EXPLORING THE ETHICAL ROLE OF ARTIFICIAL INTELLIGENCE IN TOURISM: A FOCUS ON USER PREFERENCES

Dajana Vukojevic Ateljević[1,*]
[1]University of East Sarajevo, Bosnia and Herzegovina
[*]Correspondence: dajanavukojevic@gmail.com

https://doi.org/10.63356/redete.2025.023

Abstract

In the era of accelerating digital transformation, artificial intelligence (AI) has become an indispensable tool in strategic decision-making, supporting organizations in achieving sustainability while upholding ethical standards. Within the tourism sector, this transformation raises complex challenges related to user security, privacy, and trust, as well as broader principles of accountability, transparency, and fairness. These dimensions form the fundamental pillars of long-term tourism sustainability and represent key priorities for all stakeholders. This study provides clear definitions of essential ethical principles—security, data protection, and trust—emphasizing both their interconnections and their distinct characteristics. The discussion is broadened to address the irreversible nature of this digital trend. Special attention is given to how AI technologies influence these principles, alongside the critical need for education, training, regulatory, and legal frameworks to ensure their protection. Furthermore, the paper argues that responsible and explainable AI must be inclusive and socially responsible, not only in tourism but also in other sectors, such as banking, where transparency and fairness are of crucial importance. Finally, the paper outlines directions for future research, with particular emphasis on ethical challenges such as bias, inaccuracy, and the demand for algorithmic explainability—underscoring the need to create fair and sustainable digital ecosystems.

Keywords: Artificial Intelligence, Tourism, Ethics, Security, Privacy, Trust, Sustainability, Transparency, Accountability, Fairness.

1. INTRODUCTION

As artificial intelligence (AI) becomes an integral part of modern society, it is profoundly reshaping business models, personal interactions, and decision-making processes (Bansal, 2025). Alongside these advancements, however, AI introduces significant ethical challenges, particularly in relation to security, privacy, algorithmic transparency, and bias (Ali et al., 2025). Given its widespread adoption across organizational activities, the ethical and responsible implementation of AI is no longer optional but essential (Papagiannidis et al., 2025). The growing reliance on AI therefore imposes a responsibility to uphold fundamental moral and ethical principles throughout its development and application. In this context, AI ethics play a critical role in strengthening security and trust, safeguarding privacy, mitigating bias, promoting transparency, and ensuring inclusivity, particularly for vulnerable groups such as individuals with disabilities. Security and privacy, identified as key user priorities (Korać & Simić, 2017, 2019), are central to the sustainability of AI systems, requiring robust data protection measures, privacy-preserving algorithms, and anonymization techniques (Holzinger et al., 2021). As organizations increasingly rely on AI for strategic decision-making that directly affects user needs, it is essential to address the complex ethical implications associated with these technologies.

The misuse of AI may lead to far-reaching consequences for stakeholders, society, and organizational reputation. Addressing these risks requires substantial investment in education, awareness, and interdisciplinary collaboration. Effective application of AI across domains depends on the involvement of diverse stakeholders, including developers, experts, policymakers, and end users. Without such collaboration, AI systems risk embedding errors and biases, particularly those arising from insufficient integration of social and ethical perspectives (Shestakova, 2021).

AI sustainability emphasizes the responsible design and management of AI systems to minimize ethical risks while reducing environmental and societal impacts (Leimanis & Palkova, 2021). The convergence of responsible AI practices and ethical frameworks represents a paradigm shift with the potential to drive sustainable development and enhance organizational performance. However, the understanding and definition of key AI ethics concepts remain inconsistent and often ambiguous, as illustrated by terms such as ethical risk and cybersecurity (Schatz et al., 2017; Douglas et al., 2024). This conceptual ambiguity further highlights the need for systematic and structured analysis.

Within the tourism sector, the application of AI raises additional ethical considerations. Existing studies emphasize the importance of privacy, transparency, and responsible data use (Gretzel et al., 2015), as well as the ethical implications of automation for user experience (Tussyadiah, 2020). AI-driven personalization and recommendation systems, while enhancing service quality, also introduce challenges related to fairness and bias (Li et al., 2019). Furthermore, research highlights the importance of responsible AI use in destination management and the role of emerging technologies such as chatbots and recommendation systems in shaping

tourism experiences (Buhalis & Sinarta, 2019; Mariani & Borghi, 2021; Guttentag, 2019).

Motivated by these challenges, this study explores AI-driven strategic decision-making in tourism through the lens of ethical principles, with a particular focus on user preferences. Given the increasing complexity of digital environments and the evolving nature of technological threats, addressing ethical and sustainability challenges in AI has become a critical priority for decision-makers in the tourism industry.

This paper contributes to the literature by clarifying key AI ethics concepts relevant to strategic decision-making, emphasizing security, privacy, and trust as core principles reflecting user preferences, examining the relationship between these principles and user-centered AI applications in tourism, and identifying current trends as well as future research directions in this domain.

The remainder of the paper is structured as follows: Section 2 reviews relevant literature, Section 3 defines key AI ethics concepts, Section 4 analyzes core ethical principles, Section 5 presents the discussion, and Section 6 concludes with future research directions.

2. RELATED LITERATURE REVIEWS

As discussed earlier, numerous comparative research based on AI ethical issues have been extensively explored in the literature. For instance, in the papers of (Papagiannidis et al., 2025; Robles-Carrillo, 2020; Taddeo and Floridi, 2018; Floridi, 2019), AI-related ethical and legal standards, principles, and technical complexities were addressed, focusing on their definition, purpose, and consequences. Several authors such as Kazim and Koshiyama (2021) explored AI ethics through engineering approaches, emphasizing principles, processes, ethical consciousness, and the application of ethics in practice, while Afroogh et al. (2018) investigated ethical issues through trustworthiness metrics, proposing future directions and potential solutions for the transition to trustworthy AI. Besides the above-mentioned research, numerous other studies focused on AI in specific contexts. For example, Meske et al. (2022) addressed explainable artificial intelligence (XAI), including stakeholders, aims, and future opportunities. Chng et al. (2025) examined ethical considerations in AI for child health, offering recommendations for child-centered medicine with a focus on the ethical principle of trustworthy AI. Hickman and Petrin (2021) explored trustworthy AI for sustainability related to reliable, transparent, and ethically aligned AI systems that support environmental goals and build trust in sustainable development. On the other side, there are research efforts (Korać et al., 2020; Bhattarai et al., 2024) focused on all three user preferences, in which security is highlighted as the primary research aim.

In the context of tourism, relevant research highlights the intersection between AI, ethics, and sustainability. Gretzel et al. (2015) provide an in-depth analysis of the ethical implications of digital and AI technologies in tourism, emphasizing the importance of privacy, transparency, and the responsible use of data. Tussyadiah (2020) discusses the ethical challenges of automation in tourism services and its impact on user experience, while Buhalis and Sinarta (2019) elaborate on the role of AI in creating personalized tourist experiences, stressing the necessity of adhering to ethical norms and sustainable destination management. Similarly, Li, Buhalis, and Weber (2019) address transparency and fairness issues in AI-based recommendation systems in tourism, underlining the importance of including multiple stakeholders in the design and implementation of such systems. More recently, Mariani and Borghi (2021) investigated algorithmic transparency and accountability in tourism platforms, arguing that opaque AI systems may undermine consumer trust and long-term destination sustainability. In addition, Gajdošík (2022) emphasized the ethical dimension of smart tourism ecosystems, highlighting risks of surveillance and bias while calling for governance frameworks that ensure fairness, inclusiveness, and responsible innovation in tourism AI applications.

While existing literature reviews have given coverage of numerous AI ethical principles from various perspectives—including tourism and service industries—there remains a need to provide their comprehensive and systematic consideration, which represents the key research aim of this study.

3. AI ETHICS BASIC TERMS AND DEFINITIONS

Given the conceptual vagueness and significant overlap among key terms used in AI-driven strategic decision-making, careful disambiguation is required (Jobin et al., 2019). To address this issue, Table 1 provides a concise overview of core concepts in AI ethics—such as ethical risks, cybersecurity, privacy, and trust—along with their respective definitions. These concepts are inherently complex, interconnected, and often subject to differing interpretations, which can create challenges for their consistent application in practice.

The concept of cybersecurity serves as a clear example of such ambiguity. Its definition has evolved over time and remains contested, with ongoing debates regarding whether it should be understood as a role, a discipline, or a multidimensional field encompassing information, operational, and network security (Schatz et al., 2017). This lack of conceptual clarity is particularly relevant in

the context of AI ethics, as it directly affects the effective, secure, and sustainable integration of AI technologies within tourism systems.

Table 1: Key terms and definitions associated with AI ethics

S.No	Term	Definition	Ref.
1	Ethical risks	The possibility that a stakeholder may neglect their ethical responsibilities toward others involved in the AI system.	Douglas et al., 2024
2	Cybersecurity	Security risk management involving strategies, policies, and tools used by organizations and states to protect data and assets in cyberspace, ensuring confidentiality, integrity, and availability.	Schatz et al., 2017
3	Privacy	The need for protection of personal and sensitive information.	Jobin et al., 2019
4	Trust	The extent to which individuals have confidence in a model's ability to handle sensitive data securely and reliably.	Humphreys et al., 2024

As emphasized by Jobin et al. (2019), the overlap and ambiguity of these terms necessitate precise interpretation in order to avoid misunderstandings and ensure their effective use in both strategic and operational contexts. This is especially important in tourism, where issues such as user trust, data protection, and transparency are fundamental to long-term sustainability and system acceptance.

To further reduce ambiguity and strengthen the conceptual foundations of ethical AI deployment, it is necessary to consider the broader set of ethical principles that guide AI development and application. Widely recognized principles in academic and policy discourse include beneficence, which emphasizes the contribution of AI to individual and societal well-being; justice and fairness, which require equitable access and the prevention of discrimination; and non-maleficence, which focuses on avoiding harm to individuals, communities, and the environment. In addition, autonomy highlights the importance of respecting user agency and informed consent, while accountability and responsibility stress the need for clearly defined obligations regarding AI outcomes. Transparency ensures that AI systems are explainable and understandable, and sustainability underscores the importance of aligning AI development with long-term environmental and societal goals.

These principles are reflected in major international policy frameworks, including the European Commission's Ethics Guidelines for Trustworthy Artificial Intelligence (2019) and UNESCO's Recommendation on the Ethics of Artificial Intelligence (2021). Both frameworks emphasize that AI systems must be lawful, ethical, and robust, while promoting human rights, inclusivity, and sustainable development.

In the context of tourism, these ethical principles provide a critical foundation for balancing user preferences with broader societal and environmental considerations. For instance, while AI enables personalized travel experiences, such personalization must not compromise user privacy or data protection. Similarly, AI-driven decision-making processes must remain transparent and accountable to prevent bias and ensure fairness in recommendation systems. Integrating sustainability into AI applications further ensures that smart tourism ecosystems support the long-term well-being of destinations, local communities, and the environment.

4. REVIEW OF BASIC AI ETHICAL PRINCIPLES

A more detailed examination of these foundational principles reveals that their interplay is far more dynamic and recursive than a simple linear model might suggest. While security and privacy indeed serve as essential prerequisites, trust is not merely a passive outcome but an active, evolving force that shapes user engagement from the very first interaction (Gillespie, 2020). Users often approach digital tourism platforms with a baseline of initial trust derived from brand reputation, third-party endorsements, or visible regulatory compliance signals such as GDPR badges, and this initial trust determines whether they will engage deeply enough to evaluate the platform's actual security and privacy practices (Mcknight & Chervany, 2001). Once established, trust operates within a continuous feedback loop: each positive, secure, and privacy-respecting interaction reinforces user confidence, creating a virtuous cycle of engagement, while even minor failures—such as a recommendation that feels uncomfortably accurate, suggesting excessive data surveillance, or a cumbersome authentication process that prioritizes security over usability—can rapidly erode trust and diminish the perceived value of the system's other ethical safeguards (Tussyadiah & Pesonen, 2018). This dynamic underscores the need to view security, privacy, and trust not as isolated concepts but as mutually reinforcing elements that must be managed holistically (Floridi et al., 2018).

Moving from conceptual principles to operational realities, the true challenge lies in how AI systems can translate these ethical commitments into concrete, user-facing mechanisms within the tourism context. In the domain of security, the central tension is the paradox between robustness and usability, where overly aggressive security measures risk alienating users who prioritize seamless experiences (Korać et al., 2022). AI addresses

this through adaptive authentication systems that employ risk-based analysis, applying minimal friction during routine activities such as booking a familiar hotel from a recognized device, while seamlessly escalating security protocols—such as multi-factor authentication—when anomalous patterns emerge, such as a login attempt from a new geographic location combined with a request to change payment details (Schneier, 2019). This approach aligns security with user preferences for both protection and convenience. Furthermore, AI enhances security proactively by analyzing network traffic to predict and neutralize cyber threats like credential stuffing or distributed denial-of-service attacks before they impact users, effectively transforming operational stability into a security feature that reinforces confidence (Kshetri, 2020).

In the realm of privacy, the growing tension between personalization and data protection demands a shift from the outdated model of static, one-time consent to a framework centered on continuous user agency and transparency (Nissenbaum, 2009). AI-driven tourism platforms rely heavily on user data to deliver tailored recommendations, yet users increasingly expect granular control over how their information is collected, processed, and shared (Acquisti, Brandimarte, & Loewenstein, 2015). This can be operationalized through dynamic privacy controls that present users with contextual choices—for instance, allowing a user to specify whether a restaurant recommendation system should utilize precise location data, past dining history, or only the current city as a general parameter (Bélanger & Crossler, 2011). Such approaches transform privacy from a passive legal obligation into an active, transparent dialogue. Additionally, privacy-enhancing technologies such as federated learning offer a powerful means of reconciling personalization with privacy by training AI models across decentralized devices without exposing individual user data; a tourism application could thus identify general patterns—for example, that a significant percentage of visitors to a particular museum also enjoy a nearby cultural tour—without any single user's data ever leaving their device, thereby delivering personalized insights while fundamentally respecting user privacy (McMahan et al., 2017; Hardt et al., 2020).

Trust, as the higher-order principle that consolidates the successful implementation of security and privacy, is increasingly contingent upon the explainability and fairness of AI systems (Ribeiro, Singh, & Guestrin, 2016). In tourism environments where users rely on unfamiliar platforms and automated recommendations, opaque decision-making undermines confidence, whereas transparent systems empower users to make informed choices (Miller, 2019). Explainable AI addresses this by providing contrastive explanations that clarify not only what is being recommended but why alternative options were deprioritized; rather than simply suggesting a particular hotel, a trustworthy system might explain that the chosen option is fifteen percent closer to the user's stated points of interest, offers higher ratings for family-friendliness, and remains within the specified budget, while noting that an alternative, though highly rated, is less conveniently located (Gunning et al., 2019). This level of clarity not only builds trust but also respects user autonomy by enabling informed override of automated suggestions (Lepri, Oliver, & Pentland, 2021). Fairness constitutes another critical dimension of trust, as users who perceive biased outcomes—such as systematic differences in accommodation recommendations based on demographic characteristics—will rapidly lose confidence in the system, making continuous auditing for disparate impact and the implementation of algorithmic fairness constraints essential components of ethical AI deployment (Barocas, Hardt, & Narayanan, 2019; Mehrabi et al., 2021).

A comprehensive analysis must also acknowledge the inherent tensions that arise when these principles are pursued simultaneously, as they are not always harmonious and often require deliberate trade-offs (Rahwan, 2018). The central paradox of personalization versus privacy captures this conflict vividly: the more data an AI system collects regarding a user's dietary restrictions, travel habits, budget, and personal preferences, the more effectively it can deliver the seamless, tailored experiences that build trust and demonstrate security value, yet this very data collection poses inherent privacy risks (Zuboff, 2019). Resolving this tension requires a transparent value exchange in which users willingly cede a degree of privacy only when the perceived benefits—such as time savings, superior recommendations, or exclusive offers—are sufficiently compelling and when trust in the platform's data stewardship is firmly established (Cranor, 2012). Similarly, the security-usability trade-off demands continuous calibration, as excessive authentication requirements can drive users away even as they provide robust protection (Weir et al., 2020). Perhaps most significantly, a tension exists between explainability and performance, as the most accurate AI models, particularly complex deep learning networks, often function as black boxes resistant to straightforward interpretation (Lipton, 2018). In high-stakes tourism applications—such as recommending medical tourism services or facilitating complex, high-value bookings—the ethical imperative for transparency may justify selecting a slightly less performant model that can offer clear, intelligible explanations over a superior but opaque alternative (Doshi-Velez & Kim, 2017).

Finally, these ethical considerations do not unfold in a vacuum but are profoundly shaped by the external context of regulation and cultural diversity. The absence of a universal binding standard for AI ethics has led to a fragmented landscape in which frameworks such as the European Union's General Data Protection Regulation (GDPR) and the proposed EU AI Act function as de facto ethical benchmarks, making principles like the right to explanation not merely aspirational but legally

enforceable requirements (European Commission, 2016; European Commission, 2021). For tourism organizations operating globally, this regulatory complexity often necessitates adopting the strictest standards as a baseline to ensure compliance and maintain trust across multiple jurisdictions (Voigt & Von dem Bussche, 2017). Simultaneously, user preferences regarding privacy exhibit significant cultural variation: travelers from regions with strong data protection traditions may expect rigorous data minimization and explicit consent mechanisms, whereas users from other cultural contexts may demonstrate greater willingness to share personal information in exchange for tangible economic benefits such as discounts or loyalty rewards (Milberg et al., 1995; Krasnova et al., 2012). A truly user-centric ethical framework must therefore be culturally adaptive, allowing for regional customization of privacy and personalization settings while maintaining a consistent core commitment to security, transparency, and user agency (Hofstede, 2001). In navigating these interconnected dimensions—the dynamic interplay of principles, the operational mechanisms that bring them to life, the inevitable trade-offs they entail, and the external regulatory and cultural forces that shape their implementation—tourism organizations can move beyond a superficial adherence to ethical ideals toward a sustainable model of AI deployment that places user confidence and long-term trust at the very center of strategic decision-making (Buhalis & Sinarta, 2019; Tussyadiah, 2020).

5. DISCUSSION

This discussion emphasizes that the role of each stakeholder is crucial in developing and implementing AI ethical principles to minimize risks, such as unethical behavior, privacy violations, or the dissemination of misinformation (Floridi et al., 2018; Jobin et al., 2019). All participants within AI-driven tourism systems—including developers, managers, and end-users—should be aware of both the potential benefits and risks associated with AI applications. Understanding AI ethics from a human-centered perspective requires addressing not only security concerns but also complex moral and ethical issues related to privacy, fairness, and accountability. The stakeholder landscape in tourism AI is notably diverse, encompassing technology developers who design algorithms, organizational leaders who determine deployment strategies, frontline staff who interact with AI tools, and end-users whose data and engagement sustain these systems, each bringing distinct perspectives, vulnerabilities, and ethical responsibilities to the ecosystem (Shneiderman, 2020). Developers, for instance, hold particular responsibility because the values and biases embedded in algorithms during the design phase become inscribed into systems that may operate at scale, affecting countless users without ongoing human oversight (Rahwan, 2018). Managers and organizational decision-makers, conversely, shape the governance structures within which AI operates, determining which ethical principles are prioritized, how transparency is implemented, and what recourse mechanisms exist when systems fail or cause harm (Mittelstadt, 2019). End-users, while often positioned as passive recipients of AI services, also play an active role in shaping ethical outcomes through their choices, feedback, and willingness to hold organizations accountable for ethical lapses (Buhalis & Sinarta, 2019).

Comprehensive AI education and awareness programs should be implemented across all organizational levels to address ethical and moral principles relevant to specific roles (Shestakova, 2021; Bansal, 2025). These programs should evolve alongside AI advancements, integrating real-world case studies and interactive learning methods to bridge the gap between ethical theory and practical application (Tussyadiah, 2020). The effectiveness of such educational initiatives depends critically on their contextual grounding; abstract ethical principles, when presented without concrete operational relevance, fail to equip stakeholders with the practical judgment needed to navigate the nuanced dilemmas that arise in daily AI deployment (Whittlestone et al., 2019). For frontline tourism staff, this may involve training on how to recognize when an AI recommendation system is producing biased or inappropriate suggestions for certain demographic groups and how to escalate such concerns appropriately. For managers, education should focus on understanding the trade-offs inherent in AI deployment—such as the tension between personalization and privacy—and developing the organizational processes necessary to audit AI systems for fairness, accuracy, and unintended consequences (Raji et al., 2020). For developers and programmers, training must extend beyond technical proficiency to incorporate human rights, inclusivity, and sustainability considerations, ensuring that ethical reflection becomes embedded within the software development lifecycle rather than treated as an external compliance exercise (Morley et al., 2020). Interactive learning methods, including scenario-based simulations and post-deployment case study analyses, have proven particularly effective in helping stakeholders internalize ethical reasoning skills that can be applied flexibly as new challenges emerge (Saltz & Dewar, 2019).

Special attention must be given to developers and programmers, ensuring their training incorporates human rights, inclusivity, and sustainability considerations. The decisions made during algorithm design—from data selection and feature engineering to model architecture and evaluation metrics—profoundly shape the ethical performance of AI systems in ways that may not become apparent until deployment (Holstein et al., 2019). Developers must be equipped to recognize how seemingly technical choices carry moral implications; for example, the selection of training data that

underrepresents certain geographic regions or socioeconomic groups can result in recommendation systems that systematically disadvantage travelers from those backgrounds, perpetuating existing inequalities within tourism (Hao, 2021). Similarly, the choice to optimize for engagement metrics without considering broader user welfare may lead systems to prioritize sensational or divisive content, a dynamic that has particular relevance for tourism platforms featuring user-generated reviews and recommendations (O'Neil, 2016). Incorporating human rights frameworks into technical education helps developers understand that issues such as non-discrimination, freedom of choice, and informational self-determination are not peripheral concerns but central to responsible AI design (United Nations, 2021). Inclusivity considerations demand that development teams themselves reflect the diversity of the populations their systems serve, as homogeneous teams are more likely to overlook bias and accessibility issues that affect marginalized user groups (West et al., 2019). Sustainability considerations, meanwhile, extend beyond environmental impact to encompass the long-term viability of AI systems in supporting resilient and equitable tourism development, ensuring that technological deployment does not come at the expense of local communities or cultural heritage (Gössling & Hall, 2019).

Beyond the core AI ethical principles of security, privacy, and trust, organizations should incorporate additional principles such as fairness, avoidance of bias, accuracy, explainability, and accountability (European Commission, 2019; UNESCO, 2021). For instance, fairness and bias principles are grounded in moral values that ensure individuals are not subjected to discrimination or inequitable treatment (Floridi & Cowls, 2019). In the tourism context, fairness concerns manifest across multiple dimensions: algorithmic bias may result in differential pricing based on perceived user characteristics, exclusionary recommendation patterns that disadvantage certain demographic groups, or unequal access to promotional opportunities based on opaque algorithmic ranking systems (Travaini et al., 2022). Addressing fairness requires not only technical interventions—such as bias audits and fairness constraints in model training—but also organizational commitments to transparency about how algorithmic decisions are made and mechanisms for users to contest outcomes they perceive as unjust (Barocas, Hardt, & Narayanan, 2019). Accuracy, while often framed as a technical metric, carries profound ethical significance in tourism where inaccurate information can have tangible consequences for traveler safety, financial outcomes, and overall experience quality (Tussyadiah, 2020). A recommendation system that consistently misrepresents accommodation locations, fails to account for seasonal closures, or overlooks critical safety information undermines user trust and may expose organizations to legal and reputational liability (Gretzel, 2021).

Transparency is a cornerstone of ethical AI, enhancing trust and enabling accountability for decisions made by AI systems (Papagiannidis et al., 2025). Transparency operates at multiple levels: disclosure to users that AI systems are in use, clarity about what data is collected and how it is utilized, and explainability regarding how specific decisions or recommendations are generated (Diakopoulos, 2020). In tourism, where users interact with AI across booking platforms, recommendation engines, chatbots, and dynamic pricing systems, transparency becomes essential for informed consent and meaningful user agency (Tussyadiah & Miller, 2019). However, transparency must be implemented thoughtfully to avoid overwhelming users with technical details that obscure rather than clarify; effective transparency provides accessible, contextually relevant information that empowers users to make informed choices about their engagement with AI systems (Rader & Gray, 2015). Accountability complements transparency by establishing clear lines of responsibility for AI system outcomes, ensuring that when systems cause harm or operate in ways inconsistent with stated ethical commitments, there are mechanisms for redress and continuous improvement (Wieringa, 2020). This requires moving beyond vague commitments to ethical principles toward concrete governance structures, including human oversight of high-stakes AI decisions, regular ethical audits, and clearly communicated channels for user feedback and complaint resolution (Rakova et al., 2021).

Sustainability in AI deployment emphasizes the long-term societal, environmental, and economic impacts of AI systems, ensuring that technology supports inclusive and responsible tourism development (Leimanis & Palkova, 2021). The environmental dimension of sustainability addresses the significant energy consumption associated with training and operating large-scale AI models, a concern that becomes amplified as tourism increasingly adopts computationally intensive applications such as real-time personalization and computer vision for destination management (Strubell, Ganesh, & McCallum, 2019). Environmentally sustainable AI requires organizations to consider not only the efficiency of their algorithms but also the broader carbon footprint of their technological infrastructure, aligning with tourism industry commitments to climate action and responsible resource management (Gössling & Hall, 2019). The social dimension of sustainability demands attention to how AI deployment affects tourism workers, local communities, and cultural practices. Automation may displace certain categories of employment while creating demand for new skills, necessitating proactive workforce development and just transition strategies (Brougham & Haar, 2018). In destinations where AI-powered platforms shape visitor flows and spending patterns, ensuring that economic benefits are distributed equitably across local communities becomes a critical ethical consideration, preventing the concentration of benefits among already-privileged actors while

exacerbating pressures such as overtourism in vulnerable destinations (Mason & Stasi, 2020). Economic sustainability, meanwhile, requires that AI investments contribute to long-term value creation rather than short-term optimization that externalizes costs onto users, communities, or the environment (Hilty & Aebischer, 2015).

Explainability is crucial for fostering understanding and trust, allowing stakeholders to comprehend how AI-driven decisions are made and ensuring that these systems can be effectively scrutinized and improved. The demand for explainability arises from the fundamental tension between the complexity of modern AI systems and the human need for understandable justifications for decisions that affect one's interests (Miller, 2019). In tourism, explainability takes on particular importance across scenarios ranging from dynamic pricing and personalized recommendations to credit-based booking decisions and automated customer service interactions (Tussyadiah, 2020). Users who receive an unexpected price increase or an unexplained recommendation decline may develop lasting negative attitudes toward the platform if they cannot understand the reasoning behind the decision, whereas clear explanations can mitigate frustration and preserve trust even when outcomes are unfavorable (Gillespie, 2020). For developers and auditors, explainability enables debugging of problematic model behavior, identification of bias, and assessment of whether systems are operating as intended (Lipton, 2018). However, the pursuit of explainability must navigate the aforementioned trade-off with performance, as the most accurate models are not always the most interpretable. Organizations must therefore make deliberate, context-sensitive choices about the level of explainability required for different applications, recognizing that in high-stakes contexts—such as safety-related recommendations or accommodations for travelers with disabilities—interpretability may justifiably take precedence over marginal gains in predictive accuracy (Doshi-Velez & Kim, 2017).

Integrating these ethical principles supports continuous improvement and the responsible use of AI, ultimately striking a balance between technical performance, user preferences, and moral responsibility in tourism. This integration cannot be achieved through checklist-based compliance or the mere publication of ethical principles; it requires embedding ethical considerations into organizational culture, governance structures, and technical practices from the earliest stages of system design through ongoing operation and refinement (Mittelstadt, 2019). Continuous improvement in AI ethics demands mechanisms for monitoring systems post-deployment to detect emergent issues that were not apparent during development, including fairness violations that arise from shifting data distributions or novel use cases (Raji et al., 2020). It also requires creating feedback loops that enable user experiences and concerns to inform system evolution, recognizing that ethical AI is not a static achievement but an ongoing process of learning and adaptation (Whittlestone et al., 2019). The balance between technical performance and moral responsibility represents a fundamental challenge that cannot be resolved through technical means alone; it demands value-based judgments about what trade-offs are acceptable and who should have the authority to make such determinations (Rahwan, 2018). In tourism, where AI systems mediate experiences that are deeply personal and culturally significant, these judgments must reflect not only organizational priorities but also the diverse values and preferences of the users and communities whose lives are affected by AI deployment (Buhalis & Sinarta, 2019).

Identifying potential threats and developing mitigation strategies is essential for fostering a safe and ethically aligned tourism environment. Among the most pressing threats are algorithmic discrimination, which can systematically disadvantage travelers based on protected characteristics such as race, gender, age, or disability, often without the knowledge of either users or platform operators (Obermeyer et al., 2019). Mitigation requires ongoing fairness auditing using both quantitative metrics and qualitative assessment methods that capture the lived experiences of affected users (Holstein et al., 2019). Privacy violations constitute another critical threat, particularly as tourism AI systems increasingly integrate data across multiple touchpoints—from search behavior and booking history to in-destination mobility patterns and biometric data—creating comprehensive surveillance profiles that users may not fully understand or consent to (Zuboff, 2019). Mitigation strategies include implementing privacy-by-design approaches, minimizing data collection to what is strictly necessary, providing users with meaningful control over their information, and adopting privacy-enhancing technologies such as differential privacy and federated learning (Acquisti et al., 2015). The dissemination of misinformation through AI systems—whether through hallucinating chatbots, manipulated visual content, or algorithmic amplification of unverified user-generated content—represents an emerging threat with particular significance for tourism, where misinformation can lead to unsafe travel decisions, economic harm to destinations, and erosion of trust in information ecosystems (Buchanan, 2020). Mitigation requires robust content moderation systems, transparent labeling of AI-generated content, and investment in digital literacy education that empowers users to critically evaluate AI-mediated information (European Commission, 2021). Finally, the concentration of market power among a small number of AI-enabled tourism platforms raises structural concerns about consumer choice, fair competition, and the capacity of destinations and local businesses to maintain autonomy in the face of algorithmic intermediation (Kenney & Zysman, 2020). Mitigation strategies may include regulatory interventions, support for open and interoperable systems, and development of public-

interest alternatives that prioritize community benefit over profit maximization.

The key discussion points emerging from this analysis underscore that stakeholder awareness and engagement are fundamental for ethical AI implementation, requiring a departure from technocentric approaches that treat ethics as a secondary consideration or compliance requirement. Education and training must be role-specific, practical, and continuously updated, recognizing that the ethical challenges of AI evolve alongside technological capabilities and deployment contexts, demanding ongoing learning rather than one-time training interventions. Ethical principles should extend beyond security, privacy, and trust to include fairness, explainability, accountability, and sustainability, forming a comprehensive framework that addresses the full scope of moral considerations raised by AI in tourism. Finally, balancing AI technical efficiency with ethical and social responsibility is essential for long-term trust and adoption in tourism, requiring organizations to recognize that ethical AI is not merely a risk management exercise but a foundational condition for sustainable value creation in an increasingly algorithmically mediated industry. As AI systems continue to advance in capability and pervasiveness within tourism, the organizations that succeed in building and maintaining user trust will be those that demonstrate not only technical excellence but also genuine commitment to ethical principles, operationalized through robust governance, meaningful stakeholder engagement, and unwavering accountability for the systems they deploy.

6. CONCLUSIONS AND FUTURE DIRECTIONS

This paper examined the fundamental ethical principles of artificial intelligence (AI) in tourism organizations, focusing on security, privacy, and trust, and their role in enhancing strategic decision-making. The study emphasizes that the ethical development and application of AI require a proactive and comprehensive approach that integrates robust governance structures, regulatory compliance, and continuous education and training.

Since AI itself is not a moral agent, responsibility and accountability remain with human stakeholders. Embedding ethical considerations from the outset—rather than responding reactively to crises—is crucial for ensuring secure, reliable, and trustworthy AI deployment. Active participation of all stakeholders throughout the development and implementation process is essential to maintaining a human-centered approach, ensuring that AI augments rather than replaces human capabilities.

The study highlights that ethical AI deployment in tourism necessitates examining both technological and educational frameworks through a moral lens. Human oversight is vital to ensure that AI operates ethically, enabling audits, interventions, and continuous monitoring of AI-driven decision-making. Such oversight mitigates unforeseen risks, aligns AI behavior with human ethical values, and ensures sustainable and socially responsible outcomes.

Furthermore, this research underscores the importance of additional ethical principles, including fairness, bias prevention, accuracy, accountability, explainability, and inclusivity. These principles warrant more in-depth investigation to develop robust AI systems that foster user trust, ensure equitable treatment, and enhance transparency in decision-making processes.

Finally, this study proposes several directions for future research, including:

- Developing methodologies for embedding ethical AI principles throughout the AI lifecycle in tourism;
- Investigating human–AI collaboration models that ensure ethical oversight without compromising operational efficiency;
- Evaluating the long-term societal and environmental impacts of AI adoption in tourism systems;
- Exploring explainable AI (XAI) frameworks to enhance transparency, trust, and informed user engagement.

Overall, the responsible and ethical use of AI in tourism is essential for promoting sustainable development, protecting user preferences, and safeguarding the common good.

REFERENCES

1. Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. Science, 347(6221), 509–514.
2. Afroogh, S., Akbari, A., Malone, E., et al. (2024). Trust in AI: Progress, challenges, and future directions. Humanities and Social Sciences Communications, 11, 1568.
3. Ali, S., Wang, J., & Leung, V. C. M. (2025). AI-driven fusion with cybersecurity: Exploring current trends, advanced techniques, future directions, and policy implications. Information Fusion, 118, 102922.
4. Bansal, C. (2025). AI ethics and sustainability: Accelerating paradigm shifts toward sustainable development. Journal of Strategy & Innovation, 36(1), 200537.
5. Bansal, P. (2025). AI literacy and ethical awareness in the tourism workforce. Journal of Sustainable Tourism, 33(2), 215–232.
6. Barocas, S., Hardt, M., & Narayanan, A. (2019). Fairness and machine learning: Limitations and opportunities. MIT Press.
7. Bélanger, F., & Crossler, R. E. (2011). Privacy in the digital age: A review of information privacy research. MIS Quarterly, 35(4), 1017–1041.

8. Bhattarai, I., Pu, C., Choo, K.-K. R., & Korać, D. (2024). Authentication protocol for the Internet of Drones. IEEE Internet of Things Journal, 11(11), 19790–19803.
9. Brougham, D., & Haar, J. (2018). Smart technology and future workplace perceptions. Journal of Management & Organization, 24(2), 239–257.
10. Buchanan, B. (2020). The AI deception problem. Center for Security and Emerging Technology.
11. Buhalis, D., & Sinarta, Y. (2019). Real-time co-creation in tourism. Journal of Travel & Tourism Marketing, 36(5), 563–582.
12. Chng, S. Y., Tern, M. J. W., Lee, Y. S., et al. (2025). Ethical considerations in AI for child health. npj Digital Medicine, 8, 152.
13. Cranor, L. F. (2012). Privacy notice and choice mechanisms. Journal on Telecommunications and High Technology Law, 10, 273.
14. Diakopoulos, N. (2020). Transparency. In The Oxford handbook of ethics of AI. Oxford University Press.
15. Doshi-Velez, F., & Kim, B. (2017). Interpretable machine learning. arXiv preprint arXiv:1702.08608.
16. Douglas, D. M., Lacey, J., & Howard, D. (2024). Ethical risk for AI. AI Ethics. https://doi.org/10.1007/s43681-024-00549-9
17. European Commission. (2016). General Data Protection Regulation (GDPR).
18. European Commission. (2019). Ethics guidelines for trustworthy AI.
19. European Commission. (2021). Proposal for an AI Act.
20. Floridi, L. (2019). Trustworthy AI. Nature Machine Intelligence, 1, 261–262.
21. Floridi, L., & Cowls, J. (2019). AI principles framework. Harvard Data Science Review, 1(1).
22. Floridi, L., Cowls, J., Beltrametti, M., et al. (2018). AI4People framework. Minds and Machines, 28(4), 689–707.
23. Gajdošík, T. (2022). Ethical dimension of smart tourism ecosystems. Journal of Sustainable Tourism, 30(9).
24. Gillespie, N. (2020). Trust in AI systems. In The Oxford handbook of AI governance.
25. Gretzel, U. (2021). AI in tourism. In Handbook of e-tourism. Springer.
26. Gretzel, U., Sigala, M., Xiang, Z., & Koo, C. (2015). Smart tourism. Electronic Markets, 25(3), 179–188.
27. Gunning, D., Stefik, M., Choi, J., et al. (2019). Explainable AI. Science Robotics, 4(37).
28. Hickman, E., & Petrin, M. (2021). Trustworthy AI and corporate governance. European Business Organization Law Review, 22, 593–625.
29. Holzinger, A., Weippl, E., Tjoa, A. M., & Kieseberg, P. (2021). AI for sustainable development. Lecture Notes in Computer Science, 12844.
30. Humphreys, D., Koay, A., Desmond, D., et al. (2024). AI hype as cybersecurity risk. AI Ethics, 4, 791–804.
31. Jobin, A., Ienca, M., & Vayena, E. (2019). AI ethics guidelines landscape. Nature Machine Intelligence, 1, 389–399.
32. Kazim, E., & Koshiyama, A. S. (2021). Overview of AI ethics. Patterns, 2(9), 100314.
33. Kenney, M., & Zysman, J. (2020). Platform economy. Cambridge Journal of Regions, Economy and Society, 13(1), 55–76.
34. Korać, D., Damjanović, B., Simić, D., & Pu, C. (2025). AI-based fusion framework. Information Processing & Management, 62(6), 104233.
35. Korać, D., Damjanović, B., Simić, D. (2022a). Digital identity model. Journal of Supercomputing, 78, 3355.
36. Korać, D., Damjanović, B., Simić, D., & Choo, K. K. R. (2022b). Hybrid XSS attack. Journal of King Saud University, 34(10), 9284–9300.
37. Korać, D., Damjanović, A., & Popović, P. (2022). AI security in tourism. Tourism Management Perspectives, 41, 100936.
38. Kshetri, N. (2020). AI in cybersecurity. Computer, 53(4), 62–66.
39. Leimanis, A., & Palkova, K. (2021). AI ethics in healthcare. European Journal of Sustainable Development, 10(1), 90–102.
40. Lipton, Z. C. (2018). Model interpretability. Queue, 16(3), 31–57.
41. Mariani, M., & Borghi, M. (2021). Responsible AI in tourism. Tourism Management, 82, 104178.
42. Meske, C., Bunde, E., Schneider, J., & Gersch, M. (2022). Explainable AI. Information Systems Management, 39, 53–63.
43. Miller, T. (2019). Explanation in AI. Artificial Intelligence, 267, 1–38.
44. Mittelstadt, B. (2019). Limits of AI principles. Nature Machine Intelligence, 1(11), 501–507.
45. Nissenbaum, H. (2009). Privacy in context. Stanford University Press.
46. Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). AI governance. Journal of Strategic Information Systems, 34(2), 101885.
47. Rahwan, I. (2018). Society-in-the-loop. Ethics and Information Technology, 20(1), 5–14.
48. Robles-Carrillo, M. (2020). AI: From ethics to law. Telecommunications Policy, 44(6), 101937.
49. Sarker, I. H., et al. (2024). Explainable AI for cybersecurity. ICT Express, 10(4), 935–958.
50. Schneier, B. (2019). Click here to kill everybody. Norton.
51. Shneiderman, B. (2020). Human-centered AI. International Journal of Human-Computer Interaction, 36(6), 495–504.
52. Taddeo, M., & Floridi, L. (2018). AI as a force for good. Science, 361, 751–752.
53. Tussyadiah, I. P. (2020). Automation in tourism. Annals of Tourism Research, 81, 102883.
54. UNESCO. (2021). Recommendation on the ethics of artificial intelligence.
55. Zuboff, S. (2019). The age of surveillance capitalism. PublicAffairs.