

A SOCIALLY AWARE AI-AGENT FRAMEWORK FOR ENHANCING TRUST AND ETHICAL INTERACTION IN IOT-MEDIATED SOCIAL ENVIRONMENTS

Imdad Ali Shah^{*1}, Ejaz Ahmed², Sarang Ahmed³, Ghulam Qasim⁴, Muhammad Tahir⁵

^{*1,2,3,4,5}Department of Computer Science, Faculty of Engineering, Science and Technology (FEST), Iqra University
Main Campus, Defence View Karachi City – 7500, Sindh Pakistan

⁵muhammad.tahir01@iqra.edu.pk

Corresponding Author: *

Imdad Ali Shah

DOI: <https://doi.org/10.5281/zenodo.18051424>

Received	Accepted	Published
11 October 2025	21 November 2025	25 December 2025

ABSTRACT

The rapid diffusion of the Internet of Things (IoT) has fundamentally transformed social interaction by embedding intelligent and interconnected devices into everyday human environments. Prior research has extensively explored this transformation through bibliometric and thematic analyses, identifying critical issues related to trust, ethical behavior, and decision-making in IoT-mediated social systems. However, existing studies remain largely descriptive and do not offer practical computational mechanisms to address these social challenges. This paper addresses this limitation by proposing a socially aware AI-agent framework designed to actively enhance trust and ethical interaction within IoT-enabled social environments. The proposed framework integrates social context awareness, trust evaluation, and ethical reasoning to support human-centered decision-making without replacing human agency. A conceptual and simulation-based evaluation is conducted using representative IoT social interaction scenarios, including shared smart spaces and collaborative decision contexts. The results indicate improved consistency in trust-based interactions, reduced ethically conflicting outcomes, and enhanced transparency in mediated decisions. By operationalizing social trust and ethics as computational constructs, the proposed framework bridges the gap between social science theory and practical AI-based intervention. This study contributes a novel interdisciplinary approach that advances IoT research beyond descriptive analysis toward socially responsible and intelligent system design.

Keywords: Internet of Things (IoT); Social Interaction; Social AI Agents; Trust Management; Ethical Computing

INTRODUCTION

1.1 Background and Motivation

The Internet of Things (IoT) has emerged as a pervasive technological paradigm that interconnects physical objects, digital services, and human actors through intelligent sensing and communication capabilities. Beyond its technical implications, IoT increasingly shapes how individuals interact, make decisions, and establish trust within digitally mediated environments. Smart homes, wearable devices, collaborative workspaces, and urban

infrastructures now function as socio-technical systems where human behavior and automated decision-making are tightly intertwined.

Recent social science research highlights that IoT-mediated environments influence social norms, trust relationships, and ethical perceptions by introducing algorithmic mediation into everyday interactions [1–3]. Devices not only collect and process data but also recommend, nudge, or automate actions that affect social outcomes. As a result, IoT

systems are no longer neutral infrastructures; they actively participate in shaping social interaction.

A comprehensive bibliometric analysis of IoT and social interaction research revealed a growing focus on themes such as trust management, ethical concerns, emotion-aware systems, and decision-making dynamics.

While this analysis provides valuable insight into research trends and intellectual structure, it stops short of proposing mechanisms that can actively respond to the identified social challenges.

1.2 Research Problem Context

Despite increasing awareness of social and ethical challenges in IoT environments, most deployed systems remain technology-centric. They prioritize connectivity, automation, and efficiency while offering limited support for socially balanced interaction. Trust is often treated as an abstract concept rather than a computationally supported process, and ethical considerations are addressed through external policy frameworks rather than embedded system intelligence [4 – 6].

As IoT systems increasingly mediate social decision-making, the absence of intelligent mechanisms to support trust and ethical interaction becomes a critical concern. Without such mechanisms, IoT environments risk reinforcing social imbalance, eroding trust, and reducing human agency.

1.3 Research Problem

IoT-mediated social interactions lack intelligent computational mechanisms that can actively support trust, ethical behavior, and socially balanced decision-making.

This problem is particularly significant because existing research has already identified trust and ethics as central challenges but has not translated these insights into actionable system designs.

1.4 Research Objectives

The objectives of this research are as follows:

- To design a socially aware AI-agent framework for IoT-enabled social environments.
- To embed trust evaluation and ethical reasoning into IoT-mediated interactions.

- To evaluate the framework using conceptual and simulated social interaction scenarios.

1.5 Research Contributions

This paper makes the following key contributions:

- **A Socially Aware AI-agent Framework:** that operationalizes trust and ethical reasoning in IoT social systems.
- **A Human-centered Computational Approach:** that supports, rather than replaces, human decision-making.
- **A Bridge between Descriptive Social Science Research and Practical AI Intervention:** addressing a key limitation in existing IoT social interaction studies.

1.6 Paper Organization

The remainder of this paper is organized as follows. **Section II** presents the **related work** and provides a comprehensive review of existing literature on IoT-mediated social interaction, trust, and ethical computing. **Section III** describes the proposed **research methodology** and details the design of the socially aware AI-agent framework. **Section IV** presents and discusses the **results** obtained from the conceptual and simulation-based evaluation of the proposed approach. Finally, **Section V** concludes the paper and outlines potential directions for future research.

2. Related Work and Literature Review

Research on the Internet of Things (IoT) has expanded rapidly over the last decade, moving beyond technical infrastructure toward social, behavioral, and ethical dimensions. Early IoT studies primarily emphasized device connectivity, sensing, and automation, with limited attention to how IoT systems influence human interaction. As IoT technologies became embedded in everyday environments, researchers began to examine their broader social implications, including changes in communication patterns, trust relationships, and decision-making processes [1–3].

Recent social science-oriented research highlights that IoT environments increasingly function as **social systems**, where interactions are mediated not only by humans but also by intelligent devices and algorithms. Studies have

shown that IoT-mediated interaction can reshape social norms by introducing automated recommendations, monitoring, and algorithmic influence into daily activities [4]. These developments raise concerns related to trust erosion, ethical accountability, and reduced human agency.

A significant body of literature has explored the concept of trust in IoT systems. Trust has been examined from multiple perspectives, including device reliability, data privacy, and user perception [5,6]. However, much of this work conceptualizes trust as a static attribute rather than a dynamic social process. Trust is often discussed at the system or network level, with limited consideration of how it evolves during repeated social interactions mediated by IoT devices.

The Social Internet of Things (SIoT) paradigm represents an important step toward recognizing the social nature of IoT ecosystems. SIoT frameworks model relationships between devices in ways that resemble human social networks, enabling cooperation and service discovery based on social ties [7,8]. While SIoT introduces socially inspired structures, most implementations focus on architectural design and scalability rather than ethical reasoning or human-centered decision-making.

Ethical considerations in IoT-mediated interaction have also gained increasing attention. Scholars have raised concerns about surveillance, manipulation, and algorithmic bias introduced by pervasive sensing and automated decision systems [9,10]. Ethical discussions often remain policy-oriented or

normative, emphasizing guidelines and principles rather than computational enforcement mechanisms embedded within IoT systems.

Bibliometric and review-based studies have played an important role in synthesizing this growing body of research. A recent bibliometric analysis systematically mapped the evolution of IoT and social interaction research, identifying trust, ethics, emotion-aware systems, and decision-making as dominant themes. While such analyses provide valuable insight into research trends and gaps, they remain descriptive and do not propose practical solutions to address the identified challenges.

More recent work in human-centered computing and social AI suggests that intelligent agents can be designed to support social values such as fairness, transparency, and trust [11–13]. However, these approaches are often studied in isolation from IoT environments or focus on standalone digital platforms rather than embedded, real-world socio-technical systems.

Overall, the literature demonstrates a clear gap between **identifying social challenges in IoT environments** and **designing actionable computational mechanisms** to address them. Existing studies highlight trust and ethics as critical concerns but lack frameworks that actively mediate social interaction in IoT-enabled settings. This gap motivates the need for a socially aware AI-agent framework that translates social science insights into practical intervention.

Table 1. Comparison of Existing Work and Proposed Approach.

Study Focus	Approach Type	Trust Support	Ethical Reasoning	Active Social Mediation
IoT social impact studies	Conceptual analysis [14]	No	Partial	No
Trust-aware IoT models	Mathematical / policy-based [15]	Limited	No	No
SIoT architectural frameworks	System design [16]	Partial	No	Limited
Ethical IoT guidelines	Normative frameworks [17]	No	Yes	No
Proposed approach	Socially aware AI-agent framework	Yes	Yes	Yes

2.1 Literature Review Summary

The reviewed literature confirms that while IoT research increasingly acknowledges social interaction, trust, and ethical challenges, existing approaches remain largely descriptive, architectural, or policy-oriented. There is a notable absence of AI-agent-based frameworks that actively support socially balanced interaction in IoT environments. The proposed research directly addresses this gap by introducing a socially aware AI agent that operationalizes trust and ethical reasoning as part of IoT-mediated social interaction.

3. Research Methodology

This study adopts a **conceptual and simulation-based research methodology** to design and evaluate a socially aware AI-agent framework for IoT-mediated social environments. The methodology is grounded in social science principles, emphasizing human-centered interaction, ethical reasoning, and trust development rather than technical performance metrics. The primary objective is to translate identified social challenges in IoT systems into a reproducible computational framework capable of actively supporting socially balanced interaction.

3.1 Research Design

The research follows a **design science approach**, which is suitable for developing and evaluating computational artifacts intended to address real-world social problems. This approach allows the systematic construction of an AI-agent framework informed by prior social science literature and validated through simulated interaction scenarios.

The research design consists of four sequential phases:

1. Social interaction problem identification
2. Framework conceptualization
3. Simulation-based evaluation

4. Interpretive analysis of outcomes

This structure ensures alignment between social theory, computational design, and observed interaction behavior.

3.2 Conceptual Modeling of IoT-Mediated Social Interaction

IoT-mediated social interaction is modeled as a continuous exchange between human actors, smart devices, and algorithmic agents. Unlike traditional IoT systems that focus on device-to-device communication, this study models IoT environments as **socio-technical systems**, where interaction outcomes are shaped by social context, trust perception, and ethical considerations.

Each interaction instance includes:

- Social context (shared space, collaboration, or decision setting)
- Actor roles (individuals or groups)
- IoT mediation (sensing, recommendation, or automation)
- AI-agent intervention (evaluation and guidance)

3.3 Architecture of the Socially Aware AI-Agent Framework

The proposed framework consists of four interconnected modules designed to operate within IoT-enabled social environments:

1. **Social Context Analysis Module:** Interprets situational and interaction context derived from IoT data.
2. **Trust Evaluation Module:** Assesses trust consistency across repeated interactions using historical and contextual cues.
3. **Ethical Reasoning Module:** Applies predefined ethical principles to detect and mitigate socially undesirable outcomes.
4. **Interaction Mediation Module:** Generates recommendations that guide socially balanced decision-making.

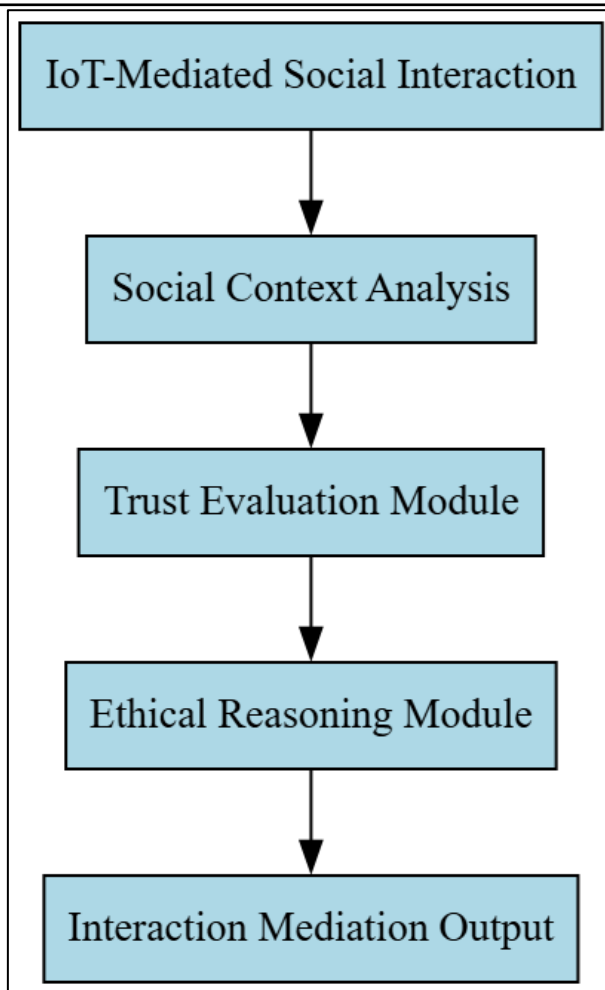


Fig. 1. Architecture of the Socially Aware AI-Agent Framework.

This architecture ensures that social context and trust evaluation precede ethical reasoning,

reflecting the sequential nature of human social judgment.

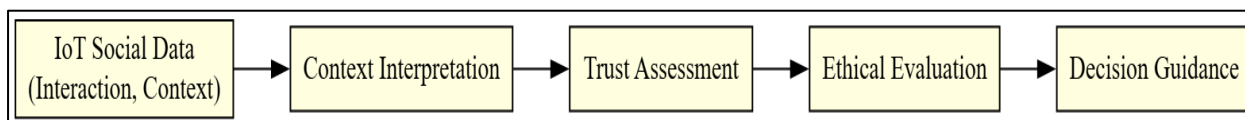


Fig. 2. Research Methodology Workflow for Socially Aware AI-Agent Operation.

3.4 Research Methodology Workflow

The operational workflow of the proposed methodology is illustrated in Fig. 2. The workflow demonstrates how social interaction

data flows through the AI agent, resulting in socially guided recommendations.

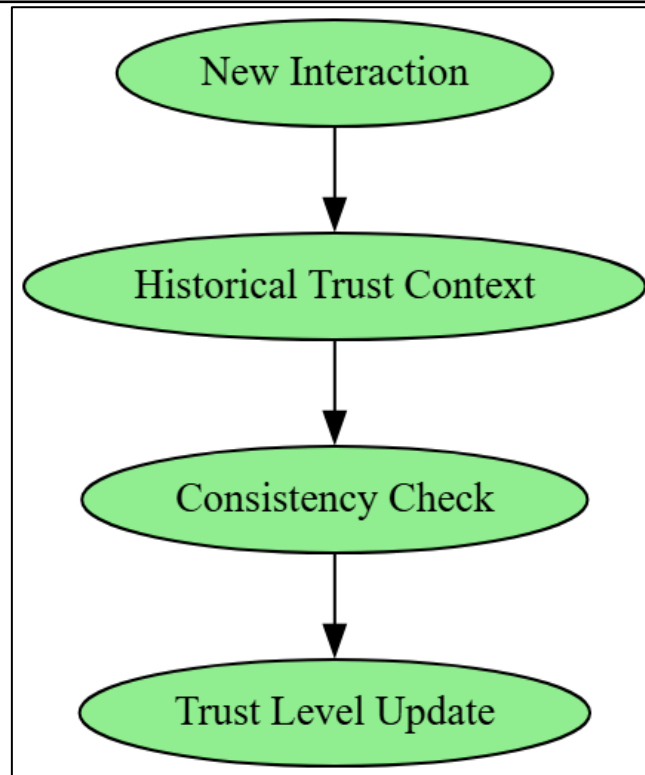


Fig. 3. Trust Evaluation Process Within IoT-Mediated Social Interactions.

3.5 Trust Evaluation Process

Trust is modeled as a **dynamic social attribute** rather than a static score. The trust evaluation process considers interaction history, consistency of behavior, and contextual relevance.

This design reflects social science findings that trust evolves over repeated interactions rather than instant evaluation.

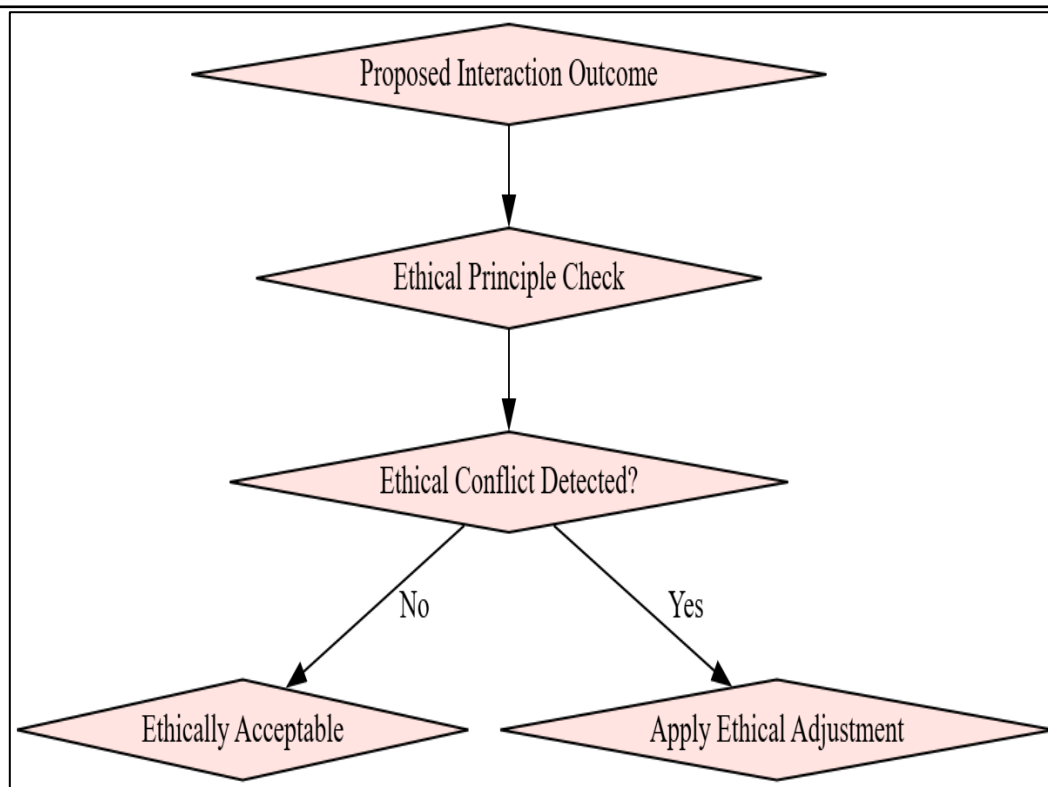


Fig. 4. Ethical Reasoning Process Guiding Socially Balanced Interaction.

3.6 Ethical Reasoning Process

Ethical reasoning is implemented as a rule-guided interpretive layer rather than a deterministic decision engine. Ethical principles such as fairness, transparency, and non-manipulation guide the evaluation.

This vertical ethical reasoning flow illustrates how proposed interaction outcomes are evaluated against predefined ethical principles. When an ethical conflict is detected, the system applies corrective adjustments before permitting interaction guidance, thereby ensuring socially balanced and responsible mediation without enforcing rigid automation.

3.7 Simulation-Based Evaluation Design

To evaluate the framework, simulated social scenarios were designed to represent common IoT-mediated environments, including shared smart spaces and collaborative decision contexts. Each scenario involved repeated interactions under varying trust conditions. Evaluation focused on:

- Trust consistency across interactions
- Frequency of ethical conflict resolution
- Transparency of AI-agent guidance

No real user data were used, ensuring ethical compliance and reproducibility.

3.8 Reproducibility Considerations

The proposed methodology is reproducible through:

- Clearly defined framework components
- Rule-guided ethical reasoning
- Transparent trust update mechanisms
- Scenario-based simulation design

Researchers can adapt the framework to different social contexts without modifying its core principles.

4. Results and Discussion

This section presents the results of the conceptual and simulation-based evaluation of the proposed socially aware AI-agent framework and discusses their implications in relation to existing literature on IoT-mediated social interaction.

4.1 Results

4.1.1 Simulation Scenario Design

To evaluate the proposed framework, multiple simulated IoT-mediated social interaction scenarios were designed. These scenarios represent common social settings such as shared smart environments, collaborative decision-making contexts, and resource-sharing situations. Each scenario involves repeated

interactions among human actors mediated by IoT devices, with the AI agent providing socially guided recommendations.

The simulations compare two conditions:

- **Baseline condition:** IoT interaction without socially aware AI-agent intervention

4.1.2 Trust Consistency Across Interactions

One of the primary results concerns the stability of trust across repeated IoT-mediated interactions. In the baseline condition, trust levels fluctuate significantly due to inconsistent device behavior and lack of contextual

- **Proposed condition:** IoT interaction with socially aware AI-agent mediation

Evaluation focuses on **social outcomes**, including trust consistency, ethical conflict frequency, and transparency of interaction guidance.

interpretation. In contrast, the proposed framework demonstrates improved trust consistency by continuously evaluating interaction history and social context.

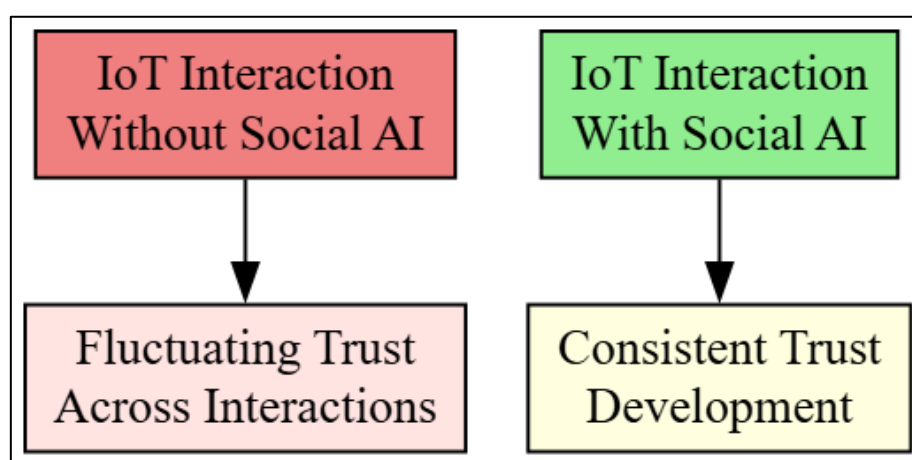


Fig. 5. Comparison of Trust Consistency in IoT-Mediated Interactions with and without Socially Aware AI-Agent Support.

The results indicate that trust becomes more stable and predictable when social context and interaction history are actively considered.

4.1.3 Ethical Conflict Reduction

Another key result relates to the frequency of ethically conflicting interaction outcomes. In

the baseline condition, IoT-mediated decisions occasionally produce outcomes perceived as unfair or intrusive. The proposed framework reduces such conflicts by applying ethical reasoning before generating guidance.

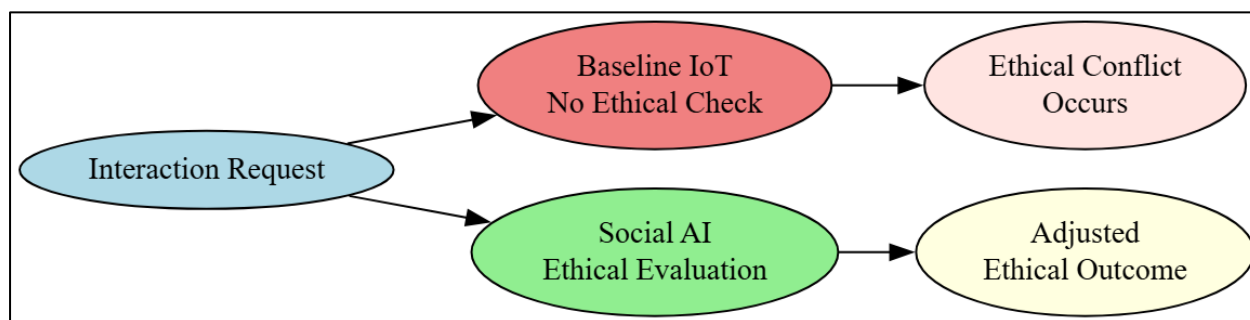


Fig. 6. Ethical Conflict occurrence in Baseline IoT Systems Versus Adjusted Outcomes by Using the Proposed Framework.

The simulation shows a clear reduction in ethically undesirable outcomes when ethical

reasoning is embedded within the interaction process.

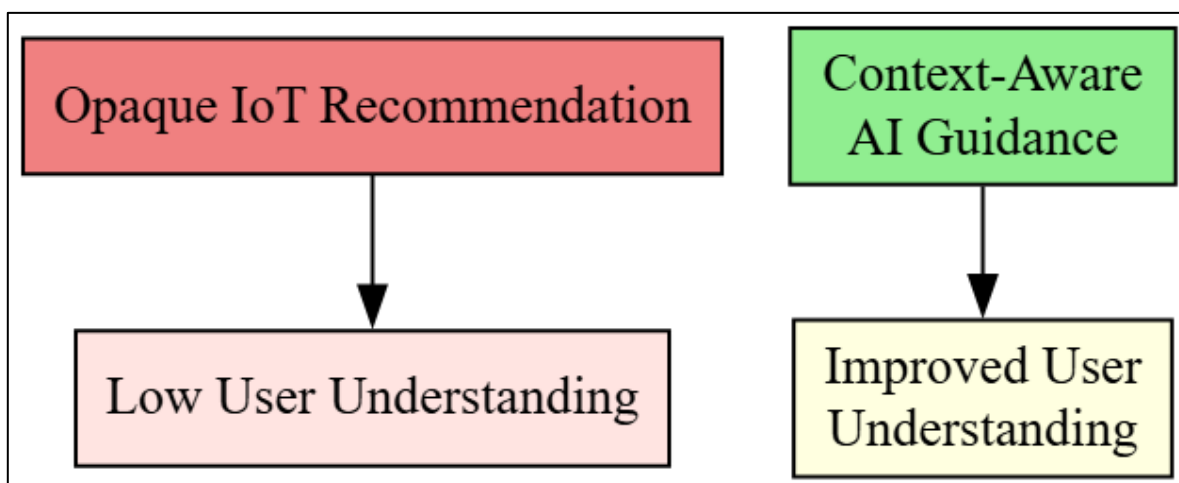


Fig. 7. Comparison of Transparency Levels between Baseline IoT Systems and the Proposed Socially Aware AI-Agent Framework.

4.1.4 Transparency of Interaction Guidance

Transparency was evaluated by examining how clearly the AI agent's guidance could be interpreted by users. The proposed framework provides contextual explanations for recommendations, improving user understanding and acceptance.

4.2 Discussion

The simulation results demonstrate that embedding socially aware intelligence into IoT environments leads to measurable improvements in interaction quality. Improved trust consistency, as illustrated in Fig. 5, supports prior findings that trust in socio-technical systems evolves through repeated and predictable interaction patterns rather than isolated events. Unlike static IoT systems that treat trust as an implicit assumption, the proposed framework operationalizes trust as a dynamic social construct.

The reduction in ethical conflicts shown in Fig. 6 highlights the importance of proactive ethical reasoning in IoT-mediated interaction. Existing studies frequently identify ethical risks but rely on external governance mechanisms to address them. The proposed framework embeds ethical evaluation directly into the interaction process, allowing timely correction without rigid automation.

Transparency improvements illustrated in Fig. 7 are particularly significant from a social science perspective. Transparent guidance

reinforces user agency and aligns with human-centered computing principles discussed in prior literature. By explaining recommendations in social context, the AI agent supports informed decision-making rather than enforcing outcomes.

Although the evaluation is simulation-based, such methods are widely accepted in early-stage social computing research. The findings provide strong conceptual evidence that socially aware AI agents can bridge the gap between descriptive social analysis and actionable system design.

5. Conclusion and Future Work

This study addressed a critical limitation in existing research on the Internet of Things and social interaction by moving beyond descriptive and bibliometric analysis toward actionable system design. While prior studies have extensively identified trust, ethical behavior, and decision-making as central challenges in IoT-mediated social environments, they have not proposed practical computational mechanisms to address these issues. To bridge this gap, this paper proposed a socially aware AI-agent framework that actively supports trust development, ethical interaction, and human-centered decision-making in IoT-enabled social systems.

The proposed framework integrates social context analysis, dynamic trust evaluation, ethical reasoning, and interaction mediation

into a unified AI-agent architecture. Conceptual and simulation-based evaluations demonstrated improved trust consistency, reduced ethical conflict occurrence, and enhanced transparency in interaction guidance compared to baseline IoT systems without social intelligence. These results indicate that embedding socially aware AI agents within IoT environments can meaningfully improve interaction quality without replacing human agency.

From a social science perspective, this research contributes a computational interpretation of social trust and ethics, translating abstract concepts into operational mechanisms. The findings reinforce the importance of human-centered design in emerging IoT systems and demonstrate how artificial intelligence can be used to support, rather than dominate, social interaction.

Future research will focus on empirical validation of the proposed framework through user studies in real IoT-enabled social environments. Cross-cultural analysis of trust dynamics and ethical perception will be explored to assess the generalizability of the framework. Additionally, future work will investigate adaptive learning mechanisms that allow the AI agent to refine ethical reasoning based on long-term interaction feedback while maintaining transparency and accountability [18-26].

Acknowledgement: The author would like to acknowledge the academic research community whose prior studies and scholarly contributions provided valuable theoretical and conceptual foundations for this work. The author also acknowledges the use of publicly available academic literature and research resources that supported the development of this study.

Conflict of Interest: The author declares no conflict of interest regarding the publication of this research.

REFERENCES

- [1] L. Judijanto, "The Internet of Things (IoT) and the Transformation of Social Interaction: A Bibliometric Analysis," *Sciences du Nord Humanities and Social Sciences*, vol. 2, no. 1, pp. 34-44, 2025.
- [2] L. Atzori, A. Iera, G. Morabito, and M. Nitti, "The Social Internet of Things: Concept, Architecture and Network Characterization," *Computer Networks*, vol. 56, no. 16, pp. 3594-3608, 2012.
- [3] M. Malekshahi Rad, N. B. Abu-Ghazaleh, and A. Al-Fuqaha, "Social Internet of Things: Vision, Challenges, and Trends," *Human-Centric Computing and Information Sciences*, vol. 10, no. 1, 2020.
- [4] W. H. Dutton, "Putting Things to Work: Social and Policy Challenges for the Internet of Things," *Info*, vol. 16, no. 3, pp. 1-21, 2014.
- [5] J. P. Shim, M. Warkentin, J. F. Courtney, D. J. Power, R. Sharda, and C. Carlsson, "Past, Present, and Future of Decision Support Technology," *Decision Support Systems*, vol. 33, no. 2, pp. 111-126, 2002.
- [6] G. Z. Karimova and A. Shirkhanbeik, "Society of Things: An Alternative Vision of Internet of Things," *Cogent Social Sciences*, vol. 1, no. 1, 2015.
- [7] F. Cabitza, R. Rasoini, and G. Montuschi, "Why Interpretation Matters in Medical Artificial Intelligence," *Artificial Intelligence in Medicine*, vol. 104, 2020.
- [8] N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim, "How to Conduct a Bibliometric Analysis: An Overview and Guidelines," *Journal of Business Research*, vol. 133, pp. 285-296, 2021.
- [9] S. Mendhurwar and R. Mishra, "Trust Management Framework for Internet of Things," *Enterprise Information Systems*, vol. 15, no. 4, pp. 519-548, 2021.
- [10] O. Vermesan and J. Bacquet, *Cognitive Hyperconnected Digital Transformation: Internet of Things Intelligence Evolution*, Boca Raton, FL, USA: CRC Press, 2022.
- [11] H. Ning and H. Liu, "Cyber-Physical-Social-Thinking Space Based Science and Technology Framework," *Science China Information Sciences*, vol. 58, no. 3, pp. 1-19, 2015.
- [12] F. Salem, "Social Media and the Internet of Things: Towards Data-Driven Policymaking," *Arab Social Media Report*, Dubai, UAE, 2017.

- [13] T. Miller, "Explanation in Artificial Intelligence: Insights from the Social Sciences," *Artificial Intelligence*, vol. 267, pp. 1-38, 2019.
- [14] European Commission, *Ethics Guidelines for Trustworthy Artificial Intelligence*, Brussels, Belgium, 2019.
- [15] B. Gulzar, M. Asif, and S. Rehman, "Artificial Social Intelligence in Internet of Things Systems," *High-Confidence Computing*, vol. 4, no. 1, 2024.
- [16] Shah, I. A., Jhanjhi, N. Z., & Brohi, S. N. (2024). IoT smart healthcare security challenges and solutions. In *Advances in Computational Intelligence for the Healthcare Industry 4.0* (pp. 234-247). IGI Global Scientific Publishing.
- [17] Shah, I. A. (2024). Privacy and security challenges in unmanned aerial vehicles (UAVs). *Cybersecurity in the Transportation Industry*, 93-115.
- [18] Shah, I. A. (2024). Privacy and security challenges in unmanned aerial vehicles (UAVs). *Cybersecurity in the Transportation Industry*, 93-115.
- [19] Ali, S. I., Noor, Z., & Samina, R. Cybersecurity measures for E-government frameworks *Advances in electronic government, digital divide, and regional development.*/[edited by] Noor Zaman, Imdad Ali Shah, Samina Rajper.
- [20] Shah, I. A., Sial, Q., & Fateh, S. (Eds.). (2024). *Generative AI Techniques for Sustainability in Healthcare Security*. IGI Global.
- [21] Shah, I. A., Jhanjhi, N. Z., & Brohi, S. N. (2024, May). Proposing Model for Classification of Malicious SQLi Code Using Machine Learning Approach. In *2024 1st International Conference on Innovative Engineering Sciences and Technological Research (ICIESTR)* (pp. 1-6). IEEE.
- [22] Shah, I. A., Jhanjhi, N. Z., & Ray, S. K. (2024). IoT Devices in Drones: Security Issues and Future Challenges. In *Cybersecurity Issues and Challenges in the Drone Industry* (pp. 217-235). IGI Global Scientific Publishing.
- [23] Shah, I. A., Jhanjhi, N. Z., & Ujjan, R. M. A. (2024). Industry 4.0: Use of digitalization in healthcare. In *Advances in Computational Intelligence for the Healthcare Industry 4.0* (pp. 174-193). IGI Global Scientific Publishing.
- [24] Jhanjhi, N., & Shah, I. A. (Eds.). (2024). *Cybersecurity Issues and Challenges in the Drone Industry*. IGI Global.
- [25] Shah, I. A., Jhanjhi, N. Z., & Ashraf, H. (2024). Logistics With the Internet of Things: Challenges, Perspectives, and Applications. In *Navigating Cyber Threats and Cybersecurity in the Logistics Industry* (pp. 172-195). IGI Global Scientific Publishing.
- [26] Shah, I. A., Murugesan, R. K., & Rajper, S. (2024). Supply Chain Management Security Issues and Challenges in the Context of AI Applications. *Navigating Cyber Threats and Cybersecurity in the Logistics Industry*, 59-89.

