Journal of Advanced Sciences and Mathematics Education
Volume 6, Issue 2, 686-700.
e_ISSN: 2798-1606
DOI: 10.58524/jasme.v6i2.1245

AI-Driven assessment of heritage competencies in balinese tourism vet schools: Resolving the tension between digital acceleration and cultural conservation

Gusti Ayu Dessy Sugiharni* Institut Pariwisata dan Bisnis Internasional, INDONESIA
Putu Herny Susanti Institut Pariwisata dan Bisnis Internasional, INDONESIA
Putu Sabda Jayendra Institut Pariwisata dan Bisnis Internasional, INDONESIA
Anak Agung Istri Putera Widiastiti Institut Pariwisata dan Bisnis Internasional, INDONESIA
Firlie Lanovia Amir Institut Pariwisata dan Bisnis Internasional, INDONESIA

Article Info

Article history:

Received: April 08, 2026
Revised: May 17, 2026
Accepted: June 17, 2026

Keywords:

Artificial Intelligence; E-Assessment; Heritage Competencies; Smart Tourism Paradox; Vocational Education.

Abstract

Background: The educational tension in vocational schools for tourism in Bali is specific. The digitalisation of assessment often conflicts with the preservation of locally valued service competencies embedded in the Balinese Hindu pedagogy of Samskara (ingrained habits), Cesta (volitional effort) and Kriya (embodied skills). At present, no validated assessment tool can measure all three constructs simultaneously.
Aims: This study develops and validates an AI-driven e-assessment instrument for them among Balinese tourism VET students. The novelty lies in operationalising culturally embedded competencies rather than generic digital skills.
Methods: It utilised a sequential mixed-methods design. Phase 1: Conduct ten semi-structured interviews with experts on culture and education to generate measurable indicators. Phase 2: Development of a 48-item AI platform for assessment using natural language processing, affective computing and computer vision. Phase 3: Pilot with 320 Grade XI learners from three schools with varying degrees of digital infrastructure.
Result: Internal consistency was acceptable (Cronbach's α = 0.89, 0.87, 0.91). Confirmatory factor analysis supported the three-factor structure (CFI=0.94, RMSEA=0.058). Across all three competences, students in moderately digitised schools outperformed students in low and highly digitised schools ($p < .01$, Cohen's d = 1.46–2.48). The quadratic regression analysis showed a negative quadratic coefficient which suggests an inverted-U relationship or the best heritage competency outcomes.
Conclusion: AI-driven assessment can validly measure culture-specific competencies in tourism VET. The inverted-U pattern supports the notion of a "smart tourism heritage tension" rather than confirming a paradox. Caution is warranted regarding generalizability, cross-sectional design, and potential algorithmic bias.

To cite this article: Sugiharni, G. A. D., Susanti, P. H., Jayendra, P. S., Widiastiti, A. A. I. P. & Amir, F. L. (2026). AI-Driven assessment of heritage competencies in balinese tourism vet schools: Resolving the tension between digital acceleration and cultural conservation. *Journal of Advanced Sciences and Mathematics Education, 6*(2), 686-700.

INTRODUCTION

Tourism education plays a fundamental role in sustaining the competitiveness and cultural identity of destinations that depend heavily on hospitality and service industries. In Bali, vocational education and training (VET) institutions are expected not only to produce technically competent graduates but also to preserve local cultural values that distinguish Balinese hospitality from other

***Corresponding author:**
Gusti Ayu Dessy Sugiharni, Institut Pariwisata dan Bisnis Internasional, Indonesia
dessy.sugiharni@ipb.intl.ac.id

tourism destinations. Among the most important competencies embedded in Balinese Hindu pedagogy are Samskara, Cesta, and Kriya, which respectively represent ingrained habits, volitional effort, and embodied practical skills (Candiasa et al., 2019; Widhiasthini et al., 2025). These competencies are deeply connected to local cultural values and contribute significantly to service excellence within Bali's tourism sector. Unlike technical competencies, these attributes are expressed through behavior, attitude, and culturally appropriate interactions in real-world hospitality settings. Consequently, tourism VET schools are responsible for ensuring that students develop and demonstrate these competencies throughout their educational experiences. However, assessing such competencies remains a significant challenge because they encompass dispositional, relational, and embodied dimensions of learning. Conventional assessment methods often prioritize cognitive achievement while providing limited opportunities to evaluate habitual conduct and culturally situated practices. Existing written examinations and teacher observation checklists have been criticized for their inability to capture the complexity of these competencies in authentic learning environments (Ferm, 2021; Paeßens et al., 2023; Purbaningrum & Arliani, 2025; Tan et al., 2024). Therefore, there is an increasing need for assessment approaches capable of evaluating culturally embedded competencies in a more valid and authentic manner.

The rapid advancement of digital technologies has transformed educational practices across vocational education systems worldwide. Artificial intelligence, learning analytics, and digital assessment platforms are increasingly being adopted to improve efficiency, scalability, and objectivity in educational evaluation. In Indonesia, the implementation of the Making Indonesia 4.0 roadmap has accelerated the integration of digital technologies into educational institutions, including tourism-focused VET schools. This transformation has also been reinforced by the post-pandemic recovery of Bali's tourism industry, which increasingly relies on smart tourism technologies and digitally mediated service experiences. Graduates entering the tourism workforce are expected to operate within environments characterized by automated booking systems, intelligent customer relationship management platforms, and data-driven decision-making processes. Consequently, educational institutions are encouraged to align their teaching and assessment practices with the technological realities of contemporary tourism industries. Artificial intelligence offers promising opportunities to assess complex competencies through real-time data collection, adaptive feedback mechanisms, and sophisticated pattern recognition capabilities. Such technologies have the potential to extend assessment beyond the limitations of paper-based examinations and subjective teacher judgments. Furthermore, digital assessment systems may provide more continuous and comprehensive evaluations of student performance across diverse learning contexts. As a result, the integration of AI into tourism education has emerged as a strategic priority for institutions seeking to prepare graduates for increasingly digitalized tourism environments (Kasemsarn & Nickpour, 2025).

Despite its potential benefits, the application of artificial intelligence in culturally grounded educational contexts has generated significant concerns among researchers and practitioners. Studies conducted in heritage destinations have demonstrated that digital technologies sometimes prioritize measurable indicators while overlooking local meanings, traditions, and cultural knowledge systems (Arianto & Hanif, 2024; Cui & Wu, 2025). This phenomenon has been described as heritage dissonance, a condition in which technological systems fail to reflect the values considered important by local communities. In educational settings, heritage dissonance may occur when assessment algorithms evaluate competencies according to externally imposed categories rather than culturally relevant standards. Such a mismatch can undermine educational authenticity and weaken efforts to preserve local cultural identity. Tourism education in Bali faces a similar risk because competencies such as Samskara, Cesta, and Kriya do not align neatly with conventional competency frameworks commonly used in AI-driven assessment systems. Generic assessment

algorithms are often designed around universal competencies such as critical thinking, communication, and collaboration, which may inadequately capture spiritual, relational, and embodied dimensions of learning. Previous studies have warned that technological innovation in heritage contexts frequently marginalizes local epistemologies and culturally specific forms of knowledge (Rizvic et al., 2024; Trivyzadakis, 2023). Consequently, the challenge is not simply technological but also epistemological, requiring assessment systems that recognize and respect local cultural categories. Rather than rejecting digital innovation, educational stakeholders must identify ways to integrate AI while maintaining cultural authenticity and educational relevance.

Although research on artificial intelligence in education has expanded substantially over the past decade, important gaps remain regarding culturally grounded competency assessment in tourism vocational education. Existing AI-based assessment frameworks primarily focus on generalized competencies that are cognitive, behavioral, or performance-oriented in nature. Very limited attention has been given to competencies that are culturally specific, spiritually informed, and embodied within local educational traditions. In the context of Bali, no validated AI-driven e-assessment instrument currently exists for measuring Samskara, Cesta, and Kriya among tourism VET students (Candiasa et al., 2019; Widhiasthini et al., 2025). Furthermore, previous studies have not empirically examined how different levels of digital integration influence the development and demonstration of these heritage-based competencies. Most research assumes that increasing digitalization automatically improves educational outcomes without considering possible cultural trade-offs. Existing AI assessment systems generally prioritize universal competency constructs while overlooking locally defined indicators of excellence (Paeßens et al., 2023; Tan et al., 2024). Consequently, there is insufficient understanding of whether technology-enhanced assessment supports or undermines culturally valued forms of learning. The possibility that moderate levels of digital integration may produce better heritage competency outcomes than either low or excessive integration has not been systematically investigated. This unresolved issue represents a significant theoretical and practical gap that warrants empirical examination within tourism vocational education.

This study aims to develop and validate an AI-based e-assessment instrument capable of measuring Samskara, Cesta, and Kriya among tourism vocational education students in Bali. The first objective is to transform these culturally embedded constructs into measurable indicators through a structured expert consensus process. The second objective is to evaluate the psychometric properties of the developed instrument by examining its internal consistency, construct validity, and criterion validity using data collected from 320 students across three tourism VET schools in Bali. The third objective is to investigate the relationship between the level of digital integration and student performance on heritage-related competencies. Particular attention is given to determining whether this relationship follows a non-linear pattern characterized by an optimal level of technological integration. The study introduces the concept of smart tourism heritage tension to explain potential trade-offs between digital innovation and cultural preservation within educational assessment practices. Theoretically, the research extends the application of paradox theory by examining how technological modernization and cultural authenticity can coexist within tourism education. Methodologically, the study contributes a culturally responsive framework for designing AI-based assessments in heritage-oriented learning environments. Practically, the findings are expected to provide evidence-based guidance for educators, policymakers, and technology developers seeking to implement AI systems without compromising local cultural values. Ultimately, the study seeks to demonstrate that digital transformation and cultural preservation can function as complementary rather than contradictory objectives in tourism vocational education.

LITERATURE REVIEW

The integration of artificial intelligence into educational assessment has attracted substantial scholarly attention due to its potential to transform how learning outcomes are measured and monitored. Contemporary AI-based assessment systems are designed to enhance efficiency, scalability, objectivity, and consistency in educational evaluation processes (Bartley et al., 2025; Daniel et al., 2023). These technologies have been applied across multiple educational domains, including automated essay scoring, adaptive testing, learning analytics, and predictive modelling of student achievement (Ma et al., 2025; Wiese et al., 2025). Through machine learning algorithms, assessment platforms can process large datasets and generate real-time feedback that supports personalized learning experiences. AI also offers opportunities to reduce administrative burdens associated with manual grading and performance monitoring. Recent developments have expanded AI capabilities beyond cognitive assessment to include sentiment analysis, behavioral tracking, and multimodal performance evaluation (Sutrisno et al., 2025). Such advancements suggest that AI may provide more comprehensive insights into student learning than traditional assessment approaches. Nevertheless, concerns remain regarding construct validity, transparency, and the appropriateness of algorithmic decision-making in diverse educational contexts. Scholars have emphasized that assessment technologies are not culturally neutral because they inherently reflect assumptions about what constitutes valuable knowledge and desirable performance. Consequently, the effectiveness of AI-based assessment depends not only on technological sophistication but also on its alignment with the sociocultural contexts in which it is implemented.

The application of AI within tourism vocational education has emerged as a particularly important area of inquiry due to the increasing digitalization of the global tourism industry. Tourism enterprises now operate within technologically advanced environments characterized by smart tourism systems, automated customer services, and data-driven management practices. As a result, vocational education institutions are expected to prepare graduates who can function effectively within these digitally mediated workplaces. Existing research has therefore focused on developing competencies related to employability, digital literacy, innovation readiness, and service quality improvement. Many studies have demonstrated that AI-supported learning environments can enhance student engagement and improve professional skill development. Digital assessment systems have also been used to evaluate communication abilities, customer service performance, and workplace problem-solving skills. However, these competency frameworks are generally designed around universal employability standards rather than culturally specific indicators of professional excellence. The emphasis on standardization often prioritizes competencies that can be measured consistently across diverse educational settings. While this approach facilitates comparability and scalability, it may overlook locally valued dimensions of hospitality performance that contribute to destination uniqueness. Therefore, the increasing adoption of AI in tourism education raises important questions regarding the extent to which culturally grounded competencies can be effectively represented within digital assessment systems.

Balinese tourism represents a unique educational context because hospitality excellence is deeply intertwined with local cultural and spiritual traditions. Educational practices in Bali are influenced by philosophical frameworks that emphasize harmony, social responsibility, moral conduct, and cultural continuity. Ethnographic and educational studies have extensively examined concepts such as Tri Hita Karana and Ajeg Bali as foundational principles guiding social and educational life within Balinese communities. These frameworks highlight the importance of balancing relationships among humans, nature, and spirituality while maintaining cultural resilience in the face of modernization. Within this broader cultural context, Samskara, Cesta, and Kriya constitute essential dimensions of learning and personal development (Candiasa et al., 2019;

Widhiasthini et al., 2025). Samskara reflects the cultivation of positive habits and internalized values that shape individual behavior over time. Cesta emphasizes purposeful effort, perseverance, and intentional action in achieving personal and professional goals. Kriya focuses on the manifestation of knowledge and values through embodied practice and observable performance. Together, these constructs represent a holistic conception of competence that integrates cognitive, affective, behavioral, and spiritual dimensions. Despite their significance within Balinese educational philosophy, these competencies have received limited attention within contemporary research on digital learning and AI-based assessment.

The growing use of digital technologies within heritage-related contexts has generated increasing concern regarding the preservation of local knowledge systems and cultural identities. Researchers have observed that technological innovations often introduce standardized frameworks that may not fully accommodate culturally specific ways of knowing and learning. This challenge has been conceptualized through the notion of heritage dissonance, which refers to a mismatch between technologically mediated representations and locally valued meanings (Cui & Wu, 2025; Li et al., 2024). Heritage dissonance becomes particularly relevant when assessment systems determine which competencies are recognized, rewarded, and legitimized within educational environments. Generic AI assessment models are commonly built around universal competency constructs such as critical thinking, communication, and collaboration. Although these competencies are important, they may not adequately capture relational, spiritual, and embodied forms of knowledge that are central to many cultural traditions. In tourism education, such limitations could unintentionally marginalize competencies that contribute to the authenticity of local hospitality practices. Previous studies have documented similar tensions in cultural heritage management, digital archiving, and technology-enhanced interpretation initiatives (Rizvic et al., 2024; Trivyzadakis, 2023). These findings suggest that technological innovation should not be evaluated solely in terms of efficiency but also in terms of its capacity to respect and preserve cultural meanings. Consequently, culturally responsive assessment frameworks are increasingly recognized as necessary for ensuring that digital transformation contributes to rather than diminishes heritage sustainability.

Paradox theory provides a valuable conceptual framework for understanding the tensions that emerge between technological modernization and cultural preservation within educational systems. The theory argues that organizations frequently encounter competing demands that cannot be resolved through simple either-or choices but instead require ongoing balancing and integration. Previous applications of paradox theory have examined tensions such as privacy versus personalization, global standardization versus local adaptation, and innovation versus stability (Geria et al., 2023). However, its application to heritage-oriented vocational education and AI-based assessment remains largely unexplored. In the context of tourism education in Bali, the adoption of AI reflects the need to respond to digital transformation and workforce modernization. Simultaneously, the preservation of competencies such as Samskara, Cesta, and Kriya reflects a commitment to maintaining local cultural identity and educational authenticity. These objectives may appear contradictory because technological systems often emphasize efficiency and standardization, whereas cultural competencies are contextual, relational, and deeply embedded within local traditions. Paradox theory suggests that sustainable educational innovation emerges not from choosing one objective over the other but from developing approaches that accommodate both simultaneously. This perspective provides a useful foundation for investigating what may be termed the smart tourism heritage tension, namely the possibility that moderate rather than extreme levels of digital integration produce the most favorable outcomes for heritage-based competencies. Accordingly, the present study seeks to bridge existing theoretical and empirical gaps by developing a culturally sensitive AI-based assessment framework and examining how digital integration influences the development of Samskara, Cesta, and Kriya within tourism vocational education.

METHOD

Research Design

This study employed a sequential three-phase research design to develop, validate, and evaluate an AI-based e-assessment system for measuring Samskara, Cesta, and Kriya competencies among tourism vocational education students in Bali (Maxwell, 2023). The first phase focused on deriving culturally meaningful and measurable indicators from expert knowledge rooted in Balinese Hindu pedagogy and hospitality practice. The second phase involved the development of an AI-assisted assessment platform based on the indicators identified during the initial phase. The third phase examined the psychometric properties of the developed instrument and investigated the relationship between digital integration and heritage competency performance among tourism VET students. This design was selected because it enabled systematic progression from construct conceptualization to technological implementation and empirical validation. The study integrated qualitative expert knowledge with quantitative psychometric evaluation to ensure both cultural relevance and measurement rigor. Furthermore, the design allowed the researchers to explore how digital transformation interacts with heritage-based competencies in authentic educational settings. The approach was particularly suitable for examining culturally embedded constructs that have not previously been operationalized within AI-driven assessment systems. Through this sequential process, the study aimed to establish both theoretical and practical foundations for culturally responsive AI assessment. The overall design reflects the interdisciplinary nature of the research, combining educational assessment, artificial intelligence, tourism education, and cultural heritage studies.

Participant

A total of 335 individuals participated across the three phases of the study. Phase 1 involved ten key informants selected through purposive sampling based on their expertise and professional experience. These informants consisted of three Balinese Hindu priests, four senior tourism vocational teachers with more than ten years of teaching experience, and three hotel executives from five-star resorts located in Ubud and Nusa Dua. The selection of these stakeholders ensured representation of cultural, educational, and industry perspectives relevant to Balinese hospitality competencies. In Phase 3, the study involved 320 Grade XI students recruited from three tourism vocational schools located in Badung (n = 108), Gianyar (n = 106), and Tabanan (n = 106) regencies. Grade XI students were selected because they had completed at least one year of hospitality training and were considered sufficiently familiar with tourism service practices. Schools were selected using a two-step procedure based on regional digital infrastructure levels. First, all regencies and cities in Bali were ranked according to a provincial digital infrastructure index that included internet bandwidth availability, computer-to-student ratios, and learning management system utilization. Subsequently, one accredited Grade A school was purposively selected from regencies representing low, moderate, and high levels of digital acceleration. Following data screening, 15 participants with incomplete assessment records were excluded, resulting in a final analytical sample of 320 students, with no evidence of attrition bias across demographic variables.

Instrument

The primary instrument was the AI-Samskara E-Assessment Platform, consisting of 48 items distributed equally across three constructs: Samskara, Cesta, and Kriya. Each construct was represented by 16 items scored on a seven-point Likert-type scale anchored by culturally specific behavioral descriptors. The Samskara module assessed students' responses to hospitality-related scenarios requiring culturally appropriate judgment and decision-making. Open-ended responses were analyzed using a fine-tuned BERT-base-Indonesian language model trained on 2,500 annotated

hospitality dialogues. The model generated cultural authenticity scores that were subsequently converted into seven-point competency ratings. The Cesta module utilized short video-based scenarios depicting service situations that required proactive responses from students. Responses were processed using an Indonesian sentiment analysis model that evaluated linguistic indicators associated with sincerity, commitment, and service orientation. The Kriya module required students to upload short videos demonstrating Balinese hospitality gestures, which were analyzed using computer vision techniques incorporating MediaPipe pose detection and a custom gesture-classification model. Additional instruments included the Digital Acceleration Exposure Scale (DAES), the Conventional Heritage Competency Test (CHCT), and a demographic questionnaire. Content validity was evaluated using Lawshe's Content Validity Ratio, criterion validity was assessed through correlations with CHCT scores, and reliability was examined using internal consistency and test–retest measures.

Research Procedure

Data collection was conducted between August and December 2025. During Phase 1, ten semi-structured interviews lasting between 60 and 90 minutes were conducted with expert informants to identify culturally relevant indicators of Samskara, Cesta, and Kriya. Interview transcripts were subjected to thematic analysis, generating an initial pool of 78 candidate assessment items. These items were subsequently reviewed during a two-day expert panel workshop and refined into a final set of 48 items using content validity criteria. Phase 2 focused on developing the AI-based assessment platform. The language processing model was trained using 2,500 annotated hospitality dialogues, while the sentiment analysis model utilized 1,200 previously collected and expert-rated student responses. In addition, the computer vision component was trained using 800 video recordings of Balinese hospitality gestures performed by volunteer participants. A pilot usability study involving 30 students was conducted to identify technical and instructional issues prior to implementation. During Phase 3, informed consent was obtained from students and their parents before participation. Students first completed the DAES, demographic questionnaire, and CHCT, followed by three AI-based assessment sessions conducted over a one-month period, after which a subsample of students completed a second assessment administration to establish test–retest reliability.

Data Analysis

All analyses were conducted using R version 4.3.2 with a significance level of $\alpha = .05$. Preliminary analyses included descriptive statistics for each assessment item, including means, standard deviations, skewness, and kurtosis values. Internal consistency was evaluated using Cronbach's alpha and McDonald's omega coefficients accompanied by bootstrap confidence intervals. Test–retest reliability was estimated using the intraclass correlation coefficient based on a two-way random-effects model with absolute agreement. Construct validity was examined through confirmatory factor analysis using robust maximum likelihood estimation. Model adequacy was evaluated using multiple fit indices, including the Comparative Fit Index, Tucker–Lewis Index, Root Mean Square Error of Approximation, and Standardized Root Mean Square Residual. Convergent validity was assessed using Average Variance Extracted values, whereas discriminant validity was evaluated using the Fornell–Larcker criterion. Criterion validity was examined through Pearson correlation coefficients between AI-generated scores and CHCT scores. To test the proposed smart tourism heritage tension hypothesis, one-way analysis of variance and Tukey post hoc comparisons were conducted across low, moderate, and high digital acceleration groups. Finally, quadratic regression, analysis of covariance, and multilevel modeling were performed to determine whether digital integration exhibited a non-linear relationship with heritage competency outcomes while controlling for relevant demographic factors.

RESULTS AND DISCUSSION

Results

The final sample consisted of 320 students of Grade XI from three schools of tourism VET in Badung (n=108; high digital infrastructure), Gianyar (n=106; moderate), and Tabanan (n=106; low). The average age was 16.8 years (SD = 0.7), and 52% were female. There were significant differences in the Digital Acceleration Exposure Scale (DAES) scores across schools ($F(2,317)=312.45$, $p<.001$, $\eta 2=0.66$), confirming the intended stratification. Badung (M=4.21, SD=0.58), Gianyar (M=2.94, SD=0.63), Tabanan (M=1.87, SD=0.71). The proportion of households with internet access at home ranged from 94 per cent in Badung to 43 per cent in Tabanan. On average, the groups used their smartphones 5.2 hours per day. Tourism was the main parental occupation (63%), followed by agriculture (22%) and others (15%). These demographic trends were similar across the three schools (see Table 1).

Table 1. Demographic Characteristics and Scores for Digital Acceleration Exposure by Three Types of Schools

Characteristic	Badung (High) n=108	Gianyar (Moderate) n=106	Tabanan (Low) n=106
Age, M (SD)	16.9 (0.6)	16.8 (0.7)	16.7 (0.8)
Female, %	51	54	51
Home internet access, %	94	71	43
DAES score (1–5), M (SD)	4.21 (0.58)	2.94 (0.63)	1.87 (0.71)

The internal consistency for all three constructs was high. Cronbach's α was 0.89 for Samskara, 0.87 for Cesta and 0.91 for Kriya; McDonald's ω produced nearly identical values (0.90, 0.88, 0.92, respectively). The two-week test-retest reliability (ICC) ranged from 0.82 to 0.88. These values support the instrument's ability to produce stable and internally consistent scores, which are needed for the subsequent interpretation of group differences.

Table 2. Internal Consistency and Test-Retest Reliability

Construct	Cronbach's α	95% CI	McDonald's ω	Test-Retest ICC (2,1)	95% CI
Samskara	0.89	[0.87 – 0.91]	0.90	0.85	[0.76 – 0.91]
Cesta	0.87	[0.84 – 0.90]	0.88	0.82	[0.72 – 0.89]
Kriya	0.91	[0.89 – 0.93]	0.92	0.88	[0.80 – 0.93]

All 48 items were above the Lawshe CVR threshold of 0.62 (mean CVR across constructs = 0.85). Items thrown away: None. Although this 100% retention rate may appear unusual, it is a result of the iterative development process in Phase 1, in which candidate items were qualitatively screened before the formal CVR panel.

Table 3. Content Validity Ratio (CVR) by Construct

Construct	Number of Items	CVR Range	Mean CVR	Items Below Threshold (CVR < 0.62)
Samskara	16	0.74 – 0.92	0.84	0
Cesta	16	0.74 – 0.89	0.81	0
Kriya	16	0.82 – 0.96	0.89	0
Total	48	0.74 – 0.96	0.85	0

Pearson correlations between AI-generated scores and traditional paper-based tests were r = 0.68 (Samskara), r = 0.64 (Cesta) and r = 0.72 (Kriya). The moderate to strong correlations indicate that the AI assessment assesses similar competencies to expert-rated paper tests, but not to the extent that there is redundancy. The correlation for Cesta (r = 0.64) was lower and may reflect the difficulty of measuring volitional effort with paper-based methods.

Table 4. Criterion validity: correlations between AI-based assessment and traditional test scores

Construct	r (Pearson)	p-value	Interpretation
Samskara	0.68	< .001	Moderate
Cesta	0.64	< .001	Moderate
Kriya	0.72	< .001	Strong
Composite	0.70	< .001	Strong

The findings from the confirmatory factor analysis provided support for the hypothesised three-factor structure (Samskara, Cesta, and Kriya as correlated factors). The model showed an acceptable fit: $\chi^2(1077)=1{,}843.42$, $p<.001$, CFI=.94, TLI=.93 and RMSEA=.058 (90% CI: 0.054–0.062) and SRMR=.049. Standardised factor loadings ranged from 0.63 to 0.89 (all above 0.60). A one-factor model (all items loading on one heritage competency factor) fit significantly worse ($\Delta\chi2(3) = 1{,}124.56$, $p < .001$) and confirmed that Samskara, Cesta and Kriya are empirically distinguishable. The bifactor model did not converge, indicating that there is no overarching general factor.

Table 5. Confirmatory Factor Analysis (CFA) Model Fit Indices

Model	χ^2	df	p	CFI	TLI	RMSEA	90% CI RMSEA	SRMR
Three-Factor Model	1,843.42	1077	< .001	0.94	0.93	0.058	[0.054 – 0.062]	0.049
One-Factor Model	3,421.78	1080	< .001	0.78	0.76	0.094	[0.091 – 0.097]	0.092
Bifactor Model	Did not converge							

Convergent validity was demonstrated using the average variance extracted (AVE) exceeded the threshold of 0.50 for all constructs (Samskara AVE=0.62; Cesta AVE=0.58; Kriya AVE=0.67). Further support for discriminant validity was provided as each construct's AVE square root value (0.79, 0.76, 0.82) was higher than its correlations with the other constructs (r = 0.58–0.62). The three competences are therefore related but not redundant. This pattern is consistent with the Balinese Hindu pedagogical framework, which views Samskara, Cesta, and Kriya as reinforcing but separate dimensions of service excellence.

Table 6. Convergent and Discriminant Validity: Average Variance Extracted (AVE) and Inter-Construct Correlations.

Construct	AVE	√AVE	Correlation with Samskara	Correlation with Cesta	Correlation with Kriya
Samskara	0.62	0.79	1.00		
Cesta	0.58	0.76	0.58*	1.00	
Kriya	0.67	0.82	0.62*	0.54*	1.00

The Smart Tourism Heritage Tension: Inverted-U Relationship. Students in moderately digitalised schools (Gianyar) outperformed those in low- (Tabanan) and high- (Badung) digitalisation schools on all three heritage competencies. Composite heritage scores (mean of three constructs) were low digitalisation (M=65.5, SD=7.4), moderate (M=78.4, SD=5.9), and high digitalisation (M=58.8, SD=8.9). One-way ANOVA results showed significant differences ($F(2,317)=176.54$, $p<.001$, $\eta2=0.53$). Findings from the Tukey HSD post hoc procedure revealed that the moderate group outperformed the low group (mean difference = 12.9, $p<.001$, Cohen's d = 1.52–1.68 across constructs) and the high group (mean difference = 19.7–20.4, $p<.001$, d = 2.15–2.48). Interestingly, the low-digitalisation group was also more performant than the high-digitalisation group (d = 0.68–0.88).

Table 7. Descriptive Statistics and One-Way ANOVA Results for Heritage Competency Scores by Digital Acceleration Level.

Competency	Low Acceleration (n=106) M (SD)	Moderate Acceleration (n=106) M (SD)	High Acceleration (n=108) M (SD)	F(2, 317)	p	η^2
Samskara	65.3 (8.2)	78.6 (6.4)	58.2 (9.5)	147.32	< .001	0.48
Cesta	62.7 (9.1)	75.3 (7.2)	56.8 (10.2)	138.65	< .001	0.47
Kriya	68.4 (7.6)	81.2 (5.9)	61.5 (8.8)	159.44	< .001	0.50
Composite	65.5 (7.4)	78.4 (5.9)	58.8 (8.9)	176.54	< .001	0.53

These very large effect sizes (d > 2.0 for moderate vs high) warrant caution: they may be specific to the particular digital environments studied and should not be generalised without replication. They do indicate that the observed differences are practically substantial, not merely statistically significant.

Table 8. Post-Hoc Tukey HSD Pairwise Comparisons

Competency	Comparison	Mean Difference	95% CI	p	Cohen's d
Samskara	Moderate vs. Low	13.3	[10.8 – 15.8]	< .001	1.52
	Moderate vs. High	20.4	[17.9 – 22.9]	< .001	2.34
	Low vs. High	7.1	[4.6 – 9.6]	< .001	0.81
Cesta	Moderate vs. Low	12.6	[9.9 – 15.3]	< .001	1.46
	Moderate vs. High	18.5	[15.8 – 21.2]	< .001	2.15
	Low vs. High	5.9	[3.2 – 8.6]	< .001	0.68
Kriya	Moderate vs. Low	12.8	[10.4 – 15.2]	< .001	1.68
	Moderate vs. High	19.7	[17.3 – 22.1]	< .001	2.48
	Low vs. High	6.9	[4.5 – 9.3]	< .001	0.88

The quadratic regression was performed to formally test whether the relationship between digital integration and heritage competency was an inverted-U curve. The predictor was the centred DAES scores. For all three constructs, the quadratic term was negative and statistically significant (Samskara β 2 = -0.34, p<.001; Cesta β 2 = -0.31, p<.001; Kriya β 2 = -0.37, p<.001), while the linear term was non-significant. The quadratic model explained 49-55 % (R^2 = 0.49-0.55) of the variance, with the quadratic term significantly adding 10-14 % to a model with linear terms only.

Table 9. Quadratic Regression Results (Inverted-U Relationship Test)

Competency	β_1 (Linear)	p	β_2 (Quadratic)	p	R^2	R^2 Change (vs. linear only)	F Change	p	Turning Point
Samskara	0.09	.21	-0.34	< .001	0.53	0.12	78.56	< .001	2.96
Cesta	0.11	.18	-0.31	< .001	0.49	0.10	62.34	< .001	2.88
Kriya	0.06	.35	-0.37	< .001	0.55	0.14	94.21	< .001	3.01

The turning points (optimal DAES scores) were 2.96 for Samskara, 2.88 for Cesta and 3.01 for Kriya – all close to the mean DAES score of the moderate-digitalisation school (Gianyar, M=2.94). In practical educational terms, this ideal range (roughly 2.9–3.0 on a 1–5 scale) corresponds to schools where students are provided regular but not all-pervasive access to digital tools, teachers incorporate technology into their teaching in a selective manner, and traditional cultural practices (e.g., daily offerings, gesture training) remain part of daily routines.

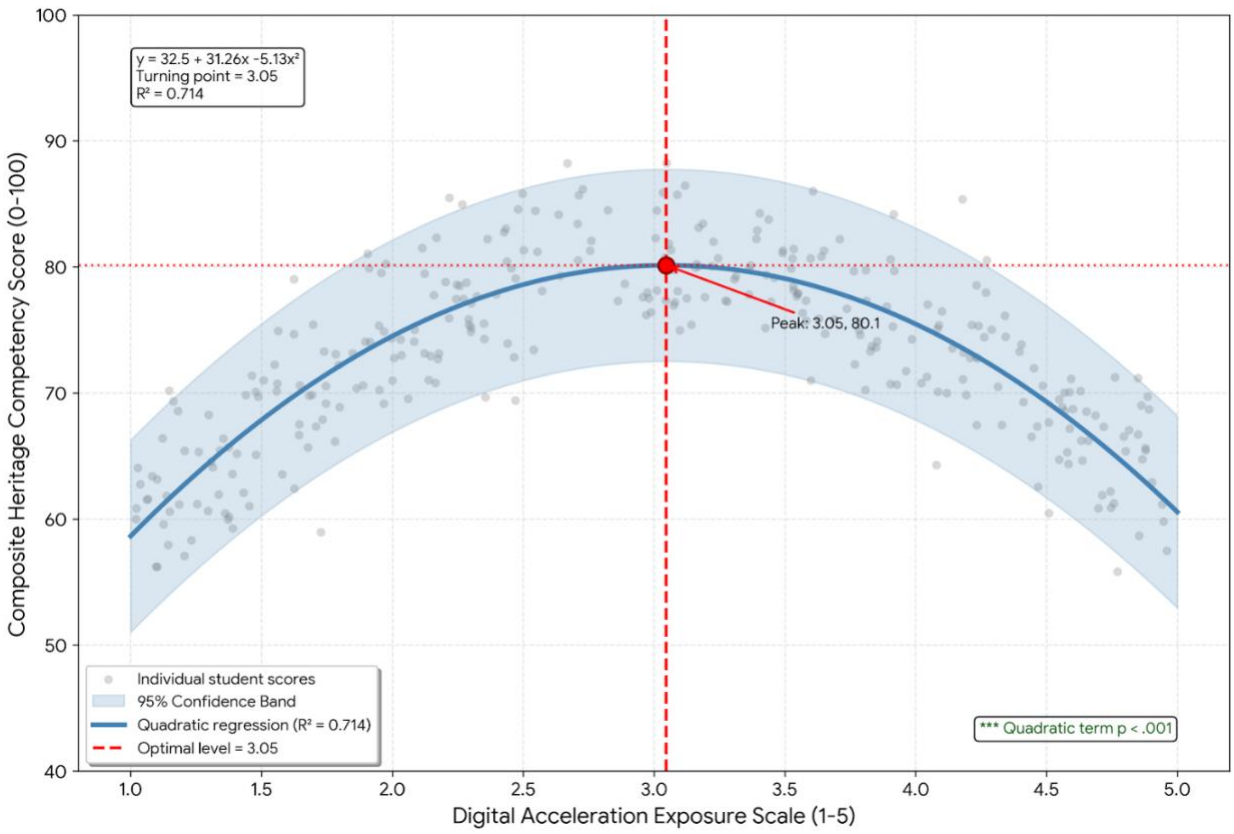

Figure 1. Inverted-U relationship between digital acceleration exposure and composite heritage competency score, shaded area = 95% confidence band.

When controlling for gender, home internet access and parental occupation, the inverted-U pattern remained significant (ANCOVA; partial η2 for digital acceleration = 0.44-0.48). Multilevel modelling (random intercepts for school; ICC = 0.07 for the composite score) also indicated a significant quadratic term (β2 = −0.31, p = .002), suggesting that this was not an artefact of school-level clustering. The test-retest subsample (n = 60) showed the same pattern. These sensitivity analyses provide us with confidence in the robustness of the inverted-U relationship; however, given the cross-sectional design, we cannot draw causal inferences.

Discussion

This research developed and validated an AI-based e-assessment tool for measuring the Samskara, Cesta, and Kriya of Balinese tourism VET students. The psychometric properties (Cronbach's α from 0.87 to 0.91, CFI = 0.94, and AVE ≥ 0.58) suggest that the instrument produces reliable and valid scores for this sample in this context with this specific AI architecture. But these values do not prove that the instrument would work just as well in other cultural contexts or with other AIs. What the evidence suggests is the feasibility of operationalising culturally embedded competencies through AI-powered assessment when the process of development is grounded in local knowledge systems (Luckin et al., 2022) on measuring previously hard-to-measure competencies.

The three-factor structure (Samskara, Cesta, Kriya as correlated factors) empirically supports the pedagogical distinction in Balinese Hindu between ingrained habits, volitional effort and embodied skills. These dimensions are related but not redundant – a pattern that is consistent with the local philosophy that service excellence emerges from their integration. The criterion-validity correlations (r = 0.64–0.72) were moderate, not perfect. It is expected that the AI assessment identifies features that paper-based tests cannot, such as tone sincerity (for Cesta) and posture

alignment (for Kriya). In other words, the two assessment methods overlap in part, but also diverge, and this is a strength, not a weakness. The key finding is an inverted-U relationship between the degree of digital acceleration and the performance of heritage competency. Students in medium-digitalisation schools achieved significantly better results than their low- and high-digitalisation peers with effect sizes large enough to be practically meaningful (d = 1.46–2.48).

This pattern provides empirical evidence of what we have termed the smart tourism heritage tension, the idea that moderate and selective digital integration might optimise outcomes of heritage competency. Not too little integration. Not too much integration. It challenges the linear assumptions often made in educational technology research. It extends the paradox theory (Khaksar et al., 2024) to the field of AI-based assessment of heritage contexts. Three mechanisms may account for this pattern. First, in low digitisation contexts (Tabanan), students continue to engage in their cultural practices but struggle to translate them into digitally mediated service encounters, and hence experience cultural discontinuity (Surata et al., 2024). Secondly, students in highly digitalised environment (Badung) adapt what can be termed as efficiency-oriented responses. They tended to give quick, formulaic answers to situational judgement tasks (e.g., the canang sari complaint scenario) rather than culturally reflective ones. Third, students in medium-digitalisation environments (Gianyar) seem to experience digital tools as a 'cultural prosthetic'-technologies that enhance, rather than replace, heritage competencies. Third, students in medium-digitalisation environments (Gianyar) seem to experience digital tools as a 'cultural prosthetic' – technologies that enhance, rather than replace, heritage competencies. Interestingly, the low-digitalisation group outperformed the high-digitalisation group (d = 0.68–0.88), suggesting that even no digital integration may be better than too much integration for the preservation of Samskara and Cesta. The cross-sectional design does not allow for causal inference. The inverted-U pattern is consistent with the hypothesised tension, but longitudinal data would be needed to establish directionality.

Implications

In this study, paradox theory is used to evaluate AI in heritage tourism education. The smart tourism heritage tension is a conceptual tool for investigating how digital acceleration and cultural conservation can coexist without either being absolute. Researchers do not claim to contribute to paradox theory (a claim for which there is no evidence). Rather, researchers suggest that the inverted-U pattern provides an empirical instantiation of how organisations deal with competing demands. This compromise approach avoids both techno-solutionism and Luddite rejectionism. This finding also supports the idea that researchers can identify and measure culturally specific competency constructs such as Samskara, Cesta, and Kriya that don't work the same way as generic skills. That casts doubt on the universalist assumption underpinning much AI education research (Wang & Liu, 2025). Two extremes should be avoided by tourism VET schools in heritage destinations: uncritical acceptance of AI (which may erode local competencies) and complete rejection of digital tools (which may put students at a disadvantage in the smart tourism job market.

The study's findings suggest that a moderate level of acceleration (around 2.9-3.0 on a 1-5 scale) is optimal. It means schools where digital tools are used in a targeted way (e.g. for formative assessment and simulation) but where traditional cultural practices (daily offerings, gesture training, face-to-face mentoring) are part of the daily routine. A rule of thumb (but only a rule of thumb) is a ratio of about two hours of culturally immersive, tech-free pedagogy for every hour of AI-enhanced assessment. We use a human-in-the-loop approach where AI performs the initial scoring and culture experts review edge cases, providing a template for culture-sensitive AI. Developers must resist the temptation to measure what is easy to measure (such as how quickly a response is produced) and instead invest in ethnographic fieldwork, expert panels, and locally built, from-the-ground-up solutions. Algorithmic bias is a real concern. The present study did not systematically investigate

measurement invariance across gender or socioeconomic subgroups, but this should be a requirement prior to any high-stakes deployment. The Indonesian Ministry of Education could employ differentiated approaches to accelerate the use of digital technologies in tourism VET schools in heritage zones. Schools located in the core heritage areas (Ubud, Kintamani, Besakih) could be encouraged to maintain moderate levels of digital integration rather than compete for maximum digital infrastructure. It is a tentative policy consideration based on one study, not a prescriptive mandate. Indicators of cultural competency could be incorporated alongside digital infrastructure metrics in accreditation standards – but this would require wider stakeholder consultation.

Limitations and Suggestions for Future Research

Several limitations should be acknowledged when interpreting the findings of this study. First, the research was conducted exclusively within tourism vocational education institutions in Bali, which limits the transferability of the findings to other heritage tourism destinations and educational contexts characterized by different cultural traditions and competency frameworks. Second, the cross-sectional nature of the study does not permit causal conclusions regarding the relationship between digital acceleration and heritage competency development, meaning that the observed inverted-U pattern should be interpreted as an association rather than evidence of causation. Third, the performance of the AI-Samskara E-Assessment platform is inherently dependent on the specific algorithms, training datasets, and model configurations employed during system development, and alternative AI architectures may produce different assessment outcomes. Fourth, because many participants had limited prior experience with AI-assisted assessment environments, novelty effects and technology-related reactivity may have influenced their responses and performance. Fifth, the measurement of digital acceleration relied partly on self-reported data, which may be subject to reporting bias and may not fully capture the complexity of students' actual exposure to digital learning environments. Sixth, the present study focused primarily on psychometric validation and did not evaluate the extent to which AI-generated competency scores predict long-term workplace performance, supervisor evaluations, guest satisfaction, or career success in the tourism industry. Seventh, measurement invariance across demographic subgroups, including gender, socioeconomic background, language use, and cultural exposure, was not comprehensively examined prior to implementation. Future research should therefore adopt longitudinal designs to investigate the developmental trajectories of heritage competencies and clarify the causal mechanisms underlying the relationship between digital integration and cultural learning outcomes. Experimental and quasi-experimental studies that systematically manipulate levels of digital acceleration would provide stronger evidence regarding the proposed smart tourism heritage tension hypothesis. Replication studies conducted in other heritage destinations with distinct cultural competency systems are also necessary to evaluate the generalizability of the proposed assessment framework across diverse sociocultural contexts. Furthermore, future investigations should incorporate objective indicators of digital engagement, such as learning management system analytics, classroom observations, and digital activity logs, to complement self-report measures. Finally, mixed-methods approaches integrating quantitative assessment data with classroom observations, teacher narratives, student experiences, and workplace performance indicators would provide deeper insight into how culturally responsive AI assessment can support both technological innovation and heritage preservation within tourism vocational education.

CONCLUSION

The purpose of this study is to develop and validate the AI-based e-assessment instrument to measure the Samskara, Cesta, and Kriya on students in three VET schools of tourism in Bali. Psychometric evidence (Cronbach's alpha = 0.87–0.91; CFI = 0.94; RMSEA = 0.058) indicates that the

instrument produces reliable and valid scores in this context and for the AI architecture studied. The inverted-U relationship between digital acceleration level and heritage competency performance is consistent with the suggestion that moderate digital integration may be associated with better competency outcomes than low or high integration. This pattern, based on cross-sectional data, provides preliminary support for what we term the smart tourism heritage tension, rather than confirming a causal paradox. In the case of tourism VET schools in heritage settings, a balanced approach of selective digital tools, combined with continued traditional cultural practices, appears to be a more promising direction than uncritical adoption or rejection of technology. Further validation studies are needed to test measurement invariance, long-term criterion validity and to replicate the inverted-U pattern in other heritage destinations in longitudinal or experimental studies.

AUTHOR CONTRIBUTIONS STATEMENT

Gusti Ayu Dessy Sugiharni conceptualized and designed the study, led the manuscript writing, and coordinated data collection. Putu Herny Susanti contributed to methodology development, data analysis, and interpretation of results. Putu Sabda Jayendra was responsible for literature review, drafting sections of the manuscript, and formatting. Anak Agung Istri Putera Widiastiti assisted in data collection, data validation, and preparation of figures and tables. Firlie Lanovia Amir contributed to manuscript editing, proofreading, and the overall supervision of research activities. All authors reviewed, edited, and approved the final version of the manuscript for submission.

REFERENCES

Arianto, F., & Hanif, M. (2024). Evaluating metacognitive strategies and self-regulated learning to predict primary school students' self-efficacy and problem-solving skills in science learning. *Journal of Pedagogical Research, 8*(3), 301-319. https://doi.org/10.33902/JPR.202428575

Bartley, M., Huemann, Z., Hu, J., Tie, X., Ross, A. B., Kennedy, T., Warner, J. D., Bradshaw, T., & Lawrence, E. M. (2025). Artificial intelligence for teaching case curation: Evaluating model performance on imaging report discrepancies. *Academic Radiology, 32*(6), 3139-3146. https://doi.org/10.1016/j.acra.2025.02.011

Candiasa, I. M., Santiyadnya, N., Sukajaya, N., & Sunu, G. K. A. (2019). Contextualization of learning models in subak system for vocational education. *Journal of Physics: Conference Series, 1165*(1), 012023. https://doi.org/10.1088/1742-6596/1165/1/012023

Cui, H., & Wu, J. (2025). Virtual exhibitions of cultural heritage: Research landscape and future directions. *Applied Sciences, 15*(22). https://doi.org/10.3390/app152212287

Daniel, S., Pacheco, M., Smith, B., Burriss, S., & Hundley, M. (2023). Cultivating writerly virtues: Critical human elements of multimodal writing in the age of artificial intelligence. *Journal of Adolescent & Adult Literacy, 67*(1), 32-38. https://doi.org/10.1002/jaal.1298

Ferm, L. (2021). Vocational students' ways of handling the academic/vocational divide. *International Journal for Research in Vocational Education and Training, 8*(1), 1-20. https://doi.org/10.13152/IJRVET.8.1.1

Geria, I. M., Nastiti, T. S., Handini, R., Sujarwo, W., Dwijendra, A., Fauzi, M. R., & Juliawati, N. P. E. (2023). Built environment from the ancient Bali: The Balinese heritage for sustainable water management. *Heliyon, 9*(11). https://doi.org/10.1016/j.heliyon.2023.e21248

Kasemsarn, K., & Nickpour, F. (2025). Digital storytelling in cultural and heritage tourism: A review of social media integration and youth engagement frameworks. *Heritage, 8*(6). https://doi.org/10.3390/heritage8060200

Khaksar, S. M. S., Shahmehr, F. S., Miah, S., Daim, T., & Ozdemir, D. (2024). Privacy concerns versus personalisation benefits in social robot acceptance by employees: A paradox theory -

contingency perspective. *Technological Forecasting and Social Change, 198*, 123034. https://doi.org/10.1016/j.techfore.2023.123034

Li, Y., Liang, J., Huang, J., Shen, H., Li, X., & Law, A. (2024). Evaluating tourist perceptions of architectural heritage values at a World Heritage Site in South-East China: The case of Gulangyu Island. *Journal of Hospitality and Tourism Management, 60*, 127-140. https://doi.org/10.1016/j.jhtm.2024.06.018

Luckin, R., Cukurova, M., Kent, C., & du Boulay, B. (2022). Empowering educators to be AI-ready. *Computers and Education: Artificial Intelligence, 3*, 100076. https://doi.org/10.1016/j.caeai.2022.100076

Ma, M., Ng, D. T. K., Liu, Z., & Wong, G. K. W. (2025). Fostering responsible AI literacy: A systematic review of K-12 AI ethics education. *Computers and Education: Artificial Intelligence, 8*, 100422. https://doi.org/10.1016/j.caeai.2025.100422

Paeßens, J., Ma, B., & Winther, E. (2023). Effectiveness of collaboration in VET: Measuring skills for solving complex vocational problems with a multidimensional authentic technology-based assessment. *International Journal for Research in Vocational Education and Training, 10*(1), 46-67. https://doi.org/10.13152/IJRVET.10.1.3

Purbaningrum, M., & Arliani, E. (2025). Mathematical literacy test on statistics for vocational high school level: Developing test items. *AIP Conference Proceedings, 3316*(1), 140001. https://doi.org/10.1063/5.0290076

Rizvic, S., Boskovic, D., & Mijatovic, B. (2024). Advanced interactive digital storytelling in digital heritage applications. *Digital Applications in Archaeology and Cultural Heritage, 33*, e00334. https://doi.org/10.1016/j.daach.2024.e00334

Surata, I. K., Sumartana, I. M., & Utama, I. G. B. R. (2024). The impact of cultural tourism on local traditions (SSRN Scholarly Paper No. 5088509). Social Science Research Network. https://doi.org/10.2139/ssrn.5088509

Sutrisno, S., Azis, A., Setyawan, M. B., Dewi, D. A., Hidayah, Y., Hakiki, M., Hamid, M. A., & Fadli, R. (2025). Assessing students' readiness for artificial intelligence-based project learning to strengthen local wisdom values in Indonesia. *Cogent Education, 12*(1), 2582948. https://doi.org/10.1080/2331186X.2025.2582948

Tan, X., Lin, X., & Zhuang, R. (2024). Development and validation of a secondary vocational school students' digital learning competence scale. *Smart Learning Environments, 11*(1), 37. https://doi.org/10.1186/s40561-024-00325-6

Trivyzadakis, N. (2023). A tool for quantitative research on the social awareness for the digitization of cultural heritage: A pilot application. *Journal of Archaeological Science: Reports, 51*, 104202. https://doi.org/10.1016/j.jasrep.2023.104202

Widhiasthini, N. W., Yanti, N. K. W., Subawa, N. S., Baykal, E., Budiana, I. N., Utami, M. S. M., & Subanda, I. N. (2025). Pentahelix model and Tri Hita Karana for Bali sustainable tourism. *Tourism and Hospitality Research*. https://doi.org/10.1177/14673584251340046

Wiese, L. J., Patil, I., Schiff, D. S., & Magana, A. J. (2025). AI ethics education: A systematic literature review. *Computers and Education: Artificial Intelligence, 8*, 100405. https://doi.org/10.1016/j.caeai.2025.100405