Human Factors in Design, Engineering, and Computing, Vol. 159, 2024, 33–43
https://doi.org/10.54941/ahfe1005565

Formalizing Trust in Artificial Intelligence for Built Environment Decision-Making

Amir H. Behzadan

Department of Civil, Environmental, and Architectural Engineering & Institute of Behavioral Science, University of Colorado, Boulder, CO 80309, USA

ABSTRACT

While artificial intelligence (AI) has transformed the planning, design, construction, and operation of physical infrastructure and spaces, it has also raised concerns about algorithmic bias, data privacy, and ethical use in built environment decision-making. Addressing these issues is crucial for designing, developing, and deploying trustworthy AI systems that promote human safety, infrastructure security, and resource allocation. This paper reviews trust issues in AI through the lens of several built environment decision scenarios, e.g., weather prediction, disaster mitigation and response, urban sensing, and bridge health monitoring. It then outlines a framework to formalize trust, aiding researchers, policymakers, and practitioners in designing AI systems that serve societal interests.

Keywords: Artificial intelligence, Trust, Built environment, Decision-making

INTRODUCTION

Recent advancements in artificial intelligence (AI) have sparked new debates on issues related to skepticism and mistrust in technology and automation. If we envision a future where humans and intelligent agents will be working alongside each other to solve complex problems at the interface of the society and the built environment, then the AI systems that enable those agents must be designed, developed, and deployed with careful consideration for certain principles to ensure that they are respectful to humans and serve the collective good, thus fostering genuine trust in their capabilities.

To date, the literature on trust in AI systems remains sparse in the built environment and its various subfields. For example, an analysis of 490 articles published in 1985–2021 revealed that trust in AI systems in the context of architecture, engineering, and construction (AEC) applications was not studied before (Emaminejad et al., 2021). In another study, involving the review of 102 articles, it was reported that the literature on trust in AI was fragmented and primarily focused on examining trust formation in experimental settings (Lockey et al., 2021). In this paper, we utilize real-world decision-making cases from the built environment domain to discuss how trust in AI may influence the quality and timeliness of resulting decision outcomes. Ultimately, we present a conceptual framework for convergent and informed discussions on this topic.

BUILDING TRUST IN AI SYSTEMS

To better understand the dynamics of acceptance and adoption of a technology intended to support and improve human decision-making, it is critical to first differentiate between trust and trustworthiness. Trust in technology is the subjective willingness of a person or group to rely on that technology (Nickel, 2012). This behavior can be driven by factors such as past experiences, reputation, and perceived reliability. This notion of trust, however, excludes critical considerations, such as the potential biases or limitations of the technology, intentions of its developers, and broader societal implications of its use. On the other hand, trustworthiness is concerned with the inherent qualities of the technology itself, encompassing aspects such as transparency, reliability, integrity, security, and ethical design (Department of Commerce, 2018). While trust may be gained or lost based on perceptions and experiences, trustworthiness requires a deeper analysis of technology's integrity and its alignment with ethical principles.

The same interplay must be considered in discussions surrounding the use of AI models for decision-making. Let's consider the application of AI in weather prediction. Trust in AI-driven outcomes, in this case, involves individuals relying on the accuracy and reliability of the model to make informed decisions, such as planning outdoor activities or preparing for severe weather events. For instance, people may trust a weather forecasting app to provide timely and accurate predictions based on positive past experiences or recommendations from meteorological experts. However, this trust can be easily challenged if (even a few) predictions deviate from actual weather conditions or fail to account for unforeseen phenomena (Burgeno and Joslyn, 2020). On the other hand, trustworthiness in AI-driven outcomes involves considerations for system's overall reliability, robustness, transparency, and ethical operation. A trustworthy weather prediction system would be built on AI algorithms trained on vast, representative, and high-quality historical data, thus enabling accurate and timely forecasts across various temporal and spatial scales. Additionally, transparent communication about system's methodologies, limitations, and uncertainty estimates enhances its trustworthiness by empowering users to understand and interpret the predictions effectively. Thus, while building trust is undoubtedly essential for encouraging acceptance and adoption of AI systems, it must be supported by a demonstrable trustworthiness of such systems to ensure sustained and responsible use. Failing to prioritize trustworthiness can lead to misplaced or eroded trust over time, particularly when unforeseen decision-making problems or ethical dilemmas arise.

The Circle of Trust: Where User Perception Meets System Capacity

In this paper, to describe trust in an AI system, we adopt the definition by the National Institute of Standards and Technology (NIST), an agency tasked with promoting U.S. innovation and industrial competitiveness. According to NIST, trust in an AI system is shaped by two interconnected elements: inherent user's characteristics, i.e., user trust profile (UTP), and perceived system trustworthiness (PST) (Stanton and Jensen, 2021). UTP captures an individual's predispositions, personal/cultural background, beliefs/values, prior experiences, and potential biases that influence their willingness to trust AI. For instance, high propensity to trust in technology may lead a person to rely more on AI decisions without questioning their validity, while those with a skeptical outlook may exhibit more caution and scrutiny. PST, on the other hand, refers to the user's assessment of the performance of the AI system. For example, a system that consistently delivers precise results, provides clear explanations for its decisions, and prioritizes data privacy and user safety is likely to be perceived as more trustworthy.

Figure 1 is a graphical representation of the interplay between UTP and PST in the circle of trust. Here, blind trust characterizes a scenario where user's reliance (if any) on AI outputs is achieved without proper consideration for system flaws or limitations. Conversely, illusory trust denotes a perceived sense of reliance in the AI system, despite its actual performance falling short of trustworthiness expectations. Reluctant trust reflects a cautious acceptance of AI recommendations or outputs, tempered by skepticism or reservations about their accuracy or efficacy. Lastly, assured trust pertains to a state of complete confidence in the capabilities of an AI system that is perceived highly trustworthy, where users rely extensively on its outputs without hesitation. While the initial trust may fall within any of these four areas, further modifications in the underpinning components of perceived trustworthiness or user's internal traits can lead to adjusted levels of trust.

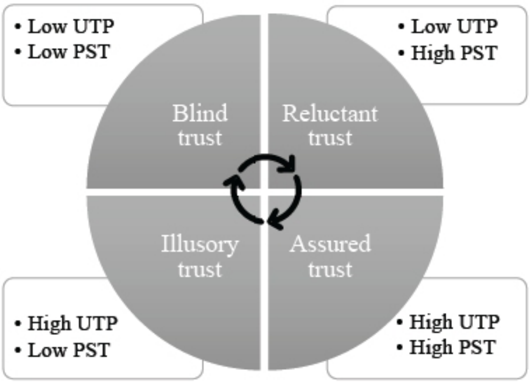

Figure 1: The circle of trust: Interplay between user trust profile (UTP) and perceived system trustworthiness (PST). The circular arrow implies the process of trust calibration.

In the built environment, where decisions are consequential to people and infrastructure, navigating the complexities of trust in AI-assisted decision-making is even more significant due to its implications on human safety, infrastructure security and integrity, and resource allocation. Within this context, blindly trusting flawed AI recommendations can lead to suboptimal outcomes or catastrophic consequences. Illusory trust may result in stakeholders overlooking potential risks or inaccuracies in AI analyses, potentially risking the integrity and safety of built structures or urban environments. Reluctant trust highlights the need for careful validation and verification of AI outputs, particularly in scenarios where human lives or large-scale investments are at stake. Finally, achieving ensured trust in AI systems requires not only robust performance and reliability but also transparent communication and accountability mechanisms to instill confidence in users.

How Does the Performance of an AI System Contribute to Trust?

According to NIST, from a technical perspective, the performance of an AI system must adhere to at least nine metrics if it is to be trusted. These include accuracy, reliability, resiliency, objectivity, explainability, accountability, security, safety, and privacy. When aggregated, these nine metrics constitute what is referred to as the perceived technical trustworthiness (PTT) (Stanton and Jensen, 2021). Table 1 lists PTT metrics and their common definitions in various domains and the literature.

Table 1. Definition of system characteristics that drive perceived technical trustworthiness.

Metric	Definition
Accuracy	How often the model correctly predicts the expected outcomes.
Reliability	How consistent and stable the model can perform in multiple runs.
Resiliency	How well the model can perform in a changing, deteriorating, or partially invisible environment.
Objectivity	How faithful (bias-free) model outcomes are to real-world facts.
Explainability	How well the model can explain why certain predictions are made.
Accountability	How closely model design, development, and deployment comply with laws and standards to ensure the proper functioning.
Security	How effective model authentication, data encryption, and access controls are in protecting data confidentiality, maintaining data integrity, and ensuring reasonable data availability.
Safety	How well the model is equipped to prevent accidents, misuse, or other harmful consequences.
Privacy	How well the model adheres to the guidelines around acquisition, analysis, and use of personal data lawfully, fairly, and transparently.

Depending on the problem context, some metrics may outweigh others. For instance, when using an AI system for stability design of slopes of rock-fill dams or embankments, reliability is a major concern due to the risks associated with slope failure, e.g., landslides or slope collapse (U.S. Army Corps of Engineers, 2003). Thus, an assistive AI system must be highly accurate in predicting potential failure mechanisms and assessing the factors contributing to slope instability. Likewise, reliability is essential to ensure that AI predictions align closely with real-world observations and can be relied upon for making critical engineering decisions, such as slope reinforcement or land use planning. Moreover, resiliency is crucial in this design context to account for uncertainties and variations in environmental conditions, e.g., changes in rainfall patterns or geological factors. The AI system should be resilient to these fluctuations by continuously updating its predictions and recommendations based on new data or changing circumstances to maintain the slope stability and safety over time. Explainability is also significant in this context, as designers and engineers must understand the driving factors and assumptions of the AI recommendations. An AI system that provides clear explanations for its analyses and decisions enables users to validate the results, identify potential limitations or sources of bias (e.g., favoring certain geological formations, historical landslide occurrences, slope stabilization measures in certain regions or terrain types), and make informed adjustments to slope stabilization strategies. Lastly, accountability ensures that the responsible parties are held liable for the design decisions made based on AI outcomes, leading to responsible and risk-aware engineering designs (Novelli et al., 2023).

Changing the problem context or parameters may alter the relative importance of the above metrics. For instance, in the immediate aftermath of a major flood event, an AI system used for coordinating emergency response must demonstrate high reliability in providing accurate information about the extent of damage, resource availability, and evacuation routes (Hillin et al., 2024). Similarly, resiliency becomes critical to ensure that the AI system can continue to function effectively despite disruptions in communication networks or infrastructure damage. In contrast, during the recovery and rebuilding phase, factors such as explainability and accountability may take on greater significance. For example, stakeholders involved in reconstruction efforts may require transparent explanations from the AI system with respect to the prioritization of resources, allocation of funds, and long-term planning decisions (Bari et al., 2023). Furthermore, accountability mechanisms are essential to ensure that AI decisions align with ethical and moral considerations, while serving the best interests of the affected communities. Therefore, the assigned weight to each performance characteristic may shift throughout the disaster management cycle, reflecting the evolving priorities and challenges faced in the real-world decision-making context.

How Does User Experience With an AI System Contribute to Trust?

The PTT metrics described in the previous section are only necessary but not sufficient for trust. Ultimately, the user's own experience with the AI system contributes to the formation and calibration of trust. Usability, a fundamental aspect of user experience (UX), comprises three key components, namely efficiency, effectiveness, and satisfaction (Stanton and Jensen, 2021). Efficiency is a measure of task completion time and overall completion time, effectiveness quantifies the number of errors encountered or the quality of task output, and user satisfaction deals with factors such as the level of frustration, engagement, or enjoyment experienced by users when interacting with the system (Frøkjær et al., 2000). If we assume that PTT and UX follow independent probability distributions, then their aggregate contribution to PST can be described as PTT$\times$UX.

Additionally, the literature suggests that trust and credibility may depend on surface features of the system interface even if they are not linked to the true capabilities of the system (Briggs et al., 1998). Therefore, for an intelligent system with anthropomorphic characteristics, relatability may also play a role in user satisfaction. An experiment with 111 participants revealed that presenting a virtual driver with human characteristics enhances human trust in the driving simulator (Verberne et al., 2015). There is, however, a delicate balance between the humanness of the AI system and the user's emotional or psychological response as a precursor to trust. Specifically, research warns against the uncanny valley effect when the technology's resemblance to a human is close but not quite perfect, potentially leading to feelings of unease or discomfort in users and reduced user satisfaction (Troshani et al., 2021).

Previous work has examined methods to obtain a single UX score by combining individual UX components (Table 2). These methods, with some modification, can be utilized to assess the usability and user experience of an AI system, identify areas for improvement, and prioritize enhancements to increase user satisfaction and efficiency. The system usability scale (SUS), for example, allows users to rate their agreement with a series of statements regarding the system's usability. User responses are then converted into a numerical score which provides a quantitative measure of overall usability, allowing for comparison between different systems or iterations of the same system (Vlachogianni and Tselios, 2022).

Table 2. Methods to assess user experience when interacting with an intelligent system.

Method	Description
System usability scale (SUS)	Users rate their agreement with a series of statements regarding the system's usability.
User experience questionnaire (UEQ)	Evaluates six dimensions of attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty.
Usefulness, satisfaction, and ease of use (USE)	Four items that measures perceived usefulness, satisfaction, and ease of use of a system.
Post-study system usability questionnaire (PSSUQ)	Evaluates the usability of software systems by assessing factors such as system usefulness, information quality, and interface quality, among others.
NASA task load index (TLX)	Evaluates the mental workload associated with completing a task by measuring six dimensions of workload, including mental demand, physical demand, temporal demand, performance, effort, and frustration.
User interface satisfaction (QUIS)	Assesses user satisfaction with specific aspects of a system's user interface, including screen design, terminology, and system capabilities.

ETHICAL CONSIDERATIONS IN THE DESIGN OF AI SYSTEMS

Beyond technical aspects and user experience, a trustworthy AI system should be also legal (i.e., adhere to all applicable laws and regulations) and ethical (i.e., align with ethical principles and moral standards). Together, these aspects must be upheld throughout the life of the AI system and apply to developers, deployers, users, and the broader society. In the built environment domain, ethical issues impacting both the workplace and society have drawn some recent attention (Adnan et al., 2012; Li et al., 2022). However, these discussions have largely overlooked the role of AI, given that the integration of human-in-the-loop intelligent systems in many of the subfields (e.g., the construction industry) is still in its early stages. In a review of 314 articles published in 2017–2022, researchers listed trust as a component of ethics when identifying nine categories of ethical issues of AI and robotics in the AEC domains, namely job loss, data privacy, data security, data transparency, decision-making conflict, acceptance and trust, reliability and safety, fear of surveillance, and liability (Liang et al., 2024).

To explain the provisions of ethical AI, let's consider the use of AI-powered urban sensing tools (e.g., facial recognition, license plate readers, biometric scanners, gunshot detectors, drones, body-worn cameras) in many smart city initiatives (Socha and Kogut, 2020; Ma, 2021). While these systems are initially deployed to deter urban crimes and improve public safety, they may also raise concerns about privacy infringement, potential misuse of sensitive information, and social inequality (Helbing et al., 2021). Within this context, the incorporation of ethical AI principles mandate that for these intelligent systems to be fully accepted and trusted by all stakeholders (i.e., people, businesses, law enforcement), they must have proper safeguards such as data anonymization and encryption, limits on data retention periods, strict access controls to prevent unauthorized use and disclosure of data, and regular audits and reviews to assess the system's compliance with privacy laws and regulations. As evidenced by this example, the application of ethical principles to the rapidly evolving AI-society interface can be complex and subject to interpretation. It is, therefore, imperative that potential implementation conflicts be identified and addressed through effective engagement with all stakeholders, and continuous evaluation and communication of tradeoffs to guarantee ethical and responsible AI implementation.

TRUST CALIBRATION

Trust is formed over time and continuously reevaluated by the trustor. The extent to which a trustor (i.e., human) trusts a trustee (i.e., technology) is influenced by a host of factors, including task complexity, domain expertise, and ethical considerations. As noted earlier, trust is defined within the specific context of a decision-making problem. As such, over-trust or under-trust that in one context may lead to a negative decision outcome, could result in a positive decision outcome under slightly (or completely) different circumstances.

In practice, for trustors with low trust resolution, a large improvement in system capabilities may not yield a large increase in user trust. Moreover, while some users calibrate their trust in response to immediate changes in the context or system's capabilities, others may wait until after observing long-term, sustained success or failure. This sensitivity of trust to temporal changes is referred to as temporal specificity (Benda et al., 2022). In addition, trust calibration can vary based on the granularity of performance observations. Trust may be calibrated in response to the success or failure of a single system component or remain intact until a system-level failure or success is achieved. While the former represents a high functional specificity, the latter is an example of low functional specificity (Wintersberger, 2023). For example, in an AI-enabled bridge health monitoring system equipped with sensors to detect structural stress and deformation, occasional data drifts in a sensor responsible for deck vibration measurement may lead an engineer with high functional specificity to adjust their trust in the system. However, an engineer with low functional specificity may reconsider their trust only in the event of a total system failure. Together, high temporal and functional specificity increase the likelihood that the level of trust will match the capabilities of a particular component of the AI system at a particular time.

CONCEPTUAL FRAMEWORK FOR FORMALIZING TRUST

In adopting an AI system for built environment decision-making, stakeholders and beneficiaries may have different expectations about system capabilities and limitations, which may influence their level of acceptance and trust (Wang and Zhou, 2022). Thus, system performance should be compared across different user groups that are on the receiving end of AI-assisted decisions. Models trained on historical data may excel with one group but can perpetuate social inequalities in other groups (Timmons et al., 2023). Figure 2 outlines our eight-step conceptual framework to formalize, calibrate, and maintain trust in AI. This framework can guide researchers, policymakers, and practitioners in designing trustworthy AI systems for built environment contexts.

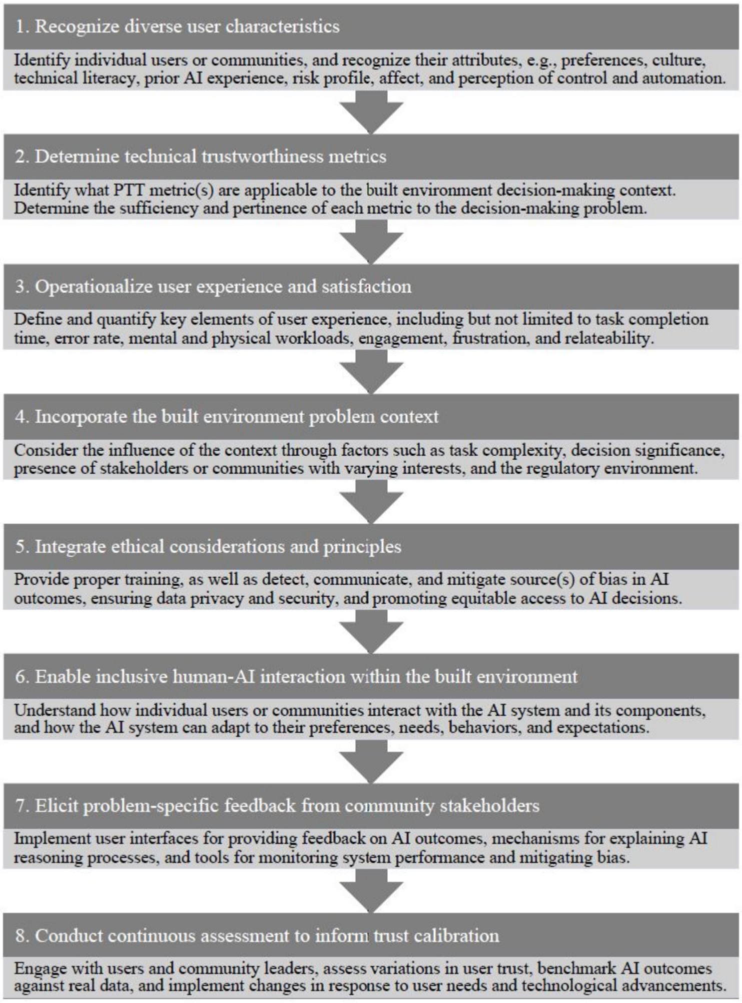

Figure 2: Proposed framework for formalizing trust in built environment decision-making.

CONCLUSION

The rapid adoption of AI in our daily lives has outpaced a thorough debate on the crucial issue of trust in AI outcomes and decisions. There is still a need for mechanisms that can objectively establish the trustworthiness of AI systems to increase user adoption and acceptance of technology. In built environment applications, understanding trust in AI systems is important to ensure the reliability, safety, and acceptance of such systems that have major societal implications. By investigating trust dynamics, designers, architects, and engineers can identify factors that influence trust, implement strategies to mitigate distrust, and ultimately promote confidence and maximize the potential benefits of AI in built environment decision-making. Defining and realizing trustworthy AI is complex, and diverse perspectives exist on what constitutes trustworthy AI, with evolving technical and non-technical aspects, and differing ethical and regulatory priorities (Alzubaidi et al., 2023). Through the lens of several built environment problem contexts (weather prediction, embankment slope stability, disaster mitigation and response, urban sensing, bridge health monitoring), this paper discussed key aspects of designing trustworthy AI, and examined the interplay between user trust and system trustworthiness in such systems. Additionally, the discussion on trust calibration highlighted the influence of trust resolution as well as temporal and functional specificity in the formation and adjustment of trust in AI systems. Lastly, a conceptual framework was proposed to formalize trust in AI and guide researchers, policymakers, and practitioners in the design and deployment of trustworthy AI systems for built environment applications.

While the study of human trust in AI systems in multi-person teams was outside the scope of this paper, in some decision-making contexts, human interaction with AI occurs within a multi-person group that shares role and responsibility for managing the AI system, and exhibit interdependencies in workflows, goals, and outcomes (Ulfert et al., 2024). With the goal of creating a convergent understanding and a unified model of trust for engineering applications, the focus of this paper, however, was on the fundamental relationship between the characteristics and experiences of a single human trustor, and the trustworthiness of an AI system.

REFERENCES

Adnan, H., Hashim, N., Mohd, N. and Ahmad, N., 2012. Ethical issues in the construction industry: Contractor's perspective. *Procedia-social and behavioral sciences*, *35*, pp. 719–727.

Alzubaidi, L., Al-Sabaawi, A., Bai, J., Dukhan, A., Alkenani, A. H., Al-Asadi, A., Alwzwazy, H. A., Manoufali, M., Fadhel, M. A., Albahri, A. S. and Moreira, C., 2023. Towards Risk-Free Trustworthy Artificial Intelligence: Significance and Requirements. *International Journal of Intelligent Systems*, *2023*(1), p. 4459198.

Bari, L. F., Ahmed, I., Ahamed, R., Zihan, T. A., Sharmin, S., Pranto, A. H. and Islam, M. R., 2023. Potential use of artificial intelligence (AI) in disaster risk and emergency health management: a critical appraisal on environmental health. *Environmental health insights*, *17*, p. 11786302231217808.

Benda, N. C., Novak, L. L., Reale, C. and Ancker, J. S., 2022. Trust in AI: why we should be designing for appropriate reliance. *Journal of the American Medical Informatics Association*, *29*(1), pp. 207–212.

Briggs, P., Burford, B. and Dracup, C., 1998. Modelling self-confidence in users of a computer-based system showing unrepresentative design. *International Journal of Human-Computer Studies*, *49*(5), pp. 717–742.

Burgeno, J. N. and Joslyn, S. L., 2020. The impact of weather forecast inconsistency on user trust. Wea. Climate Soc., 12, 679–694.

Department of Commerce, 2018. *NIST Special Publication 800-37: Risk management framework for information systems and organizations: A system life cycle approach for security and privacy*, Washington, D. C.

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*, pp. 295–303.

Frøkjær, E., Hertzum, M. and Hornbæk, K., 2000, April. Measuring usability: are effectiveness, efficiency, and satisfaction really correlated?. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 345–352).

Helbing, D., Fanitabasi, F., Giannotti, F., Hänggli, R., Hausladen, C. I., van den Hoven, J., Mahajan, S., Pedreschi, D. and Pournaras, E., 2021. Ethics of smart cities: Towards value-sensitive design and co-evolving city life. *Sustainability*, *13*(20), p. 11162.

Hillin, J., Alizadeh, B., Li, D., Thompson, C. M., Meyer, M. A., Zhang, Z. and Behzadan, A. H., 2024. Designing user-centered decision support systems for climate disasters: what information do communities and rescue responders need during floods?. *Journal of Emergency Management (Weston, Mass.)*, *22*(7), pp. 71–85.

Li, Y., Soomro, M. A., Khan, A. N., Han, Y. and Xue, R., 2022. Impact of ethical leadership on employee turnover intentions in the construction industry. *Journal of Construction Engineering and Management*, *148*(7), p. 04022054.

Liang, C. J., Le, T. H., Ham, Y., Mantha, B. R., Cheng, M. H. and Lin, J. J., 2024. Ethics of artificial intelligence and robotics in the architecture, engineering, and construction industry. *Automation in Construction*, *162*, p. 105369.

Lockey, S., Gillespie, N., Holm, D. and Asadi Someh, I., 2021. A review of trust in artificial intelligence: Challenges, vulnerabilities and future directions. In: *54th Hawaii International Conference on System Sciences*, International Network for Social Network Analysis (INSNA), Maui.

Ma, C., 2021. Smart city and cyber-security; technologies used, leading challenges and future recommendations. *Energy Reports*, *7*, pp. 7999–8012.

Nickel, P. J., 2012. Trust in technological systems. In *Norms in technology* (pp. 223–237). Dordrecht: Springer Netherlands.

Novelli, C., Taddeo, M. and Floridi, L., 2023. Accountability in artificial intelligence: what it is and how it works. *AI & Society*, pp. 1–12.

Socha, R. and Kogut, B., 2020. Urban video surveillance as a tool to improve security in public spaces. *Sustainability*, *12*(15), p. 6210.

Stanton, B. and Jensen, T., 2021. *Trust and artificial intelligence*. National Institute of Standards and Technology (NIST), Gaithersburg.

Timmons, A. C., Duong, J. B., Simo Fiallo, N., Lee, T., Vo, H. P. Q., Ahle, M. W., Comer, J. S., Brewer, L. C., Frazier, S. L. and Chaspari, T., 2023. A call to action on assessing and mitigating bias in artificial intelligence applications for mental health. *Perspectives on Psychological Science*, *18*(5), pp. 1062–1096.

Troshani, I., Rao Hill, S., Sherman, C. and Arthur, D., 2021. Do we trust in AI? Role of anthropomorphism and intelligence. *Journal of Computer Information Systems*, *61*(5), pp. 481–491.

Ulfert, A. S., Georganta, E., Centeio Jorge, C., Mehrotra, S. and Tielman, M., 2024. Shaping a multidisciplinary understanding of team trust in human-AI teams: a theoretical framework. *European Journal of Work and Organizational Psychology*, *33*(2), pp. 158–171.

U.S. Army Corps of Engineers, 2003. *Slope stability*. Department of the Army, Washington, D. C.

Verberne, F. M., Ham, J. and Midden, C. J., 2015. Trusting a virtual driver that looks, acts, and thinks like you. *Human factors*, *57*(5), pp. 895–909.

Vlachogianni, P. and Tselios, N., 2022. Perceived usability evaluation of educational technology using the System Usability Scale (SUS): A systematic review. *Journal of Research on Technology in Education*, *54*(3), pp. 392–409.

Wang, X. and Zhou, R., 2023. Impacts of user expectation and disconfirmation on satisfaction and behavior intention: the moderating effect of expectation levels. *International Journal of Human–Computer Interaction*, *39*(15), pp. 3127–3140.

Wintersberger, P., 2023. Team at Your Service: Investigating Functional Specificity for Trust Calibration in Automated Driving with Conversational Agents. *International Journal of Human–Computer Interaction*, *39*(16), pp. 3254–3267.