Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 7, 1436-1451
2025
Publisher: Learning Gate
DOI: 10.55214/25768484.v9i7.8936
© 2025 by the authors; licensee Learning Gate

Generative AI for storytelling in cultural tourism: enhancing visitor engagement through AI-driven narratives

Rungtiva Saosing[1], Sooksawaddee Nattawuttisit[2*]
[1]Faculty of Science and Technology, Rajamangala University of Technology Krungthep, Thailand; rungtiva.s@mail.rmutk.ac.th (R.S.).
[2]Faculty of Information Technology, Sripatum University, Thailand; sooksawaddee.na@spu.ac.th (S.N.).

Abstract: This study develops and evaluates a generative AI-powered storytelling model designed to enhance cultural tourism by providing personalized, multilingual, and emotionally engaging visitor experiences. Grounded in a theoretical model integrating semiotic theories, human-centered AI design, and multimodal interaction, the model conceptualizes AI as a co-partner in culturally meaningful storytelling rather than merely a passive recommender. Addressing a critical gap in tourism technologies—the lack of adaptive, narrative-based interpretation tools—the study introduces a hybrid architecture integrating the GPT-4o model for dynamic storytelling, Retrieval-Augmented Generation (RAG) for context-sensitive recommendations, and a custom image-generation pipeline. A mobile application deploying this model was tested across four heritage sites in Bangkok with 400 international tourists from Thailand, China, Japan, Europe, and ASEAN. Model training occurred over 100 epochs using an 80/20 split, achieving an F1-score of 89.94%, classification accuracy of 87.39%, and semantic similarity scores of up to 0.95. Empirical findings indicate significant improvements in emotional engagement, cultural understanding, satisfaction, recommendation intentions, and memory retention. These findings reinforce the model's efficacy, offering pragmatic guidelines for tourism authorities and cultural enterprises to modernize visitor services and appeal to global audiences through intelligent, adaptive storytelling.

Keywords: *Cultural tourism, Personalized storytelling, Generative artificial intelligence, GPT, Human-centered AI. retrieval-augmented generation (RAG).*

1. Introduction

In recent years, the importance of Artificial Intelligence (AI) in modern businesses and its applications has grown rapidly. Nowadays, the value of AI-driven innovation in creating and enhancing competitive advantages requires a comprehensive understanding of its impact on processes and platform integration across all types of organizations [1]. The effectiveness of operations within firms is increasingly influenced by AI-driven innovation.

Tourism is among the industries where these technological advancements have become particularly significant. The demand for personalized, inclusive, and interactive visitor experiences has increased considerably. While traditional interpretive methods—such as signage, printed brochures, and pre-recorded audio guides—continue to provide essential information, they often fall short of meeting the evolving expectations of contemporary tourists [1]. These methods typically deliver static and generic content that lacks the flexibility to adapt to individual preferences, situational contexts, and linguistic diversity [2, 3].

International visitors, in particular, frequently encounter language barriers and limited access to culturally tailored information, which can diminish emotional engagement and reduce the quality of their experiences. Moreover, while human tour guides can offer valuable personalized services, their

© 2025 by the authors; licensee Learning Gate
History: Received: 7 May 2025; Revised: 24 June 2025; Accepted: 27 June 2025; Published: 18 July 2025
* Correspondence: sooksawaddee.na@spu.ac.th

scalability and availability remain limited, especially at remote or overcrowded cultural heritage sites [4].

The rapid advancement of artificial intelligence (AI), particularly in the area of generative models and large language models (LLMs), offers a compelling solution to these challenges [5, 6]. Generative AI technologies such as GPT-4o enable dynamic narrative generation, multilingual communication, and context-aware recommendations [7]. When integrated into mobile applications, these models have the potential to act as virtual tour guides, capable of providing visitors with real-time responses, recommending points of interest (POIs), suggesting accommodations, and even facilitating in-depth cultural conversations [8, 9]. Unlike traditional guides or static media, emerging technologies such as hybrid and XR narrative applications have been shown to deliver more personalized and engaging experiences, better adapting to user preferences and situational contexts [10, 11]. The integration of custom image generation enhances these narratives, making visitors active participants in visually engaging cultural stories [12].

To address these possibilities, this study investigates the application of generative AI in enhancing cultural tourism through interactive, multilingual, and personalized storytelling. A mobile application was developed and deployed, leveraging GPT-4o and a custom image generation model to generate dynamic narratives based on user-selected POIs and preferred genres [13]. To ensure cultural relevance and conversational fluency, the AI model was fine-tuned using the Survey Questionnaire Dataset, supplemented with local cultural datasets. The research employed a mixed-methods approach, combining experimental field deployment with survey-based evaluations [14]. Data were collected from a stratified random sample of 400 domestic and international tourists at four major cultural heritage sites in Thailand. Their experiences were assessed through structured questionnaires and interviews focused on emotional engagement, cultural learning, and satisfaction [15].

Preliminary findings indicate that AI-generated narratives significantly enhance emotional connection, cultural comprehension, and the likelihood of visitors recommending the sites to others. Furthermore, personalized storytelling and multilingual interactions contribute to higher levels of satisfaction and memory retention. This research makes a notable contribution to the emerging field of AI-enhanced tourism by demonstrating how generative AI can be ethically and practically applied to foster deeper cultural connections and improve interpretive services. The results offer valuable insights for tourism authorities and cultural heritage managers aiming to modernize visitor experiences through scalable, adaptive, and inclusive digital solutions.

2. Literature

Generative AI refers to models capable of creating text, images, or other media based on input data. Models like GPTs are widely used for tasks such as image, audio, and text generation [16] forming the foundation of modern Large Language Models (LLMs), which produce coherent, context-aware outputs [17]. Variational Autoencoders (VAEs) further enhance generative performance by learning complex data distributions to produce diverse, realistic content [5].

Advances in object detection, driven by mobile device proliferation, have also contributed to generative tasks. For example, YOLOv8 incorporates sophisticated feature extraction for improved real-time content generation [18-20]. Within the tourism industry, AI has emerged as a promising tool for addressing longstanding challenges. As demand grows for personalized, immersive travel experiences, AI technologies—particularly machine learning (ML), recommender models, and conversational agents—are being adopted to enhance visitor engagement and provide context-sensitive guidance [21, 22]. Beyond tourism, AI's proven impact on business intelligence and decision-making [23] reinforces its relevance for designing intelligent, adaptive service ecosystems, including tourism guidance platforms.

© 2025 by the authors; licensee Learning Gate

2.1. Theoretical Foundation and Model Innovation

This study builds on an integrated theoretical framework that combines semiotic theory, human-centered AI design, and emerging views of generative AI as a narrative agent in cultural tourism. Together, these elements enable a multilingual, emotionally engaging, and culturally specific storytelling model. From a semiotic perspective, the framework follows Baudrillard's theory of symbolic representation, emphasizing the need to preserve cultural authenticity in digital interpretation. The model addresses this by training on relevant cultural datasets and fine-tuning with structure-aware language generation, aligning with Landini [4] call for ethical, sustainable AI [4].

The model conceptualizes generative AI not merely as a recommendation tool, but as an active co-creator in the construction of cultural narratives. This aligns with recent research on AI-generated folklore [2] and GPT-powered heritage storytelling [1, 9] reflecting a broader recognition of AI's ability to create both cultural and business value through adaptive, user-centered technologies [23]. Such perspectives are particularly relevant to tourism, where intelligent models must support personalized, real-time engagement in diverse cultural settings. The framework also incorporates multimodal interaction theory, combining GPT-4o for multilingual text generation, DALL·E and Hugging Face APIs for personalized image creation, and Retrieval-Augmented Generation (RAG) with PostgreSQL for context-aware recommendations. The importance of visual storytelling in AI applications is well established [3] while co-creative, emotionally rich design approaches have proven effective in heritage education and museum interpretation [6].

Technically, the model employs a hybrid architecture consisting of (1) GPT-4o for adaptive narrative generation, (2) a PostgreSQL-based Retrieval-Augmented Generation (RAG) database for personalized recommendations, (3) custom image generation for visual enrichment, and (4) API integration for real-time environmental and event data. This architecture builds on recent advances in generative AI, particularly in visual storytelling [3], multimodal interaction [9] and hybrid XR narrative design [10] which have demonstrated the potential of AI to deliver more engaging, culturally nuanced experiences in heritage and tourism contexts. By integrating these foundations, the proposed model addresses key gaps identified in tourism AI research [22, 24] moving beyond static information delivery toward adaptive, emotionally resonant, and culturally specific storytelling. It supports inclusive AI innovation in cultural tourism [7] and contributes to emerging concepts of AI-generated imaginative geographies [5].

2.2. AI for Emotion Analysis and User-Centric Recommendation

Emotion-aware AI and intelligent recommendation models are increasingly central to modern tourism design. For example, supervised learning approaches such as Naïve Bayes and SVM have demonstrated high precision in detecting emotional cues from travel blogs [25] highlighting the potential of sentiment analysis to enhance AI-driven recommendations. Recent research has expanded beyond basic emotion detection to enable context-specific, real-time suggestions. Notably, AI models for tourist guidance [18] bio-inspired mobile recommendation models in religious tourism [19] and advanced POI personalization frameworks [20, 21] all demonstrate the growing sophistication of AI in understanding user preferences and situational needs.

Gamification has further enhanced AI-driven engagement. Studies show that AI-powered gamified apps boost tourist participation, while AI attraction mapping and smart tourism platforms improve decision-making and service personalization [26, 27]. Conversational AI, particularly chatbots, plays a key role in facilitating real-time tourist assistance. Research highlights how AI chatbots reduce loneliness, improve emotional well-being [28] and enable seamless multilingual interactions via NLP-powered voice assistants [29]. Hybrid models combining deep learning and collaborative filtering have achieved high accuracy in addressing tourist queries and guiding city exploration [30, 31].

Despite these advances, few studies explore generative AI as a cultural storyteller. Addressing this gap, the present study introduces a generative AI-powered tourism guidance model that integrates GPT-4o's narrative capabilities with DALL·E and Hugging Face image generation. The model serves

© 2025 by the authors; licensee Learning Gate

as a culturally adaptive digital guide, delivering real-time, personalized storytelling that fosters deeper emotional connection, cultural understanding, and visitor satisfaction at heritage sites.

In recent years, Generative AI has gained significant momentum in tourism research, particularly after the introduction of advanced models like ChatGPT-4o and image generation models [1, 3]. Bibliometric analysis by To and Yu [22] highlights a sharp increase in AI publications in tourism and hospitality since 2023, with Generative AI emerging as a key focus area. Similarly, Fouad, et al. [24] provide a comprehensive review identifying Generative AI's potential for enhancing user experiences, but also note the persistent gap in multimodal, narrative-based applications for cultural tourism [24]. These findings underscore the need for integrated, adaptive AI models that deliver personalized, context-aware storytelling, addressing both linguistic and cultural diversity—a gap this study directly targets.

Table 1.
Summary of Recent Literature Review on Generative AI in Cultural Tourism in 2025.

Author (Year)	Methods	Contribution	Gap/ Limitation
Mogre, et al. [1]	GPT-based narrative architecture	AI storytelling for built heritage	Lacks emotional and interactive elements
Oztabak [2]	Folklore text generation	Culturally resonant folklore content	No multilingual or mobile deployment
Martinez, et al. [3]	AI-based visual storytelling analysis	Visual narratives through AI-generated imagery	Lacks real-time, context-aware interaction
Alkhateeb, et al. [5]	Generative AI for landscape representation	Cultural scenic generation	No link to interactive tour experiences
George and Mattathil [7]	Qualitative design with AI	Inclusive AI for cultural tourism	Business-centric, limited focus on user experience
Massidda, et al. [10]	Hybrid XR and multisensory design	Authoring tools for cultural storytelling	Focused on authoring tools, not real-time adaptive personalization
Alharbi, et al. [19]	Chatbot user experience survey	AI-guided religious tourism	Contextual only, lacks generative storytelling
Sidiq and Sahman [25]	Model-driven AI personalization	Smart, tailored tourism services	No integration of visual cultural narratives
To and Yu [22]	Mapping analysis of AI in tourism & hospitality	Identified a sharp increase in AI research—with a clear rise in generative AI (e.g. ChatGPT)	Existing models lack interactive, itinerary-linked, narrative-based functionalities.

2.3. Conceptual Framework

The conceptual foundation of this study adopts an integrated, human-centered AI approach to tourism guidance, emphasizing the role of multimodal artificial intelligence in enriching visitor experiences. By combining advanced natural language processing for text and speech, visual content generation, dynamic content management, personalized recommendations, and intuitive user interaction, the proposed model addresses key challenges faced by modern tourists.

The framework positions AI not merely as a passive information provider but as an active co-creator of personalized, culturally specific narratives, offering real-time, linguistically inclusive, and emotionally resonant guidance [32]. At its core, the model integrates OpenAI's GPT-4o for natural language and speech processing [1, 9]. DALL·E and Hugging Face APIs for personalized image generation [3, 13] and a PostgreSQL-powered Retrieval-Augmented Generation (RAG) model for context-aware knowledge retrieval and adaptive recommendations [33]. This human-centered, AI-driven framework enhances cultural tourism by delivering contextually appropriate, personalized, and visually enriched storytelling that fosters deeper visitor engagement, cultural understanding, and satisfaction.

© 2025 by the authors; licensee Learning Gate

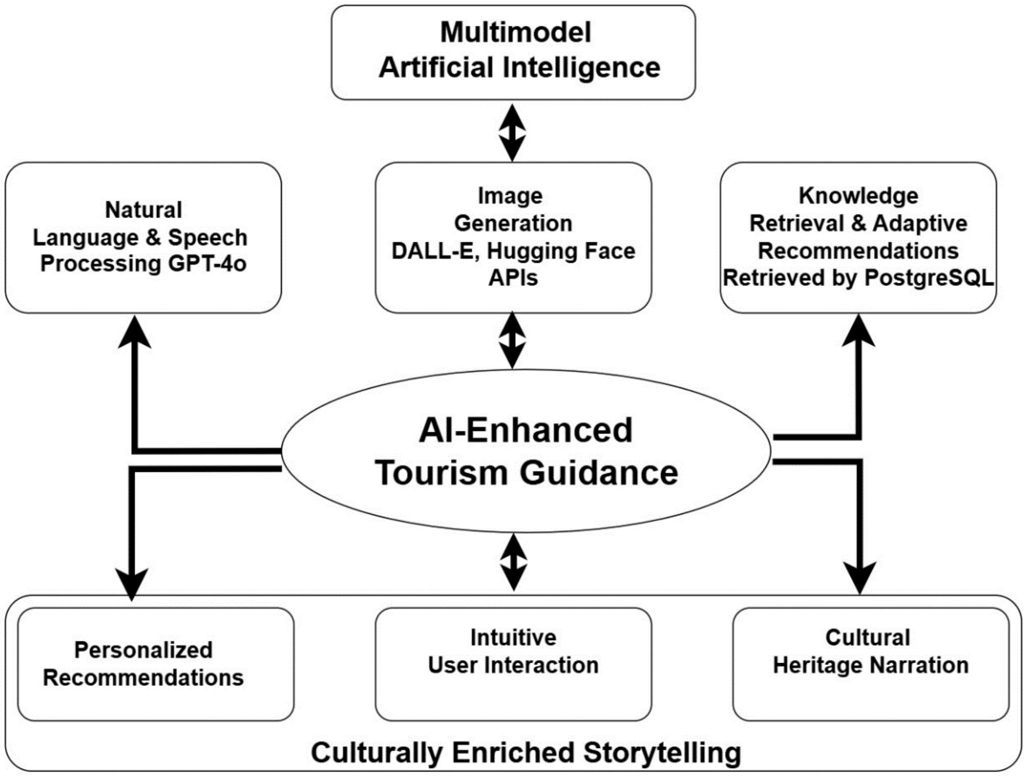

Figure 1.
The Conceptual Framework of AI-Enhanced Multimodal Tourism Guidance.

3. Methodology

The development of generative AI-powered storytelling model for cultural tourism was grounded in a comprehensive, data-driven methodology designed to ensure the model's practical effectiveness and reliability in real-world settings.

The model architecture comprises several interconnected AI components operating seamlessly to provide personalized, real-time support to tourists:

- Large Language Model (LLM, GPT-4o): Enables natural language understanding and multilingual response generation
- Retrieval-Augmented Generation (RAG) with PostgreSQL: Facilitates knowledge retrieval, personalization, and context adaptation based on user queries, preferences, and location data
- Custom Image Generation Pipeline (DALL·E, Hugging Face APIs): Produces visually appealing, culturally relevant content to complement textual responses

To ensure the delivery of accurate, adaptive, and culturally appropriate information, the model accesses dynamically curated databases of points of interest, accommodations, events, and other tourism-related content through real-time Application Programming Interfaces (APIs). Additionally, custom actions enable the chatbot to retrieve up-to-date information such as weather conditions, hotel availability, and event schedules from external sources.

A human-centered, systematic approach was applied throughout model development, including:

- Structured data collection from 400 tourists representing diverse cultural backgrounds
- An 80/20 train-test split to facilitate unbiased evaluation
- Iterative model training over 100 epochs, with continuous monitoring of performance metrics such as loss, accuracy, and F1-score

© 2025 by the authors; licensee Learning Gate

- Sentiment analysis to assess user acceptance and guide model refinements

This integrated methodology ensures that the AI-enhanced tourism guidance model can deliver reliable, engaging, and personalized experiences, aligning with the evolving expectations of modern, culturally diverse tourists.

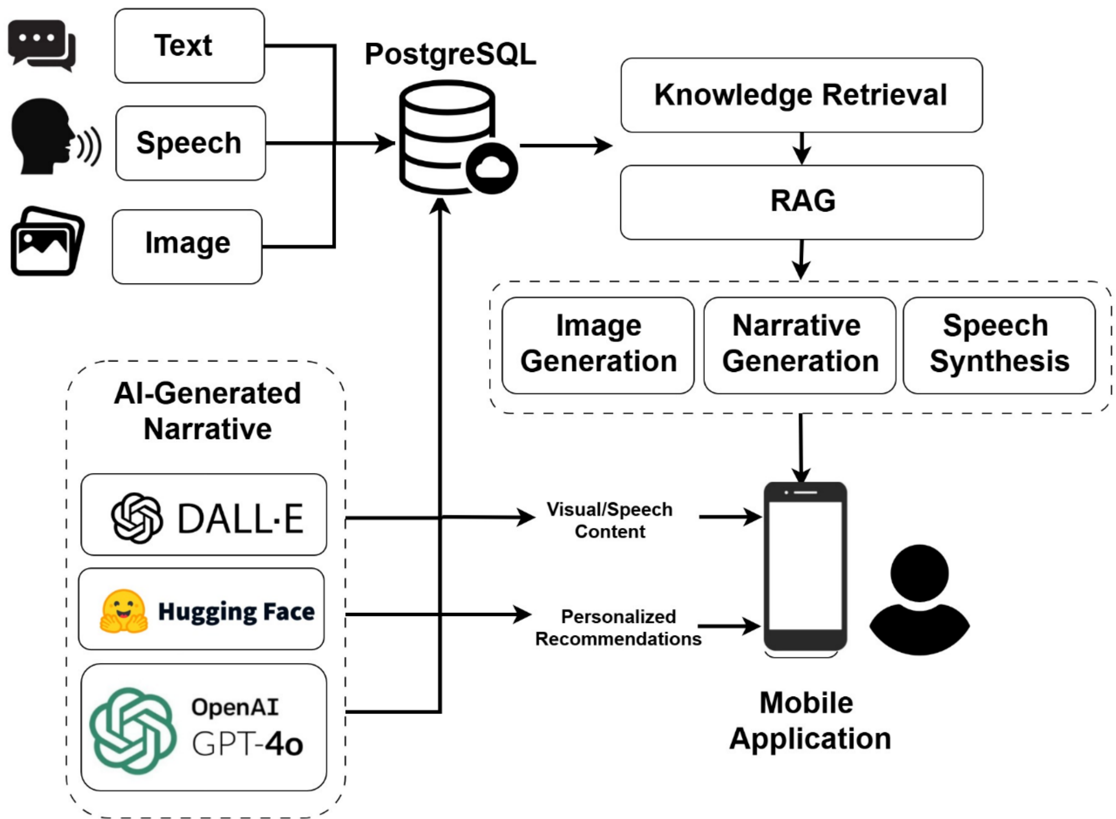

Figure 2.
Architecture of AI based Multi-Model Tourism Guidance.

3.1. Dataset Description

This study utilized two complementary datasets to develop and evaluate the AI-powered tourism guidance model. The first dataset was collected from 400 tourists during controlled field trials at four major cultural heritage sites in Bangkok, Thailand. Participants were recruited through stratified random sampling to ensure diversity in nationality, age, and language proficiency. The group included both Thai and international tourists from China, Japan, Europe, ASEAN, and other regions. The sample size was calculated using [34] with a 95% confidence level and a 5% margin of error. The dataset included:

- Demographics: Age, gender, nationality, language proficiency
- Travel behavior: Frequency of cultural site visits, preferred information formats (text, audio, visual)
- Attitudes toward AI: Expectations, concerns, and willingness to engage with AI-generated narratives

© 2025 by the authors; licensee Learning Gate

- Interaction data: Text and speech inputs, AI-generated images, sentiment feedback (positive, neutral, negative), and AI-generated captions evaluated for semantic similarity to human references

The second dataset was collected during the 12-month real-world deployment of the mobile application at the same heritage sites. It included anonymized interaction logs from thousands of tourists, covering:

- Click-throughs on AI recommendations
- Selection of points of interest (POIs)
- Engagement with AI-generated narratives and images
- Usage patterns such as language preference and interaction duration

The structured dataset was used for model training and validation (see Section 3.2), while the behavioral dataset supported real-world platform evaluation (see Section 5).

Table 2.
Descriptive Statistics of Participants.

Group	Samples (n)	%	Mean Age (Years)	SD (Age)
Domestic (Thailand)	120	30.00%	29.5	7.2
China	85	21.25%	27.3	6.5
Japan	50	12.50%	28.1	5.9
Europe	60	15.00%	32.4	8
Neighboring Countries (ASEAN)	55	13.75%	30	6.8
Other Regions	30	7.50%	31.7	7.5
Total	**400**	**100%**	—	—

Note: Neighboring Countries include visitors from ASEAN member countries such as Malaysia, Singapore, and Vietnam.

3.2. Data Preparation

A systematic data preparation process was undertaken to ensure reliable model training and robust performance evaluation. The structured dataset, collected from 400 tourists as outlined in Section 3.1, served as the foundation for AI model development.

The dataset was partitioned using an 80/20 train-test split, enabling the model to learn from a diverse set of user inputs while reserving a portion of the data for unbiased assessment of its generalization capability.

The data preparation steps included:

- Text and speech inputs were cleaned, tokenized, and encoded to ensure compatibility with the GPT-4o model.
- Image caption data was preprocessed to facilitate training of the image generation and captioning pipeline, leveraging DALL·E and Hugging Face APIs.
- Sentiment labels, derived from participant feedback, were categorized as positive, neutral, or negative to guide model refinement and support the development of emotionally resonant model responses.
- Semantic similarity scores between AI-generated captions and human-authored references were computed to support subsequent evaluation of narrative quality and contextual relevance.

To prevent data leakage and ensure the validity of applied model evaluation, all interaction logs and behavioral data obtained during real-world deployment were strictly excluded from the training process. These data were reserved exclusively for performance evaluation in the applied case study (see Section 5).

This structured and controlled approach ensured the AI model's capacity to deliver personalized, adaptive, and culturally relevant interactions while maintaining the integrity and independence of real-world model validation.

© 2025 by the authors; licensee Learning Gate

3.3. Model Training

Following data preparation, the AI model was trained to enable personalized, adaptive, and culturally relevant interactions with tourists. The model architecture integrated several key AI components to deliver a comprehensive, multimodal experience:

- Large Language Model (LLM): Facilitates natural language understanding and response generation
- Retrieval-Augmented Generation (RAG) System: Powered by PostgreSQL for real-time knowledge retrieval and contextual adaptation
- Custom Image Generation Pipeline: Utilizes DALL·E and Hugging Face APIs to produce visually enriched and culturally appropriate content

The model was trained over 100 epochs, using a combined loss function to optimize both textual and visual outputs:

- Mean Squared Error (MSE): Minimized discrepancies between AI-generated responses or image captions and human references
- Kullback-Leibler (KL Divergence): Encouraged the model to generate diverse yet contextually consistent content, enhancing the quality of AI interactions

Throughout the training process, performance metrics such as loss, accuracy, and F1-Score were closely monitored. This ensured the model could deliver reliable, high-quality responses that met the expectations of tourists in real-world cultural tourism scenarios.

3.4. Model Evaluation

Following model training, a series of controlled experiments were conducted to assess the model's effectiveness in real-world tourism scenarios. These tests simulated typical tourist interactions ranging from simple greetings and general inquiries to more complex tasks such as:

- Providing personalized, location-specific recommendations
- Generating AI-assisted cultural narratives
- Delivering visually enriched content through AI-generated images and descriptions

The evaluation process focused on two key dimensions, aligning with prior research highlighting the importance of machine learning-driven personalization and intelligent decision-making in complex, user-centric environments [35, 36].

3.4.1. Conversational and Recommendation Performance

1. Accuracy: Measures the proportion of correctly predicted outcomes (both positive and negative) relative to the total number of predictions.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

2. Precision: Indicates the proportion of true positive predictions out of all positive predictions made by the model. High precision reflects fewer false positives.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

3. Recall: Reflects the model's ability to correctly identify all actual positive cases, representing its sensitivity to relevant outcomes.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

F1-Score: Represents the harmonic mean of precision and recall, providing a balanced measure of the model's overall accuracy, especially useful in cases with class imbalance.

$$\text{F1-Score} = 2 \times \frac{(Precision\ x\ Recall)}{(Precision+Recall)} \quad (4)$$

where:

TP : True Positives — Correct positive predictions

TN : True Negatives — Correct negative predictions

ISSN: 2576-8484
Vol. 9, No. 7: 1436-1451, 2025
DOI: 10.55214/25768484.v9i7.8936
© 2025 by the authors; licensee Learning Gate

FP : False Positives — Incorrect positive predictions
FN : False Negatives — Incorrect negative predictions

3.4.2. Visual Content Generation Quality

In addition to conversational evaluation, the quality of AI-generated image descriptions—which support visual storytelling for cultural tourism—was assessed using cosine similarity, comparing AI-generated captions with human-authored references from the test dataset. The formula for cosine similarity is as follows:

$$\text{Cosine similarity} = \frac{A.B}{\|A\|.\|B\|} \quad (5)$$

where:

A and B represent the TF-IDF vector representations of two text descriptions

$A \cdot B$ is the dot product of the two vectors

$\|A\|.\|B\|$ are the magnitudes (norms) of the respective vectors

A higher cosine similarity score indicates greater semantic alignment between the AI-generated and human-authored descriptions, reflecting the model's ability to generate contextually appropriate and meaningful visual content.

4. Findings

To evaluate the performance of the proposed generative AI model for personalized storytelling in cultural tourism, an empirical experiment was conducted using data collected from 400 participants, comprising both domestic and international tourists who visited four major cultural heritage sites in Bangkok. These participants represented diverse backgrounds, including visitors from Thailand, China, Japan, Europe, neighboring ASEAN countries, and other regions. The dataset was divided into training and testing sets using an 80/20 split approach, ensuring a robust evaluation of the model's generalization capability.

4.1. Model Training and Learning Progress

The model was trained over 100 epochs, with its learning progress continuously monitored through loss and accuracy metrics. As illustrated in Figure 3, the loss values exhibited a consistent downward trend, reflecting the model's ability to minimize prediction errors during training. In parallel, the accuracy steadily increased and stabilized at approximately 95%, indicating the model's growing proficiency in understanding and classifying complex, multimodal tourism data.

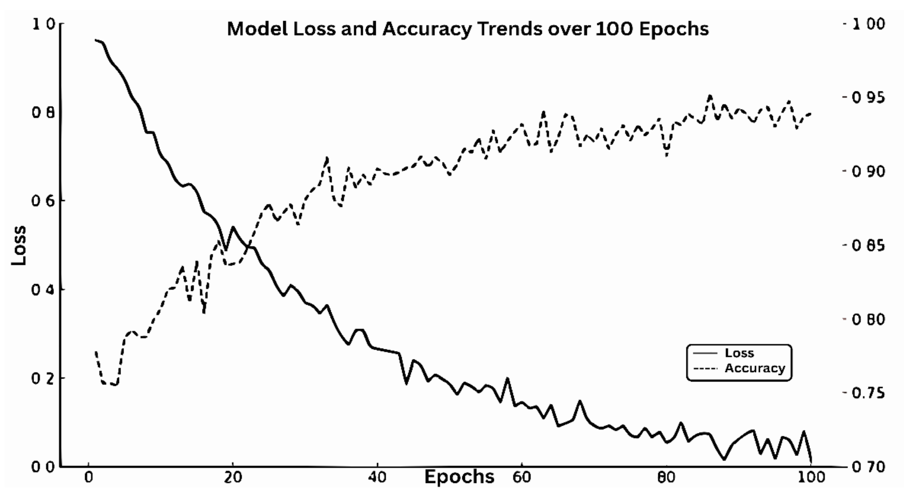

Figure 3.
Model Loss and Accuracy Trends over 100 Epochs of Training.

ISSN: 2576-8484
Vol. 9, No. 7: 1436-1451, 2025
DOI: 10.55214/25768484.v9i7.8936
© 2025 by the authors; licensee Learning Gate

4.2. Model Performance Evaluation

Following the completion of training, the model was evaluated using the reserved test dataset. The evaluation was conducted based on 880 interaction instances, representing text inputs, speech queries, and image prompts submitted by users during model testing. The classification results are summarized in the confusion matrix below:

Table 3.
Classification Outcomes and Performance Metrics of Generative AI Model

Item	Value
True Positives (TP)	496
True Negatives (TN)	273
False Positives (FP)	74
False Negatives (FN)	37
Precision	87.02%
Recall	93.06%
F1-Score	89.94%
Accuracy	87.39%

These results demonstrate the model's high reliability in accurately identifying user intents and generating relevant, context-aware responses across diverse interaction scenarios. In addition, sentiment analysis of user feedback achieved 85.4% classification accuracy, providing reliable insights into tourist emotional responses and supporting model refinement for improved narrative delivery.

4.3. Caption Generation Evaluation

Beyond classification performance, the model's capability to generate meaningful and contextually appropriate image captions was assessed by comparing AI-generated captions to human-labeled references using semantic similarity analysis. The similarity scores ranged from 0.80 to 0.95, highlighting the model's ability to produce coherent and culturally relevant visual descriptions.

Table 4.
Caption Generation Similarity Scores Compared to Human-Labeled Data.

Test Sample	Human-Labeled Caption	AI-Generated Caption	Similarity Score
1	"Tourists explore the temple courtyard"	"Visitors walk through the temple area"	0.92
2	"Traditional dance performance on stage"	"Cultural dance show at the stage"	0.89
3	"Street market with local food vendors"	"Local street vendors sell traditional food"	0.85
4	"Boats along the riverside during sunset"	"Small boats seen at riverside at dusk"	0.9
5	"Visitors viewing ancient mural paintings"	"Tourists observe mural art on walls"	0.95
6	"Children playing near the cultural center"	"Kids enjoy games close to cultural building"	0.8
7	"Tourists taking photos of the pagoda"	"Visitors capture images of the pagoda"	0.91
8	"Locals and tourists enjoy night festival"	"Festival-goers participate in night event"	0.88

Note: Similarity scores were calculated using semantic similarity measures between human-labeled and AI-generated captions.

4.4. Model Comparison and Performance Analysis

To further validate the effectiveness of the proposed generative AI-powered storytelling model, a comparative evaluation was conducted against baseline models, including Random Forest and Gradient Boosting, commonly used in classification tasks. This approach aligns with recent research emphasizing the importance of benchmarking AI models in applied decision-making contexts, where Random Forest and Gradient Boosting remain competitive baselines in classification tasks [37, 38]. Moreover, comparative studies show the growing effectiveness of generative AI models in outperforming these traditional classifiers under specific conditions [39].

© 2025 by the authors; licensee Learning Gate

As illustrated in Figure 4, the Receiver Operating Characteristic (ROC) curves compare the proposed model with the baseline models. The x-axis represents the False Positive Rate, while the y-axis represents the True Positive Rate. A larger area under the curve (AUC) reflects superior discrimination capability.

The results demonstrate that the proposed AI model (depicted with a solid black line) outperforms the baseline methods, achieving superior classification performance across various decision thresholds.

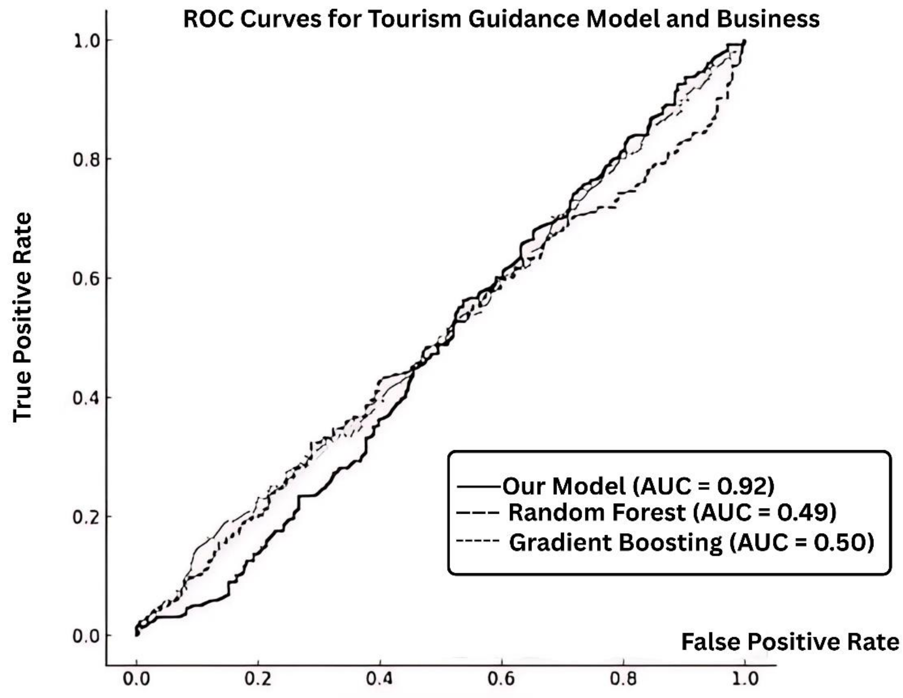

Figure 4.
ROC Curves for Multimodal AI Tourism Guidance Model Compared to Baseline Methods.

In addition, a subgroup analysis was performed to assess model performance across different tourist segments. The results are presented in Table 5, highlighting key metrics, including AUC-ROC, Precision, Recall, F1-Score, and Accuracy for each group.

Table 5.
Performance Metrics Across Tourist Segments.

Metric	Overall	Thai Tourists	International Tourists	Frequent Visitors	First-time Visitors
AUC-ROC	0.92	0.94	0.91	0.93	0.89
Precision	0.87	0.89	0.86	0.88	0.85
Recall	0.91	0.93	0.89	0.92	0.87
F1-Score	0.89	0.91	0.87	0.9	0.86
Accuracy	0.88	0.9	0.87	0.89	0.86

4.5. Case Study: Application of the Model in Real Tourism Scenarios

To assess the real-world applicability of the AI-powered multimodal tourism guidance model, an observational field study with pre-post comparisons was conducted over a 12-month period at four major cultural heritage sites in Bangkok, Thailand. This deployment followed the controlled field trials and model training with 400 tourists, as described in Section 3.1.

© 2025 by the authors; licensee Learning Gate

During the study period, anonymized interaction data were collected from approximately 12,400 unique tourist sessions (averaging roughly 1,000 sessions per site per quarter). These interactions included multilingual text queries, speech inputs, engagement with AI-generated visual content, and click-through events on AI-recommended points of interest (POIs).

To evaluate model effectiveness, key performance indicators (KPIs) were compared between the initial deployment phase (first month) and subsequent periods. Additionally, follow-up survey responses were obtained from a randomly selected sub-sample of 600 tourists, providing self-reported measures of cultural understanding and visitor satisfaction.

The following results summarize the observed improvements relative to the pre-deployment baseline:

- Attraction Recommendation Click-Through Rate increased by 17.6% ($p < 0.05$)
- Venue Visit Conversion Rate rose by 12.8% ($p < 0.05$)
- Cultural Understanding Self-Assessment Scores improved by 9.4% ($p < 0.05$)
- Visitor Retention Intention (likelihood of future visit or recommendation) increased by 11.2% ($p < 0.05$)

These findings align with prior research on AI-driven tourism personalization Andrianto, et al. [18] and Alharbi, et al. [19] and indicate the model's potential to enhance tourist engagement and satisfaction in real-world heritage settings.

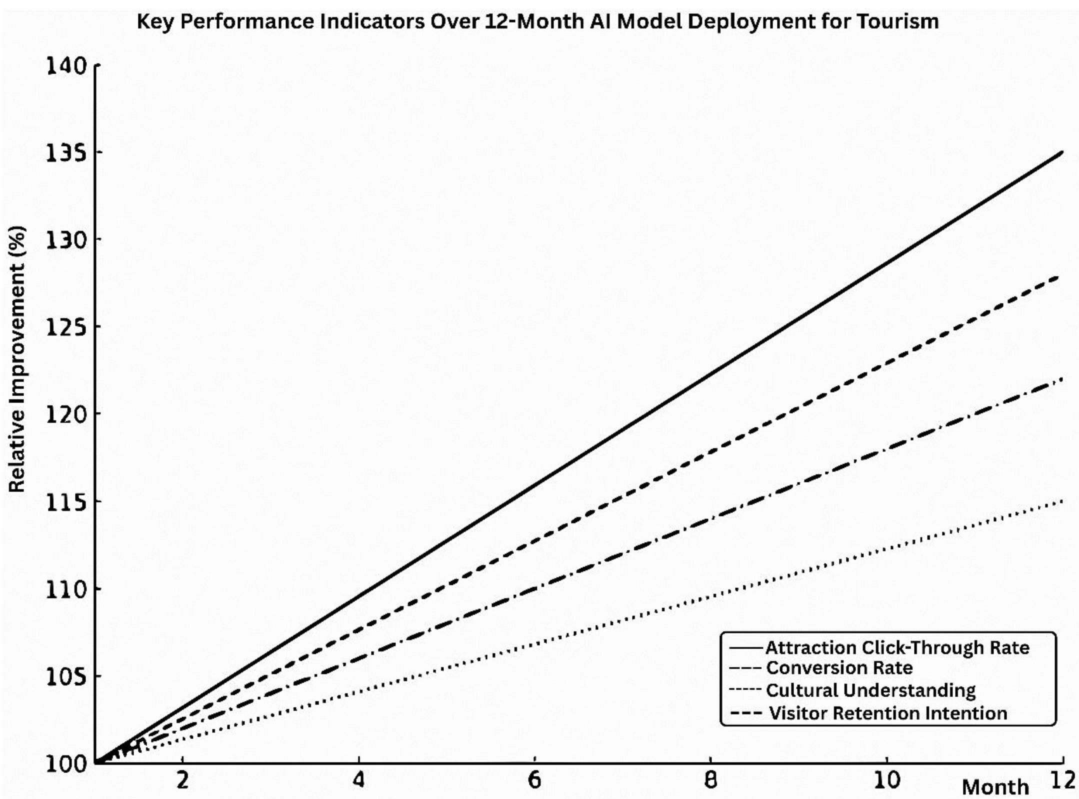

Figure 5.
Key Performance Indicators Over 12-Month AI Model Deployment for Tourism.

Further analysis of the model's effectiveness across distinct tourist segments revealed consistent and statistically significant performance improvements, as summarized in Table 6. The findings indicate

© 2025 by the authors; licensee Learning Gate

that the model positively influences key engagement and cultural indicators, including click-through rate, conversion rate, cultural understanding, and retention intention, across both domestic and international tourist groups.

Table 6.
Performance Metrics Improvement Across Tourist Segments

Tourist Segment	Click-Through Rate	Conversion Rate	Cultural Understanding	Retention Intention
Thai Tourists	+18.5% ($p < 0.05$)	+13.2% ($p < 0.05$)	+10.3% ($p < 0.05$)	+12.5% ($p < 0.05$)
International Tourists	+16.8% ($p < 0.05$)	+12.4% ($p < 0.05$)	+8.7% ($p < 0.05$)	+10.6% ($p < 0.05$)
Repeat Visitors	+15.9% ($p < 0.05$)	+11.1% ($p < 0.05$)	+9.2% ($p < 0.05$)	+10.4% ($p < 0.05$)
First-Time Visitors	+18.1% ($p < 0.05$)	+13.5% ($p < 0.05$)	+9.6% ($p < 0.05$)	+11.8% ($p < 0.05$)

Note: Improvements were measured against model usage data from the first month of deployment. Statistical significance was confirmed using paired t-tests at a 95% confidence level.

While the results highlight the model's potential to enhance short-term tourist engagement and cultural understanding, they primarily reflect initial user interactions. Longitudinal research is required to assess its long-term effectiveness, including knowledge retention, sustained behavior change, and cross-cultural applicability. Future studies should also consider how factors such as prior cultural exposure and digital literacy may influence these outcomes.

5. Discussion

The applicability of generative AI in enhancing immersive cultural tourism experiences through adaptive, multimodal, and tailored storytelling was a central focus of this research. The empirical results provide clear support for this objective, demonstrating that AI-based digital storytelling can effectively increase visitor engagement and promote deeper cultural understanding in heritage environments. These findings align with the growing body of research emphasizing AI's role in modernizing tourism services [18, 19].

Beyond its demonstrated technical capabilities, the implementation of generative AI in this context raises important theoretical and practical considerations. The model's performance, as evidenced by precision (87.02%), recall (93.06%), F1-score (89.94%), and high semantic similarity scores for AI-generated captions (0.80–0.95), highlights its ability to process user input and generate contextually relevant narratives. These results correspond with existing research highlighting the contributions of adaptive AI and deep learning algorithms to real-time content delivery and enhanced user experiences in tourism [40, 41] and business applications [36].

Building upon previous research, the present work positions generative AI not merely as a functional tool for itinerary planning or information delivery [19, 21] but as an active agent in the co-creation of immersive, narrative-driven cultural experiences. Through the integration of multimodal inputs (text, audio, and image prompts) and outputs (visual content and personalized recommendations), the model demonstrates how AI can facilitate interactive and emotionally resonant heritage interpretation. This development reflects a broader technological shift in which AI increasingly enables personalized and engaging tourism experiences [18, 23].

However, the application of AI in cultural tourism introduces considerable ethical and epistemological challenges. While the model successfully generated adaptive content, concerns remain regarding the risk of algorithmic storytelling oversimplifying or misrepresenting complex cultural meanings—issues that have been widely discussed in the literature on AI and cultural authenticity [42]. To mitigate these risks, the model incorporated a semiotic-inspired framework informed by Baudrillard's theory to help preserve the symbolic richness of cultural expressions. This approach is consistent with contemporary academic discourse, which emphasizes the importance of preserving cultural integrity and ensuring that cultural richness is maintained in AI-driven heritage tourism experiences [6, 12].

© 2025 by the authors; licensee Learning Gate

In addition to its theoretical contributions, the model's practical deployment yielded consistent improvements in key performance indicators, including click-through rates, conversion rates, cultural understanding, and visitor retention across diverse tourist groups. These empirical results reinforce the potential of generative AI as both a scientific innovation and a practical tool for advancing tourism development strategies [20, 26, 29].

Finally, this work contributes to a broader understanding of AI's transformative role in creating business value and supporting informed decision-making. The results underline the potential of generative AI to deliver personalized, context-aware cultural interpretation, while also emphasizing the importance of responsible implementation to safeguard cultural integrity [23, 43].

6. Conclusion

This study demonstrates the feasibility of leveraging generative AI to deliver culturally immersive tourism experiences. Through the integration of GPT-4o, a RAG database, and custom image generation pipelines, the model provided personalized, real-time narratives and visuals that enriched visitor experiences at four heritage sites in Bangkok, Thailand.

The results confirmed that the AI-based platform effectively enhanced user engagement, cultural learning, and overall visitor satisfaction. Furthermore, the model demonstrated strong performance and consistent improvements across key real-world indicators, validating its potential to support more meaningful and inclusive cultural tourism experiences.

Despite these promising outcomes, several limitations must be acknowledged. The model evaluation was conducted through short-term deployments at selected heritage sites, leaving long-term impacts on cultural knowledge retention and visitor behavior unexamined. In addition, although cultural sensitivity was considered during AI content generation, challenges related to authenticity, bias, and broader ethical implications remain and require further exploration.

Future research should investigate the longitudinal effects of AI-enhanced storytelling on cultural learning, expand testing across diverse cultural and linguistic contexts, and develop comprehensive ethical guidelines to ensure responsible use of generative AI in tourism. Such efforts are essential to ensure that AI technologies not only enhance visitor experiences but also contribute to sustainable, culturally sensitive, and inclusive tourism development in Thailand and beyond.

Transparency:

The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

Copyright:

© 2025 by the authors. This open-access article is distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

References

[1] S. C. Mogre, J. S. Nikam, P. Damani, and A. Banik, "Generative AI for cultural heritage preservation using ar and data science," in *2025 International Conference on Machine Learning and Autonomous Systems (ICMLAS)*, 2025: IEEE, pp. 1068-1073, doi: https://doi.org/10.1109/ICMLAS64557.2025.10968619. https://doi.org/10.1109/ICMLAS64557.2025.10968619

[2] C. Oztabak, "AI-generated folklore for culturally resonant storytelling in the digital age: Bridging tradition and technology in storytelling," in Understanding Generative AI in a Cultural Context: Artificial Myths and Human Realities: IGI Global Scientific Publishing, 2025, pp. 161-184. https://doi.org/10.4018/979-8-3693-7235-7.ch007

[3] L. Martinez, B. Caramiaux, and S. Fdili Alaoui, "Generative AI in documentary photography: Exploring opportunities and challenges for visual storytelling," in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 2025, pp. 1-13, doi: https://doi.org/10.1145/3706598.3714200. https://doi.org/10.1145/3706598.3714200

[4] G. G. Landini, "Traditional Knowledge, Environmental Challenges, and Artificial Intelligence: Ethical Generative AI Use and Sustainable Approaches," in The routledge handbook of artificial intelligence and international relations: Routledge, 2025, pp. 304-321. https://doi.org/10.4324/9781003518495-29

[5] G. Alkhateeb, J. Storie, S. Bell, and M. Suškevičs, "Virtual imaginative geographies: Generative ai and the representation of landscape imagery," Geographies, vol. 5, no. 1, p. 9, 2025. https://doi.org/10.3390/geographies5010009

[6] F. Condorelli and F. Berti, "Creativity and awareness in co-creation of art using artificial intelligence-based systems in heritage education," Heritage, vol. 8, no. 5, p. 157, 2025. https://doi.org/10.3390/heritage8050157

[7] B. George and A. P. Mattathil, "Empowering African American Tourism Entrepreneurs with Generative AI: Bridging Innovation and Cultural Heritage," Societies, vol. 15, no. 2, p. 34, 2025. https://doi.org/10.3390/soc15020034

[8] E. Creely and J. Blannin, "Creative partnerships with generative AI. Possibilities for education and beyond," Thinking Skills and Creativity, vol. 56, p. 101727, 2025. https://doi.org/10.1016/j.tsc.2024.101727

[9] B. Dilaveroglu, "The architecture of visual narrative: Can text-to-image algorithms enhance the power of stylistic narrative for architecture," International Journal of Architectural Computing, vol. 22, no. 3, pp. 432-457, 2024. https://doi.org/10.1177/14780771241234449

[10] M. Massidda, L. Travaglini, and S. Pescarin, "Designing hybrid and XR narrative applications: A framework to describe authoring tools in cultural domain," Digital Applications in Archaeology and Cultural Heritage, p. e00424, 2025. https://doi.org/10.1016/j.daach.2025.e00424

[11] C. H. Chuan, "A critical review of news framing of artificial intelligence," Handbook of Critical Studies of Artificial Intelligence, pp. 266–276, 2023. https://doi.org/10.4337/9781803928562.00029

[12] K. Feher and A. Generative, *Media, and society*, 1st ed. London: Routledge, 2025. https://doi.org/10.4324/9781003591023

[13] S. Barbosu and P. Khashabi, "Dali or DALL-E? Popper or …? The implications of emerging generative AI on the future of creative work," SSRN Electron Journal, 2023. https://doi.org/10.2139/ssrn.4586863

[14] A. Asrifan, M. F. Hasa, S. Syafryadin, and H. Pelu, "Narrative machines: The evolution of storytelling in the age of generative AI," in Generative AI Foundations, Developments, and Applications: IGI Global Scientific Publishing, 2025, pp. 81-110. https://doi.org/10.4018/979-8-3693-6255-6.ch004

[15] V. N. Antony and C.-M. Huang, "ID. 8: Co-creating visual stories with generative AI," ACM Transactions on Interactive Intelligent Systems, vol. 14, no. 3, pp. 1-29, 2025. https://doi.org/10.1145/3672277

[16] V. Attaluri and L. N. R. Mudunuri, "Generative AI for creative learning content creation: project-based learning and art generation," in Smart Education and Sustainable Learning Environments in Smart Cities: IGI Global Scientific Publishing, 2025, pp. 239-252. https://doi.org/10.4018/979-8-3693-7723-9.ch014

[17] F. Lys, *Creating stories: Generative artificial intelligence tools as writing tutors," in AI in language teaching, learning, and assessment.* Hershey, PA, USA: IGI Global, 2024, pp. 222–243. https://doi.org/10.4018/979-8-3693-0872-1.ch011

[18] T. Andrianto, T. M. Tangit, and N. C. Minh, "Adoption of artificial intelligence (AI) technology in enhancing tourist experience: A conceptual model," Journal of Tourism, Hospitality and Travel Management, vol. 3, no. 1, pp. 53-66, 2025. https://doi.org/10.58229/jthtm.v3i1.302

[19] A. Alharbi, A. Pandit, P. J. Rosenberger III, and S. Miah, "Understanding AI-enabled conversational agent customer experiences in religious tourism," Journal of Islamic Marketing, 2025. https://doi.org/10.1108/jima-07-2024-0309

[20] W. Chiengkul, P. Kumjorn, T. Tantipanichkul, and K. Suphan, "Engaging with AI in tourism: A key to enhancing smart experiences and emotional bonds," Asia-Pacific Journal of Business Administration, 2025. https://doi.org/10.1108/apjba-09-2024-0488

[21] L. Benaddi, C. Ouaddi, A. Jakimi, and B. Ouchao, "A systematic review of chatbots: Classification, development, and their impact on tourism," IEEE Access, 2024. https://doi.org/10.1109/access.2024.3408108

[22] W. M. To and B. T. Yu, "Artificial intelligence research in tourism and hospitality journals: trends, emerging themes, and the rise of generative AI," Tourism and Hospitality, vol. 6, no. 2, p. 63, 2025. https://doi.org/10.3390/tourhosp6020063

[23] I. M. Enholm, E. Papagiannidis, P. Mikalef, and J. Krogstie, "Artificial intelligence and business value: A literature review," Information Systems Frontiers, vol. 24, no. 5, pp. 1709-1734, 2022. https://doi.org/10.1007/s10796-021-10186-w

[24] A. M. Fouad, I. E. Salem, and E. A. Fathy, "Generative AI insights in tourism and hospitality: A comprehensive review and strategic research roadmap," Tourism and Hospitality Research, vol. 0, no. 0, p. 14673584241293125, 2024. https://doi.org/10.1177/14673584241293125

[25] Z. F. Sidiq and Z. Sahman, "The role of artificial intelligence in transforming smart tourism: Enhancing customer experience and service personalization," Journal of Sharia Economy and Islamic Tourism, vol. 5, no. 2, pp. 82-98, 2025. https://doi.org/10.31764/jseit.v5i2.30705

[26] A. R. Herawati, A. E. Pradana, and N. Shahira, "Integrating artificial intelligence (AI) into sustainable tourism management in Indonesia and the impact on customer experience: A literature review," in *International Conference on Medical Imaging, Electronic Imaging, Information Technologies, and Sensors (MIEITS 2025)*, 2025, vol. 13631: SPIE, pp. 24-35, doi: https://doi.org/10.1117/12.3058730. https://doi.org/10.1117/12.3058730

© 2025 by the authors; licensee Learning Gate

[27] H. Tosyali, F. Tosyali, and E. Coban-Tosyali, "Role of tourist-chatbot interaction on visit intention in tourism: the mediating role of destination image," Current Issues in Tourism, vol. 28, no. 4, pp. 511-526, 2025. https://doi.org/10.1080/13683500.2023.2288937

[28] S. Mishra, M. Anifa, and J. Naidu, "AI-driven personalization in tourism services: Enhancing tourist experiences and business performance," in Integrating Architecture and Design Into Sustainable Tourism Development: IGI Global Scientific Publishing, 2025, pp. 1-20.

[29] E. J. Bedu, P. Foroudi, and M. Palazzo, "Artificial intelligence (AI) systems, customer experience and preference in the tourism industry in Ghana: Focus on Accra Metropolis," in Contemporary Marketing and Consumer Behaviour in Sustainable Tourism: Routledge, 2024, pp. 208-222. https://doi.org/10.4324/9781003388593-15

[30] F. Alhashmi, M. Alhefeiti, S. B. Mirza, and F. L. Ridouane, "Efficient dorsal fin-based classification of Risso's and common Bottlenose dolphins using YOLOv7 and YOLOv8 models for real-time applications," International Journal of Advanced Technology and Engineering Exploration, vol. 11, no. 115, p. 875, 2024. https://doi.org/10.19101/ijatee.2023.10102512

[31] S. Chanrueang, S. Thammaboosadee, and H. Yu, "A personalized hybrid tourist destination recommendation system: An integration of emotion and sentiment approach," International Journal of Advanced Computer Science & Applications, vol. 15, no. 8, 2024. https://doi.org/10.14569/ijacsa.2024.0150803

[32] H. Pallathadka, E. H. Ramirez-Asis, T. P. Loli-Poma, K. Kaliyaperumal, R. J. M. Ventayen, and M. Naved, "Applications of artificial intelligence in business management, e-commerce and finance," Materials Today: Proceedings, vol. 80, pp. 2610-2613, 2023. https://doi.org/10.1016/j.matpr.2021.06.419

[33] P. Gomes, L. Verçosa, F. Melo, V. Silva, C. B. Filho, and B. Bezerra, "Artificial intelligence-based methods for business processes: A systematic literature review," Applied Sciences, vol. 12, no. 5, p. 2314, 2022. https://doi.org/10.3390/app12052314

[34] T. Yamane, *Statistics: An introductory analysis*, 2nd ed. New York: Harper and Row, 1967.

[35] M. O. Ayemowa, R. Ibrahim, and M. M. Khan, "Analysis of recommender system using generative artificial intelligence: A systematic literature review," SSRN Electron Journal, 2024. https://doi.org/10.2139/ssrn.4922584

[36] M. Hamzehi and S. Hosseini, "Business intelligence using machine learning algorithms," Multimedia Tools and Applications, vol. 81, no. 23, pp. 33233-33251, 2022. https://doi.org/10.1007/s11042-022-13132-3

[37] L. Grinsztajn, E. Oyallon, and G. Varoquaux, "Why do tree ensemble approximators not outperform the recursive-rule extraction algorithm?," Machine Learning and Knowledge Extraction, vol. 6, no. 1, pp. 658-678, 2024. https://doi.org/10.3390/make6010031

[38] M. Raparthi, D. Dhabliya, T. Kumari, R. Upadhyaya, and A. Sharma, "Implementation and performance comparison of gradient boosting algorithms for tabular data classification," in *Deep Learning and Visual Artificial Intelligence*, Singapore, V. Goar, A. Sharma, J. Shin, and M. F. Mridha, Eds., 2024// 2024: Springer Nature Singapore, pp. 461-479, doi: https://doi.org/10.1007/978-981-97-4533-3_36. https://doi.org/10.1007/978-981-97-4533-3_36

[39] M. A. Roshani *et al.*, "Generative large language model—powered conversational ai app for personalized risk assessment: Case study in covid-19," JMIR AI, vol. 4, no. 1, p. e67363, 2025. https://doi.org/10.2196/67363

[40] S. Agarwal, "Deep learning-based sentiment analysis: Establishing customer dimension as the lifeblood of business management," Global Business Review, vol. 23, no. 1, pp. 119-136, 2022. https://doi.org/10.1177/0972150919845160

[41] N. Rane, M. Paramesha, S. Choudhary, and J. Rane, "Business intelligence and business analytics with artificial intelligence and machine learning: Trends, techniques, and opportunities," Techniques, and Opportunities, 2024. https://doi.org/10.2139/ssrn.4831920

[42] B. Bosma and A. van Witteloostuijn, "Machine learning in international business," Journal of International Business Studies, vol. 55, no. 6, pp. 676-702, 2024. https://doi.org/10.1057/s41267-024-00687-6

[43] A. R. Doshi, J. J. Bell, E. Mirzayev, and B. S. Vanneste, "Generative artificial intelligence and evaluating strategic decisions," Strategic Management Journal, vol. 46, no. 3, pp. 583-610, 2025. https://doi.org/10.1002/smj.3677

© 2025 by the authors; licensee Learning Gate