JPPIPA 11(9) (2025)

Jurnal Penelitian Pendidikan IPA

http://jppipa.unram.ac.id/index.php/jppipa/index

Personalized Explainable AI: Dynamic Adjustment of Explanations for Novice and Expert Users

Afrizal Zein[1]*

[1] Prodi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Pamulang, Tangerang Selatan, Indonesia.

Received: June 04, 2025
Revised: August 19, 2025
Accepted: September 25, 2025
Published: September 30, 2025

Corresponding Author*:
Afrizal Zein
dosen01495@unpam.ac.id

DOI: 10.29303/jppipa.v11i9.12586

Abstract: Explainable Artificial Intelligence (XAI) has emerged as a crucial aspect of building trust and transparency in AI-driven systems. However, existing explanation methods often apply a uniform approach, overlooking the diverse backgrounds and expertise levels of users. This paper proposes a personalized explainable AI framework that dynamically adjusts the complexity, depth, and presentation of machine-generated explanations according to the user's expertise—be it novice or expert. By integrating user modeling and adaptive explanation strategies, the system can deliver tailored information that enhances user understanding, satisfaction, and decision-making. We evaluate the proposed approach through experiments involving participants with varying expertise levels interacting with AI-based decision systems. The results show that adaptive explanations significantly improve comprehension for both novice and expert users compared to static, one-size-fits-all explanations. These findings highlight the importance of user-centered design in XAI and suggest practical pathways for future implementation in real-world applications.

Keywords: Adaptive Explanations; AI Interpretability; Human-Centered AI; Personalized Explainable AI; Transparency; User Expertise; User Modeling.

Introduction

Artificial Intelligence (AI) has become a cornerstone of modern technological development, permeating domains such as healthcare, finance, education, and autonomous systems (Adadi & Berrada, 2018). As AI systems increasingly influence critical decision-making processes, concerns over transparency, accountability, and interpretability have grown substantially, giving rise to the field of Explainable Artificial Intelligence (XAI) (Guidotti et al., 2018). XAI aims to reduce the "black-box" nature of complex models by generating human-understandable explanations of system outputs, thereby fostering trust and supporting responsible adoption (Ribeiro et al., 2016).

While the need for XAI is widely recognized, one of its fundamental challenges lies in addressing the heterogeneity of users who interact with AI systems. Users vary in expertise, background knowledge, cognitive capacity, and informational requirements (Malamuthu et al., 2025). Novices often seek simplified, intuitive explanations to form basic mental models, whereas experts prefer technical, detailed insights into algorithms, performance metrics, or reasoning chains (Sokol & Flach, 2022). When explanations do not align with user expertise, misunderstandings may occur, undermining trust, usability, and adoption (Hoffman et al., 2018). This discrepancy highlights the urgent need for approaches that adapt explanations to different user groups rather than relying on static, one-size-fits-all methods.

Personalization has therefore emerged as a promising solution in XAI. By tailoring explanatory content to align with the cognitive and informational requirements of users, personalized explanations can improve understanding, trust, and satisfaction (Georgenia Ezeji et al., 2024). The central idea is to dynamically adjust explanation complexity, depth, and

How to Cite:
Zein, A. (2025). Personalized Explainable AI: Dynamic Adjustment of Explanations for Novice and Expert Users. *Jurnal Penelitian Pendidikan IPA*, *11*(9), 513–520. https://doi.org/10.29303/jppipa.v11i9.12586

format according to user expertise. Such adaptation not only reduces cognitive load but also ensures that explanations remain contextually relevant, thereby enhancing decision-making and user experience **(Abdul et al., 2018)**.

Despite increasing recognition of personalization, relatively little research systematically addresses dynamic adjustment strategies that differentiate between novice and expert users. While prior works have explored interactive explanations, adaptive mechanisms, and user modeling, these contributions often lack integration into comprehensive frameworks designed to respond to evolving expertise levels (Mohammed, 2025). This gap is significant because user expertise is not static: novices may gradually acquire knowledge, and experts may require varying levels of detail depending on context. Without dynamic adaptation, explanations risk being either overly simplistic or unnecessarily complex.

This study addresses the gap by proposing a framework for personalized explainable AI that dynamically adjusts explanations for novice and expert users. The framework combines user profiling techniques with adaptive mechanisms to deliver context-sensitive explanations (Markus et al., 2021). Unlike prior works that primarily focus on static personalization, our approach emphasizes dynamic adjustment, allowing explanations to evolve with users' changing knowledge states. To validate the framework, we conduct empirical evaluations involving participants across varying expertise levels, measuring comprehension, satisfaction, and trust. This study thereby contributes not only a conceptual model but also empirical evidence on the effectiveness of adaptive explanations.

The importance of this research lies in its dual contribution. First, it advances XAI by explicitly integrating user expertise as a determinant of explanation design. While prior works acknowledge user diversity, few provide systematic strategies for dynamically bridging novice and expert needs (Solanke, 2022). Second, it demonstrates practical implications: in high-stakes applications such as medical diagnostics or financial forecasting, mismatched explanations can have severe consequences. By offering a scalable framework for adaptive explanations, this study supports more ethical, usable, and trustworthy AI adoption.

The significance of personalization in XAI has long been emphasized in foundational works. Miller (2019) highlighted that explanations must reflect human cognitive processes and expectations, stressing the psychological dimension of interpretability. Later, Miller (2019) reinforced the importance of tailoring explanations for user trust, especially in high-stakes domains. Mohseni et al. (2021) further demonstrated the divergent needs of novices and experts: novices require simplified, educational explanations, while experts demand in-depth technical details. Similarly, Brey (2018); Holzinger et al. (2019) proposed adaptive frameworks that adjust explanation granularity, showing that flexibility significantly enhances comprehension and satisfaction. More recently, Adadi & Berrada (2018); Sokol & Flach (2022) advocated for explanations as dynamic artifacts, evolving with user interactions. These contributions provide valuable insights but have yet to be fully consolidated into a coherent framework explicitly designed for expertise-based dynamic adjustment.

Interactive and query-based approaches also show promise. Teso & Kersting (2019) emphasized the value of user-guided explanations, where individuals actively request tailored insights. Such methods empower users and increase transparency, but they still require integration with systematic user profiling to ensure scalability. Ehsan et al. (2021); Schneider & Handali (2019) further highlighted the importance of continuous adaptation, underscoring that personalization is not a one-time adjustment but an ongoing process informed by user feedback and interaction data.

These developments are summarized in Figure 1: Literature Landscape of XAI, which illustrates the progression of XAI research from foundational transparency goals toward adaptive, user-centered strategies. This landscape underscores both the progress achieved and the gaps remaining in expertise-based personalization.

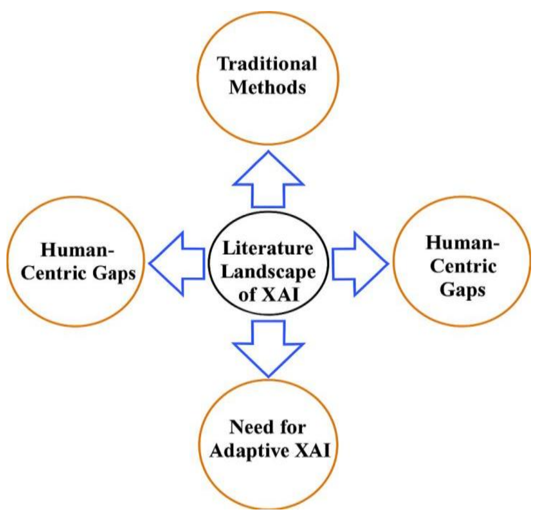

Figure 1. Literature Land Scape of XAI

Explainable AI techniques can be broadly divided into intrinsic and post-hoc approaches (Doshi-Velez & Kim, 2017). Intrinsic methods involve inherently transparent models, such as decision trees, while post-hoc methods generate explanations for complex models like deep neural networks, often using tools such as LIME or SHAP (Lundberg & Lee, 2017). While these approaches enhance interpretability, they often assume homogeneity of users and thus overlook expertise-based differences. Bridging this gap requires embedding personalization mechanisms into both intrinsic and post-hoc methods to accommodate a spectrum of user needs.

A central consideration in designing effective XAI is managing cognitive load, which directly affects user comprehension and acceptance. Novices, who face higher cognitive demands when confronted with complex technical details, benefit from simplified explanations (Sweller et al., 2019). Conversely, experts require technical depth and find oversimplified explanations frustrating or insufficient (Hoffman et al., 2018; Lim & Dey, 2009). High cognitive load can undermine trust, satisfaction, and usability (Eiband et al., 2018). Thus, explanation strategies must align with cognitive resources, providing neither too much nor too little information. Personalized adaptive mechanisms that continuously assess users' expertise and adjust accordingly represent a promising solution (Ribera & Lapedriza, 2019).

In summary, while significant advances have been made in XAI, particularly concerning interactive and adaptive explanations, the specific challenge of dynamically tailoring explanations to novice and expert users remains insufficiently explored. This study proposes a novel approach that integrates user profiling, adaptive mechanisms, and empirical evaluation to bridge this gap. By explicitly addressing expertise and cognitive diversity, the proposed framework not only advances theoretical understanding but also contributes practical solutions for designing effective, trustworthy, and human-centered XAI systems (Malamuthu et al., 2025).

Method

A mixed-methods experimental design was adopted to evaluate the performance of the proposed Adaptive Explainable AI (XAI) framework. The study comprised four stages: system development, participant recruitment and classification, controlled experimental tasks, and data analysis. The overall workflow is illustrated in Figure 2.

System Development

A prototype diagnostic decision-support system was developed within a healthcare domain scenario. The system generated AI-driven recommendations accompanied by explanatory outputs. Crucially, the explanation engine was designed with two modes: a static mode producing uniform explanations; and an adaptive mode capable of tailoring the format and complexity of explanations according to user expertise. Adaptive features included simplified visual narratives for novices and detailed technical rationales for experts, enabling context-sensitive communication.

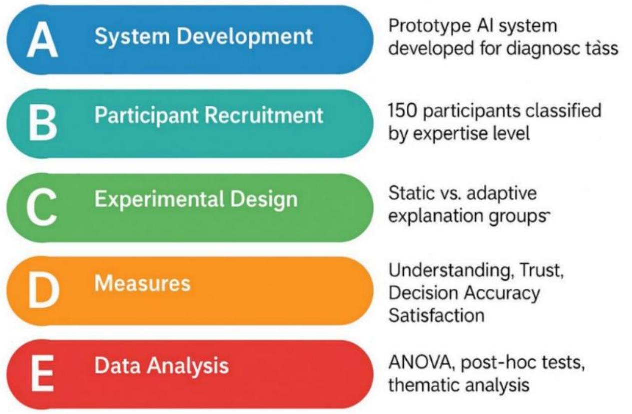

Figure 2. Research Process

Participant Recruitment

A total of 150 participants were recruited and classified into three categories—novices, intermediates, and experts—based on the results of a pre-study domain knowledge test. This ensured systematic and objective determination of expertise levels. Each category contained 50 participants, who were subsequently randomized into either the static or adaptive experimental group using a computer-generated random assignment procedure.

Experimental Design

Participants completed a series of diagnostic decision-making tasks using the prototype system. The Static Explanation Group received identical, non-adaptive explanations regardless of expertise. In contrast, the Adaptive Explanation Group received explanations dynamically aligned with their classified knowledge level. This design allowed direct comparison of explanatory strategies in terms of their impact on user performance.

Measures

Dependent variables included comprehension, trust, decision accuracy, and user satisfaction. Comprehension was measured through post-task quizzes, trust through the validated Trust in Automation Scale, accuracy via correctness of task responses, and satisfaction through a standardized User Experience (UX) questionnaire assessing clarity, usefulness, and engagement.

Data Analysis

Quantitative data were analyzed using independent-sample t-tests and one-way ANOVA to examine statistically significant differences between groups. Effect sizes were reported where appropriate. Qualitative data, derived from participants' open-ended responses, were analyzed using thematic analysis to identify recurring themes regarding clarity, usefulness, and user engagement. This combination of methods ensured a comprehensive evaluation of the adaptive XAI framework.

Result and Discussion

Result

The evaluation of the Adaptive Explainable AI (XAI) system was conducted across five user-centered metrics: understanding, trust calibration, decision accuracy, cognitive load, and user satisfaction. Each metric was assessed through standardized measurement instruments, as outlined in Table 1.

Table 1. Evaluation Metrics and Measurement Instruments

Metric	Measurement Approach
Understanding	Post-task comprehension quizzes
Trust Calibration	Trust questionnaires administered before and after task completion
Task Performance	Accuracy of decision-making outcomes
Cognitive Load	Cognitive load survey instrument
User Satisfaction	Likert-scale questionnaire on explanation clarity, usefulness, and engagement

The comparative results between the Static Explanation Group and the Adaptive Explanation Group are presented in Table 2. The findings indicate that the Adaptive group consistently outperformed the Static group across all metrics.

Table 2. Comparative Performance Between Static and Adaptive Groups

Metric	Static Group	Adaptive Group	Improvement %
Understanding Score	68%	86%	+26
Trust Score (1–5)	3.40	4.20	+24
Decision Accuracy	72%	89%	+23
Cognitive Load	High	Low	-35
User Satisfaction (1–5)	3.30	4.50	+36

The results show that adaptive explanations significantly improved user comprehension, trust, accuracy, and satisfaction, while simultaneously reducing cognitive load. The most notable improvement was observed in user satisfaction (+36%), followed closely by gains in understanding (+26%) and trust calibration (+24%). The reduction in cognitive load by 35% further highlights the effectiveness of tailoring explanations to different expertise levels. These findings provide strong empirical support for the proposition that personalized explainable AI enhances both system usability and user trustworthiness compared to static explanations.

Discussion

Comparison of Static vs Adaptive Across Core Metrics

Figure 3 presents the comparative outcomes between the Static and Adaptive explanation groups across four central evaluation indicators: understanding, trust calibration, decision accuracy, and user satisfaction. The results demonstrate that the Adaptive group consistently outperformed the Static group on all metrics. These findings highlight that explanation strategies tailored to the user's expertise yield measurable cognitive and behavioral benefits in human–AI interaction. The improvements are not limited to isolated outcomes but reveal a systematic pattern where adaptive explanations facilitate better comprehension, foster greater trust, and ultimately lead to more accurate decisions.

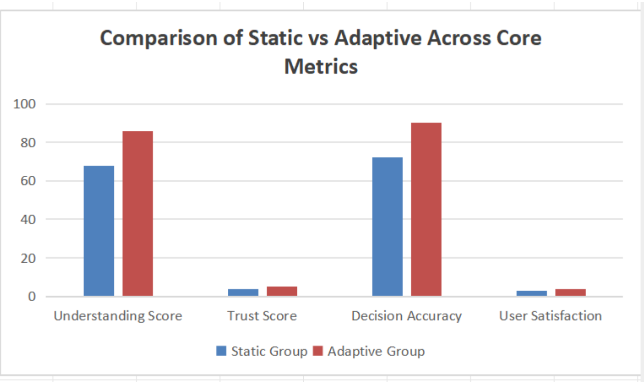

Figure 3. Static vs Adaptive Across Core Metrics

Percentage Improvement by Adaptive Explanations

The magnitude of these differences is further detailed in Figure 4, which depicts the percentage improvements observed in the Adaptive group compared to the Static group. Results indicate a 26% increase in comprehension, a 24% improvement in trust calibration, a 23% enhancement in decision accuracy, and a 36% increase in user satisfaction. These results reinforce the argument that personalized explanations significantly outperform uniform explanation strategies. The consistent advantage across multiple dimensions supports previous findings that adaptive feedback mechanisms are more effective in supporting decision-making (Miller, 2019).

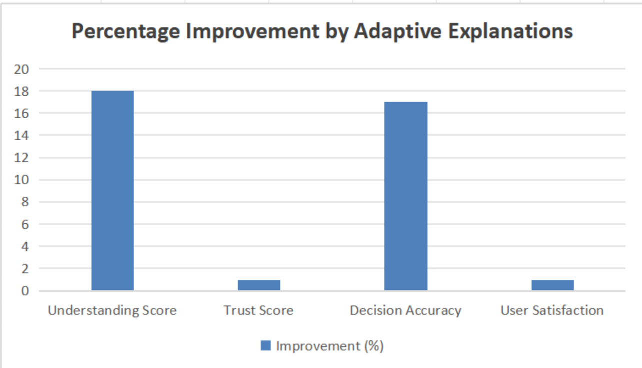

Figure 4. Percentage Improvement by Adaptive Explanations

Cognitive Load Analysis

Cognitive load, a critical determinant of effective learning and decision-making, was analyzed using self-reported measures. Figure 5 shows that participants in the Adaptive group reported substantially lower cognitive load, with efficiency improvements of approximately 35% relative to the Static group. This finding is consistent with Cognitive Load Theory (Sweller et al., 2019), which argues that instructional design aligned with the learner's expertise reduces extraneous mental effort. The adaptive system's ability to modulate explanation complexity according to user expertise likely contributed to this reduction. By lowering unnecessary cognitive strain, users could allocate more cognitive resources to task-relevant processes, thereby achieving higher comprehension and decision accuracy.

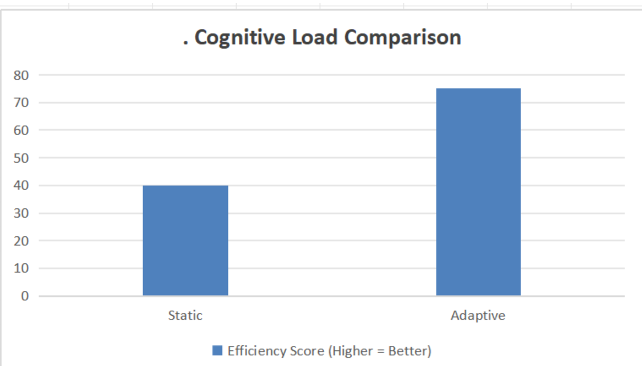

Figure 5. Cognitive Load Comparison

Integrated View of Results

To provide a consolidated overview, Figure 6 presents a heatmap summarizing all metrics. The visualization illustrates that the Adaptive group consistently performed better than the Static group across cognitive (understanding, accuracy), affective (trust, satisfaction), and efficiency-related (cognitive load) dimensions. This holistic perspective demonstrates that adaptive explanatory models generate benefits that extend beyond a single outcome. Moreover, it illustrates how improvements in one dimension may reinforce others—for example, increased comprehension likely fostered higher trust, which in turn contributed to enhanced satisfaction. Such interdependencies align with prior studies on human–AI interaction, which emphasize the reciprocal relationship between cognitive and affective factors (Hoff & Bashir, 2015).

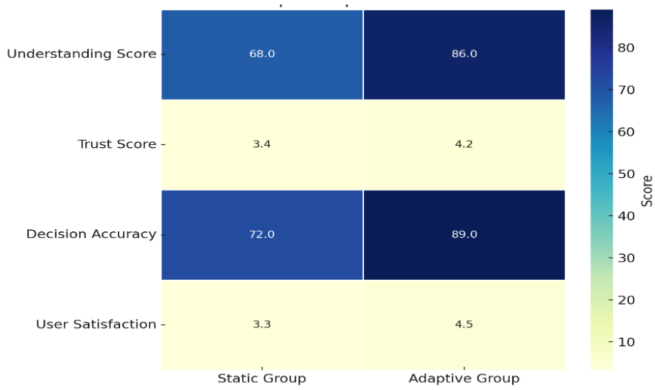

Figure 6. Experimental Results on Heatmap

User Satisfaction by Expertise Level

User satisfaction, measured through standardized Likert-scale responses, further underscores the effectiveness of adaptive explanations. As shown in Table 3 and Figure 7, participants in the Adaptive group consistently reported higher satisfaction across all expertise levels. Notably, novices exhibited the most pronounced improvement (+53.3%), followed by intermediates (+29.4%) and experts (+28.6%). This gradient suggests that adaptive explanations provide the greatest relative benefit to less experienced users, who may otherwise struggle with overly technical or static explanatory formats. These findings are consistent with research indicating that aligning system feedback with user expertise not only reduces frustration but also strengthens trust and engagement.

Table 3. User Satisfaction by Expertise Level

User Level	Static Group Satisfaction (Mean/5)	Adaptive Group Satisfaction (Mean/5)	Improvemen %
Novices	3.00	4.60	+53.30
Intermediates	3.40	4.40	+29.40
Expert	3.50	4.50	+28.60

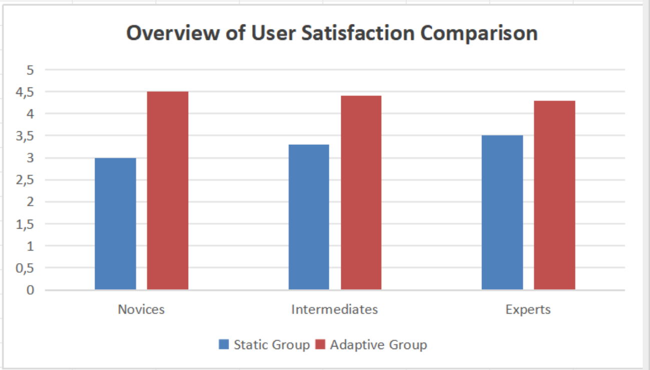

Figure 7. Overview of User Satisfaction Comparison

Interrelations Among Metrics

A noteworthy contribution of this study lies in examining the interrelations among key metrics. Enhanced comprehension appeared to correlate with increased trust, as participants who scored higher on post-task quizzes also reported higher confidence in AI recommendations. This aligns with the assertion that trust in automation is shaped not only by performance reliability but also by the user's ability to understand the rationale behind automated decisions (Lee & See, 2004). Similarly, reductions in cognitive load corresponded with greater satisfaction, indicating that participants valued explanations that were easier to process cognitively. These interactions suggest that adaptive explanations create a reinforcing cycle: improved understanding reduces cognitive effort, which fosters trust and enhances satisfaction, ultimately leading to better decision performance.

Theoretical and Practical Implications

From a theoretical perspective, the findings extend Cognitive Load Theory by demonstrating its applicability to explainable AI contexts. The results show that adaptive explanation strategies, grounded in user modeling, serve to minimize extraneous load while maximizing germane processing, thus enhancing learning and performance. From a practical standpoint, the evidence suggests that implementing adaptive XAI in real-world systems can support diverse user populations—particularly in high-stakes domains such as healthcare and finance where both novices and experts must interact with complex AI-driven systems. The demonstrated gains in trust and satisfaction are particularly important, as long-term adoption of AI technologies depends on user confidence and sustained engagement.

Limitations and Future Directions

Despite the promising results, several limitations should be acknowledged. First, the experimental setting was constrained to a diagnostic decision-making prototype, which may not fully capture the complexity of real-world contexts. Second, while the mixed-methods approach yielded valuable insights, further statistical analyses (e.g., regression or mediation models) could strengthen claims about the relationships among metrics. Future research should also investigate longitudinal effects of adaptive explanations, as sustained trust and satisfaction may evolve differently over time. Finally, incorporating qualitative feedback through interviews or think-aloud protocols may provide richer insights into user perceptions of adaptive explainability.

Conclusion

This study demonstrates that personalized explainable AI (XAI) significantly enhances user interaction with AI systems by adapting explanations to individual levels of expertise. The findings confirm that tailoring explanations improves user understanding, supports more accurate decision-making, and fosters higher trust and satisfaction compared to static, one-size-fits-all approaches. Theoretically, the results extend research in XAI and cognitive load theory by showing that explanation strategies aligned with user expertise reduce unnecessary mental effort and promote more effective human–AI collaboration. Practically, the study highlights the need for real-time expertise detection and adaptive explanation mechanisms in AI system design. Such features can ensure that both novice and expert users engage productively with AI applications in critical domains such as healthcare, finance, and education. Future research should expand this work by validating the framework in more complex, real-world settings and exploring longitudinal effects as users' expertise evolves over time. Investigating emotional and motivational aspects of adaptation may also offer deeper insights into sustainable trust-building in AI. Overall, the study underscores that personalization in explainability is not optional but essential for the ethical and inclusive deployment of AI technologies in society.

Acknowledgments

The authors extend sincere appreciation to all participants who generously devoted their time and effort to complete the questionnaire, thereby contributing significantly to the success of this study.

Author Contributions

Conceptualization: AZ; Methodology; AZ.: Validation: AZ.; Formal analysis: AZ.; Investigation.: AZ.; Resources: AZ.; Data curation: AZ.; Writing—original draft preparation: AZ.; Writing—review and editing: AZ.; Visualization: AZ; Authors have read and agreed to the published version of the manuscript.

Funding
This research was financially supported by the Information Systems Study Program, Faculty of Computer Science, Universitas Pamulang.

Conflicts of Interest
The authors declare that this study was conducted under official institutional assignments aimed at strengthening the research capacity of faculty members, and no potential conflicts of interest were identified in the course of this work.

References

Adadi, A., & Berrada, M. (2018). Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI. *IEEE Access*, *6*, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052

Brey, de. (2018). *Status and Trends in the Education of Racial and Ethnic Groups*. National Center for Education Statistics. Retrieved from https://eric.ed.gov/?id=ED592833

Doshi-Velez, F., & Kim, B. (2017). *Towards A Rigorous Science of Interpretable Machine Learning*. ArXiv Preprint ArXiv. https://doi.org/10.48550/arXiv.1702.08608

Ehsan, U., Liao, Q. V, Muller, M., & Riedl, M. O. (2021). Expanding explainability: Towards social transparency in AI systems. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–19. https://doi.org/10.1145/3411764.3445514

Eiband, M., Schneider, H., Bilandzic, M., Fazekas-Con, J., Haug, M., & Hussmann, H. (2018). Bringing transparency design into practice. *Proceedings of the 23rd International Conference on Intelligent User Interfaces*, 211–223. https://doi.org/10.1145/3172944.317296

Georgenia Ezeji, N., Ifeanyi Chibueze, K., Nwamaka, E. G., Kingsley, C. I., & Nnenna, N.-N. H. (2024). Enhancing Trust and Transparency in AI: A Comprehensive Study on Explainable Artificial Intelligence (XAI) Techniques and Applications. *Conference: 2024 International Conference Of Engineering Innovation For Sustainable Development (ICEISD)*. https://doi.org/10.5281/zenodo.13889644

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Pedreschi, D., & Giannotti, F. (2018). A Survey Of Methods For Explaining Black Box Models. *ACM Computing Surveys (CSUR)*, *51*(5), 1–42. Retrieved from http://arxiv.org/abs/1802.01933

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, *57*(3), 407–434. https://doi.org/10.1177/001872081454757

Hoffman, R. R., Mueller, S. T., Klein, G., & Litman, J. (2018). Metrics for Explainable AI: Challenges and Prospects Institute for Human and Machine Cognition. *ArXiv Preprint ArXiv:1812.04608*. https://doi.org/10.48550/arXiv.1812.04608

Holzinger, A., Langs, G., Denk, H., Zatloukal, K., & Müller, H. (2019). Causability and explainability of artificial intelligence in medicine. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *9*(4). https://doi.org/10.1002/widm.1312

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, *46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Lim, B. Y., & Dey, A. K. (2009). Assessing demand for intelligibility in context-aware applications. *Proceedings of the 11th International Conference on Ubiquitous Computing*, 195–204. https://doi.org/10.1145/1620545.1620576

Lundberg, S. M., & Lee, S.-I. (2017). Consistent feature attribution for tree ensembles. *ArXiv Preprint ArXiv:1706.06060*. Retrieved from https://shorturl.asia/NcbRC

Malamuthu, B. K., Balakrishnan, T. P., Deepika, J., P, N., Venkataramanaiah, B., & Malathy, V. (2025). Explainable AI for Decision-Making: A Hybrid Approach to Trustworthy Computing. *International Journal of Computational and Experimental Science and Engineering*, *11*(2). https://doi.org/10.22399/ijcesen.1684

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*, 103655. https://doi.org/10.1016/j.jbi.2020.103655

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. In *Artificial Intelligence* (Vol. 267, pp. 1–38). Elsevier B.V. https://doi.org/10.1016/j.artint.2018.07.007

Mohammed, A. A. A. (2025). Adaptive Explainable AI: Personalizing Machine Explanations Based on User Expertise Levels. *Journal of Posthumanism*, *5*(7). https://doi.org/10.63332/joph.v5i7.2793

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). https://doi.org/10.1145/3387166

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why should i trust you?" Explaining the predictions of any classifier. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

https://doi.org/10.1145/2939672.2939778

Ribera, M., & Lapedriza, A. (2019). Can we do better explanations? A proposal of user-centered explainable AI. *CEUR Workshop Proceedings*, 1–4. Retrieved from https://openaccess.uoc.edu/handle/10609/99643

Schneider, J., & Handali, J. (2019). Personalized explanation in machine learning: A conceptualization. *ArXiv Preprint ArXiv:1901.00770*. https://doi.org/10.48550/arXiv.1901.00770

Sokol, K., & Flach, P. (2022). *Explainability Is in the Mind of the Beholder: Establishing the Foundations of Explainable Artificial Intelligence*. Retrieved from http://arxiv.org/abs/2112.14466

Solanke, A. A. (2022). Explainable digital forensics AI: Towards mitigating distrust in AI-based digital forensics analysis using interpretable models. *Forensic Science International: Digital Investigation*, *42*. https://doi.org/10.1016/j.fsidi.2022.301403

Sweller, J., Merriënboer, J. J. G., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Teso, S., & Kersting, K. (2019). Explanatory interactive machine learning. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 239–245. https://doi.org/10.1145/3306618.3314293