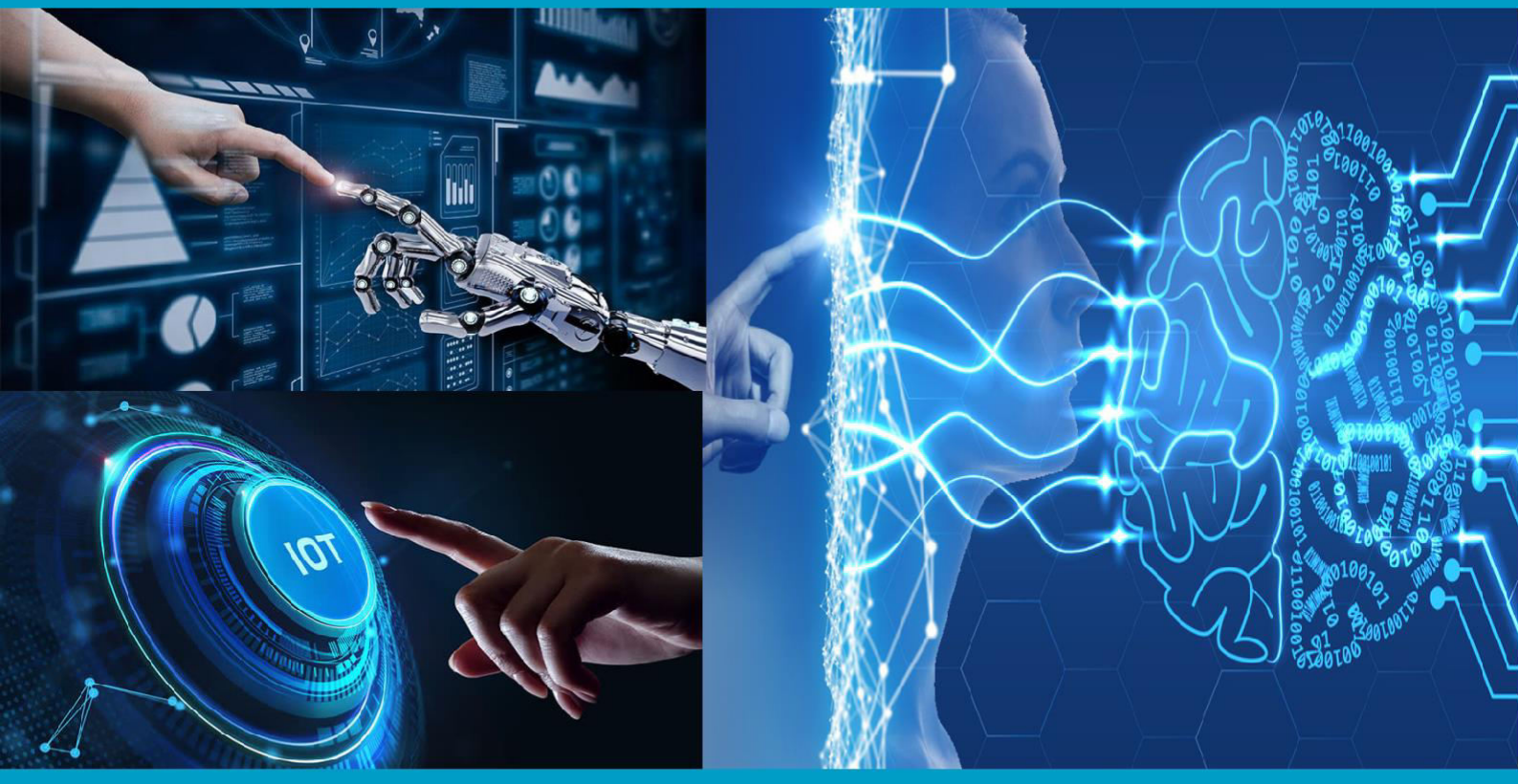
IJMRSETM

e-ISSN: 2395 - 7639




# INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

## IN SCIENCE, ENGINEERING, TECHNOLOGY AND MANAGEMENT

Volume 10, Issue 1, January 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.580

International Journal of Multidisciplinary Research in Science, Engineering, Technology & Management (IJMRSETM)

| ISSN: 2395-7639| www.ijmrsetm.com| Impact Factor: 7.580 | A Monthly Double-Blind Peer Reviewed Journal |

|| Volume 10, Issue 1, January 2023 ||

# Human Trust and AI: Psychological Factors

**Dr. Shivya Saxena**

Assistant Professor, Bharathi College of Education, Kandri, Mandar, Ranchi, Jharkhand, India

**ABSTRACT:** In the rapidly evolving landscape of artificial intelligence (AI), the concept of trust plays a pivotal role, influencing how individuals interact with and perceive AI systems. This interaction is not merely technical but deeply intertwined with psychological factors that shape human-AI relationships. Understanding these dynamics is crucial as AI increasingly integrates into various aspects of daily life, from healthcare to finance, from education to entertainment. Human trust in AI is a multifaceted phenomenon that draws from psychological principles such as perception, cognition, and social influence. It is shaped by several key factors that intertwine to determine the extent and nature of trust placed in AI systems.

**KEYWORDS**: Human Trust, Artificial Intelligence, Psychological Factors

## I. INTRODUCTION

In the rapidly evolving landscape of artificial intelligence (AI), the concept of trust plays a pivotal role, influencing how individuals interact with and perceive AI systems. This interaction is not merely technical but deeply intertwined with psychological factors that shape human-AI relationships. Understanding these dynamics is crucial as AI increasingly integrates into various aspects of daily life, from healthcare to finance, from education to entertainment. Human trust in AI is a multifaceted phenomenon that draws from psychological principles such as perception, cognition, and social influence. It is shaped by several key factors that intertwine to determine the extent and nature of trust placed in AI systems. This exploration delves into the nuanced interplay between human psychology and AI technologies, aiming to elucidate how and why individuals place trust in these systems [1].

## II. REVIEW STUDY

**Pieters (2011)** explores the pivotal link between explanation and trust in AI and information security. His synthesis of system and actor-network theories illuminates how justifying decisions fosters user trust, critical in both domains. By applying this framework to expert systems and electronic voting, Pieters underscores its universal utility, urging ethical considerations in user consent and dissent.

**Lyons & Stokes (2012)** delve into human reliance amidst human-automation interactions, crucially noting shifts in reliance patterns under varying risk scenarios. Their Convoy Leader experiment unveils nuanced dynamics where automated tools influence decision-making, impacting both behavioral and subjective reliance indices.

**Yagoda & Gillan (2012)** develop a multidimensional trust scale for human-robot interaction (HRI), highlighting its role in fostering acceptance and usage. Their studies validate factors crucial for establishing trust within diverse HRI contexts, offering insights vital for designing trustworthy robotic systems.

**Verberne et al. (2012)** scrutinize trust dynamics in smart systems, emphasizing how shared goals enhance perceived trustworthiness and acceptability. Their findings underscore the role of automation levels in shaping user trust perceptions, crucial for advancing adaptive technologies.

**Touré-Tillery & McGill (2015)** investigate the influence of messenger type on persuasion, revealing differential responses based on interpersonal trust levels. Their studies underline how message delivery impacts persuasion efficacy, with implications for effective advertising strategies.

**Blöbaum (2016)** theorizes on trust in academic discourse, distinguishing its multifaceted components and evolving manifestations in digital contexts. His analysis of trust dynamics underscores the evolving nature of trust in contemporary digital environments.

**Lahijanian & Kwiatkowska (2016)** advocate for a comprehensive understanding of social trust in human-robot relationships, stressing its pivotal role in enhancing collaborative outcomes. Their call for interdisciplinary research seeks to formalize trust theories essential for future robotic advancements.

| ISSN: 2395-7639| www.ijmrsetm.com| Impact Factor: 7.580 | A Monthly Double-Blind Peer Reviewed Journal |

|| Volume 10, Issue 1, January 2023 ||

**Lee (2018)** explores perceptions of algorithmic versus human decision-making in managerial contexts, revealing nuanced biases in fairness and trust attributions. His findings highlight the complexities of human-computer interactions in organizational settings.

**Neigel et al. (2018)** investigate the impact of trust in automated decision support systems on task performance, emphasizing its critical role in enhancing operational efficacy. Their study underscores the symbiotic relationship between trust and system performance in intelligence tasks.

**Følstad et al. (2018)** address factors influencing user trust in chatbots for customer service, emphasizing the role of bot quality, human-likeness, and contextual factors in shaping trust perceptions. Their insights provide a roadmap for designing effective and trusted chatbot interactions in service environments.

**Glikson, E., & Woolley, A. W. (2020).** Artificial intelligence (AI) characterizes a new generation of technologies capable of interacting with the environment and aiming to simulate human intelligence. The success of integrating AI into organizations critically depends on workers' trust in AI technology. This review explains how AI differs from other technologies and presents the existing empirical research on the determinants of human "trust" in AI, conducted in multiple disciplines over the last 20 years. Based on the reviewed literature, we identify the form of AI representation (robot, virtual, and embedded) and its level of machine intelligence (i.e., its capabilities) as important antecedents to the development of trust and propose a framework that addresses the elements that shape users' cognitive and emotional trust.

**Riedl, R. (2022).** Artificial intelligence (AI) refers to technologies which support the execution of tasks normally requiring human intelligence (e.g., visual perception, speech recognition, or decision-making). Examples for AI systems are chatbots, robots, or autonomous vehicles, all of which have become an important phenomenon in the economy and society. Determining which AI system to trust and which not to trust is critical, because such systems carry out tasks autonomously and influence human-decision making. This growing importance of trust in AI systems has paralleled another trend: the increasing understanding that user personality is related to trust, thereby affecting the acceptance and adoption of AI systems. We developed a framework of user personality and trust in AI systems which distinguishes universal personality traits (e.g., Big Five), specific personality traits (e.g., propensity to trust), general behavioral tendencies (e.g., trust in a specific AI system), and specific behaviors (e.g., adherence to the recommendation of an AI system in a decision-making context).

## III. TRUST FORMATION AND PERCEPTION

Trust in AI is not a mere transactional concept but a complex psychological process influenced by factors such as transparency, reliability, and explainability. Humans tend to trust AI systems that are transparent about their decision-making processes and demonstrate reliability over time. Perceptions of AI's capabilities and intentions significantly impact how trust is formed and maintained [2].

## IV. ETHICAL CONSIDERATIONS AND BIAS

Ethical concerns surrounding AI algorithms, particularly biases in data and decision-making, profoundly affect trust. Instances of AI systems perpetuating or amplifying societal biases can erode trust and exacerbate disparities. Addressing these issues requires proactive measures to ensure fairness, accountability, and inclusivity in AI development and deployment [3].

## V. HUMAN-AI INTERACTION DYNAMICS

The nature of interactions between humans and AI influences trust dynamics. Factors such as user experience, interface design, and personalized interactions play crucial roles in fostering or hindering trust. Effective design that prioritizes user-centric principles and clear communication enhances user confidence and engagement with AI technologies [4].

## VI. PSYCHOLOGICAL SAFETY AND RISK PERCEPTION

Human perceptions of risk associated with AI technologies significantly impact trust levels. Concerns about privacy, security, and the potential for AI to replace human jobs contribute to varying levels of trust. Establishing psychological safety through robust security measures, regulatory frameworks, and transparent policies is essential to mitigate these concerns and build trust [5].

| ISSN: 2395-7639| www.ijmrsetm.com| Impact Factor: 7.580 | A Monthly Double-Blind Peer Reviewed Journal |

|| Volume 10, Issue 1, January 2023 ||

## VII. CULTURAL AND SOCIETAL INFLUENCES

Trust in AI is also shaped by cultural norms, societal values, and historical experiences. Cultural differences in expectations of technology and varying levels of acceptance towards AI innovations underscore the importance of context-specific approaches in fostering trust. Understanding cultural nuances can guide more effective AI implementation strategies globally [6-7].

## VIII. BUILDING AND SUSTAINING TRUST

Establishing initial trust in AI systems is challenging but essential for long-term acceptance and adoption. Continuous efforts to build trust through ethical practices, user empowerment, stakeholder engagement, and clear communication about AI's benefits and limitations are critical. Sustaining trust requires ongoing monitoring, adaptation to user feedback, and responsiveness to evolving societal expectations [8-10].

## IX. CONCLUSION

The complex interplay between human psychology and AI technologies reveals that trust is a foundational element in fostering acceptance and successful integration of AI into society. The factors influencing trust in AI transparency, reliability, ethical considerations, human-AI interaction dynamics, psychological safety, and cultural influences highlight the need for comprehensive strategies in AI development and deployment. To effectively build and sustain trust, stakeholders must prioritize transparency in AI decision-making processes, mitigate biases through rigorous ethical frameworks, and enhance user experience through intuitive design and clear communication. Addressing societal concerns such as privacy, security, and job displacement requires proactive measures and collaborative efforts across industries and regulatory bodies. Moreover, recognizing and adapting to cultural nuances in different global contexts is essential for designing AI systems that resonate with diverse populations and enhance trust universally. By fostering an environment of openness, accountability, and continuous improvement, stakeholders can navigate the complexities of human-AI relationships and maximize the benefits of AI while mitigating potential risks. Ultimately, as AI technologies continue to evolve, nurturing trust remains foundational not only for technological advancement but also for ensuring that these advancements align with human values and societal needs. Through ongoing dialogue, research, and ethical reflection, we can collectively shape a future where AI enhances human well-being and fosters a more inclusive and trustworthy digital era.

## REFERENCES

1. **Pieters, W. (2011).** Explanation and trust: what to tell the user in security and AI?. *Ethics and information technology*, *13*, 53-64.
2. **Lyons, J. B., & Stokes, C. K. (2012).** Human–human reliance in the context of automation. *Human factors*, *54*(1), 112-121.
3. **Lee, M. K. (2018).** Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*, *5*(1), 2053951718756684.
4. **Blöbaum, B. (2016).** Key factors in the process of trust. On the analysis of trust under digital conditions. *Trust and communication in a digitized world: Models and concepts of trust research*, 3-25.
5. **Neigel, A. R., Caylor, J. P., Kase, S. E., Vanni, M. T., & Hoye, J. (2018).** The role of trust and automation in an intelligence analyst decisional guidance paradigm. *Journal of Cognitive Engineering and Decision Making*, *12*(4), 239-247.
6. **Yagoda, R. E., & Gillan, D. J. (2012).** You want me to trust a ROBOT? The development of a human–robot interaction trust scale. *International Journal of Social Robotics*, *4*, 235-248.
7. **Lahijanian, M., & Kwiatkowska, M. (2016, September).** Social trust: a major challenge for the future of autonomous systems. In *2016 AAAI Fall Symposium Series*.
8. **Touré-Tillery, M., & McGill, A. L. (2015).** Who or what to believe: Trust and the differential persuasiveness of human and anthropomorphized messengers. *Journal of Marketing*, *79*(4), 94-110.
9. **Følstad, A., Nordheim, C. B., & Bjørkli, C. A. (2018).** What makes users trust a chatbot for customer service? An exploratory interview study. In *Internet Science: 5th International Conference, INSCI 2018, St. Petersburg, Russia, October 24–26, 2018, Proceedings 5* (pp. 194-208). Springer International Publishing.
10. **Verberne, F. M., Ham, J., & Midden, C. J. (2012).** Trust in smart systems: Sharing driving goals and giving information to increase trustworthiness and acceptability of smart systems in cars. *Human factors*, *54*(5), 799-810.
11. **Glikson, E., & Woolley, A. W. (2020).** Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, *14*(2), 627-660.
12. **Riedl, R. (2022).** Is trust in artificial intelligence systems related to user personality? Review of empirical evidence and future research directions. *Electronic Markets*, *32*(4), 2021-2051.

Impact Factor: 7.580

# INTERNATIONAL JOURNAL
## OF MULTIDISCIPLINARY RESEARCH
### IN SCIENCE, ENGINEERING, TECHNOLOGY AND MANAGEMENT

+91 99405 72462 +91 63819 07438 ijmrsetm@gmail.com