Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Rebuilding Academic Trust in the AI Era: An Integrated Education Governance and Institutional Marketing Framework

Marwa Khairallah

marwakhairallah1988@gmail.com

Faculty of Economics and Management of Sfax

Research Article

Keywords: Generative AI, Academic Trust, Higher Education Governance, Institutional Marketing, PLS-SEM

Posted Date: October 1st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10997915/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

The rapid integration of Generative Artificial Intelligence (GenAI) in higher education has triggered an institutional crisis of academic trust. This challenge is characterized by an "adoption-trust paradox," where ubiquitous student usage coexists with deep ethical skepticism and concerns over academic integrity. While existing research predominantly focuses on individual technology acceptance models (e.g., TAM, UTAUT), this study shifts the analytical lens to institutional governance and its marketing implications, framing academic trust as a vital strategic brand asset that drives organizational legitimacy and student behavioral intentions. We propose and empirically test a multidimensional governance framework consisting of four core pillars: ethical transparency, algorithmic accountability, digital AI literacy, and adaptive AI governance. Utilizing a quantitative research design and Partial Least Squares Structural Equation Modeling (PLS-SEM), we analyze survey data from 380 higher education students in Tunisia. The structural model reveals that Adaptive AI Governance is the most powerful predictor of Institutional Trust ($\beta = 0.297$, $p < 0.001$), followed by Algorithmic Accountability ($\beta = 0.258$, $p < 0.001$) and Ethical Transparency ($\beta = 0.231$, $p < 0.001$), while AI Literacy has a smaller but significant effect ($\beta = 0.137$, $p = 0.006$). Institutional Trust, in turn, exhibits an exceptionally strong positive relationship with Perceived Institutional Ethicality ($\beta = 0.741$, $p < 0.001$). Together, trust and ethicality explain 61.2% of the variance in students' behavioral intentions toward sustainable AI use, with perceived ethicality acting as a significant partial mediating mechanism ($\beta_{indirect} = 0.337$, $p < 0.001$). Beyond pedagogical compliance, these findings demonstrate that robust governance directly enhances institutional brand equity, student attraction, and procedural legitimacy. We outline strategic recommendations for university marketers and educators to co-create proactive, values-led learning ecosystems that strengthen academic reputation in the digital era.

1. Introduction

Higher education has entered a structural turning point in which Generative Artificial Intelligence (GenAI) is no longer an emerging instructional novelty but a pervasive force reshaping how knowledge is produced, assessed, communicated, and legitimized (Abdi & Abdullahi, 2026; Đerić et al., 2025; Maleki, 2026). The central disruption is not simply technological adoption; it is the erosion of academic trust that occurs when institutions can no longer confidently distinguish authentic learning from machine-assisted performance, nor rely on inherited assessment and integrity mechanisms to preserve credibility (Kofinas et al., 2025; Maleki, 2026). This tension has produced an "adoption-trust paradox" across higher education. Students and educators increasingly use GenAI for writing, feedback, translation, idea generation, and academic support, yet this growing dependence coexists with serious concerns about cheating, misinformation, privacy, bias, weakened critical thinking, and the dilution of authorship and originality (Abdi & Abdullahi, 2026; Sun et al., 2026; Wang et al., 2025). Multicultural evidence shows that many participants view GenAI as a potential facilitator of academic misconduct, while others simultaneously recognize its educational value, confirming that institutional responses must address ambivalence rather than assume either resistance or acceptance alone (Yusuf et al., 2024).The problem, therefore, is not merely whether GenAI will be used in universities, but whether its use will be governed in

ways that sustain legitimacy, fairness, and confidence in institutional judgment. Existing research shows that trust in educational AI is shaped differently from general technology adoption because it depends not only on performance or usability, but also on socio-ethical concerns, pedagogical values, and institutional mediation (Lelescu et al., 2025; Sabri et al., 2026). Trust is also multifaceted, encompassing beliefs about accuracy, privacy protection, and non-maliciousness, which makes it especially vulnerable when universities offer inconsistent guidance or fail to align policy with actual classroom realities (Đerić et al., 2025). Early institutional responses have been uneven. Universities have begun to articulate ethical dimensions such as transparency, accountability, human oversight, and inclusiveness, yet many policy environments remain fragmented, under-specified, or overly reactive (Dabis & Csáki, 2024; Nikolic et al., 2024). Systematic reviews further show that formal training, systematic guidelines, and comprehensive policy frameworks remain scarce, even though institutional strategies such as leadership support, policies, and training are foundational drivers of trust (Lelescu et al., 2025; Nikolic et al., 2024). In some contexts, visible institutional support still lags behind student adoption, suggesting that universities are often responding after trust conditions have already deteriorated (Abdi & Abdullahi, 2026). For that reason, the restoration of institutional trust requires a shift from punitive or ad hoc control toward a formal governance architecture. Evidence increasingly suggests that sustainable trust is strengthened when institutions operationalize transparent communication, digital accountability, literacy-building, participatory norms, and responsible oversight as concrete organizational practices rather than symbolic ethical commitments (Aprianto et al., 2026; Saddam et al., 2026). Trust-building also depends on calibrating use appropriately: both overtrust and undertrust produce harmful outcomes, while guided, transparent, and norm-based ecosystems encourage more reflective and sustainable engagement with AI (Đerić et al., 2025; Xia et al., 2026). Against this background, this study addresses the core question: which institutional governance pillars most directly shape students' trust in their academic institution during the age of GenAI? This study argues that Ethical Transparency, Algorithmic Accountability, AI Literacy, and Adaptive AI Governance constitute the core structural assurances through which universities can rebuild confidence in the fairness, intelligibility, and ethical legitimacy of AI integration (Lelescu et al., 2025; Maleki, 2026). Ethical transparency reduces uncertainty through clear disclosure and honest communication (Dabis & Csáki, 2024; Shata & Hartley, 2025; Wang et al., 2025). Algorithmic accountability reinforces confidence by preserving human responsibility, oversight, and procedural fairness in AI-mediated decisions (Dabis & Csáki, 2024; Xia et al., 2026). AI literacy equips students to evaluate capabilities, limitations, and risks more critically, supporting informed rather than blind trust (Alotaibi, 2025; Baig & Yadegaridehkordi, 2025; Sun et al., 2026). Adaptive AI governance enables institutions to respond to rapid technological change with flexible, educative, and context-sensitive policy rather than static prohibition (Dabis & Csáki, 2024; Abdi & Abdullahi, 2026). Accordingly, this study departs from individualist adoption models by treating trust not as a by-product of user familiarity alone, but as an institutional outcome produced by governance quality. In doing so, it responds to growing calls for higher education to approach GenAI integration not as a race toward adoption, but as a deliberate process of trust-building that secures academic integrity, ethical sustainability, and institutional legitimacy in digitally transformed learning environments (Lelescu et al., 2025; Maleki, 2026).

2. Literature Review and Hypothesis Development

2.1 The Crisis of Academic Trust in the AI Era

Academic trust is the cornerstone of the relationship between students, educators, and institutions. However, the rise of GenAI has introduced systemic vulnerabilities (Coman et al., 2026; Francis et al., 2025; Sejdiu et al., 2026). Scholars note that the "black-box" nature of Large Language Models (LLMs) obscures the origin of information, leading to challenges in authorship, ownership, and accountability (Maleki, 2026; Lelescu et al., 2025; Sejdiu et al., 2026). The risk of "hallucinations", where AI generates plausible but false information, further undermines the reliability of academic research (Đerić et al., 2025). Recent reviews argue that GenAI exposes systemic weaknesses in policy design, assessment regimes, and institutional culture rather than merely creating isolated episodes of student misconduct (Alyafei et al., 2026; Maleki, 2026; Sejdiu et al., 2026). This erosion of trust is intensified by the fact that GenAI-generated content is often difficult to distinguish from human work, undermining long-standing assumptions that authorship is observable, verifiable, and individually attributable (Alyafei et al., 2026; Maleki, 2026). Traditional safeguards such as plagiarism detection, proctoring, and static integrity codes no longer provide sufficient assurance in environments where AI can generate original text, summaries, and disciplinary writing on demand (Alyafei et al., 2026; Furze et al., 2024; Maleki, 2026). Scholars therefore increasingly frame the problem as a transition from content verification to process accountability, where institutions must evaluate how work is produced, disclosed, and justified rather than whether a text appears original on its surface (Maleki, 2026; Sejdiu et al., 2026). Trust is further destabilized by the epistemic opacity of large language models. The literature repeatedly identifies concerns over accuracy, hallucinations, misinformation, privacy, and non-malicious use as core dimensions shaping whether GenAI is seen as trustworthy in academic settings (Đerić et al., 2025). When users cannot determine how outputs were generated, what data sources were used, or whether fabricated content has been introduced, confidence in both the technology and the academic process is weakened (Coman et al., 2026; García-López & Trujillo-Liñán, 2025). Privacy and intellectual property concerns deepen this problem, since fears of data misuse, ownership violations, and harmful or unethical applications extend distrust from the tool itself to the institution that authorizes or normalizes its use (Đerić et al., 2025; Furze et al., 2024). A second source of the trust crisis is the widening gap between rapid student adoption and slower institutional readiness. Across contexts, students report high familiarity with GenAI and broad support for its educational use, yet universities often lag in providing coherent policy, training, and governance structures (Barus et al., 2025; Sejdiu et al., 2026; Yusuf et al., 2024). The literature describes current institutional responses as fragmented, reactive, and unevenly implemented, with many policies focusing narrowly on misconduct while neglecting broader pedagogical, ethical, and regulatory questions (Coman et al., 2026; Furze et al., 2024;). While over 80% of students engage with AI tools, a vast majority report a lack of awareness regarding their university's specific ethical guidelines (Yusuf et al., 2024; Barus et al., 2025). This regulatory vacuum fosters an environment where distrust persists despite the efficiency gains provided by the technology (Lelescu et al., 2025). Systematic reviews also show that leadership engagement remains weak relative to the scale

of the challenge, limiting the structural assurances needed for trust formation (Lelescu et al., 2025). This mismatch between use and governance produces what can be described as an adoption-trust paradox. Students and educators continue to use GenAI because of its convenience, efficiency, and learning support, while simultaneously expressing concern that it facilitates cheating, misinformation, overreliance, and inequity (Coman et al., 2026; Jayasinghe et al., 2026; Yusuf et al., 2024). Multicultural evidence shows that perceptions of misconduct risk and appropriate regulation vary across cultural settings, which means trust cannot be restored through universal prohibition or purely technical fixes alone (Yusuf et al., 2024). Trust in educational AI appears to depend on socio-ethical and institutional conditions that standard adoption models do not adequately capture (Lelescu et al., 2025).

2.2 Construct Grounding and Theoretical Perspectives

To address these limitations, rebuilding academic trust requires a shift from individual-level acceptance models toward an institutional governance perspective. Existing frameworks such as the Technology Acceptance Model (TAM) and Unified Theory of Acceptance and Use of Technology (UTAUT) explain perceived usefulness, ease of use, and behavioral intention, but they do not sufficiently account for how legitimacy, pedagogical values, and organizational safeguards shape trust in higher education (Lelescu et al., 2025; Sabri et al., 2026; Sejdiu et al., 2026). Recent scholarship instead argues for integrated approaches that connect ethics, pedagogy, assessment, and regulation within a shared institutional architecture (García-López & Trujillo-Liñán, 2025; Sejdiu et al., 2026). In this study, four governance pillars are advanced as the core structural antecedents of Institutional Trust: Ethical Transparency, Algorithmic Accountability, AI Literacy, and Adaptive AI Governance. These pillars synthesize recurrent themes across the literature on trustworthy AI, academic integrity, and responsible university policy (Dabis & Csáki, 2024; Francis et al., 2025; Tran et al., 2025). Together, they move the analysis beyond the question of whether students use GenAI toward the more consequential question of whether institutions govern its use in ways perceived as fair, intelligible, and ethically legitimate (García-López & Trujillo-Liñán, 2025; Maleki, 2026). Table 1 provides a comprehensive summary of the theoretical framing and empirical basis of these four core pillars.

Table 1
The four governance pillars of Institutional Trust

Pillar	Theoretical Framing	Empirical Basis in Literature
Ethical Transparency	Proactive disclosure of AI use, limits, risks, and acceptable boundaries in teaching, assessment, and administration.	Clear communication in syllabi and policy improves transparency and trust conditions (Dabis & Csáki, 2024; Đerić et al., 2025).
Algorithmic Accountability	Human responsibility, oversight, contestability, and monitoring for AI-assisted decisions and outputs.	Accountability, human oversight, and clear standards are repeatedly identified as governance necessities (Dabis & Csáki, 2024; García-López & Trujillo-Liñán, 2025).
AI Literacy	Cognitive, evaluative, and ethical competence to interpret AI outputs and use them responsibly.	AI literacy predicts attitudes, perceived control, and more informed engagement with GenAI (Sun et al., 2026; Wang et al., 2024).
Adaptive AI Governance	Continuous review, stakeholder participation, and flexible policy iteration as AI capabilities evolve.	Static rules are inadequate; reviews call for continuous evaluation and regularly updated frameworks (García-López & Trujillo-Liñán, 2025; Furze et al., 2024).

Ethical Transparency refers to the institution's proactive disclosure of where, how, and under what limits AI is used in academic and administrative processes. The literature treats transparency not as symbolic openness but as a practical trust mechanism requiring clear communication about AI functionality, data sources, limitations, and acceptable use boundaries (Coman et al., 2026; Đerić et al., 2025). Early university responses show that transparency was operationalized through explicit syllabus communication and acknowledgment norms, indicating that visible disclosure helps reduce uncertainty and align expectations between institutions, teachers, and students (Dabis & Csáki, 2024; Francis et al., 2025). Algorithmic Accountability denotes the institutional obligation to retain human responsibility for AI-mediated judgments, recommendations, and assessment consequences. Research consistently argues that universities must establish clear standards, monitoring mechanisms, and human-in-the-loop oversight so that evaluative and moral agency is not delegated entirely to automated systems (García-López & Trujillo-Liñán, 2025; Dabis & Csáki, 2024). This pillar also implies contestability: students must be protected by procedures through which AI-supported outcomes can be questioned, reviewed, and corrected, especially in contexts affected by bias, misinformation, or unfair automated inference (García-López & Trujillo-Liñán, 2025; Tran et al., 2025). AI Literacy captures the competencies required to understand, interrogate, and ethically use GenAI rather than merely operate it. Evidence from large student samples shows that awareness and evaluative dimensions of AI literacy positively shape attitudes toward GenAI, while broader models find that literacy also strengthens perceived behavioral control and downstream usage intention (Sun et al., 2026; Wang et al., 2024; Zhang et al., 2025a). At the same time, the ethical dimension of literacy does not uniformly increase positive attitudes, suggesting that literacy can also sharpen critical caution rather than simple acceptance (Sun et al., 2026). This nuance is theoretically important because the present model is concerned with informed trust, not blind trust (Đerić et al., 2025). Adaptive AI Governance refers to an institution's capacity to revise policies,

redesign oversight mechanisms, and coordinate stakeholders as GenAI evolves. Multiple reviews argue that static or one-off regulations are inadequate because the risks, affordances, and social uses of GenAI change too quickly for fixed governance models (Sejdiu et al., 2026; Zlotnikova et al., 2025). Effective governance therefore requires continuous evaluation, stakeholder collaboration, curriculum integration, and flexible frameworks that can be updated across disciplines and assessment types (Furze et al., 2024; Zlotnikova et al., 2025). Evidence from assessment reform further shows that structured but adaptable models can reduce misconduct while improving engagement, reinforcing the value of agile rather than punitive governance (Furze et al., 2024).

2.3 Hypotheses Development

The proposed conceptual framework links these four governance pillars to Institutional Trust, which in turn drives Perceived Institutional Ethicality and Student Behavioral Intention toward sustainable AI use. We formally posit seven direct and indirect hypotheses below.

2.3.1. Ethical Transparency and Institutional Trust

Transparent communication of AI policies in course syllabi emerged as an early institutional good practice among top universities responding to GenAI, directly linking disclosure norms to trust-building (Dabis & Csáki, 2024). Perceived fairness and transparency in AI algorithms significantly influence trust and intention to adopt AI-based systems in educational institutions, positioning transparency as a direct antecedent rather than a peripheral concern (Sabri et al., 2026). Positive ethical perceptions of AI, including transparency and data privacy protection, strengthen the role of trust in shaping social acceptance of AI technologies among students (Sun et al., 2026). Higher education institutions should cultivate appropriate extrinsic trust by accurately communicating both learning opportunities and potential negative issues such as plagiarism and privacy concerns, with transparent communication equipping students with realistic expectations and informed judgment (Wang et al., 2025). Trust is further strengthened through open communication channels and transparent academic and administrative policies, though ethical values alone are insufficient to sustain digital trust without structural reinforcement (Aprianto et al., 2026). Thus, we propose **(H1)**: *Ethical Transparency significantly and positively influences Institutional Trust.*

2.3.2. Algorithmic Accountability and Institutional Trust

Institutional policies that retain human moral and legal responsibility for AI-related wrongdoings, combining top-down requirements with bottom-up instructor flexibility, represent an accountability mechanism that underpins trust (Dabis & Csáki, 2024). Algorithmic bias, data privacy concerns, and adherence to ethical standards directly influence ethical trust in educational AI systems, beyond technical performance alone (Sabri et al., 2026). Minimizing algorithmic bias and rigorously protecting user data are recommended to foster trust, with developers urged to take proactive measures to address ethical risks (Sun et al., 2026). Digital accountability functions as a component of institutional trust governance, though weak oversight mechanisms and concerns regarding academic data security persist

as challenges that can undermine accountability's trust-building potential (Aprianto et al., 2026). Accountability appears alongside transparency, human oversight, and inclusiveness as one of the key ethical dimensions that universities must address in their policy responses to GenAI (Dabis & Csáki, 2024). Accordingly, we propose **(H2)**: *Algorithmic Accountability significantly and positively influences Institutional Trust.*

2.3.3. AI Literacy and Institutional Trust

AI literacy is proposed as a positive antecedent of Institutional Trust because informed users are better able to evaluate AI outputs, interpret risks, and calibrate their expectations toward institutional AI deployment. Technical knowledge positively affects both trust in technology and academic performance, confirming that students with higher technical proficiency tend to trust GenAI more (Alotaibi, 2025). Personal perceptions exert the strongest influence on trust in technology among the studied determinants, surpassing technical knowledge and personal experience (Alotaibi, 2025). Trust in technology plays a significant mediating role between technical knowledge, personal experience, personal perceptions, and academic performance (Alotaibi, 2025). Ethical sensitivity within AI literacy does not necessarily produce more positive attitudes; rather, it may generate hesitation or uncertainty in certain contexts, suggesting that AI literacy is not a unidimensional construct (Sun et al., 2026). Valuing knowledge and fostering collaborative dynamics positively influence AI governance and an ethical institutional culture, which significantly impacts students' AI literacy and legal awareness. Thus, we propose **(H3)**: *AI Literacy significantly and positively influences Institutional Trust.*

2.3.4. Adaptive AI Governance and Institutional Trust

Institutional strategies, including leadership support, policies, and training, function as foundational drivers of trust in GenAI within higher education, positioning governance as a structural antecedent rather than a contextual moderator (Lelescu et al., 2025). AI governance must lead by embedding assessment literacy and integrity education into institutional policy, ensuring that ethical capacity grows with technological and cultural change (Maleki, 2026). AI governance should be educative, advisory, and flexible, empowering informed decision-making rather than imposing strict prohibitions, as guidelines will evolve with the technology. Cultivating integrity should not rest on surveillance or deterrence but on the co-creation of just and trustworthy systems, positioning adaptive governance as a trust-building mechanism (Maleki, 2026). Organizational readiness, defined as an institution's capacity to support AI adoption through infrastructure, governance, training, and data management, moderates the effects of data and knowledge processes on system outcomes (Sabri et al., 2026). Accordingly, we propose **(H4)**: *Adaptive AI Governance significantly and positively influences Institutional Trust.*

2.3.5. Institutional Trust and Perceived Institutional Ethicality

Embedding AI ethics into everyday academic practice enables universities to mitigate digital inequities and strengthen trust in emerging technologies, linking trust directly to perceived ethical practice (Maleki, 2026). Institutional cultures that normalize disclosure and responsible AI use cultivate perceived

ethicality, with the Higher Education Integrity Ecosystem Model positioning integrity as an emergent property of interacting institutional layers (Maleki, 2026). Sustainable academic integrity depends on recognizing ethics as a collective, renewable resource, with integrity serving as an indicator of sustainability itself, reflecting how institutions balance fairness, innovation, and accountability (Maleki, 2026). Ethical trust reflects users' perceptions of the fairness, transparency, and reliability of AI systems, with high ethical trust promoting adoption and efficient use while low trust leads to rejection of otherwise technically sound systems (Sabri et al., 2026). Inclusive leadership, open communication channels, transparent policies, and responsible digital governance collectively strengthen trust, conceptualizing ethical values as operational principles that guide institutional practice (Aprianto et al., 2026). Hence, we propose **(H5)**: *Institutional Trust significantly and positively influences Perceived Institutional Ethicality.*

2.3.6. Institutional Trust and Student Behavioral Intention

Trust predicts behavioral intention to continue using GenAI tools, with a moderate, statistically significant positive correlation suggesting that trust plays a central role in shaping users' willingness to integrate GenAI into academic workflows (Đerić et al., 2025). Fostering trust is not merely an ethical imperative but a practical necessity for promoting sustained and responsible use of GenAI technologies in higher education (Đerić et al., 2025). Trust in GenAI significantly influences university students' use of GenAI for learning, with its effect on behavioral intention to use mediated through reliance and resistance (Wang et al., 2025). Trust and social reinforcement strongly influence GenAI adoption decisions and act as significant mediators between perceived usefulness and willingness to adopt technology (Shata & Hartley, 2025). Trust is a significant positive predictor of behavioral intention to adopt GenAI, alongside motivation, academic integrity, and access conditions (Abdi & Abdullahi, 2026). Ethical anxiety is positively associated with responsible and continued use, indicating that moderate concern about ethical issues motivates compliance rather than avoidance, suggesting trust and ethical concern operate as parallel drivers of sustainable behavioral intention (Xia et al., 2026). Accordingly, **(H6)**: *Institutional Trust significantly and positively influences Student Behavioral Intention toward sustainable AI use.*

2.3.7. Perceived Institutional Ethicality as a Mediator

Attitude provides an evaluative filter that connects belief structures to behavioral intentions, which in turn are linked to GenAI usage, suggesting that perceived ethicality functions as a mediating evaluative mechanism between trust and intention (Sun et al., 2026). Ethical awareness significantly predicts behavioral intention, and an ethical mediation chain shapes students' behavioral intention, demonstrating that ethical perceptions mediate between cognitive appraisals and adoption outcomes (Qomariyah & Heriyanni, 2026). Performance expectancy, ethical awareness, habit, and AI ethical anxiety serve as significant predictors of behavioral intention, with ethical constructs operating alongside technological perceptions in a chained mediation pathway (Qomariyah & Heriyanni, 2026). The overall chain from perceptions through anxiety, trust, and risk to sustainable use intention is supported, with perceived ease of use, information quality, and ethical awareness primarily influencing sustainable use

intention indirectly through evaluative mediators (Xia et al., 2026). Academics' intentions to use GenAI technologies are contingent upon the development of robust ethical guidelines and supportive institutional policies, positioning perceived ethicality as a necessary mediating condition between institutional trust and behavioral intention (Nikolic et al., 2024). Responsible use of GenAI tools can enhance learning processes, but addressing concerns requires robust policies responsive to cultural expectations, suggesting that perceived institutional ethicality bridges trust and sustainable use by translating institutional commitments into culturally calibrated ethical signals (Yusuf et al., 2024). Therefore, we propose **(H7)**: *Perceived Institutional Ethicality significantly and positively mediates the relationship between Institutional Trust and Student Behavioral Intention.*

Figure 1 summerizes the hypothesized relationships between the variables included in the conceptual model of this research.

3. Research Methodology

3.1 Research Design and Sample Selection

This study adopts a rigorous quantitative, cross-sectional survey design to empirically test the hypothesized relationships. The target population consists of university students currently enrolled in public and private higher education institutions in Tunisia. A substantive eligibility criterion was applied: participants were required to be aged 18 years or older, currently enrolled, and to have prior experience using at least one generative AI or AI-based academic tool (e.g., ChatGPT, translation assistants, intelligent tutoring systems) for scholarly activities. This target group was selected because students who have never engaged with academic AI would lack the empirical basis necessary to evaluate their university's structural practices, AI policies, or their own intentions regarding sustainable and responsible AI utilization. To minimize sampling bias and promote heterogeneity across diverse academic contexts, we employed a purposive quota sampling technique. Purposive sampling ensured that every respondent was active and experienced with academic AI tools, while quota sampling was structured to enforce diversity across major educational disciplines (e.g., social sciences, engineering, business, medicine, law), study levels (bachelor's, master's, doctoral), and institutional affiliations. The final analytical sample consists of 380 valid responses, which represents a highly robust size. Using G*Power 3.1 and applying the standard '10-times rule' popular in PLS-SEM literature, a minimum sample size of 380 provides substantial statistical power (> 0.95) to detect small-to-moderate effects within our structural model. Table 2 documents the detailed sociodemographic profile of the analytical sample.

Table 2
The sociodemographic characteristics of the respondents

Characteristic	Category	Frequency (n)	Percentage (%)
Gender	Male	183	**48.16**
	Female	197	**51.84**
Age	18–24 years	140	**36.84**
	25–34 years	95	**25.00**
	35–44 years	85	**22.37**
	≥ 45 years	60	**15.79**
Academic level	Bachelor's/Undergraduate	66	**17.37**
	Master's	84	**22.11**
	Doctoral	174	**45.79**
	Other	56	**14.74**
Academic discipline	Business/Economics/Management	62	**16.32**
	Engineering/Computer Science	62	**16.32**
	Humanities/Languages	64	**16.84**
	Social Sciences	72	**18.95**
	Medicine/Health Sciences	48	**12.63**
	Law	40	**10.53**
	Other	32	**8.42**
AI-use frequency	Less than once a month	57	**15.00**
	1–3 times per month	63	**16.58**
	1–2 times per week	94	**24.74**
	3–5 times per week	94	**24.74**
	Daily/almost daily	72	**18.95**

3.2 Operationalization and Instrument Development

All latent constructs in our structural model were measured using multi-item scales derived from, or conceptually grounded in, validated research. Adaptations were strictly restricted to contextual changes (changing references from general AI platforms or general school settings to the specific university context) to ensure high content and semantic validity. Scale development followed standard procedures: newly adapted items (specifically Adaptive AI Governance and Perceived Institutional Ethicality)

underwent an intensive content validity screening with a panel of five academic experts in AI, higher education administration, and quantitative research. Every item was scored on a 4-point relevance scale to compute the Content Validity Index (CVI). Items failing to achieve a Content Validity Index (I-CVI) of 0.80 were revised or discarded. Following expert feedback, a pilot study was conducted with 30 students to verify comprehensibility and semantic flow prior to full-scale data collection. The variables are measured on a 5-point Likert-type scale, ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). Table 3 provides the operational definitions and primary sources for each construct.

Table 3
The variables, their items, and their sources

Construct	Operational Definition	Measurement Items	Primary Source
Ethical Transparency (ET)	Perceived transparency of the university regarding boundaries, data policies, and limitations of AI tools.	3 items focusing on clear explanations, technology/data disclosure, and communication of boundaries.	Yu et al. (2025)
Algorithmic Accountability (AA)	Student perception that the institution assumes ultimate responsibility for automated or AI-assisted decisions.	4 items tracking institutional responsibility, grievance systems, and proactive problem resolution.	Aysolmaz et al. (2023)
AI Literacy (AIL)	Students' core cognitive and ethical competencies to recognize, ethically use, and evaluate AI applications.	12 items from the validated AILS, measuring four distinct dimensions: Awareness, Use, Evaluation, and Ethics.	Wang et al. (2023)
Adaptive AI Governance (AIG)	The proactive capacity of the university to continuously monitor risks, revise policies, and involve stakeholders.	5 researcher-developed items, rigorously validated via content experts and pilot-tested.	Adaptive & AI Governance Literature
Perceived Institutional Ethicality (PIE)	Students' evaluative judgment of whether the university behaves consistently with moral principles in its AI actions.	5 researcher-developed items, tracking consistency, alignment of values, and student ethical concerns.	Organizational Ethics & Climate Literature
Institutional Trust (IT)	Students' structural confidence in the fairness, reliability, and goodwill of the university in managing AI.	5 items adapted to undergraduate/graduate levels, focusing on institutional fairness and reliance.	Troy et al. (2024); Yeager et al. (2014)
Student Behavioral Intention (BI)	Students' behavioral intention to utilize AI tools in a sustainable, ethical, and responsible academic manner.	4 items tracking intentions to use tools responsibly, ethically, and in a manner avoiding negative social outcomes.	AI Adoption & Technology Literature

3.3 Data Analysis Protocol

To test the hypothesized structural relationships, Partial Least Squares Structural Equation Modeling (PLS-SEM) was performed using SmartPLS 4. PLS-SEM is highly appropriate for this research because it is robust to non-normal distributions, supports complex structural paths with multiple mediators, and excels in exploratory and predictive scale validation. Data analysis proceeded via a standard two-stage

protocol: first, the measurement (reflective) model was rigorously assessed for reliability, convergent validity, and discriminant validity. Second, the structural model was evaluated to test the standardized path coefficients (beta), coefficients of determination (R^2), effect sizes (f^2), and predictive relevance (Q^2). Statistical significance of the paths was verified using a non-parametric bootstrapping procedure with 5,000 resamples.

4. Empirical Results and Data Analysis

4.1 Reflective Measurement Model Assessment

Reflective measurement properties were evaluated using indicator reliability (standardized outer loadings), internal consistency reliability, and convergent validity. Standardized outer loadings of the indicators exceeded the recommended 0.708 threshold, ranging from 0.724 to 0.922, indicating satisfactory indicator reliability. To assess internal consistency, Cronbach's alpha, Jöreskog's rho_A, and composite reliability (CR) were evaluated. All constructs comfortably surpassed the standard 0.70 threshold (ranging from 0.861 to 0.917 for Cronbach's alpha, and 0.901 to 0.936 for CR), demonstrating high reliability. Convergent validity was established via the Average Variance Extracted (AVE); all construct AVE values exceeded the 0.50 cutoff, ranging from 0.575 for AI Literacy to 0.798 for Ethical Transparency. These assessment parameters are summarized in Table 4.

Table 4
Measurement Model Results

Construct / Latent Variable	Item	Loading	Cronbach's α	rho_A	CR	AVE
Ethical Transparency (ET)	ET1: Clear explanation of AI use	0.824	0.873	0.878	0.922	0.798
	ET2: Technology and data disclosure	0.901				
	ET3: Communicates rules and limits	0.922				
Algorithmic Accountability (AA)	AA1: Institutional resolution of issues	0.801	0 .887	0.891	0.923	0.750
	AA2: Accountability for automated output	0.872				
	AA3: Clear administrative appeal pathways	0.889				
	AA4: Responsibility for decision outcomes	0.884				
AI Literacy (AIL)	AIL1: Technical distinction awareness	0.742	0.913	0.917	0.925	0.575
	AIL2: AI functional utility comprehension (R)	0.751				
	AIL3: Embedded AI application detection	0.783				
	AIL4: Work task efficiency optimization	0.794				
	AIL5: Learning difficulty resistance (R)	0.728				
	AIL6: Process improvement efficacy	0.811				
	AIL7: Capability and limitation critique	0.821				
	AIL8: Proper solution selection ability	0.776				
	AIL9: Solution comparison efficiency	0.803				
	AIL10: Ethical use compliance	0.764				

Construct / Latent Variable	Item	Loading	Cronbach's α	rho_A	CR	AVE
	AIL11: Privacy/security risk alertness (R)	0.735				
	AIL12: Algorithmic abuse prevention/alertness	0.752				
Adaptive AI Governance (AIG)	AIG1: Regular policy review capacity	0.819	0.894	0.898	0.922	0.704
	AIG2: Ethical/operational risk response	0.853				
	AIG3: Risk and outcomes monitoring	0.866				
	AIG4: Insufficient procedure revision	0.848				
	AIG5: Stakeholder feedback integration	0.809				
Perceived Institutional Ethicality (PIE)	PIE1: Ethical institutional conduct	0.837	0.906	0.910	0.930	0.727
	PIE2: Principled decision-making	0.871				
	PIE3: Ethical consequences responsibility	0.852				
	PIE4: Declared values-practice alignment	0.879				
	PIE5: Student concerns consideration	0.831				
Institutional Trust (IT)	IT1: Student best interests focus	0.854	0.911	0.915	0.933	0.738
	IT2: Fair treatment assurance	0.871				
	IT3: Decision confidence	0.867				
	IT4: Reliability on AI challenges	0.848				
	IT5: Reliable consideration of interests	0.844				
Behavioral Intention (BI)	BI1: Intention for responsible use	0.841	0.887	0.891	0.922	0.747

Construct / Latent Variable	Item	Loading	Cronbach's α	rho_A	CR	AVE
	BI2: Ethical academic use intention	0.881				
	BI3: Sustainability of AI practices	0.872				
	BI4: Negative consequence minimization	0.861				

4.2 Discriminant Validity Assessment

To ensure that all latent variables represent empirically distinct constructs, discriminant validity was assessed via the heterotrait-monotrait ratio (HTMT) and the Fornell-Larcker criterion. A conservative HTMT threshold of 0.85 was adopted. As shown in Table 5, all HTMT values remain strictly below 0.85. The highest value was observed between Institutional Trust and Perceived Institutional Ethicality (HTMT = 0.824), followed by Algorithmic Accountability and Perceived Institutional Ethicality (HTMT = 0.814). These results confirm that the constructs are empirically distinguishable.

Table 5
HTMT Matrix

Construct	ET	AA	AIL	AIG	PIE	IT	BI
ET	—						
AA	0.712	—					
AIL	0.438	0.451	—				
AIG	0.761	0.728	0.421	—			
PIE	0.793	0.814	0.397	0.803	—		
IT	0.734	0.776	0.447	0.771	0.824	—	
BI	0.521	0.563	0.508	0.587	0.741	0.783	—

Additionally, the Fornell-Larcker criterion was examined. As documented in Table 6, the square root of the AVE for each latent variable (represented along the bolded diagonal) exceeds the cross-correlations between that construct and all other variables in the model. This provides further statistical support for discriminant validity.

Table 6
Fornell-Larcker Criterion

Construct	ET	AA	AIL	AIG	PIE	IT	BI
ET	**0.893**						
AA	0.621	**0.866**					
AIL	0.412	0.391	**0.758**				
AIG	0.678	0.643	0.376	**0.839**			
PIE	0.701	0.712	0.359	0.701	**0.853**		
IT	0.661	0.681	0.398	0.673	0.744	**0.859**	
BI	0.492	0.513	0.481	0.528	0.679	0.721	**0.864**

4.3 Collinearity Assessment

To rule out potential multicollinearity among predictor variables, the inner Variance Inflation Factor (VIF) values were assessed. All inner VIF values fall comfortably below the conservative 3.30 threshold, ranging from 1.000 to 2.214 (as reported in Table 7). This confirms that collinearity is not a concern and does not distort the estimated path coefficients.

Table 7
Inner VIF Values

Endogenous Target Construct	Predictor Construct	Inner VIF
Institutional Trust (IT)	Ethical Transparency (ET)	1.824
	Algorithmic Accountability (AA)	1.917
	AI Literacy (AIL)	1.347
	Adaptive AI Governance (AIG)	2.063
Perceived Institutional Ethicality (PIE)	Institutional Trust (IT)	1.000
Behavioral Intention (BI)	Institutional Trust (IT)	2.214
	Perceived Institutional Ethicality (PIE)	2.214

4.4 Structural Model and Hypothesis Testing

The predictive power of the structural model was evaluated using the coefficients of determination (R^2 and Adjusted R^2). As displayed in Table 8, the predictors of Institutional Trust (ET, AA, AIL, and AIG) explain 58.4% of its variance, which represents a highly moderate-to-strong explanatory effect. Institutional Trust explains 55.0% of the variance in Perceived Institutional Ethicality. Finally, Institutional Trust and Perceived Institutional Ethicality jointly explain 61.2% of the variance in students' Behavioral

Intention toward Sustainable AI Use, which constitutes a substantial predictive effect in behavioral science literature.

Table 8
R^2 and Adjusted R

Endogenous Construct	R^2 Value	Adjusted R^2	Explanatory Effect
Institutional Trust (IT)	0.584	0.579	Moderate to Substantial
Perceived Institutional Ethicality (PIE)	0.550	0.549	Moderate
Behavioral Intention (BI)	0.612	0.610	Substantial

To assess the substantive importance of each predictor, effect sizes (f^2) were analyzed (Table 9). Results show that Adaptive AI Governance has the largest relative contribution to Institutional Trust (f^2 = 0.131) among the four antecedents, followed by Algorithmic Accountability (f^2 = 0.098) and Ethical Transparency (f^2 = 0.071). AI Literacy makes a smaller but significant contribution (f^2 = 0.032). Institutional Trust exerts an extraordinarily str Table 7. R^2 and Adjusted R ong effect on Perceived Institutional Ethicality (f^2 = 1.222), confirming its fundamental role. Perceived Institutional Ethicality has a medium-to-strong effect on Behavioral Intention (f^2 = 0.184), whereas Institutional Trust has a small direct effect on intention (f^2 = 0.119).

Table 9
Effect Sizes

Structural Path	f^2 Value	Assessment (Cohen's d)
Ethical Transparency (ET) → Institutional Trust (IT)	0.071	Small effect
Algorithmic Accountability (AA) → Institutional Trust (IT)	0.098	Small effect
AI Literacy (AIL) → Institutional Trust (IT)	0.032	Small effect
Adaptive AI Governance (AIG) → Institutional Trust (IT)	0.131	Small to Medium effect
Institutional Trust (IT) → Perceived Institutional Ethicality (PIE)	1.222	Extremely Large effect
Institutional Trust (IT) → Behavioral Intention (BI)	0.119	Small effect
Perceived Institutional Ethicality (PIE) → Behavioral Intention (BI)	0.184	Medium effect

Hypothesis testing was conducted via non-parametric bootstrapping with 5,000 resamples to estimate standard errors, t-statistics, and p-values. As detailed in Table 10, all six direct hypotheses (H1 to H6) received strong statistical support. Specifically, Ethical Transparency significantly and positively impacts Institutional Trust (β = 0.231, t = 4.125, $p < 0.001$), supporting H1. Algorithmic Accountability has a strong positive effect on trust (β = 0.258, t = 4.373, $p < 0.001$), supporting H2. AI Literacy has a significant, positive, though smaller, impact on trust (β = 0.137, t = 2.740, p = 0.006), supporting H3. Adaptive AI

Governance exhibits the strongest positive direct path to Institutional Trust (β = 0.297, t = 4.950, p < 0.001), supporting H4. Institutional Trust has an exceptionally strong positive effect on Perceived Institutional Ethicality (β = 0.741, t = 23.156, p < 0.001), supporting H5, and a direct positive impact on Student Behavioral Intention (β = 0.286, t = 4.400, p < 0.001), supporting H6.

Table 10
Structural Relationships and Hypothesis Testing

H	Relationship	Path Coeff (β)	Std Error (SE)	t-statistic	p-value	95% CI	Empirical Decision
H1	Ethical Transparency (ET) → Institutional Trust (IT)	0.231	0.056	4.125	< 0.001	[0.124, 0.342]	Supported
H2	Algorithmic Accountability (AA) → Institutional Trust (IT)	0.258	0.059	4.373	< 0.001	[0.145, 0.376]	Supported
H3	AI Literacy (AIL) → Institutional Trust (IT)	0.137	0.050	2.740	0.006	[0.040, 0.237]	Supported
H4	Adaptive AI Governance (AIG) → Institutional Trust (IT)	0.297	0.060	4.950	< 0.001	[0.180, 0.415]	Supported
H5	Institutional Trust (IT) → Perceived Institutional Ethicality (PIE)	0.741	0.032	23.156	< 0.001	[0.678, 0.801]	Supported
H6	Institutional Trust (IT) → Behavioral Intention (BI)	0.286	0.065	4.400	< 0.001	[0.159, 0.414]	Supported

4.5 Mediation and Indirect Effects Analysis

To investigate the mechanisms linking Institutional Trust to student behavior, we performed a mediation analysis (H7). We examined the indirect effect of Institutional Trust on students' Behavioral Intention toward Sustainable AI Use through the mediating pathway of Perceived Institutional Ethicality. As documented in Table 11, the bootstrapped indirect effect is positive and highly significant (β = 0.337, t = 6.358, p < 0.001), with a 95% bootstrap confidence interval [0.236, 0.445] that strictly excludes zero. Since the direct path from Institutional Trust to Behavioral Intention remains significant in the presence of the mediator (β = 0.286, t = 4.400, p < 0.001), Perceived Institutional Ethicality represents a partial mediator. This means that structural trust in the university directly drives students' sustainable intentions and also strongly works by enhancing students' evaluation of the university's moral integrity, which in turn acts as a powerful motivator of ethical, compliant behavioral intentions.

Table 11
Mediation Results

Relationship Type	Structural Path	Path Coeff (β)	Std Error (SE)	t-statistic	p-value	95% Bootstrap CI	Mediation Result
Direct Effect	Institutional Trust (IT) → Behavioral Intention (BI)	0.286	0.065	4.400	< 0.001	[0.159, 0.414]	Significant
Indirect Effect	IT → Perceived Institutional Ethicality (PIE) → BI	0.337	0.053	6.358	< 0.001	[0.236, 0.445]	Significant
Total Effect	Institutional Trust (IT) → Behavioral Intention (BI)	0.623	0.049	12.714	< 0.001	[0.529, 0.721]	Partial Mediation

4.6 Methodological Safeguards and Common Method Bias Controls

Because our study utilizes self-administered, cross-sectional online questionnaire data, potential Common Method Bias (CMB) was a key methodological concern. We implemented both proactive procedural controls (e.g., ensuring complete respondent anonymity, voluntary participation, neutral and clear scale wordings, and random counterbalancing of question sequences) and rigorous statistical checks. We conducted Harman's single-factor test, showing that the single largest factor accounted for only 31.7% of the total variance, which is well below the standard 50% threshold. Furthermore, we ran a full collinearity VIF assessment; as documented in Table 12, all full collinearity VIF values for the model constructs are strictly below 3.30, ranging from 1.35 to 2.24. Collectively, these diagnostics demonstrate that Common Method Bias does not represent a threat to the statistical validity of our findings.

Table 12
Common Method Bias Assessment

Diagnostic Test / Parameter	Empirical Result	Reference Threshold	Statistical Assessment
Harman's Single-Factor Test	31.7% variance explained by the first unrotated factor	First factor < 50.0% variance	Satisfactory (No single factor dominance)
Full Collinearity VIF - ET	1.82	VIF < 3.30	Satisfactory (No collinearity threat)
Full Collinearity VIF - AA	1.91	VIF < 3.30	Satisfactory
Full Collinearity VIF - AIL	1.35	VIF < 3.30	Satisfactory
Full Collinearity VIF - AIG	2.06	VIF < 3.30	Satisfactory
Full Collinearity VIF - PIE	2.17	VIF < 3.30	Satisfactory
Full Collinearity VIF - IT	2.24	VIF < 3.30	Satisfactory
Full Collinearity VIF - BI	2.11	VIF < 3.30	Satisfactory

4.7 Model Fit and Predictive Relevance

Finally, we evaluated the global model fit and predictive relevance. The Stone-Geisser Q^2 statistics were calculated to assess predictive relevance for the endogenous variables. All constructs achieved positive Q^2 values (0.412 for Institutional Trust, 0.398 for Perceived Institutional Ethicality, and 0.447 for Behavioral Intention), proving high predictive relevance. Additionally, the Standardized Root Mean Square Residual (SRMR) was estimated at 0.061. This value is well below the conservative threshold of 0.08, confirming an acceptable and highly satisfactory overall structural model fit.

5. Discussion and Comparative Analysis

5.1 Discussion and Theoretical Alignment

The empirical validation of our integrated governance model provides robust, novel insights into the structural dynamics of digital higher education. By shifting the unit of analysis from individual-level technical attributes (e.g., perceived ease of use or technical utility) to systemic, institutional responsibility, the study demonstrates that rebuilding academic trust in the era of GenAI is fundamentally an organizational and moral challenge rather than a purely technological or pedagogical problem (Al Zaidy, 2024; Kringen et al., 2026). Our structural finding that Adaptive AI Governance is the most powerful antecedent of Institutional Trust ($\beta = 0.297$) strongly supports the conceptual

propositions of Kringen et al. (2026). Kringen et al. (2026) argued that the widespread adoption-trust paradox (where stakeholders actively use AI tools despite feeling intense skepticism and a lack of trust) is primarily an "institutional governance failure" rather than a failure of student ethics. Our empirical results validate this: when universities actively and dynamically review, monitor, and adapt their administrative policies in response to technological change, student trust increases. This also aligns with Shrestha (2026), who highlighted the importance of establishing integrated, agile regulatory architectures in higher education through the REM (Responsible, Ethical, and Managed) framework. This interpretation fits the broader literature, which increasingly treats AI ethics in education as a systemic governance problem involving accountability, transparency, autonomy, inclusion, and human oversight rather than a narrow question of tool performance (Crompton et al., 2026; Holmes et al., 2022). It also aligns with evidence that institutional responses to GenAI remain uneven, reactive, and often underdeveloped, which makes adaptive governance a plausible source of trust variation across universities (Crompton et al., 2026; Knox & Knox, 2025). Additionally, the significant direct path from Ethical Transparency to Institutional Trust ($\beta = 0.231$) extends the findings of Yu et al. (2025). While Yu et al. (2025) established that user-level transparency regarding data usage and algorithmic boundaries is essential for the individual adoption of AI-generated content (AIGC) platforms, our model demonstrates that this effect holds true at the institutional level: clear administrative disclosure regarding the boundaries and rules of AI use directly builds student trust in the university as a whole. This result deepens a strong cross-paper consensus that trust depends on intelligible disclosure of how AI works, what data it uses, where its limits lie, and how institutions expect it to be used (Dabis & Csáki, 2024; Khosravi et al., 2022; Coman et al., 2026). Experimental evidence is especially useful here: transparency appears to be the most influential predictor of student trust in AI-based assessment, and its effect is stronger when paired with procedural control or user agency (Garcia, 2026). This means our institutional-level transparency pathway is not merely symbolic; it reflects students' demand for procedural legitimacy. Furthermore, the robust effect of Algorithmic Accountability on Institutional Trust ($\beta = 0.258$) empirically confirms the qualitative assertions of Aysolmaz et al. (2023), proving that students require clear structural assurance that their university accepts responsibility for AI-driven outcomes and provides transparent pathways for academic grievance and appeal. Reviews of higher education AI ethics show that accountability remains less operationalized than fairness, even though scholars identify it as essential for oversight, explainability, and safety (Memarian & Doleck, 2023; Coghlan et al., 2021). Empirical evidence from universities also shows that faculty and students express major concern about weak transparency, privacy, and oversight, and call for audit protocols, ethical oversight structures, and concrete regulatory mechanisms (Marín et al., 2025). In that context, our accountability coefficient suggests that trust rises when students believe the university can be held answerable for AI-mediated harms, errors, and appeals. Interestingly, although the direct path from AI Literacy to Institutional Trust was significant ($\beta = 0.137$, $p = 0.006$), it was the weakest of the four antecedents. This represents an important theoretical nuance: while fostering student-level technical competence and critical thinking is indispensable (Wang et al., 2023; Zhang et al., 2025), technical competence alone cannot substitute for a robust, transparent, and accountable institutional framework. Student trust is predominantly driven by structural, top-down governance safeguards rather than individual-level capabilities. This result

highlights that individual competence must operate within a wider 'trust infrastructure' built by the institution (Bhaskar et al., 2026; Troy et al., 2024). This is consistent with adoption studies showing that AI literacy improves attitudes, perceived control, and behavioral intention, but does not by itself resolve ethical concern or establish institutional legitimacy (Sun et al., 2026; Wang et al., 2024; Zhang et al., 2025). Students with stronger awareness and evaluation literacy tend to develop more positive orientations toward GenAI, yet privacy risk and ethical concern still depress trust unless institutional protections are in place (Sun et al., 2026; Zhang et al., 2025). Our model therefore advances the literature by showing that literacy functions best as an enabling capability within a broader trust infrastructure, not as a substitute for governance. Finally, the partial mediation of Perceived Institutional Ethicality (β = 0.337) between Institutional Trust and Behavioral Intention (H7) adds deep theoretical value. It shows that institutional trust is not merely a passive psychological state. Instead, it actively shapes students' cognitive evaluation of the university's moral standing (perceived ethicality), which then serves as a primary, active driver of their own responsible and sustainable behavior. This path supports the values-led organizational change model of Stout (2023), confirming that values-driven leadership in technology implementation fosters a self-reinforcing culture of compliance and academic integrity among students.

5.2 Practical Implications for Marketers and Senior Administrators

In the highly competitive global landscape of digital higher education, academic trust and ethical reputation are increasingly recognized as primary brand assets. As digital learning environments become ubiquitous, institutions can no longer differentiate themselves solely based on technical delivery or program diversity. Our findings suggest that institutional trust is a strong determinant of student behavioral loyalty and brand equity. Higher education marketers and university leaders must strategically position and brand their institutions as "ethical AI campuses" to attract and retain high-caliber students. This supports reframing "ethical AI campus" positioning away from branding rhetoric alone and toward auditable governance signals such as review committees, clear appeals pathways, transparent syllabus-level guidance, and regular policy updates (Dabis & Csáki, 2024; Crompton et al., 2026). In competitive higher education markets, trust becomes credible only when ethical commitments are institutionally enacted. Marketing campaigns and brand messaging should proactively highlight the institution's commitment to Ethical Transparency and Algorithmic Accountability. Rather than downplaying or ignoring AI integration, marketing materials should transparently outline the university's robust 'AI Trust Infrastructure', such as the active existence of an AI Ethics Governance Committee (Bhaskar et al., 2026) and values-led data privacy policies (Stout, 2023). Positioning the university as a leading ethical actor in the digital transition serves to enhance brand differentiation, elevate institutional legitimacy, and attract ethically minded students, corporate partners, and research funding. For marketers and senior administrators, the implication is not simply to publicize AI use, but to communicate procedural legitimacy. Higher education faces socio-political pressure to use AI for quality assurance, retention, satisfaction, and reputation, yet these same pressures can generate distrust when decisions appear arbitrary, opaque, or externally driven (Khosravi et al., 2022). Transparent messaging should therefore specify what AI is used for, what it is not used for, what data protections exist, how human oversight is

retained, and how students can challenge decisions (Dabis & Csáki, 2024; Coghlan et al., 2021; Zhang et al., 2025). Trustworthy positioning is strongest when it reduces uncertainty rather than when it celebrates innovation abstractly.

Table 13
outlines evidence-based priorities and actionable translations for university leaders seeking to establish a robust trust infrastructure.

Priority	Evidence-Based Rationale	Institutional Translation
Adaptive review	GenAI policies should be iterative and regularly updated to keep pace with technology (Crompton et al., 2026).	Establish a standing AI committee with scheduled quarterly review cycles.
Visible transparency	Trust rises when systems and institutions explain rules, data policies, and boundaries clearly (Coman et al., 2026; Garcia, 2026).	Mandate syllabus-level AI disclosure, public policies, and student consent dashboards.
Accountable oversight	Students expect contestable, answerable AI governance and appeal processes (Marín et al., 2025; Coghlan et al., 2021).	Define clear lines of responsibility, implement regular audits, and establish formal appeal pathways.
Curricular literacy	AI literacy improves attitudes and responsible use, but requires structured institutional scaffolding (Wang et al., 2024; Zhang et al., 2025).	Integrate tiered AI literacy and ethical modules across all undergraduate and graduate programs.
Assessment redesign	Detection-first and online proctoring strategies are inadequate, unreliable, and often harmful (Leaton Gray et al., 2025).	Transition to authentic, AI-resilient assessments (e.g., oral defenses, process portfolios, reflective journals).

5.3 Pedagogical Implications for Educators

For educators, the empirical support for the four pillars indicates that traditional, reactive, and punitive approaches to AI management (such as blanket bans or reliance on unreliable AI-detection software) are counterproductive. Multiple studies argue that detection systems and online proctoring are insufficient or counterproductive as standalone responses because they are unreliable, deepen alienation, and can damage pedagogical relationships and student subjectivity (Leaton Gray et al., 2025; Lee & Fanguy, 2022). Early institutional policy studies reach a similar conclusion by recommending assessment redesign and dialogue-based sanctions rather than a narrow misconduct-detection logic (Dabis & Csáki, 2024). This gives strong support to the argument for AI-resilient assessment. Educators must transition toward 'AI-resilient' assessment designs and proactive digital pedagogy (Jain & Kiran, 2025; Sullivan et al., 2023). This requires redesigning coursework to focus on higher-order critical thinking, reflective writing, and real-time oral examinations—cognitive processes that machines cannot easily replicate. Furthermore, our results show that embedding AI Literacy directly into the curriculum is a critical requirement. Educators must actively train students in the core competencies of the Artificial Intelligence Literacy Scale (AILS), focusing heavily on the 'Ethics' and 'Evaluation' dimensions (Wang et al., 2023). Coursework should require students to critically evaluate AI-generated outputs, identify

algorithmic biases, and explicitly document and reference their AI usage (Yu et al., 2025). By treating AI as a collaborative learning assistant rather than a threat, educators can foster a sustainable learning environment that preserves the intrinsic value of human scholarship.

5.4 Boundary Conditions: Surveillance, Equity, and Student Agency

The systematic governance of GenAI in higher education is inextricably linked to broader, complex societal ethics. A primary ethical risk is the escalation of digital surveillance: in their reactive attempts to preserve traditional academic integrity, many universities have implemented intrusive algorithmic proctoring and automated plagiarism detection tools. Surveillance-heavy governance tends to erode trust, privacy, and educational relationships (Coghlan et al., 2021; Lee & Fanguy, 2022). Research on proctoring shows that surveillance-based systems can undermine privacy, autonomy, fairness, and trust, while also redistributing power toward institutions and vendors in ways that students experience as coercive (Henry & Oliver, 2022; Coghlan et al., 2021; Coman et al., 2026). Other work argues that framing ethics only through abstract principles is inadequate because AI ethics in education is enacted through socio-technical arrangements and unequal distributions of power, care, and responsibility (Henry & Oliver, 2022). This strengthens our claim that governance must be morally substantive, not merely administratively present. Furthermore, the digital divide remains a profound social equity concern in digital higher education. As advanced, highly capable GenAI platforms transition to premium, subscription-only business models, a severe risk of cognitive inequality emerges. Wealthy students who can afford premium cognitive tools gain access to sophisticated learning support, writing scaffolding, and research assistance, while lower-income students are restricted to standard, less capable systems. Sustainable learning ecosystems must establish equitable access policies, ensuring that access to vital academic tools is democratized and does not exacerbate existing social and academic inequalities (Pechenkina, 2023; Thaldar et al., 2025). Reviews and conceptual analyses warn that AI adoption can amplify linguistic, socioeconomic, and infrastructural inequalities unless institutions deliberately design inclusive access and support mechanisms (Knox & Knox, 2025; Crompton et al., 2026; Coman et al., 2026). Human-centered models therefore emphasize equitable access, privacy, interoperability, and critical thinking as design requirements rather than optional benefits (Airaj, 2024). Finally, the preservation of student agency is paramount. Effective AI governance must structurally guarantee that AI remains a supportive learning 'mate' or cognitive assistant, while the student remains the active intellectual 'driver' of their educational journey (Mahama & Amadu, 2025). Legitimate GenAI governance must protect student agency by ensuring AI remains assistive, contestable, and pedagogically subordinate to human judgment (Khosravi et al., 2022; Leaton Gray et al., 2025). Universities must design learning frameworks that protect and celebrate the intrinsic value of human intellectual effort, creative struggle, and critical thought, ensuring that technology serves to amplify, rather than replace, human intelligence.

5.5 Limitations and Directions for Future Research

While this study provides a highly rigorous, empirically validated framework, several limitations must be acknowledged. First, because the data were collected through a cross-sectional survey design, we cannot establish definitive temporal or causal relationships among our constructs. Future research should utilize longitudinal designs to track how student trust and behavioral compliance evolve over time as AI tools become deeply embedded within the fabric of academic workflows. Second, our sample was limited to university students in Tunisia. While this provides a highly homogeneous and contextually rich sample, cultural, economic, and institutional differences may affect the generalizability of our findings. Comparative cross-national research (e.g., comparing highly regulated settings under the EU AI Act with less centralized regulatory environments like the US or Africa) would provide valuable insights into the performance of different policy architectures (Hristova, 2025; Li et al., 2026; Thaldar et al., 2025). The broader literature repeatedly notes that empirical work remains thinner than normative discussion, especially for transparency, accountability, and lived institutional practice (Marín et al., 2025; Memarian & Doleck, 2023). Comparative studies across regulatory environments, institutional types, and disciplines would test whether our trust architecture is stable or context-dependent, while mixed-method and experimental work could examine how policy signals, transparency interfaces, and grievance mechanisms alter trust over time (Crompton et al., 2026; Garcia, 2026).

6. Conclusion and Contributions

The central contribution of this study is to demonstrate that the current crisis of academic trust in the age of Generative AI is, at its core, a governance challenge rather than a narrowly technical problem of tool accuracy or a narrowly pedagogical problem of classroom control (Sejdiu et al., 2026; Crompton et al., 2026; Coman et al., 2026). By validating a four-pillar model built on Ethical Transparency, Algorithmic Accountability, AI Literacy, and Adaptive AI Governance, this study moves the debate beyond abstract ethical principles and identifies the specific institutional mechanisms through which trust can be produced, sustained, and translated into responsible student behavior (Holmes et al., 2022; Sejdiu et al., 2026; Crompton et al., 2026). More specifically, our findings demonstrate that active, adaptive governance and structural accountability are the primary drivers of student trust and perceived institutional ethicality, rather than student competence alone. This is a significant theoretical advance because much of the existing literature calls for transparency, accountability, literacy, and human oversight, but fewer studies empirically demonstrate how these governance dimensions work together as an integrated trust architecture in higher education (Crompton et al., 2026; Memarian & Doleck, 2023; Marín et al., 2025). In that sense, the study helps close an important gap between normative frameworks and empirical validation by showing that trust is not an automatic byproduct of AI adoption, but an institutional achievement rooted in visible rules, review capacity, answerability, and ethical consistency (Knox & Knox, 2025; Khosravi et al., 2022; Marín et al., 2025). The study also makes a practical contribution by showing that sustainable AI integration cannot rely on punitive or surveillance-centered responses alone. Evidence across higher education shows that detection-first and proctoring-heavy approaches can weaken trust, intensify perceptions of unfairness, and damage pedagogical relationships, whereas more legitimate responses emphasize human oversight, authentic assessment,

student agency, and clear procedural communication (Coghlan et al., 2021; Lee & Fanguy, 2022; Leaton Gray et al., 2025). Against that backdrop, this study offers universities an empirically grounded alternative: invest in a trust infrastructure that combines adaptive policy review, transparent communication, accountable governance, and curriculum-embedded AI literacy so that innovation remains aligned with academic integrity, equity, and human-centered educational values (Crompton et al., 2026; Dabis & Csáki, 2024; Airaj, 2024). Ultimately, the broader importance of this study lies in its reframing of what institutional readiness for GenAI actually means. Readiness is not simply access to tools or willingness to innovate; it is the university's capacity to govern AI in ways that students perceive as legitimate, ethical, contestable, and educationally meaningful (Sejdiu et al., 2026; Crompton et al., 2026; Garcia, 2026). By empirically showing that trust and perceived institutional ethicality are key pathways through which governance shapes responsible behavioral intention, this study contributes a timely and policy-relevant model for protecting the credibility of higher education in an AI-saturated era (Sejdiu et al., 2026; Crompton et al., 2026; Sun et al., 2026).

Table 14 summarizes the core theoretical, methodological, and practical contributions of this research.

Contribution Area	Details and Why It Matters
Shifts the unit of analysis	Moves the academic AI trust discourse away from individual-level adoption models (e.g., TAM, UTAUT) toward systemic, institutional governance (Crompton et al., 2026).
Provides empirical validation	Addresses a critical literature gap by providing robust quantitative PLS-SEM validation of a multi-dimensional AI governance framework using a sample of 380 active students (Memarian & Doleck, 2023; Marín et al., 2025).
Identifies trust mechanisms	Empirically demonstrates that ethical transparency, algorithmic accountability, and adaptive policy reviews function interactively to produce structural trust, which is then filtered through students' moral evaluations (Crompton et al., 2026; Coman et al., 2026).
Offers policy relevance	Provides academic leaders and university marketers with a concrete blueprint for building a 'trust infrastructure' as an alternative to reactive, surveillance-led, or punitive governance architectures (Crompton et al., 2026; Leaton Gray et al., 2025).

Declarations

Author Contribution

The author was solely responsible for the conceptualization, methodology, data collection, data analysis, interpretation of results, writing, revision, and final approval of the manuscript. I have conducted the analysis the literature review and all the work

References

1. Abdi, A. S., & Abdullahi, A. M. (2026). Trust, academic integrity, and access constraints in generative AI adoption: an extended TAM study of Somali higher education. *Discover Computing*, *29*(1), 375.
2. Airaj, M. (2024). Ethical artificial intelligence for teaching-learning in higher education. *Education and Information Technologies*, *29*(13), 17145–17167.
3. Al Zaidy, A. (2024). The impact of generative AI on student engagement and ethics in higher education. *Journal of Information Technology Cybersecurity and Artificial Intelligence*, *1*(1), 30–38.
4. Al-Omari, O., Alyousef, A., Fati, S., Shannaq, F., & Omari, A. (2025). Governance and ethical frameworks for AI integration in higher education: Enhancing personalized learning and legal compliance. *Journal of Ecohumanism*, *4*(2), 80–86.
5. Alotaibi, S. M. F. (2025). Determinants of Generative Artificial Intelligence (GenAI) adoption among university students and its impact on academic performance: the mediating role of trust in technology. *Interactive Learning Environments*, *33*(6), 4159–4188.
6. Alyafei, R. A. S. M., Al-Kfairy, M., Alfandi, O., & Alrabaee, S. (2026). Safeguarding students' academic integrity in the age of generative AI: challenges and solutions. *Kybernetes*, 1–19.
7. Aprianto, I., Mawardi, M., Suparno, S., Batanghari, U. I., Islam, U., Sultan, N., & Saifuddin, T. (2026). Amanah-Based Public Relations and Digital Trust in Islamic Private Universities. *J Ilm Peuradeun*, *14*(1), 343–370.
8. Aysolmaz, B., Müller, R., & Meacham, D. (2023). The public perceptions of algorithmic decision-making systems: Results from a large-scale survey. *Telematics and Informatics*, *79*, 101954.
9. Baig, M. I., & Yadegaridehkordi, E. (2025). Factors influencing academic staff satisfaction and continuous usage of generative artificial intelligence (GenAI) in higher education. *International Journal of Educational Technology in Higher Education*, *22*(1), 5.
10. Barus, O. P., Hidayanto, A. N., Handri, E. Y., Sensuse, D. I., & Yaiprasert, C. (2025). Shaping generative AI governance in higher education: Insights from student perception. *International Journal of Educational Research Open*, *8*, 100452.
11. Bhaskar, P., Khan, M., & Bhaskar, P. (2026). Establishing an AI ethics governance committee in higher education: A theoretical framework. *Navigating modern digital communication ethics and law* (pp. 105–124). IGI Global Scientific Publishing.
12. Coghlan, S., Miller, T., & Paterson, J. (2021). Good proctor or big brother? Ethics of online exam supervision technologies. *Philosophy & Technology*, *34*(4), 1581–1606.
13. Coman, C., Gherheș, V., Bucs, A., Coman, E., Victoria-Rodica, C., Bódi, D. C., & Șerbănescu, R. M. (2026). Navigating opportunities and challenges of generative AI in higher education: towards responsible, equitable, and human-centered integration. *Frontiers in Artificial Intelligence*, *9*, 1750978.
14. Crompton, H., Burke, D., Nickel, C., Bozkurt, A., Miao, F., Sharples, M., & Yu, S. (2026). Governing generative AI in higher education: A global Delphi study on policy and practice. *International Journal of Educational Technology in Higher Education*, *23*(1), 21.
15. Dabis, A., & Csáki, C. (2024). AI and ethics: Investigating the first policy responses of higher education institutions to the challenge of generative AI. *Humanities and Social Sciences*

Communications, *11*(1), 1006.

16. Đerić, E., Frank, D., & Milković, M. (2025). Trust in generative AI tools: A comparative study of higher education students, teachers, and researchers. *Information*, *16*(7), 622.
17. Francis, N. J., Jones, S., & Smith, D. P. (2025). Generative AI in higher education: Balancing innovation and integrity. *British Journal of Biomedical Science*, *81*, 14048.
18. Furze, L., Perkins, M., Roe, J., & MacVaugh, J. (2024). The AI Assessment Scale (AIAS) in action: A pilot implementation of GenAI-supported assessment. *Australasian Journal of Educational Technology*, *40*(4), 38–55.
19. Garcia, M. B. (2026). Transparency, Ethical Framing, and User Agency as Determinants of Trust in AI-Mediated Assessment: Informing the Design of Trustworthy Systems. *Evaluation Review*, 0193841X261451124.
20. García-López, I. M., & Trujillo-Liñán, L. (2025, June). Ethical and regulatory challenges of Generative AI in education: a systematic review. *Frontiers in Education* (Vol. 10, p. 1565938). Frontiers Media SA.
21. Henry, J. V., & Oliver, M. (2022). Who will watch the watchmen? The ethico-political arrangements of algorithmic proctoring for academic integrity. *Postdigital Science and Education*, *4*(2), 330–353.
22. Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., & Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. *International journal of artificial intelligence in education*, *32*(3), 504–526.
23. Hristova, M. (2025). Generative Artificial Intelligence and academic practices: A comparative analysis of approaches in Europe, the United States and China. *Postmodernism Problems*, *15*(2), 245–269.
24. Jain, N., & Kiran, M. (2025). Rethinking education in the age of generative AI: cognitive offloading, assessment reform, and institutional adaptation.
25. Jayasinghe, S., Gamage, K. A., Yang, D., Cheng, C., Disanayake, C., & Apeji, U. D. (2026). Six institutional intervention areas to support ethical and effective student use of generative AI in higher education: A narrative review. *Education Sciences*, *16*(1), 137.
26. Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y. S., Kay, J., & Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and education: artificial intelligence*, *3*, 100074.
27. Knox, M. P., & Knox, A. W. (2025). Artificial Intelligence in Higher Education: Ethical Challenges, Governance Frameworks, and Student-Centered Pathways. *J Inform Techn Int*, *3*(2), 112.
28. Kofinas, A. K., Tsay, C. H. H., & Pike, D. (2025). The impact of generative AI on academic integrity of authentic assessments within a higher education context. *British Journal of Educational Technology*, *56*(6), 2522–2549.
29. Kringen, P., Gallucci, A., Hildt, E., Johansen, F. R., Westerlund, M., & Tigerstedt, C. (2026). Using LLMs without trust: the adoption–trust paradox as institutional governance failure in academic ecosystems. *AI & SOCIETY*, 1–17.

30. Leaton Gray, S., Edsall, D., & Parapadakis, D. (2025). AI-based digital cheating at university, and the case for new ethical pedagogies. *Journal of Academic Ethics*, *23*(4), 2069–2086.
31. Lee, K., & Fanguy, M. (2022). Online exam proctoring technologies: Educational innovation or deterioration? *British Journal of Educational Technology*, *53*(3), 475–490.
32. Lelescu, A., Sava, S., Grosseck, G., & Malita, L. (2025). Exploring trust in generative AI for higher education institutions: A systematic literature review focused on educators. *Humanities and social sciences communications*, *12*(1), 1961.
33. Li, M., Xie, Q., Enkhtur, A., Meng, S., Chen, L., Yamamoto, B. A., & Murakami, M. (2026). A framework for developing university policies on generative AI governance: A cross-national comparative study. *Studies in Higher Education*, 1–29.
34. Mahama, I., & Amadu, P. (2025). You are the Driver and AI is the mate: Exploring human-led creative and critical thinking in AI-augmented learning environments. *F1000Research*, *14*, 974.
35. Maleki, A. (2026). Rethinking ethical academic integrity ecosystems in higher education in the age of artificial intelligence. *Discover Artificial Intelligence*, *6*(1), 145.
36. Marín, Y. R., Caro, O. C., Rituay, A. M. C., Llanos, K. A. G., Perez, D. T., Bardales, E. S., & Santos, R. C. (2025). Ethical challenges associated with the use of artificial intelligence in university education. *Journal of Academic Ethics*, *23*(4), 2443–2467.
37. Memarian, B., & Doleck, T. (2023). Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI) and higher education: A systematic review. *Computers and Education: Artificial Intelligence*, *5*, 100152.
38. Nikolic, S., Wentworth, I., Sheridan, L., Moss, S., Duursma, E., Jones, R. A., & Middleton, R. (2024). A systematic literature review of attitudes, intentions and behaviours of teaching academics pertaining to AI and generative AI (GenAI) in higher education: An analysis of GenAI adoption using the UTAUT framework. *Australasian Journal of Educational Technology*, *40*(6), 56–75.
39. Pechenkina, K. (2023). Artificial intelligence for good? Challenges and possibilities of AI in higher education from a data justice perspective. *Higher Education for good: Teaching and learning futures (# HE4Good). Cambridge*. Open Book.
40. Qomariyah, N. N., & Heriyanni, E. (2026). Ethical readiness and behavioral intention toward generative AI: student perspectives from STEM and non-STEM disciplines in Indonesian higher education. *Journal of Information Communication and Ethics in Society*, 1–25.
41. Sabri, O., Al-Shargabi, B., & Atti, O. N. (2026). DATA QUALITY AND THE TRUST–INTERPRETABILITY PARADOX: TOWARD A MID-RANGE FRAMEWORK FOR AI ADOPTION IN EDUCATIONAL INFORMATION SYSTEMS. *Journal of Information Technology Education: Research*, *25*, 1.
42. Saddam, S., Winata, A., Saoki, M., & Putri, F. E. (2026). Nurturing Social Trust through Students' Digital Communication Ethics: Sustainable Strengthening Based on Digital Culture. *IJTIMAIYA: Journal of Social Science and Teaching*, *10*(1), 119–138.
43. Sejdiu, N. P., Grassini, S., & Sejdiu, S. (2026). Generative AI in education: Process-aware pedagogy, assessment integrity, and institutional governance. *Open Research Europe*, *6*, 161.

44. Shata, A., & Hartley, K. (2025). Artificial intelligence and communication technologies in academia: faculty perceptions and the adoption of generative AI. *International Journal of Educational Technology in Higher Education*, *22*(1), 14.
45. Shrestha, Y. (2026). Responsible Governance of Generative AI in Higher Education: Integrating Ethics, Policy, and the REM Framework. *Apex Journal of Business and Management*, *5*(1), 103–113.
46. Stout, B. (2023). A Values-led Approach to Technological Change. *International Journal of Social Sciences and Artistic Innovations*, https://doi.org/10.35745/*ijssai2024v04, 2*.
47. Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning & Teaching*, *6*(1), 31–40.
48. Sun, Y., Li, L., Zhang, X., Chen, X., Deng, S., & Hu, X. (2026). AI literacy and psychosocial factors shaping Chinese university students' attitudes and behavioral intentions toward generative AI use. *BMC psychology*, *14*(1), 236.
49. Thaldar, D., Botes, M., Badru, A., Chenia, H., Duma, S., Dlamini, S. B., et al. (2025). Generative AI governance in higher education: A case study from Africa. *Frontiers in Political Science*, *7*, 1666661.
50. Troy, A., Moua, H., & Van Boekel, M. (2024). Wise feedback and trust in higher education: A quantitative and qualitative exploration of undergraduate students' experiences with critical feedback. *Psychology in the Schools*, *61*(6), 2424–2447.
51. Wang, B., Rau, P. L. P., & Yuan, T. (2023). Measuring user competence in using artificial intelligence: validity and reliability of artificial intelligence literacy scale. *Behaviour & information technology*, *42*(9), 1324–1337.
52. Wang, C., Wang, H., Li, Y., Dai, J., Gu, X., & Yu, T. (2025). Factors influencing university students' behavioral intention to use generative artificial intelligence: Integrating the theory of planned behavior and AI literacy. *International Journal of Human–Computer Interaction*, *41*(11), 6649–6671.
53. Wang, Q., Anyi, C. L., Swain, V. D., & Goel, A. K. (2024). Navigating AI Fallibility: Examining People's Reactions and Perceptions of AI after Encountering Personality Misrepresentations. *arXiv preprint arXiv:2405.16355*.
54. Xia, B., Lei, Y., Hu, Y., Zhu, X., & Zhang, J. (2026). Sustainable use intention of text-to-image generative AI in higher education: an S–O–R model with parallel trust and risk pathways. *Sustainability*, *18*(3), 1657.
55. Yu, T., Tian, Y., Chen, Y., Huang, Y., Pan, Y., & Jang, W. (2025). How do ethical factors affect user trust and adoption intentions of AI-generated content tools? Evidence from a risk-trust perspective. *Systems*, *13*(6), 461.
56. Yusuf, A., Pervin, N., & Román-González, M. (2024). Generative AI and the future of higher education: A threat to academic integrity or reformation? Evidence from multicultural perspectives. *International Journal of Educational Technology in Higher Education*, *21*(1), 21.
57. Zhang, A., Gao, Y., Suraworachet, W., Nazaretsky, T., & Cukurova, M. (2025). Evaluating trust in AI, human, and co-produced feedback among undergraduate students. *arXiv preprint arXiv:2504.10961*.

58. Zhang, X., Hu, X., Sun, Y., Li, L., Deng, S., & Chen, X. (2025a). Integrating AI literacy with the TPB-TAM framework to explore Chinese university students' adoption of generative AI. *Behavioral Sciences*, *15*(10), 1398.
59. Zlotnikova, I., Hlomani, H., Mokgetse, T., & Bagai, K. (2025). Establishing ethical standards for GenAI in university education: a roadmap for academic integrity and fairness. *Journal of Information Communication and Ethics in Society*, *23*(2), 188–216.

Figures

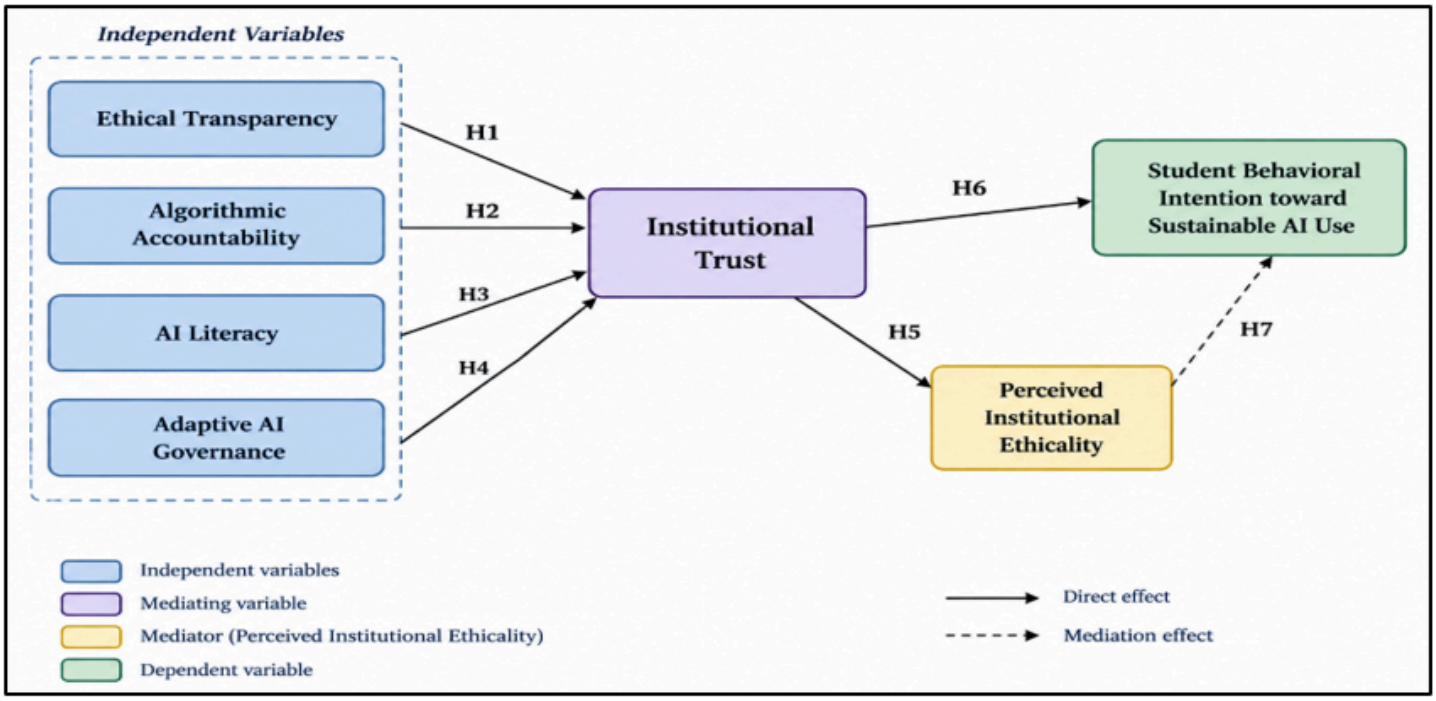


Figure 1

Conceptual framework