Trust in Human-AI Interaction: Review of Empirical Research on Trust in AI-powered Smart Home Ecosystems

Tianzhi He,[1] Farrokh Jazizadeh, Ph.D.[2]

[1]Graduate Research Assistant, Department of Civil and Environmental Engineering, Virginia Tech, 315 Patton Hall, 750 Drillfield, Blacksburg, VA, 24061; email: tianzhi@vt.edu
[2]Associate Professor, Department of Civil and Environmental Engineering, Virginia Tech, 200 Patton Hall, 750 Drillfield, Blacksburg, VA, 24061; email: jazizade@vt.edu

ABSTRACT

The growing socio-technical influence of artificial intelligence (AI) is increasing in various dimensions of our daily lives. The monitoring and decision-making role of smart home AI agents could go beyond simple operations, and a lack of users' trust could lead to the underuse or misuse of their capabilities. As such, lying at the intersection of trust in human-AI interactions and smart home ecosystems, this paper attempts to explore the driving factors of building trustworthy AI-powered smart home ecosystems. To this end, through a systematic literature review, we have identified the characteristics of human-AI trust in intelligent environments, factors that can influence users' trust in AI, features that can be implemented to improve users' trust, and challenges and opportunities that need to be considered. The findings can help develop a pathway for building long-term users' trust in AI-powered smart homes for integrated human-AI interactions.

Key Words: Trust in AI, Human-AI Interaction, Smart Homes, AI agents, Virtual Assistants

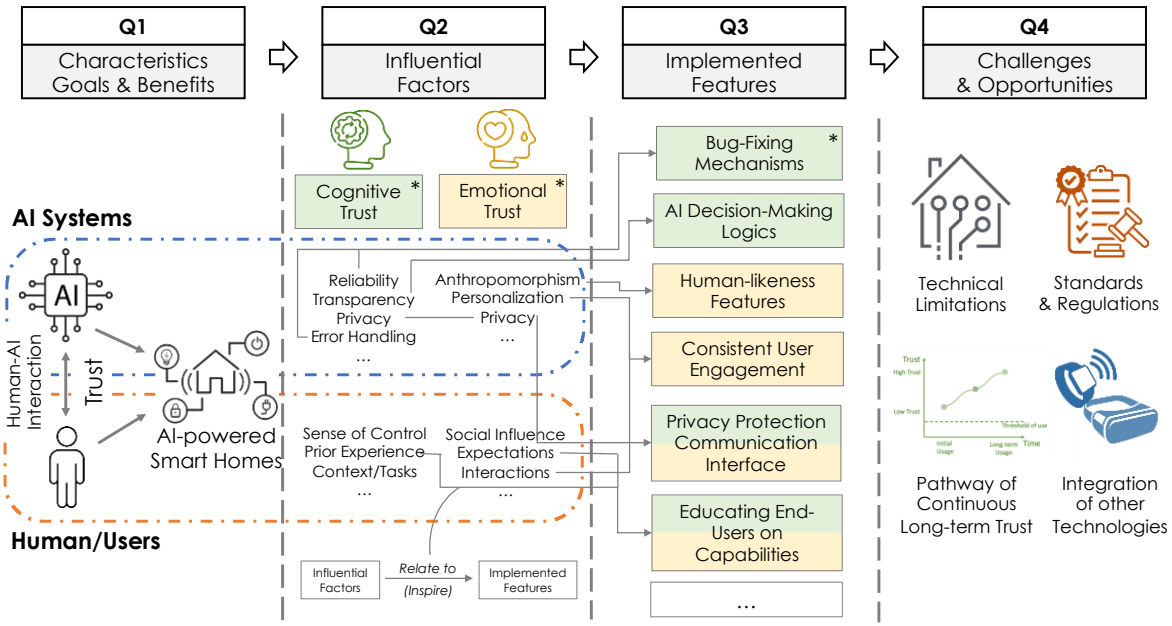
INTRODUCTION

Over the past decade, there has been a rapid development of Artificial Intelligence (AI) technologies in support of human activities. Various application domains (e.g., healthcare, transportation, and construction) have incorporated AI to perform tasks, such as image recognition, natural language processing, and decision-making. With the ubiquitous adoption of AI-infused systems, human-AI interaction has become more common and plays an increasingly important role, motivating the need to maintain trust between human users and AI (Glikson and Woolley 2020). As a critical component of human-AI interaction, trust in AI is the foundation for the successful and sustainable deployment of AI systems. Due to the black-box characteristic of AI systems, users show hesitation toward AI-based technologies, indicating a relatively low level of trust in AI in today's society (Ueno et al. 2022). As such, researchers have conducted various investigations in the field of trust in human-AI interaction. However, despite extensive research on human-AI trust using different theories and models, little is known about human trust in AI-powered intelligent environments (e.g., smart buildings and smart homes) (Liu et al. 2021). Compared with the trust between human and individual AI tools or agents (e.g., chatbots or robots), human trust in AI-powered intelligent environments pertains to the trust people place in environments where multiple AI systems work together to create a seamless, interconnected experience. While there may be some overlap between them, they have distinct characteristics and nuances in terms of their focus, scope, user control, and complexity.

Intelligent environments, such as smart homes, are environments that integrate the Internet of Things (IoT) and AI technologies to provide a more convenient and sustainable living

environment for residents (He and Jazizadeh 2022). By integrating smart devices such as sensors, smart thermostats, and actuators, AI agents can collect and analyze data about the environment and users, contextualize and optimize the operations to meet users' needs (Cannizzaro et al. 2020). The market penetration of smart central hubs, such as smart speakers (e.g., Amazon Echo, Google Home) has further accelerated the adoption of AI in smart homes, allowing users to interact with their smart environment through voice-based conversations communication, making human-AI interaction more intuitive and user-friendly (He et al. 2022). The widespread adoption of AI-powered smart homes with voice assistants has accentuated the importance of building human-AI trust, as it is critical to ensuring that users feel secure and willing to rely on AI agents performing daily tasks and making decisions on their behalf (Liu et al. 2021). Lack of trust can lead to limited adoption of AI in intelligent environments, hindering its potential to help users improve their lives and reducing the benefits of human-AI interaction (Glikson and Woolley 2020).

Research and development efforts are exploring various strategies to build and maintain Human-AI trust in intelligent environments, with the goal of creating trustworthy and responsive AI systems that users can rely on. This paper aims to provide an understanding of how trust in human-AI interaction could be accounted for in intelligent environments. We sought to identify the driving factors that could lead to a successful implementation of AI-powered smart home ecosystems through the lens of trust in human-AI interaction. To this end, we conducted a systematic literature review to identify these factors by answering four questions (Figure 1). We first present a comprehensive definition of human-AI trust in intelligent environments, including its terminologies, different dimensions, goals, and benefits (***Q1***). Then, we describe how different factors affect trust (***Q2***), and what features can be implemented to promote trust in AI-powered smart homes (***Q3***). Finally, we discuss the challenges and opportunities (***Q4***).



* The green and yellow background colors are associated with cognitive and emotional trust, respectively.

Figure 1. Pathway to trust in AI-powered smart home ecosystems

QUESTIONS AND ANSWERS

Following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) systematic literature review guideline (Page et al. 2021), a comprehensive keyword search was

conducted on electronic databases (Google Scholar, Scopus, and ASCE Library) to identify the relevant studies published in the past five years (2018-2023). The search terms included variations of keywords such as: “Trust in AI”, “Trustworthy AI”, “Trust in Smart Homes”, “Trust in Voice Assistants”, and “Trust in Smart Speakers”. The relevant studies in the reference list of collected articles were also examined. Through manual filters, studies not tied to the factors and features related to trust in AI or smart homes were excluded, and studies including the analysis of factors and features in human trust were selected. For studies that meet the inclusion criteria, findings in these studies were synthesized into the answer of the following four questions:

Q1: How is human-AI trust in intelligent environments characterized?

Drawing on the definition of interpersonal trust, trust in human-AI interaction includes two fundamental elements: vulnerability and anticipation (Jacovi et al. 2021). Human users, believing that AI-powered systems will act in the users' best interests, delegate the responsibility and consequences of an action to the AI, which also indicates the users' willingness to be vulnerable to the AI's actions (Ueno et al. 2022). Moreover, trust in AI is an attempt to anticipate the impact of AI actions under risk, which includes undesirable and adverse events that may, but not certainly occur (Jacovi et al. 2021). According to studies of trust in human-AI interaction, we summarize that human-AI trust is defined as the willingness of human users to rely on AI systems to make critical decisions or take actions with anticipated consequences.

Trust in AI is a dynamic concept that depends on the characteristics of AI agents (Ueno et al. 2022). Therefore, trust in human-AI interaction in intelligent environments could be different from the general concept of trust in AI. In this context, we discuss trust in an envisioned AI-powered intelligent environments with virtual assistants that can not only interact directly (either passively or proactively (He et al. 2022) with users and provide recommendations but also control various devices and systems in the environment, including HVAC (heating, ventilation, and air conditioning), lighting, and security systems. This requires a higher level of trust as AI agents become more integrated into people's daily routines and activities. In this context, various terms and metrics are used to assess the level of trust, such as general trust, perceived risk, intention to use, adoption, and acceptability.

Prior studies have considered different dimensions of trust and focused on two main types when it comes to trust in AI: cognitive and emotional trust (Glikson and Woolley 2020). Cognitive trust is a rational and logical form of trust that involves users’ objective evaluation of the AI agents' cognitive capabilities, such as reliability, accuracy, competence, and processing speed. Through positive interaction experiences, users can gradually develop the belief that the AI has the cognitive ability to sense their needs and adjust accordingly. Emotional trust, on the other hand, is an intuitive and affective form of trust based on users’ emotional perception of AI systems (Glikson and Woolley 2020). Users would emotionally trust an AI virtual assistant to provide helpful recommendations because they would perceive the AI as friendly, empathetic, and genuinely interested in their needs and preferences. Although these two types of trust are distinct dimensions, they can influence each other. A positive emotional experience can increase an individual's cognitive trust in a new AI system. Likewise, a negative emotion (e.g., fear and concern) can decrease cognitive trust, even if the individual has tested the system's capabilities.

Q2: How do different factors affect trust in AI-powered smart homes?

Different dimensions of trust in AI with varied antecedents can be affected by different factors (Glikson and Woolley 2020). As such, we compiled empirical research that examined the factors that can affect both cognitive and emotional trust and further categorized them from two

perspectives of AI systems and human users as shown in Table 1. It is worth noting that some of the factors are not unique to one type of trust. For example, privacy and interaction experiences can affect users' both cognitive and emotional trust.

Table 1. Factors affecting trust in AI-powered smart homes

Perspectives	Cognitive Trust	Emotional Trust
AI Systems	• Reliability/Accuracy • Capability • Security • Automation • Transparency • Compatibility • Error handling • Privacy preservation • Expectation/capability communication • Interaction interfaces	• Anthropomorphism • Emotion expressions • Adaptability/Personalization • Ethical considerations • Institution reputation • Interaction interfaces
Human/Users	• Interaction experiences • Sense of control • Prior experience and knowledge of AI • Context/Tasks	• Interaction experiences • Perceived privacy • Perceived expectation • Social influence

AI Systems' Perspective: Many characteristics of AI systems can have a considerable effect on users' cognitive trust in AI-powered smart homes. The reliability, accuracy, and capability of the system are the most important factors that can affect trust. Prior studies have found that the system quality of voice-enabled smart homes (e.g., accurate and timely responses to users' commands) has a significant positive impact on users' perceived trust (Liu et al. 2021). Another important factor is privacy. AI system privacy risks can affect users' concern about privacy and trust in AI, thus affecting their behavioral intention and adoption of AI-based voice assistants (Sharif and Tenbergen 2020). In terms of security, many security breaches of AI systems can negatively affect users' trust, especially in the IoT context. For example, single-point access to IoT devices and poor authentication of third-party services make AI-powered smart homes vulnerable to cyber-physical attacks and the theft of user identity and sensitive data (Sharif and Tenbergen 2020). AI systems must be secure enough against attacks by unauthorized agents to be trusted by users (Emaminejad et al. 2021). The automation level of AI-powered smart homes may affect users' sense of control and perceived risks, and thus affect their cognitive trust in AI systems. Also, the automation-augmentation paradox requires the AI systems to take a human-in-the-loop approach to engage users in the development of the smart home automation (Lockey et al. 2021). Transparency of the AI systems has also been found to be effective in improving cognitive trust by providing explanations of the decision-making processes (Glikson and Woolley 2020). When users understand how the AI system works and why it makes certain decisions, they are more likely to trust it with rational evaluation. Compatibility and usefulness of an application were found to be critical determinants of the use intention in smart homes (Hubert et al. 2019). If an AI system integrates well with other devices and services, users are more likely to trust it. How the system handles errors and malfunctions can also impact trust. If AI systems respond well to errors (e.g., correction after making mistakes) and quickly resolve issues (e.g., fixing device integration bugs), users would hold the belief that the AI systems are trustworthy even when they encounter errors (Cuadra et al. 2021; Wang et al. 2022).

Other dimensions of the AI systems can specifically affect users' emotional trust. For example, anthropomorphic characteristics of the AI agents (e.g., gender, accent, dialect, human-like expressions) can affect users' trust by developing emotional connections with users (Rheu et al. 2021). When the AI system with a humanoid voice and a certain "personality" interacts with users, it may help users to feel more emotionally attached to it. Previous studies have also shown that emotional expressions of the AI systems (e.g., empathetic voices, tone of expressions) play an important role in the interactions between human and voice-based AI agents (Seaborn et al. 2021). In terms of adaptability and personalization, the extent to which the AI agents can adapt to the user's preferences and needs can affect users' emotional attachment, with which the AI systems can gain emotional trust from users (Glikson and Woolley 2020). It was also identified in previous studies that the ethical considerations of the AI systems and the reputation of the AI systems institutions or manufacturers (on safety, reliability, or privacy) can affect how users emotionally trust the AI systems (Ryan 2020; Seymour et al. 2022).

Human/Users' Perspective: Many human subjective factors can influence their cognitive and emotional trust in AI-powered smart homes. One important factor from the human side is the users' sense of control, which may affect users' perceptions and behaviors in smart homes (Glikson and Woolley 2020). Prior experience and knowledge of AI can also influence users' trust. A UK study found that awareness of the IoT and smart home, as well as prior experience and use of smart home devices, were positively correlated with general trust in the smart home and intention of adoption (Cannizzaro et al. 2020). Another critical factor in cognitive trust in AI can be the characteristics of the tasks and the contexts in which they are performed (Glikson and Woolley 2020). Users would evaluate whether the AI system is trustworthy enough based on the tasks it would perform. Some other human factors can also be influential in users' emotional trust. Social norms and public opinion are very important factors in users' decisions to adopt smart home devices (Liu et al. 2021). Perceived expectation is also a critical factor that can affect users' cognitive and emotional trust. Based on the trust trajectory theory (Glikson and Woolley 2020), if users have high expectations for AI capabilities and find that the reality of the technology does not match their expectations, their trust in AI systems would drop dramatically, and could even lead to distrust or abandonment of the AI systems. The interaction experiences with the AI systems such as users' perceived usefulness or ease of use can also directly affect users' trust in the AI-powered smart homes (Liu et al. 2021).

Q3: What can be implemented to promote trust in AI-powered smart homes?

With the various factors mentioned above, prior studies have applied different methods and strategies to promote trust in AI-powered smart homes. From the perspective of AI systems, users' trust can be improved by implementing different features on the AI agents and the smart home IoT. Wang et al. (2022) investigated the root causes, trigger conditions, and fixing methods of the device integration bugs of smart homes to improve the error-handling mechanism of the AI systems. When a smart home device behaves abnormally, such as failing to configure, incorrectly scheduling tasks, or losing connectivity, the AI system should be able to identify and detect the errors and provide clear and informative error messages that help users understand what went wrong and how to fix it. In this way, users' cognitive trust can be enhanced through increased durability and reliability of the system, and their emotional trust can be promoted through reduced frustration caused by malfunction or failure during human-AI interactions (Cuadra et al. 2021).

Another effective strategy to increase user trust is to improve the transparency and explainability of AI systems. A transparent and explainable AI system in smart homes should be capable of informing users about the activities (e.g., data collection and device control), as well as

explaining their actions and recommendations (Emaminejad et al. 2021). In terms of improving the privacy aspect, Jin et al. (2022) found that users prefer the privacy protection concept of a third-party application with privacy diagnostics or data collection live monitor that can show users the status of data collection, evaluate the privacy settings on existing devices, and give instructions on how to improve. Previous studies have also found that incorporating a mechanism to summarize and communicate the privacy policy of voice interaction skills can help prevent malicious activity by developers or third parties and misuse of users' trust (Liao et al. 2020).

Integration of the anthropomorphism features in the AI agents has also been applied in previous studies to improve users' trust and acceptability of AI systems. Previous studies have found that the human-likeness of virtual agents can help increase the trust resilience of the AI systems (Lockey et al. 2021). In other words, compared with machine-like AI agents, the decline in trust for anthropomorphic agents is less pronounced when the performance of the AI systems doesn't meet the needs of their users. Adding personalities to the AI agents or matching the personalities of the AI agents with the users (e.g., introverted agents for introverts) can also be a potential way of building an emotional connection between the users and AI agents (Rheu et al. 2021; Seaborn et al. 2021). However, it should not be neglected that implementing human-like features on AI systems can also potentially cause a negative effect on the users' emotional trust (Glikson and Woolley 2020). Human-like AI can lead to an experience of eeriness, discomfort, and anxiety emotions for some users during the initial interactions. It is therefore necessary to consider both positive and negative effects when implementing trust-promoting features.

Actions can also be taken from the human/users' perspective to promote their trust in AI-powered smart homes. Education and training about the capabilities and limitations of AI systems can help users set more reasonable expectations for them. Lowering unrealistically high initial trust in AI can help improve the actual interaction experience and prevent trust erosion, which can be beneficial for building long-term trust (Glikson and Woolley 2020). Consistent user engagement in human-AI interactions, such as continuously providing feedback can also be beneficial to promoting trust. Not only can the engagement help users foster a sense of collaboration and perceive the AI systems as more intelligent and reliable, but it can also develop a positive reinforcement that helps make the AI systems more personalized to the users' preferences and needs, and thus improving the interaction experiences and trust.

Q4: What are the challenges and opportunities for trust in AI-powered smart homes?

Although previous studies have identified the influential factors and features for trust in AI-powered smart homes, several other challenges and opportunities still exist for expanding the research and development of trust. One major challenge in building users' trust is the fact that current AI-powered smart homes are not "smart" enough to meet users' expectations and needs. The technical limitations of the current AI systems and IoT reduce the reliability and capability of smart homes. Users generally have a low level of satisfaction with the current smart homes and the ability of AI agents (Cannizzaro et al. 2020; Rheu et al. 2021). The unexpected errors and malfunctions, the uncertainty of consequences, and the uncertain long-term effects of AI systems can reduce user confidence in adopting AI systems in their living environments. Improving the usability and reliability of AI-powered smart homes is a challenge that needs to be overcome and is critical in the development of the continuous trust (Siau and Wang 2018).

The standards and guidelines for building trustworthy AI in intelligent environments have not been considered in previous studies. The lack of standards for building trustworthy AI can make it difficult for users to know what to expect and for researchers and developers to design AI systems that inspire trust. In addition, there is a lack of legal regulation of AI systems, such as

accountability for the loss caused by the malfunction of the AI system, the leakage of personal data, or the conflict of interests. Previous studies have found that some skills (applications) on voice assistant marketplaces that violate content policies are still being verified for use, raising ethical concerns among users of virtual assistants (Seymour et al. 2022).

While some efforts have been made to investigate the trajectories of trust in AI (Glikson and Woolley 2020), limited studies have explored the pathway of building long-term continuous trust in AI-powered smart homes. Both the initial trust formation and the continuous trust development need to be focused on as users' trust is not a static condition but a dynamic subjective perception (Siau and Wang 2018). In addition, the interrelationship between different influencing factors should be considered in developing continuous trust. For example, users' sensitivity to privacy may be higher when it comes to security management with home security cameras compared to the tasks of remote controlling the lights. In addition, investigating the continuous trust in AI can also help reveal the long-term impact of human-AI trust on the energy efficiency of AI-powered smart homes. The implementation of cutting-edge advanced technologies can also bring about new opportunities. The integration of wearable sensors can enhance the context awareness of the AI systems with additional user data, including health and wellness metrics, which can be used to create more personalized and responsive intelligent environments. The application of AR (Augmented Reality) and VR (Virtual Reality) technologies in smart home simulations can be beneficial for educating users and investigating human-AI interactions in smart homes, helping researchers explore the possibilities of building trustworthy AI-powered smart homes.

CONCLUSION

This study explored the driving factors that could lead to a successful implementation of AI-powered smart home ecosystems through the lens of trust in human-AI interaction. Based on a systematic literature review, the presented study provides a definition of human-AI trust in intelligent environments with its various terms, different dimensions (cognitive and emotional trust), goals, and benefits. The impact of different influencing factors that can affect trust was described from the dimensions of cognitive and emotional trust, and from the perspectives of AI systems and human users separately. Different methods and strategies for promoting trust and the challenges and opportunities that need to be considered were then presented. Future studies can refer to the findings in this study in building trustworthy AI-powered smart home ecosystems.

ACKNOWLEDGMENT

This material is based upon work supported in part through a 4VA grant. Any findings and conclusions are those of the authors and do not necessarily reflect the views of 4VA.

REFERENCES

Cannizzaro, S., Procter, R., Ma, S., and Maple, C. (2020). "Trust in the smart home: Findings from a nationally representative survey in the UK." *Plos one*, 15(5), e0231615.

Cuadra, A., Li, S., Lee, H., Cho, J., and Ju, W. (2021). "My bad! repairing intelligent voice assistant errors improves interaction." *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1), 1-24.

Emaminejad, N., Maria North, A., and Akhavian, R. (2021). "Trust in AI and Implications for AEC Research: A Literature Analysis." *Computing in Civil Engineering 2021*, 295-303.

Glikson, E., and Woolley, A. W. (2020). "Human trust in artificial intelligence: Review of empirical research." *Academy of Management Annals*, 14(2), 627-660.

He, T., and Jazizadeh, F. "Nudging Occupants for Energy-Saving through Voice-Based Proactive Virtual Assistants." *Proc., Construction Research Congress 2022*, 402-411.

He, T., Jazizadeh, F., and Arpan, L. (2022). "AI-powered virtual assistants nudging occupants for energy saving: proactive smart speakers for HVAC control." *Building Research & Information*, 50(4), 394-409.

Hubert, M., Blut, M., Brock, C., Zhang, R. W., Koch, V., and Riedl, R. (2019). "The influence of acceptance and adoption drivers on smart home usage." *European journal of marketing*, 53(6), 1073-1098.

Jacovi, A., Marasović, A., Miller, T., and Goldberg, Y. "Formalizing trust in artificial intelligence: Prerequisites, causes and goals of human trust in AI." *Proc., Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*, 624-635.

Jin, H., Guo, B., Roychoudhury, R., Yao, Y., Kumar, S., Agarwal, Y., and Hong, J. I. "Exploring the needs of users for supporting privacy-protective behaviors in smart homes." *Proc., Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 1-19.

Liao, S., Wilson, C., Cheng, L., Hu, H., and Deng, H. "Measuring the effectiveness of privacy policies for voice assistant applications." *Proc., Annual Computer Security Applications Conference*, 856-869.

Liu, Y., Gan, Y., Song, Y., and Liu, J. (2021). "What influences the perceived trust of a voice-enabled smart home system: An empirical study." *Sensors*, 21(6), 2037.

Lockey, S., Gillespie, N., Holm, D., and Someh, I. A. (2021). "A review of trust in artificial intelligence: Challenges, vulnerabilities and future directions."

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., and Brennan, S. E. (2021). "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews." *International journal of surgery*, 88, 105906.

Rheu, M., Shin, J. Y., Peng, W., and Huh-Yoo, J. (2021). "Systematic review: Trust-building factors and implications for conversational agent design." *International Journal of Human–Computer Interaction*, 37(1), 81-96.

Ryan, M. (2020). "In AI we trust: ethics, artificial intelligence, and reliability." *Science and Engineering Ethics*, 26(5), 2749-2767.

Seaborn, K., Miyake, N. P., Pennefather, P., and Otake-Matsuura, M. (2021). "Voice in human–agent interaction: a survey." *ACM Computing Surveys (CSUR)*, 54(4), 1-43.

Seymour, W., Zhan, X., Cote, M., and Such, J. (2022). "A Systematic Review of Ethical Concerns with Voice Assistants." *arXiv preprint arXiv:2211.04193*.

Sharif, K., and Tenbergen, B. (2020). "Smart home voice assistants: a literature survey of user privacy and security vulnerabilities." *Complex Systems Informatics and Modeling Quarterly*(24), 15-30.

Siau, K., and Wang, W. (2018). "Building trust in artificial intelligence, machine learning, and robotics." *Cutter business technology journal*, 31(2), 47-53.

Ueno, T., Sawa, Y., Kim, Y., Urakami, J., Oura, H., and Seaborn, K. "Trust in human-ai interaction: Scoping out models, measures, and methods." *Proc., CHI Conference on Human Factors in Computing Systems Extended Abstracts*, 1-7.

Wang, T., Zhang, K., Chen, W., Dou, W., Zhu, J., Wei, J., and Huang, T. "Understanding device integration bugs in smart home system." *Proc., Proceedings of the 31st ACM SIGSOFT International Symposium on Software Testing and Analysis*, 429-441.