (2021) argue that XAI should be developed based on stakeholder needs because each actor has different interests, responsibilities, and standards of explanation. Rong et al. (2024) show that user studies in XAI still need to be strengthened through a human-centered approach because evaluations of model explanations have not fully integrated cognitive, social, and contextual aspects. Kim et al. (2024) also emphasize the importance of human-centered XAI evaluation so that AI explanations are not only technically

correct, but also meaningful for users. Hase and Bansal (2020) found that not all forms of explanation help users understand model behavior. Therefore, the effectiveness of XAI should be examined through users' experiences in specific task contexts.

Based on this background, this study aims to develop an Explainable AI framework in information systems to enhance user trust and system effectiveness. The study focuses on users' experiences in understanding AI explanations, factors that shape trust, forms of explanation considered relevant, and the process of using AI to support organizational decision-making. Theoretically, this study contributes to the development of human-centered XAI, trust calibration, and AI-based information systems research. Ridley (2025) states that human-centered XAI places humans at the center of explanation, not merely as recipients of algorithmic outputs. Practically, this study can help system developers, organizational managers, and digital policymakers design AI-based information systems that are transparent, understandable, auditable, and effective for users. Aoki (2024) shows that algorithmic explanation design needs to be adjusted to decision contexts and affected stakeholders. Ferrario and Loi (2022) also emphasize that explainability should support justified trust, not blind trust. Thus, this study has social, organizational, and technological relevance in promoting more responsible AI use.

2. Materials and Methods

This study uses a qualitative approach with an exploratory case study design. This approach is selected because the study focuses on users' experiences, meanings, and social processes when interacting with information systems based on Explainable AI. An exploratory case study is appropriate because the XAI phenomenon in information systems cannot be separated from organizational context, system type, user needs, workflow, and decision-making culture. Lim (2025) explains that qualitative research is suitable for understanding complex phenomena through the experiences and social contexts of research subjects. In this study, XAI is not viewed only as a technical feature. It is viewed as a mechanism that shapes the relationship between systems, users, and organizational decisions. Braun and Clarke (2021) state that qualitative approaches allow researchers to identify patterns of meaning in data in a reflective and systematic way. Therefore, the exploratory case study design is used to produce an in-depth understanding of how AI explanations influence trust, understanding, and the effectiveness of information system use.

This research is designed to be conducted in organizations in Indonesia that have used AI-based information systems in work processes or decision-making. The research sites may include higher education institutions, digital service units, technology companies, or organizations that use recommendation systems, AI-based chatbots, decision support systems, or predictive analytics. The research period is planned for three months, covering instrument preparation, data collection, data validation, data analysis, and the development of the XAI framework. The selection of research sites is based on three criteria. First, the organization has used an AI-based information system for at least six months. Second, the organization has active users who directly interact with system outputs or recommendations. Third, the organization has a need to improve transparency, trust, and the effectiveness of system use. Judijanto et al. (2025) show that the Indonesian context still faces challenges in XAI implementation, especially in relation to user understanding, resource readiness, and resistance to technology. Therefore, the research sites are selected to capture users' real experiences in organizations that are adapting to AI.

The research subjects consist of end users of AI-based information systems, system administrators, developers or data analysts, unit managers, and parties involved in AI-assisted decisions. Informants are selected using purposive sampling because the study requires participants who have direct experience with AI systems. Campbell et al. (2020) explain that purposive sampling helps researchers obtain rich data because informants are selected based on the relevance of their experience to the research objectives. The informant criteria include having used or managed an AI-based information system for at least three months, understanding the

basic functions of the system, being involved in work or decisions supported by AI, and being willing to participate in interviews and data validation. The planned number of informants is 15 to 20 people, but the final number will be determined based on data adequacy and saturation. Snowball sampling is also used to identify additional relevant informants, especially those with technical or strategic roles in system development and use. Suresh et al. (2021) emphasize that AI explanation needs differ across stakeholders, so this study involves actors with diverse roles.

Data are collected through semi-structured interviews, non-participant observation, documentation, and data triangulation. Semi-structured interviews are used to explore users' experiences in understanding AI recommendations, evaluating the clarity of system explanations, determining their level of trust, and using AI outputs in their work. Adeoye-Olatunde and Olenik (2021) explain that semi-structured interviews allow researchers to maintain the focus of questions while still giving participants space to describe their experiences in depth. The interview guide is developed based on several main themes, including understanding of AI outputs, perceptions of system transparency, experiences of accepting or rejecting AI recommendations, needs for explanation formats, trust-forming factors, and the impact of XAI on work effectiveness. Non-participant observation is conducted when informants use the system to observe interaction patterns, responses to system explanations, and how users verify AI outputs. Documentation includes user manuals, organizational SOPs, screenshots of explanation features, system usage reports, and technology policy documents. Kaur et al. (2020) show that observation of interpretability tool use is important because users' actual understanding does not always match developers' assumptions.

Data validation is carried out through source triangulation, method triangulation, member checking, and an audit trail. Source triangulation is conducted by comparing information from end users, system administrators, developers, and unit managers. Method triangulation is conducted by comparing interview data, observation results, and documentation. Ahmed (2024) explains that trustworthiness in qualitative research can be strengthened through credibility, transferability, dependability, and confirmability. Member checking is conducted by sending interview summaries or initial interpretations to informants to ensure that the researcher's interpretation matches their experiences. The audit trail is conducted by documenting the data collection process, analytical decisions, code changes, theme development, and the researcher's reflective notes. This procedure is important because research on XAI concerns users' interpretation of systems. Jacovi et al. (2021) argue that trust in AI should be understood through its prerequisites, causes, and goals. Therefore, data validation in this study is directed at ensuring that the meaning of trust truly comes from informants' experiences, not only from the researcher's assumptions.

Data are analyzed using reflexive thematic analysis combined with the Miles and Huberman interactive model, which includes data condensation, data display, and conclusion drawing. The analysis process begins with interview transcription, repeated reading, identification of meaning units, initial coding, code grouping, theme development, theme review, and narrative construction. Initial codes may include "understanding AI recommendations," "trust in the system," "doubt toward AI outputs," "need for simple explanations," "human control," "decision transparency," "system accountability," and "work effectiveness." Final themes are then used to develop an XAI framework in information systems that includes transparency, explanation relevance, ease of understanding, user control, trust calibration, organizational support, and system effectiveness. Zhang et al. (2020) show that trust calibration is important so users know when to accept or question AI recommendations. Leichtmann et al. (2023) also show that AI explanations can help users develop a more appropriate level of trust in high-risk decision tasks. Thus, data analysis is directed at producing an Explainable AI framework that not only strengthens user trust, but also supports critical, contextual, and effective system use.

3. Results and Discussions

The qualitative analysis identified five main themes that explain how users understand, evaluate, and use Explainable AI in information systems. These themes are: limited understanding of AI decision logic, trust shaped by explanation clarity, the need for contextual and role-based explanations, the importance of human control, and the link between explainability and work effectiveness. The findings show that users did not reject AI because they opposed technology. They hesitated because they could not always understand how the system produced recommendations, scores, or risk labels. One participant stated, "I can see the recommendation, but I do not always know why the system suggests it. I still need someone to explain the logic behind it" (P3). This statement shows that users need more than output accuracy. They need understandable reasons that connect system results with their work context.

The first theme concerns users' limited understanding of AI decision logic. Most participants said that AI-based information systems helped them process data faster, but they still viewed the system as difficult to interpret. Users often understood the final output but struggled to identify which data, rules, or variables influenced the result. Observation also showed that users tended to recheck AI outputs manually when the system produced recommendations that differed from their expectations. One system administrator explained, "Users usually ask whether the system reads all data correctly. They trust the result more when we can show which indicators affect the output" (P7). This finding indicates that explanation must help users connect AI output with visible evidence, not only present technical model information.

The second theme shows that trust was shaped by the clarity, consistency, and relevance of explanations. Participants expressed higher trust when the system provided simple reasons, displayed the main contributing factors, and allowed users to trace the source of recommendations. However, trust decreased when the explanation used technical language, statistical terms, or visual elements that users could not interpret. One end user noted, "The explanation is useful when it tells me the main reason. If it only shows numbers and technical charts, I still feel unsure" (P5). This finding shows that explanation quality depends on user interpretation. A clear explanation does not mean a long explanation. Users preferred short, direct, and task-relevant explanations.

The third theme relates to contextual and role-based explanation needs. Users, managers, and technical staff required different types of explanation. End users needed simple operational explanations to support daily tasks. Managers needed explanations that linked AI recommendations to organizational targets, risk control, and accountability. Technical staff needed deeper information about model behavior, data quality, and system limitations. One manager stated, "For me, the most important thing is not the algorithm name. I need to know whether the recommendation is safe to use for a decision and what risk follows it" (P11). This finding shows that XAI in information systems must serve different stakeholders. A single explanation format cannot meet all needs.

The fourth theme highlights the importance of human control in AI-assisted decisions. Participants did not want AI to replace human judgment. They wanted AI to support decision-making while still allowing human review, correction, and final approval. Several participants stated that they trusted the system more when they could compare AI recommendations with historical data, organizational rules, and professional judgment. One participant said, "I feel more comfortable when the system gives a recommendation, but the final decision remains with us" (P2). Observation also showed that users became more confident when the system allowed them to review input data, see explanation details, and override recommendations with reasons. This finding indicates that user trust grows when the system preserves human agency.

The fifth theme shows that explainability influenced perceived work effectiveness. Participants reported that AI explanations helped them reduce uncertainty, speed up verification,

and communicate decisions to supervisors or colleagues. Documentation review showed that organizations with clearer system guidelines had smoother adoption because users could understand how AI outputs should be used in workflow. One participant explained, "When the system gives a reason, I can explain my decision faster to my team. It saves time because I do not need to justify everything from the beginning" (P9). This finding suggests that XAI supports effectiveness not only by improving system usability, but also by improving communication, accountability, and confidence in organizational decision-making.

Overall, the findings indicate that an effective Explainable AI framework in information systems should include six core components: transparent output, simple explanation, contextual relevance, role-based explanation, human control, and feedback mechanisms. Transparent output helps users understand what the system recommends. Simple explanation helps users understand why the recommendation appears. Contextual relevance connects explanation with real tasks. Role-based explanation adjusts information depth to user responsibility. Human control protects users from overreliance. Feedback mechanisms allow users to report unclear explanations, incorrect outputs, or contextual factors that the system does not capture. These components form the basis for developing a human-centered XAI framework that enhances user trust and system effectiveness.

Discussion

The findings show that user trust in AI-based information systems depends strongly on how users understand the system's reasoning. This result supports the argument that XAI is not only a technical mechanism, but also a socio-technical process. Arrieta et al. (2020) explain that XAI aims to make AI systems more understandable, transparent, and accountable. The present study extends this view by showing that understandability must be linked to users' daily work context. Users did not demand full access to algorithmic complexity. They wanted explanations that helped them answer practical questions, such as why a recommendation appeared, whether it could be trusted, and how it should be used in decision-making.

The finding on limited understanding of AI decision logic aligns with Kaur et al. (2020), who found that even data scientists may struggle to use interpretability tools effectively. In this study, the challenge became more complex because most end users did not have technical backgrounds. They needed explanations that translated system logic into operational meaning. This finding also supports Liao et al. (2020), who argue that XAI design should begin with users' questions. The present study adds that user questions differ across organizational roles. End users ask how to act on the recommendation. Managers ask how to justify the decision. Technical staff ask how the model works and where the risk of error appears.

The study also shows that explanation clarity shapes trust. Users trusted AI outputs more when explanations were direct, consistent, and relevant to their tasks. This finding is consistent with Shin (2021), who found that explainability affects perception, trust, and acceptance of AI. However, this study also shows that explanation can fail when it uses technical language that users cannot translate into action. Therefore, explainability should not be measured only by the presence of explanation features. It should be measured by whether users can use the explanation to understand, evaluate, and communicate AI-assisted decisions. This point strengthens the human-centered XAI perspective discussed by Rong et al. (2024).

The finding on contextual and role-based explanation needs supports Langer et al. (2021), who emphasize that XAI should consider stakeholder needs. This study confirms that users do not need identical explanations. They need explanations that match their responsibilities, risks, and decision authority. For example, end users need concise explanations that guide action. Managers need explanations that support accountability. Developers need explanations that support model monitoring and improvement. This result offers a practical contribution because it suggests that XAI frameworks in information systems should use layered explanations. A basic layer can serve general users, while advanced layers can serve managers, auditors, and technical teams.

The theme of human control provides an important critical insight. Participants trusted AI more when they could review, question, or override recommendations. This finding aligns with Jacovi et al. (2021), who argue that trust in AI must be understood through its prerequisites, causes, and goals. In this study, trust emerged when users felt that AI supported their judgment rather than removed their authority. This finding also relates to Buçinca et al. (2021), who argue that system design should reduce overreliance on AI. Users need explanation not only to accept AI outputs, but also to know when to question them. Therefore, effective XAI should promote calibrated trust, not blind trust.

The findings also show that explainability contributes to work effectiveness by improving verification, communication, and decision justification. This result supports Brasse et al. (2023), who state that XAI in information systems should be studied in relation to organizational use and future research directions. In the present study, users considered AI explanations useful when they helped them complete tasks faster and explain decisions to others. This means that XAI effectiveness should not be limited to individual comprehension. It should include broader organizational outcomes, such as workflow clarity, decision accountability, and coordination among users. This perspective expands the discussion of XAI from model transparency to organizational effectiveness.

However, the findings also reveal a possible tension between explanation and efficiency. Some participants wanted detailed explanations, while others preferred short explanations that did not slow down their work. This finding differs from the assumption that more explanation always improves trust. Zhang et al. (2020) found that explanation does not automatically improve decision quality in AI-assisted decision-making. The present study supports this caution. Users need the right level of explanation, not the maximum amount of explanation. Too little explanation creates uncertainty. Too much explanation creates cognitive burden. Therefore, XAI design should allow users to access explanation progressively, starting from simple reasons and moving to detailed information when needed.

Theoretically, this study contributes to human-centered XAI by showing that trust and effectiveness emerge through interaction between system features, user interpretation, organizational context, and decision responsibility. The proposed framework strengthens the view that explainability must include transparency, contextual relevance, role-based explanation, human control, and feedback. This finding is consistent with Kim et al. (2024), who emphasize the need for human-centered evaluation of XAI applications. The study also contributes to trust calibration theory by showing that users need explanations that help them decide when to trust, when to verify, and when to reject AI recommendations.

Practically, the study suggests that organizations should not implement AI-based information systems without explanation governance. Developers need to design explanations in simple language, use clear indicators, and provide layered information for different users. Managers need to prepare SOPs that explain how AI recommendations should be used, reviewed, and documented. Organizations also need training programs that improve AI literacy, especially for users who rely on AI outputs in routine decisions. Ferrario and Loi (2022) argue that explainability should support justified trust. The present study confirms this point by showing that users need explanations that make trust reasonable, accountable, and connected to work practice.

Future studies can deepen this research in three ways. First, future research can test the proposed XAI framework in different sectors, such as healthcare, education, banking, and public services. Second, future studies can compare users with different levels of AI literacy to understand how explanation needs change across user groups. Third, mixed-method research can measure whether the proposed framework improves trust calibration, decision quality, and system effectiveness. This study has provided a qualitative foundation for understanding how users

experience XAI in information systems. Further research can refine the framework through broader empirical testing and sector-specific application.

4. Conclusions

This study concludes that Explainable AI in information systems plays a central role in shaping user trust and system effectiveness. The findings show that users need more than accurate AI outputs. They need clear, simple, contextual, and role-based explanations that help them understand why a recommendation appears, when it can be trusted, and how it should be used in decision-making. Five main themes emerged from the analysis: limited understanding of AI decision logic, trust shaped by explanation clarity, the need for contextual and role-based explanations, the importance of human control, and the link between explainability and work effectiveness. These findings strengthen the view that XAI should not be treated only as a technical feature. It should be understood as a socio-technical mechanism that connects system transparency, user interpretation, organizational responsibility, and practical decision-making.

Theoretically, this study contributes to human-centered XAI and trust calibration research by showing that user trust develops through meaningful interaction between explanation quality, task context, user role, and decision authority. Practically, the study suggests that organizations should design AI-based information systems with layered explanations, clear user guidance, human review mechanisms, and feedback channels. From a policy perspective, the findings support the need for organizational standards that regulate how AI recommendations are explained, reviewed, and documented. Future research can expand this study by testing the proposed XAI framework in different sectors, such as education, healthcare, banking, and public services. Further studies can also use mixed methods to measure how XAI affects trust calibration, decision quality, and long-term system effectiveness.

References

Adeoye-Olatunde, O. A., & Olenik, N. L. (2021). Research and scholarly methods: Semi-structured interviews. *Journal of the American College of Clinical Pharmacy, 4*(10), 1358-1367. https://doi.org/10.1002/jac5.1441

Ahmed, S. K. (2024). The pillars of trustworthiness in qualitative research. *Journal of Medicine, Surgery, and Public Health, 2*, Article 100051. https://doi.org/10.1016/j.glmedi.2024.100051

Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., Guidotti, R., Del Ser, J., Díaz-Rodríguez, N., & Herrera, F. (2023). Explainable artificial intelligence (XAI): What we know and what is left to attain trustworthy artificial intelligence. *Information Fusion, 99*, Article 101805. https://doi.org/10.1016/j.inffus.2023.101805

Aoki, N. (2024). Explainable AI for government: Does the type of explanation matter to the accuracy, fairness, and trustworthiness of an algorithmic decision as perceived by those who are affected? *Government Information Quarterly, 41*(4), Article 101965. https://doi.org/10.1016/j.giq.2024.101965

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58*, 82-115. https://doi.org/10.1016/j.inffus.2019.12.012

Bach, T. A., Khan, A., Hallock, H., Beltrão, G., & Sousa, S. (2024). A systematic literature review of user trust in AI-enabled systems: An HCI perspective. *International Journal of*

Human-Computer Interaction, 40(5), 1251-1266. https://doi.org/10.1080/10447318.2022.2138826

Bansal, G., Wu, T., Zhou, J., Fok, R., Nushi, B., Kamar, E., Ribeiro, M. T., & Weld, D. S. (2021). Does the whole exceed its parts? The effect of AI explanations on complementary team performance. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-16). Association for Computing Machinery. https://doi.org/10.1145/3411764.3445717

Brasse, J., Broder, H. R., Förster, M., Klier, M., & Sigler, I. (2023). Explainable artificial intelligence in information systems: A review of the status quo and future research directions. *Electronic Markets, 33*, Article 26. https://doi.org/10.1007/s12525-023-00644-5

Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in reflexive thematic analysis? *Qualitative Research in Psychology, 18*(3), 328-352. https://doi.org/10.1080/14780887.2020.1769238

Buçinca, Z., Malaya, M. B., & Gajos, K. Z. (2021). To trust or to think: Cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making. *Proceedings of the ACM on Human-Computer Interaction, 5*(CSCW1), Article 188. https://doi.org/10.1145/3449287

Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S., Bywaters, D., & Walker, K. (2020). Purposive sampling: Complex or simple? Research case examples. *Journal of Research in Nursing, 25*(8), 652-661. https://doi.org/10.1177/1744987120927206

Choung, H., David, P., & Ross, A. (2023). Trust in AI and its role in the acceptance of AI technologies. *International Journal of Human-Computer Interaction, 39*(9), 1727-1739. https://doi.org/10.1080/10447318.2022.2050543

Ehsan, U., Liao, Q. V., Muller, M., Riedl, M. O., & Weisz, J. D. (2021). Expanding explainability: Towards social transparency in AI systems. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-19). Association for Computing Machinery. https://doi.org/10.1145/3411764.3445188

Ferrario, A., & Loi, M. (2022). How explainability contributes to trust in AI. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 1457-1466). Association for Computing Machinery. https://doi.org/10.1145/3531146.3533202

Hase, P., & Bansal, M. (2020). Evaluating explainable AI: Which algorithmic explanations help users predict model behavior? In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 5540-5552). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.491

Jacovi, A., Marasović, A., Miller, T., & Goldberg, Y. (2021). Formalizing trust in artificial intelligence: Prerequisites, causes and goals of human trust in AI. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 624-635). Association for Computing Machinery. https://doi.org/10.1145/3442188.3445923

Judijanto, L., Achmady, S., Fadhilah, S. A. N., & Vandika, A. Y. (2025). The impact of explainable AI on user trust and ethical AI adoption in Indonesia. *The Eastasouth Journal of Information System and Computer Science, 2*(03), 344-351. https://doi.org/10.58812/esiscs.v2i03.531

Kaur, H., Nori, H., Jenkins, S., Caruana, R., Wallach, H., & Wortman Vaughan, J. (2020). Interpreting interpretability: Understanding data scientists' use of interpretability tools for machine learning. In *Proceedings of the 2020 CHI Conference on Human Factors in*

Computing Systems (pp. 1-14). Association for Computing Machinery. https://doi.org/10.1145/3313831.3376219

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

Langer, M., Oster, D., Speith, T., Hermanns, H., Kästner, L., Schmidt, E., Sesing, A., & Baum, K. (2021). What do we want from explainable artificial intelligence (XAI)? A stakeholder perspective on XAI and a conceptual model guiding interdisciplinary XAI research. *Artificial Intelligence, 296*, Article 103473. https://doi.org/10.1016/j.artint.2021.103473

Leichtmann, B., Humer, C., Hinterreiter, A., Streit, M., & Mara, M. (2023). Effects of explainable artificial intelligence on trust and human behavior in a high-risk decision task. *Computers in Human Behavior, 139*, Article 107539. https://doi.org/10.1016/j.chb.2022.107539

Liao, Q. V., Gruen, D., & Miller, S. (2020). Questioning the AI: Informing design practices for explainable AI user experiences. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-15). Association for Computing Machinery. https://doi.org/10.1145/3313831.3376590

Lim, W. M. (2025). What is qualitative research? An overview and guidelines. *Australasian Marketing Journal, 33*(2), 199-229. https://doi.org/10.1177/14413582241264619

Ridley, M. (2025). Human-centered explainable artificial intelligence: An Annual Review of Information Science and Technology (ARIST) paper. *Journal of the Association for Information Science and Technology, 76*(1), 98-120. https://doi.org/10.1002/asi.24889

Rong, Y., Leemann, T., Nguyen, T.-T., Fiedler, L., Qian, P., Unhelkar, V., Seidel, T., Kasneci, G., & Kasneci, E. (2024). Towards human-centered explainable AI: A survey of user studies for model explanations. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 46*(4), 2104-2122. https://doi.org/10.1109/TPAMI.2023.3331846

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

Suresh, H., Gomez, S. R., Nam, K. K., & Satyanarayan, A. (2021). Beyond expertise and roles: A framework to characterize the stakeholders of interpretable machine learning and their needs. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-16). Association for Computing Machinery. https://doi.org/10.1145/3411764.3445088

Zhang, Y., Liao, Q. V., & Bellamy, R. K. E. (2020). Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 295-305). Association for Computing Machinery. https://doi.org/10.1145/3351095.3372852