Machine Learning Advances in Technology Applications: Cultural Heritage Tourism Trends in Experience Design

Meihua Deng
School of International Communication, Hunan Mass Media Vocational and Technical College,
Changsha 410100, Hunan, China

***Abstract*—This study investigates the evolving trends in cultural heritage tourism experience design and examines how machine learning technologies are being applied to enhance visitor engagement and heritage preservation. Using bibliometric data from the Web of Science (WoS) and visualization tools such as VOSviewer, the research identifies key themes, author collaborations, and keyword clusters from 2016 to 2025. The analysis reveals a shift in focus from traditional conservation and display methods to user-centered experiences supported by advanced technologies. Machine learning techniques—such as deep learning, natural language processing, and multimodal data fusion—are increasingly used to personalize tours, analyze tourist behavior, restore damaged artifacts, and improve decision-making in resource management. Tools like CNNs and BERT models enable smart guiding systems and interactive Q&A features, while sentiment analysis enhances feedback mechanisms. The study also highlights several ongoing challenges, including data privacy issues, algorithmic bias, and unequal access to technological infrastructure, especially in developing regions. Ethical considerations and the need for human-centered design principles are emphasized to ensure that technological innovation aligns with cultural values and sustainability goals. In conclusion, this research provides a comprehensive overview of academic progress in cultural heritage tourism and illustrates the growing importance of AI and machine learning in creating immersive, efficient, and culturally respectful tourism experiences. The findings offer practical insights for scholars, heritage site managers, and policymakers seeking to leverage digital tools for both preservation and enhanced visitor satisfaction.**

Keyword—Heritage tourism; tourism experience; machine learning; VOSviewer; bibliometric data

I. Introduction

Driven by both globalization and digitization, tourism has gradually become an important link between history and modernity, and between preservation and dissemination. Cultural heritage is not only a carrier of national memory but also an important resource for cultural identity and sustainable development in contemporary society [1]. However, with the rapid growth of tourism demand, it has become a common challenge for both academia and industry practice to provide tourists with both in-depth and interesting experiences under the premise of preserving the authenticity of tourism heritage [2]. Traditional display methods such as static exhibits and one-way explanations have made it difficult to meet tourists' needs for interactivity, personalization, and immersion, while the rise of artificial intelligence (AI) and machine learning technologies has injected new vitality into this field [3]. Through natural language processing, computer vision, data mining, and other technologies, machine learning can not only optimize the analysis of visitor behavior and resource management, but also promote digital preservation and innovative displays, and even reconfigure the interaction mode between visitors and culture.

The paradox of cultural heritage tourism lies in the balance between conservation and utilization. The fragility of tourism resources requires strict environmental control and physical protection; in addition, tourism development needs to attract public participation to realize its social value [4]. Traditional means such as interactive screens in physical museums or simple virtual reality (VR) experiences, although enhancing the sense of participation to a certain extent, still face problems such as homogenization of content and insufficient technical adaptability [5]. Some VR devices cause users to experience a sense of vertigo due to technical limitations, while AR applications often lack long-term appeal due to single content [6]. In recent years, breakthroughs in machine learning technology have provided a new path to solve these problems. Image recognition technology based on deep learning can carry out high-precision classification and repair of cultural relics, and natural language processing technology can realize personalized guided tours through intelligent question-and-answer systems, and even optimize the feedback mechanism for tourists by combining sentiment analysis [7]. The application of these technologies not only enhances the intelligent level of experience but also through data-driven dynamic adjustment.

At the theoretical level, this study integrates the multidisciplinary perspectives of heritage, tourism management, and computer science to build a trinity analytical framework of "technology-experience-protection", which makes up for the insufficiency of the existing literature that separates the application of technology and humanistic care [8]. The application of Laboratory Information Management System (LIMS) in heritage monitoring not only improves data management efficiency but also realizes preventive protection through environmental sensors and early warning mechanisms, reflecting the dual value of technological empowerment and cultural sustainability. At the practical level, the research findings can provide operable technical solutions for scenic spots [9]. Visitor evaluation models based on sentiment analysis can help managers identify service shortcomings, while multimodal data fusion techniques (e.g., combining image and text data) can assess the quality of experience more comprehensively. In addition, the case studies show that the successful application of AI technologies needs to be based on cross-disciplinary cooperation.

To systematically explore these topics, this paper is organized as follows: Section II provides a comprehensive review of existing literature on cultural heritage tourism and the integration of machine learning technologies. Section III details the methodology, including data collection strategies and the analytical framework used. Section IV presents the findings derived from bibliometric and visual analyses, supported by relevant case studies. Finally, Section V discusses the conclusions drawn from the research, outlines practical implications, and proposes directions for future studies.

II. Research Review

A. *Overview of Cultural Heritage Tourism Research*

Cultural heritage tourism is developing rapidly, driven by globalization and digitization. Its core objective is to realize cultural dissemination, education, and social value through the interaction between tourists and tourism resources [10]. This model of tourism not only emphasizes cultural and educational aspects, but also focuses on tourists' sense of participation and immersion, and seeks to find a balance between conservation and development [11]. In recent years, with the diversification of tourism needs and technological advances, related research has gradually shifted from traditional conservation and display to focus on visitor experience and technological innovation. In terms of the definition and characteristics of tourism, scholars generally agree that its uniqueness lies in the combination of culture, education, experience, and sustainability [12]. Cultural heritage tourism is not only an economic activity but also a means of cultural dissemination and education. In recent years, research hotspots have focused on visitor experience optimization, digital preservation, community participation and cultural identity, and tourism impact assessment [13]. The application of virtual reality (VR) and augmented reality (AR) technologies has provided visitors with immersive experiences, but how to avoid "de-culturalization" caused by the misuse of technology remains a challenge. The digitization project of the Mogao Grottoes in Dunhuang has reduced the loss of physical artifacts and provided visitors with a richer experience through high-precision scanning and virtual reconstruction [14]. In addition, the importance of community participation in heritage tourism is becoming more and more evident. In Japan, Shirakawa-go Hapo Village has successfully preserved its traditional architecture and culture through community-led tourism development [15]. However, related tourism resources still face challenges such as the contradiction between conservation and development, the limitations of technology application, and the conflict between globalization and localization. Over-commercialization may lead to "westernization" and weaken its cultural value, while misuse of technology may lead to "dehumanization" of the cultural experience.

In the future, research related to tourism will pay more attention to multidisciplinary integration and technological innovation. Optimizing tourists' behavior prediction and resource management through artificial intelligence and big data analysis, or realizing digital rights and protection through blockchain technology will become important directions [16]. At the same time, research should also pay attention to cultural ethics and community participation, and explore the path of "people-centered" sustainable development [17]. All in all, the research in this field not only provides rich theoretical perspectives for academics but also provides important guidance for conservation and development in industry practice.

B. *Review of Machine Learning Research*

Machine learning, as one of the core technologies of artificial intelligence, has shown great potential in the field of tourism in recent years. Its core lies in the algorithmic model to learn the laws from the data and use them for prediction, classification, and decision-making [18]. Machine learning is mainly divided into three categories: supervised learning, unsupervised learning, and reinforcement learning, in which supervised learning trains models through labeled data, unsupervised learning discovers potential laws from unlabeled data, and reinforcement learning optimizes decision-making through trial-and-error and reward mechanisms. Deep learning, as an important branch of machine learning, processes complex data through multi-layer neural networks and shows unique advantages in cultural heritage tourism [19]. Convolutional neural networks (CNN) can be used for cultural relics image classification and restoration, while recurrent neural networks (RNN) are suitable for tourists' behavior sequence analysis. The application of machine learning mainly focuses on the analysis and prediction of tourists' behavior, intelligent guiding and interactive experience, as well as resource management and decision support [20]. By analyzing tourists' browsing trajectories, consumption records, and comment data, machine learning can identify tourists' preferences and behavioral patterns and provide tourists with personalized tour routes. In terms of digital protection, machine learning uses deep learning algorithms to virtually repair damaged cultural relics, or automatically classify and label cultural relics using image recognition technology. Natural Language Processing (NLP) technology provides support for the intelligent tour system, and the NLP-based intelligent Q&A system can answer visitors' questions in real-time and enhance the interactive experience. In addition, the application of machine learning in scenic resource scheduling and risk assessment is gradually maturing, predicting visitor flow through time series analysis, or identifying safety hazards using anomaly detection algorithms.

However, the application of machine learning in tourism still faces many challenges. Data quality and privacy are the first obstacles, as data in the tourism sector is often noisy and missing, and the privacy protection of tourists' data is also an urgent issue to be solved. Algorithmic bias and interpretive issues should also not be ignored [21]. Recommender systems based on historical data may reinforce tourists' inherent preferences, limiting the experience of cultural diversity, while the "black-box" nature of deep-learning models makes it difficult to explain their decision-making process, affecting user trust. In addition, the application of machine learning technology requires high hardware and software investment, which limits its popularization in small and medium-sized scenic spots.

The application of machine learning in tourism will focus more on multimodal data fusion and interdisciplinary cooperation. Combining image, text, and sensor data to build a more comprehensive analysis model of tourist behavior. At the same time, research should also pay attention to technical ethics and user experience, and explore the "human-centered" technology adaptation path [22]. In conclusion, machine learning provides a new technical tool and research paradigm for tourism, but its successful application should be based on data quality, algorithmic fairness, and user needs, and seek a balance between technological innovation and humanistic values.

III. Methodology

A. Data Collection

In this study, the Web of Science (WoS) database was used as the data source, and VOSviewer software was used to visualize and analyze the literature data to reveal the research trends and hotspots in the field of cultural heritage tourism experience design. First, the WoS database was searched with keywords such as "cultural heritage" and "tourism" to filter out high-quality literature related to the topic, spanning from 2016 to 2025. By setting the type of literature as "Article" and "Review", and excluding the literature not related to the topic, we finally obtained an effective literature data set.

In the data pre-processing stage, the title, abstract, keywords and other information of the literature are cleaned and standardized, and the expression form of keywords is unified (e.g., "VR" and "virtual reality" are merged) to ensure the accuracy of the analysis. Subsequently, the processed data were imported into VOSviewer software, and its powerful network visualization function was used to construct the keyword co-occurrence network, author cooperation network, and literature co-citation network [23]. In the keyword co-occurrence analysis, high-frequency keywords were screened out by setting the minimum occurrence frequency threshold, and cluster maps were generated to visualize the research hotspots and their relevance [24]. In author collaboration network analysis, the core research power in the field is revealed by identifying high-output authors and their collaboration teams. In addition, literature co-citation analysis is used to identify high-impact literature and its research themes to further explore the knowledge base and evolution path in the field.

Through the visual analysis of VOSviewer, this study can not only identify the research hotspots in the field of cultural heritage tourism experience design (e.g., "digital preservation", "visitor experience optimization", "community participation", etc.) but also reveal the correlation between different research themes and their trends over time. This study not only identifies the research hotspots in the field of cultural heritage tourism experience design (e.g. "digital preservation", "tourist experience optimization", "community participation", etc.) but also reveals the correlation between different research themes and their trends over time. While early studies focused on the application of technological tools, in recent years more attention has been paid to user experience and sustainability. This visualization method provides an important literature base and theoretical support for subsequent research on the application of machine learning techniques.

B. Research Methods in Machine Learning

In the research method section of machine learning, this study adopts a data-driven approach, combining with the actual needs of the tourism field, to design and implement a series of machine learning models to optimize the design and protection of the tourist experience. First, data collection and preprocessing are the basis of machine learning research [25]. This study obtains data from multiple sources, including literature data in the WoS database, visitor behavior data (browsing trajectory, comment data) from cultural heritage scenic spots, and digitized data (images of cultural relics, 3D models) from tourism resources. High-quality training datasets are constructed through data cleaning, feature extraction, and normalization.

In the model selection and training phase, this study uses a variety of machine learning algorithms according to the specific task requirements. In the task of tourist behavior analysis and prediction, supervised learning algorithms (Random Forest, Support Vector Machine) are used to classify and regressively analyze tourists' preferences and behavioral patterns; in the task of digital protection of cultural heritage, deep learning algorithms (Convolutional Neural Networks CNN) are used to automatically classify and repair cultural relics images; and in the task of intelligent tour guiding and interactive experience, the natural language processing technology (BERT model) is used to construct an intelligent Q&A system to realize personalized tour guiding [26]. BERT model is used to build an intelligent Q&A system and realize personalized tours [27]. In addition, this study also tries to apply reinforcement learning to scenic resource scheduling, optimizing the visitor experience by dynamically adjusting the distribution of people flow. Among them, are the machine learning methods in the basic class, as shown in Table I.

In the model evaluation and optimization stage, cross-validation, confusion matrix, and ROC curve are used to evaluate the model performance, and the model precision is further improved by hyperparameter tuning and feature selection. In the tourist satisfaction prediction task, the optimal model is selected by comparing the accuracy and recall of different algorithms. Meanwhile, this study also focuses on the interpretability of the model and utilizes methods such as Shapley Additive Explanations (SHAP) values to explain the model decision-making process to enhance user trust. Among the machine learning methods in the application category, as shown in Table II.

TABLE I. MACHINE LEARNING METHODS FOR THE BASE CLASS

Research Methodology	Specific Techniques/Algorithms	Application Scenario	Data Sources	Assessment of Indicators
Supervised learning	Random Forest	Tourist behavior classification (e.g., preference analysis, satisfaction prediction)	Visitor comment data, browsing track data	Accuracy, Recall, F1 Score
	Support Vector Machines (SVM)	Tourist flow forecasts, scenic area resource demand forecasts	Historical visitor data, scenic area operation data	Mean Square Error (MSE), R² value
	Logistic Regression	Visitor satisfaction dichotomy (satisfied/dissatisfied)	Visitor questionnaire data	ROC curves, AUC values
unsupervised learning	K-means Clustering (K-means Clustering)	Clustering of tourist behavior patterns (e.g., high-spending groups, cultural preference groups)	Tourist consumption data, behavioral track data	Silhouette Score
	Principal Component Analysis (PCA)	Data dimensionality reduction and feature extraction for visitor behavior analysis	Multi-dimensional visitor data	Explained Variance Ratio (EVR)
deep learning	Convolutional Neural Network (CNN)	Classification and Restoration of Cultural Heritage Images	Artifact image data, 3D scanning data	Classification accuracy, image reconstruction error
	Recurrent Neural Networks (RNN)	Visitor behavior sequence analysis (e.g., browsing path prediction)	Visitor track data	Sequence prediction accuracy, loss function value
	Long Short-Term Memory Network (LSTM)	Tourist flow time series forecasting	Historical Visitor Traffic Data	Mean Square Error (MSE), Mean Absolute Error (MAE)

TABLE II. MACHINE LEARNING METHODS FOR APPLICATION CLASS

Research Methodology	*Specific techniques/algorithms*	*Application scenario*	*Data sources*	*Assessment of indicators*
Natural Language Processing (NLP)	BERT model	Intelligent Q&A system, visitor comment sentiment analysis	Visitor comment data, cultural heritage interpretation texts	Accuracy, F1 scores, BLEU scores (Q&A system)
	TF-IDF + Sentiment Analysis	Theme Extraction and Sentiment Tendency Analysis of Visitor Comments	Visitor comment data	Sentiment classification accuracy, subject coverage
Model Interpretive Approach	SHAP values (Shapley Additive Explanations)	Explaining the decision-making process in predictive models of tourist behavior	Model Output and Feature Data	Feature Importance Ranking, Interpretive Consistency
	LIME (Local Interpretable Model-agnostic Explanations)	Localized interpretation of model predictions (e.g., satisfaction predictions)	Model output with local feature data	Interpretation stability, local fit
Multimodal data fusion	Image+ Text fusion model (e.g. CLIP)	Multimodal analysis combining images of artifacts with explanatory texts	Cultural heritage image data, explanatory text data	Multimodal classification accuracy, feature alignment effect
	Sensor Data+ Behavioral Data Fusion	Integrated analysis of visitor behavior and environmental data (e.g., impact of temperature and humidity on visitor experience)	Environmental sensor data, visitor behavior data	Data fusion effect, model prediction accuracy

Finally, this study combines the practical application effect of the machine learning model with the practical needs of tourism and verifies its feasibility and effectiveness through case studies. In conclusion, through the data-driven machine learning research method, this study not only provides new technical tools for cultural heritage tourism experience design but also provides an important reference for technical landing and optimization in the industry practice.

IV. FINDINGS AND DISCUSSION

A. Visual Analysis of Authors

The authors' visualization analysis is obtained by using "tourism" and "machine learning" as keywords on WOS and using VOS software. Fig. 1, shows a visual graph generated by VOSviewer, which shows the collaboration network of related authors under the themes of "tourism" and "machine learning". The position and size of the nodes in the graph reflect the authors' research impact and collaboration intensity, while the lines between the nodes indicate the collaboration relationship between the authors [28]. From the Fig. 1, it can be seen that some authors have larger nodes, indicating that their research results in this field are more prominent and their cooperation with other authors is more intense. Some scholars have larger nodes and more connected lines, indicating that they are more active in collaborative research in the field of machine learning and tourism. In addition, the graph shows some smaller nodes of authors who may have less research output or less collaboration in the field but are still part of the research network [29]. Overall, this map reveals the core research strengths and their collaboration patterns in the field of machine learning and tourism, providing an important reference for subsequent research.

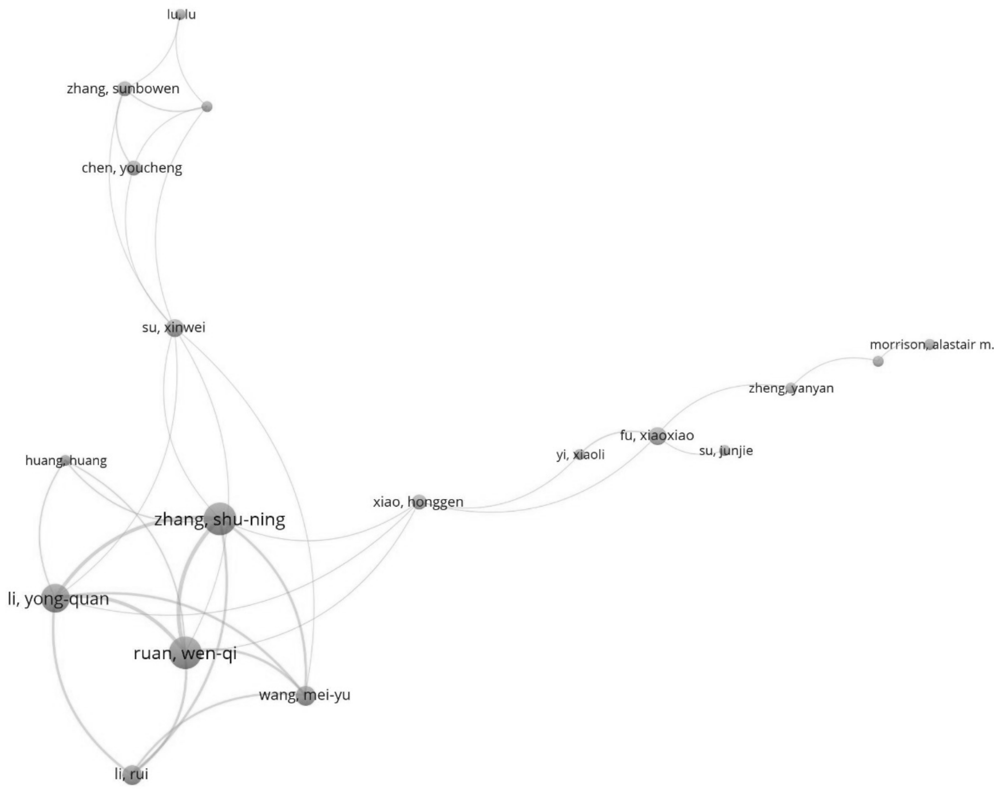

Fig. 1. Author's visualization.

Table III further refines the results of the VOSviewer analysis by listing the relevant authors' documents, citations, and total link strength under the themes of "tourism" and "machine learning". The results of the study are listed as documents, citations, and total link strength under the themes of "tourism" and "machine learning". From the data, Ruan Wen-Qi and Zhang Shu-Ning have the most prominent research results, with 11 documents and 97 citations respectively, and their collaboration strength is 31, indicating that their research in this field has a greater influence and close cooperation. Some scholars' research results are also more significant, with 9 and 5 documents, 75 and 60 citations, respectively, and 29 and 20 intensity of cooperation, respectively. In addition, some scholars' research results are fewer (all 4 documents), but the number of citations is 54 and 35, respectively, indicating that the quality of their research is higher [30]. Overall, the document data further validates the visualization results in the picture, reveals the core authors and their research impact in the field of machine learning and tourism, and provides an important literature base for subsequent research.

TABLE III. FURTHER REFINEMENT OF VOS VIEWER ANALYSIS RESULTS

Id	*Author*	*Documents*	*Citations*	*Total link strength*
46	al-ANSI, amr	4	36	0
772	fu, Xiao Xiao	4	54	0
1288	Lee, Timothy J.	4	65	0
1333	Li, Rui	5	52	19
1348	Li, Yong-Quan	9	75	29
1587	Mishra, Smriti	5	22	5
1655	Nag, Aditi	5	22	5
1987	Ruan, Wen-Qi	11	97	31
2224	Su, Xinwei	4	35	4
2303	Timothy, Dallen J.	4	21	0
2409	Wall, Geoffrey	4	38	0
2428	Wang, Mei-Yu	5	60	20
2652	Zhang, Mu	4	121	0
2660	Zhang, Shu-Ning	11	97	31

Combined with the results of the charts, the following conclusions can be drawn: under the themes of "tourism" and "machine learning", some of the authors are the core researchers in the field, with rich and highly cited research outputs, and close collaborative networks. Some authors are the core researchers in this field, with rich research results, high citation counts, and close collaborative networks [31]. The research direction of these authors may cover multiple application scenarios of machine learning in tourism, such as tourist behavior analysis, intelligent recommendation systems, and so on. In addition, although some authors have fewer research results, the quality of their research is higher, indicating that their research in this field has high academic value. Overall, the results of these analyses provide important literature support and collaborative references for subsequent research in the field of machine learning and tourism [32]. Further analysis of the research directions of these core authors reveals that their applications in the field of machine learning and tourism are mainly focused on the following aspects: firstly, tourists' behavior analysis and prediction is one of the hotspots of the research, and tourists' browsing trajectories, consumption records, and other data can be analyzed by machine learning algorithms, which can identify the tourists' preferences and behavioral patterns, and thus provide decision-making support for scenic spot management [33]. Secondly, an intelligent recommendation system is also an important research direction, through collaborative filtering algorithms or deep learning models, it can provide tourists with personalized travel route recommendations, and enhance the satisfaction and experience of tourists.

B. Visualization and Analysis of Cultural Heritage and Tourism

The visual analysis of cultural heritage and tourism is obtained by using "tourism" and "cultural heritage" as keywords on WOS and using VOS software. In Table IV, the keywords related to cultural heritage tourism, their occurrences, and total link strength are listed, which reflect the research hotspots and their relevance in this field.

"Tourism" and "cultural heritage" are the most frequent keywords, 271 and 225 times respectively, and the total link strength is 431 and 304 respectively, which indicates that "tourism" and "cultural heritage" are the core research themes in this field, and the correlation between them is very strong. Other high-frequency keywords such as "heritage tourism" (99 times), "cultural tourism" (158 times) and "heritage" (158 times). “Heritage" (heritage, 140 times) further confirms the centrality of this keyword research. The high frequency of these keywords suggests that research on this tourism model focuses not only on the preservation and presentation of its heritage but also on its integration with tourism activities to achieve the dual goals of cultural dissemination and economic development. The keywords also listed in Table IV reveal several research hotspots in the field of this mode of tourism. "Authenticity" (99 times) and "intangible cultural heritage" (64 times) reflect researchers' concern for the authenticity of heritage and intangible cultural preservation. Culture protection; "sustainability" (68 times) and "sustainable tourism" (70 times) indicate that sustainability is an important direction for tourism research in this mode. In addition, the keywords "management" (103 times) and "impact" (41 times) reveal that the management and impact assessment of this mode of tourism is also one of the hot spots of research [34]. The high frequency of these keywords indicates that the research on this mode of tourism not only focuses on the protection and display of cultural heritage but also on its integration with tourism activities to achieve the dual goals of cultural dissemination and economic development.

TABLE IV. VISUAL MAPPING OF "TOURISM" AND "CULTURAL HERITAGE"

Id	***Keyword***	***Occurrences***	***Total link strength***
204	authenticity	99	243
378	China	40	78
527	conservation	68	161
687	cultural heritage	225	304
770	cultural tourism	158	246
798	cultural-heritage	76	182
1122	experience	44	85
1360	Heritage	140	233
1416	Heritage tourism	99	185
1534	identity	43	87
1557	impact	41	96
1632	intangible cultural heritage	64	113
1911	management	103	249
2029	model	55	117
2316	perceptions	51	124
2740	satisfaction	68	145
2842	sites	45	128
3055	sustainability	68	165
3065	sustainable development	47	91
3090	sustainable tourism	70	118
3196	tourism	271	431
3652	world heritage	43	97

The total link strength reflects the relevance of the keywords. The high link strength between "cultural heritage" and "tourism" (304 vs. 431) indicates a strong association between the two in the study. In addition, the link strength between "heritage tourism" and "cultural tourism" is also high (185 vs. 246), indicating that these two forms of tourism are often discussed side by side in research. The strength of links between other keywords such as "authenticity" and "sustainability" (243 vs. 165) suggests that authenticity and sustainability of cultural heritage are also often discussed together in research. These linkage analyses reveal that the study is not a simple one. These linkage analyses reveal the multidimensional character of this model of tourism research, i.e., researchers are not only concerned with the preservation and presentation of heritage but also with its integration with tourism activities to achieve the dual goals of cultural dissemination and economic development. The occurrence of the keyword "China" (China, 40 times) indicates that China occupies an important position in related tourism research. This may be related to China's rich cultural heritage resources and its need for cultural dissemination in the context of globalization. In addition, the occurrence of "world heritage" (world heritage, 43 times) indicates that the protection and tourism development of world heritage sites are also important directions of research. The occurrence of these keywords reflects the globalization feature of related tourism research, i.e., researchers not only focus on the heritage of a specific region but also heritage conservation and tourism development on a global scale. The occurrence of the keywords "experience" (44 times) and "satisfaction" (68 times) indicates that tourist experience and satisfaction are important directions in tourism research. By analyzing tourists' experience and satisfaction, the researchers explore how to improve the quality and attractiveness of cultural heritage tourism. The immersion of tourists is enhanced through virtual reality (VR) technology or the engagement of tourists is enhanced through intelligent guiding systems [35]. The emergence of these keywords indicates that research on cultural heritage tourism not only focuses on the protection and display of heritage but also on tourists' experience and satisfaction, to realize the dual goals of cultural dissemination and economic development.

Fig. 2 shows the keyword density map generated based on VOSviewer, which reflects the distribution density and importance of each keyword in the research under the themes of "tourism" and "cultural heritage". The color of the density map indicates the degree of concentration of the keywords, and the darker the color, the higher the frequency of the keywords in the study, and the greater the heat of the study [36]. As can be seen from the graph, "tourism" and "cultural heritage" are the keywords with the highest density and the darkest color, indicating that they are the core themes of the study. Other high-density keywords such as "heritage tourism", "sustainability" and "authenticity" also show a high density of keywords. The "authenticity" also shows high research intensity. The high-density distribution of these keywords indicates that tourism-related research not only focuses on the preservation and display of tourism heritage but also its integration with tourism activities to realize the dual goals of cultural dissemination and economic development [37]. In addition, the density of keywords such as "community", "experience" and "satisfaction" is also high, indicating that the tourism research is not only concerned with the preservation and display of tourism heritage but also focuses on its integration with tourism activities to achieve the dual goals of cultural dissemination and economic development.) also have a high density, indicating that community participation, tourist experience, and satisfaction are important directions for research.

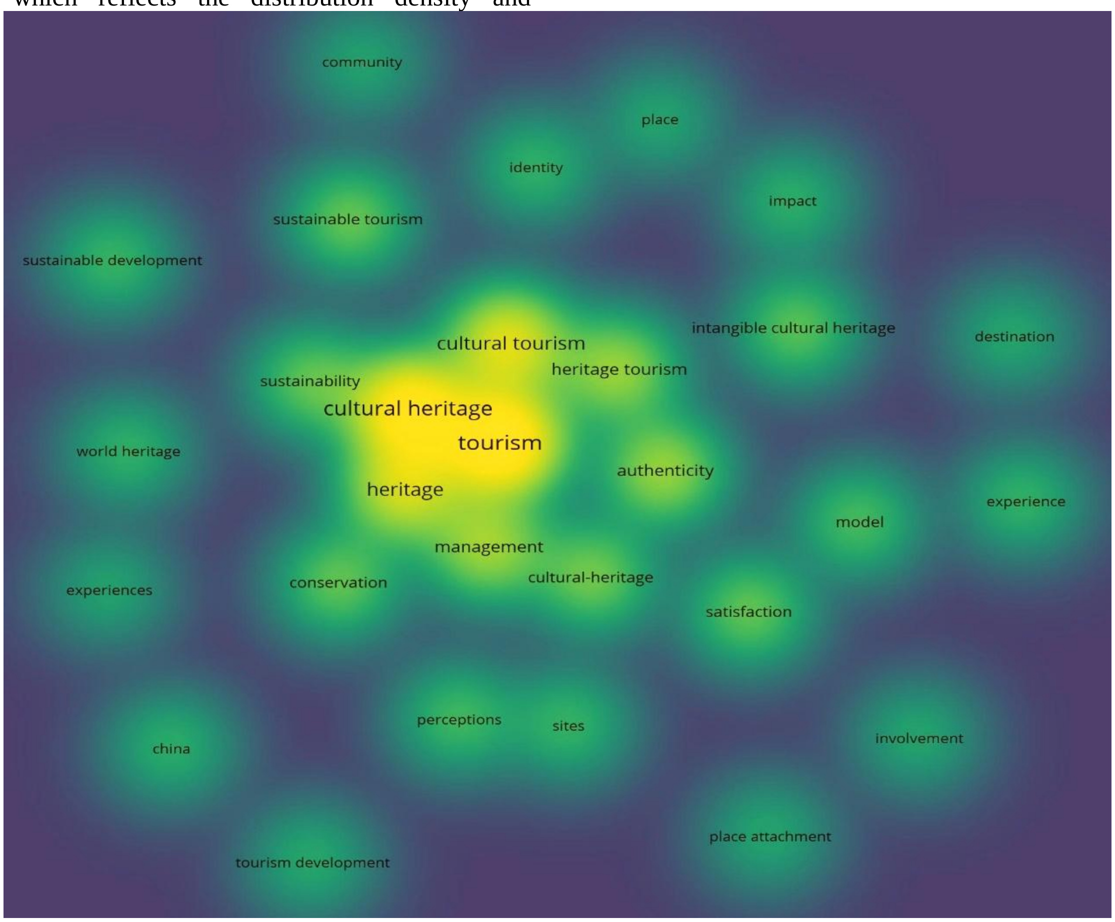

Fig. 2. Density mapping of cultural heritage and tourism.

Fig. 3 is a keyword clustering diagram generated based on VOSviewer, reflecting the correlation between keywords and their clustering characteristics under the themes of "tourism" and "cultural heritage". Different colors in the clustering diagram indicate different research themes or directions, and keywords within the same color have a strong correlation. From the figure, it can be seen that the keywords are categorized into multiple clusters, and each cluster represents a research topic or direction. For example, the red cluster may represent "cultural heritage conservation and sustainability", including keywords such as "conservation", "sustainability" and "sustainability", “(conservation)”, "sustainability" and "sustainable development". The green cluster might represent "visitor experience and satisfaction", including keywords such as "experience", "satisfaction" and "sustainability", “satisfaction" and "perceptions". The blue cluster may represent "community engagement and place attachment", including keywords such as "community", "identity" and "place attachment". These clusters reveal the multidimensional character of tourism research, i.e., researchers not only focus on heritage preservation and display but also tourists' experience and satisfaction, community participation, and place attachment [38]. For example, the keywords in the red clusters reflect the importance of cultural resource conservation and sustainable development, while the keywords in the green clusters reveal the centrality of tourist experience and satisfaction in cultural resource tourism.

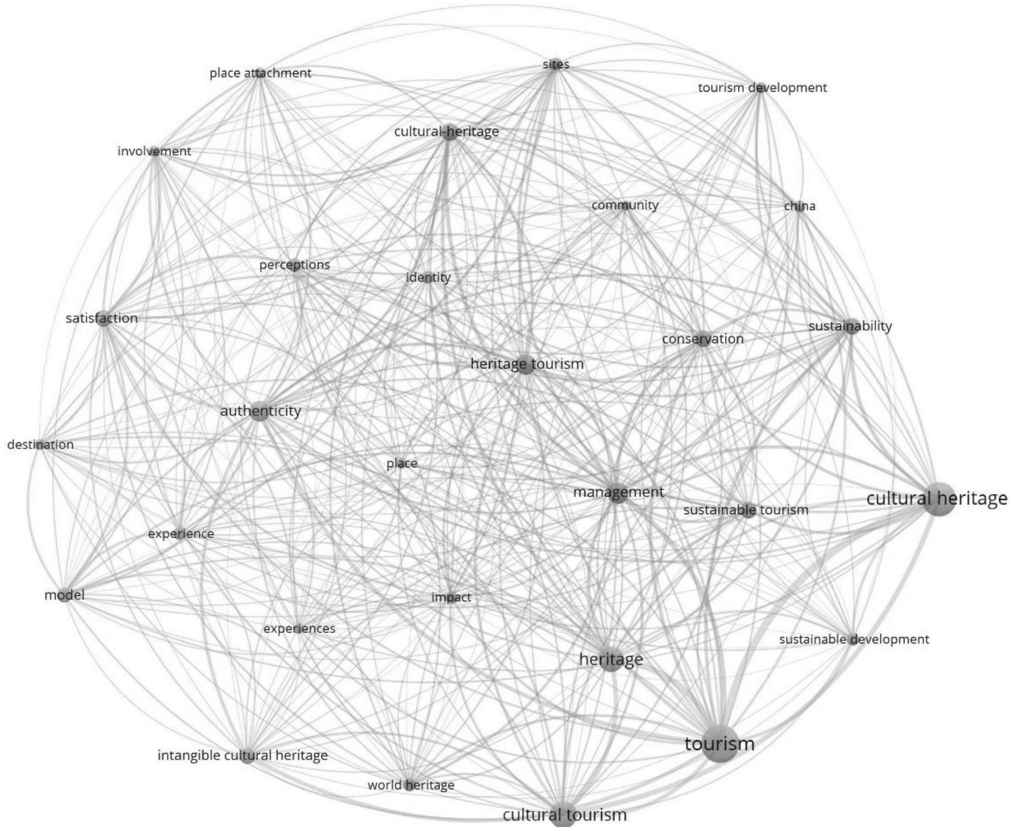

Fig. 3. Basic clustering map of cultural heritage and tourism.

C. Machine Learning and Visual Analytics for Tourism

The visual analysis of machine learning and tourism is obtained by using "tourism" and "machine learning" as keywords on WOS and using VOS software. The keywords related to the application of machine learning in tourism, their occurrences, and total link strength are listed in Table V and Table VI, which reflect the research hotspots and their relevance in this field.

From Table V and Table VI, it can be seen that "machine learning" and "tourism" are the keywords with the highest frequency of 282 and 147 times respectively, and the total link strength is 433 and 285 respectively. This indicates that "machine learning" and "tourism" are the core research topics in this field, and the correlation between them is very strong. Other high-frequency keywords such as "big data" (58 times), "sentiment analysis" (61 times) and "artificial intelligence" (61 times) are also used. “Intelligence" (artificial intelligence, 42 times) further confirms the wide application of machine learning in the field of tourism. The high frequency of these keywords indicates that the application of machine learning techniques in tourism research is not only limited to the traditional analysis of tourists' behaviors but also covers a variety of aspects such as sentiment analysis and social media data processing. The keywords listed in the table reveal multiple research hotspots of machine learning in tourism. The “big data" and "deep learning" reflects researchers' focus on big data analysis and deep learning algorithms; "sentiment analysis" and "social media" indicate that sentiment analysis and social media data processing are important directions for research. In addition, the keywords "prediction" (28 times) and "model" (53 times) reveal the application of machine learning in tourism demand prediction and model construction. The high frequency of these keywords indicates that the application of machine learning technology in tourism research is not only limited to the traditional analysis of tourists' behaviors but also covers a variety of aspects such as sentiment analysis and social media data processing.

TABLE V. VISUALIZATION OF "TOURISM" AND "MACHINE LEARNING" (I)

Id	*Keyword*	*Occurrences*	*Total link strength*
147	arrivals	20	37
151	artificial intelligence	42	95
238	behavior	22	52
260	big data	58	166
382	classification	40	75
543	covid-19	23	52
668	deep learning	47	96
680	demand	40	80
688	destination	20	53
1208	hospitality	55	159
1290	impact	48	105
1526	machine learning	282	433

The total link strength reflects the relevance of the keywords. The high link strength between "machine learning" and "tourism" (433 vs. 285) indicates a strong association between the two in the study. In addition, the link strength between "big data" and "deep learning" is also high (166 vs. 96), indicating that these two techniques are often discussed side by side in research. Other keywords such as "sentiment analysis" and "social media" have high link strengths (150 vs. 152), suggesting that sentiment analysis and social media data processing are often discussed side by side in research. These correlation analyses reveal that machine analysis and social media data processing are often explored simultaneously in research [20]. These linkage analyses reveal the multidimensional character of machine learning in tourism research, i.e., researchers not only focus on the traditional analysis of tourists' behaviors but also pay attention to various aspects such as sentiment analysis and social media data processing. The occurrence of keywords such as "natural language processing" (natural language processing, 22 times) and "random forest" (random forest, 23 times) suggests that the natural language processing and the random forest algorithms are gradually increasing in tourism research. Natural language processing techniques can be used to analyze tourists' review data, while random forest algorithms can be used for tourists' behavior prediction. In addition, the occurrence of "COVID- 19" (new crown epidemic, 23 times) suggests that machine learning techniques also play an important role in coping with the impact of new crown epidemics on tourism. The occurrence of these keywords reflects the wide application and innovation of machine learning techniques in tourism research. The appearance of the keywords "satisfaction" (satisfaction, 43 times) and "reviews" (reviews, 25 times) indicates that tourist experience and satisfaction are important directions for machine learning in tourism research. By analyzing tourists' review data and satisfaction, the researcher explores how to improve the quality and attractiveness of the tourism experience. Tourists' satisfaction is identified through sentiment analysis techniques, or tourists' engagement is enhanced through recommendation systems [39]. The appearance of these keywords indicates that the application of machine learning techniques in tourism research not only focuses on technical implementation but also tourists' experience and satisfaction.

TABLE VI. VISUALIZATION OF "TOURISM" AND "MACHINE LEARNING" (II)

Id	*Keyword*	*Occurrences*	*Total link strength*
1550	management	26	59
1647	model	53	101
1722	natural language processing	22	54
1804	online reviews	25	73
1887	performance	23	63
1980	prediction	28	61
2070	random forest	23	40
2174	reviews	25	78
2236	satisfaction	43	109
2314	sentiment analysis	61	150
2401	social media	51	152
2645	tourism	147	285

Fig. 4 is the keyword clustering graph generated based on VOSviewer, which shows the correlation between keywords and their clustering characteristics under the topics of "tourism" and "machine learning". Through the clusters of different colors, we can see the research hotspots of machine learning in tourism and its multi-dimensional characteristics [40]. The red clusters in the figure may represent "machine learning technology and algorithm application", including "deep learning", and "artificial intelligence”. The red clusters may represent "machine learning techniques and algorithm applications", including keywords such as "deep learning", "artificial intelligence", "natural language processing" and "random forest". These keywords reflect the core application of machine learning technology in tourism research, especially the wide application of deep learning and natural language processing technology in the analysis of tourists' behavior and sentiment analysis. The green clusters may represent "tourism demand and prediction", including "demand", and "prediction". The keywords "demand", "prediction" and "model" indicate that the application of machine learning technology in tourism demand prediction and model construction is gradually increasing, especially in tourism flow prediction and resource scheduling optimization. The blue clusters may represent "tourist experience and satisfaction", including "satisfaction", "reviews" and "sentiment analysis". The keywords "satisfaction", "reviews" and "sentiment analysis" reveal the important role of machine learning in visitor experience optimization and satisfaction enhancement, such as identifying visitor satisfaction through sentiment analysis techniques or enhancing visitor engagement through recommendation systems.

In addition, the keywords in the cluster diagram reflect the innovative application of machine learning techniques in tourism research. The keywords "big data" and "social media" indicate that big data analysis and social media data processing are important directions for research. Analyzing social media data through machine learning techniques can identify tourists' preferences and behavioral patterns and provide decision support for tourism management [37]. The appearance of the keyword "covid-19" (new crown epidemic) indicates that machine learning techniques also play an important role in coping with the impact of the new crown epidemic on tourism, such as analyzing the impact of the epidemic on tourism demand through predictive models or evaluating tourists' responses to the epidemic through sentiment analysis techniques. The keywords "destination" and "management" suggest that machine learning techniques are increasingly being used in destination management, for example, through predictive models to optimize the allocation of tourism resources, or through sentiment analysis techniques to improve the performance of tourists [41]. The use of machine learning techniques in tourism destination management is increasing, such as optimizing the allocation of tourism resources through predictive modeling, or improving tourist satisfaction through sentiment analysis techniques.

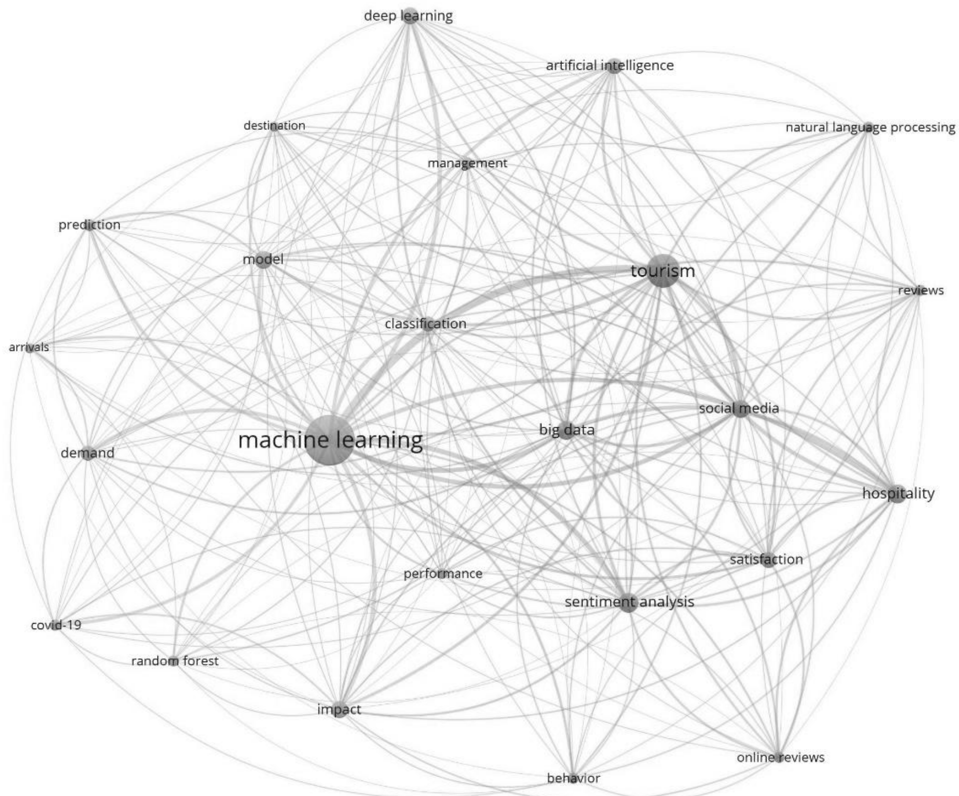

Fig. 4. Basic clustering mapping for machine learning and tourism.

V. Conclusion

This study reveals the research hotspots, current status of technology applications, and future development direction in the field of cultural heritage tourism experience design by systematically combing the literature trends and the progress of machine learning technology applications in this field. It is found that the research focus of tourism experience design has shifted from traditional conservation and display to visitor experience optimization and technological innovation, especially the introduction of machine learning technology has injected new vitality into related tourism. Through big data analysis, deep learning, natural language processing, and other technologies, machine learning can not only optimize the analysis of tourist behavior and resource management, but also promote the digital protection and innovative display of tourism heritage, and even reconfigure the interaction mode between tourists and culture. In addition, the study shows that the success of tourism cannot be achieved without community participation and cultural identity, and machine learning technologies show great potential in enhancing tourist satisfaction and personalized experience. Overall, this study provides an important theoretical basis and technical support for academic exploration and industry practice of cultural heritage tourism experience design.

There are some limitations in this study. First, the source of the literature data mainly relies on the Web of Science database, which may have the problem of incomplete data coverage, and future research can combine with other databases (e.g. Scopus, CNKI) for a more comprehensive analysis. Second, the application cases of machine learning technology are mostly concentrated in developed countries or regions, with fewer practice cases in developing countries or regions, and future research can further expand the geographical scope and explore the technology's adaptability in different cultural contexts. In addition, the ethical issues of machine learning technology (e.g. data privacy, algorithmic bias) have not yet been fully discussed, so future research needs to find a balance between technological innovation and humanistic values and explore the path of "human-centered" technology adaptation. In the future, with the deepening of multimodal data integration and interdisciplinary cooperation, the application of machine learning technology in tourism will be more extensive and precise, providing more possibilities for the protection, dissemination, and sustainable development of cultural heritage.

REFERENCES

[1] Del Vecchio P, Secundo G, Garzoni A. Phygital technologies and environments for breakthrough innovation in customers' and citizens' journey. A critical literature review and future agenda [J]. *Technol Forecasting Social Change*, 2023, 189: 122342.

[2] Li X, Chen C, Kang X. Research on relevant dimensions of tourism experience of intangible cultural heritage lantern festival: Integrating generic learning outcomes with the technology acceptance model [J]. *Front Psychol*, 2022, 13: 943277.

[3] Gurcan F, Boztas G, Dalveren G. Digital transformation strategies, practices, and trends: A large-scale retrospective study based on machine learning [J]. *Sustainability*, 2023, 15(9): 7496.

[4] Hou Y, Kenderdine S, Picca D. Digitizing intangible cultural heritage embodied: State of the art [J]. *J Comput Cult Herit*, 2022, 15(3): 1–20. doi: 10.1145/3494837.

[5] Liu S, Pan Y. Exploring trends in intangible cultural heritage design: a bibliometric and content analysis [J]. *Sustainability*, 2023, 15(13): 10049.

[6] Nannelli M, Capone F, Lazzeretti L. Artificial intelligence in hospitality and tourism. State-of-the-art and future research avenues [J]. *European Planning Studies*, 2023, 31(7): 1325–1344. doi: 10.1080/09654313.2023.2180321.

[7] Lu S, Moyle B, Reid S. Technology and museum visitor experiences: a four-stage model of evolution [J]. *Inf Technol Tourism*, 2023, 25(2): 151–174. doi: 10.1007/s40558-023-00252-1.

[8] Allal-Chérif O. Intelligent cathedrals: using augmented reality, virtual reality, and artificial intelligence to provide an intense cultural, historical, and religious visitor experience [J]. *Technological Forecasting and Social Change*, 2022, 178: 121604.

[9] Hu H, Li C. Smart tourism products and services design based on user experience under the background of big data [J]. *Soft Comput*, 2023, 27(17): 12711–12724. doi: 10.1007/s00500-023-08851-0.

[10] Alsahafi R, Alzahrani A, Mehmood R. Smarter sustainable tourism: data-driven multi-perspective parameter discovery for autonomous design and operations [J]. *Sustainability*, 2023, 15(5): 4166.

[11] Lukita C, Pangilinan G, Chakim M. Examining the impact of artificial intelligence and internet of things on smart tourism destinations: A comprehensive study [J]. *Aptisi Transactions on Technopreneurship (ATT)*, 2023, 5(2sp): 135–145.

[12] Pisoni G, Díaz-Rodríguez N, Gijlers H. Human-centered artificial intelligence for designing accessible cultural heritage [J]. *Applied Sciences*, 2021; 11(2): 870.

[13] Doborjeh Z, Hemmington N, Doborjeh M. Artificial intelligence: A systematic review of methods and applications in hospitality and tourism [J]. *Int J Contemp Hosp Manag*, 2022, 34(3): 1154–1176.

[14] Skublewska-Paszkowska M, Milosz M, Powroznik P. 3D technologies for intangible cultural heritage preservation—literature review for selected databases [J]. *Heritage Sci*, 2022, 10(1): 3. doi: 10.1186/s40494-021-00633-x.

[15] Ammirato S, Felicetti A, Linzalone R. Digital business models in cultural tourism [J]. *International Journal of Entrepreneurial Behavior & Research*, 2022, 28(8): 1940–1961.

[16] Su P, Hsiao P, Fan K. Investigating the relationship between users' behavioral intentions and learning effects of VR system for sustainable tourism development [J]. *Sustainability*, 2023, 15(9): 7277.

[17] Jia S (Jasper), Chi O, Martinez S. When "old" meets "new": Unlocking the future of innovative technology implementation in heritage tourism [J]. *Journal of Hospitality & Tourism Research*, 2025, 49(3): 640–661. doi: 10.1177/10963480231205767.

[18] Poulopoulos V, Wallace M. Digital technologies and the role of data in cultural heritage: The past, the present, and the future [J]. *Big Data Cogn Comput*, 2022, 6(3): 73.

[19] Asif M, Fazel H. Digital technology in tourism: a bibliometric analysis of transformative trends and emerging research patterns [J]. *Journal of Hospitality and Tourism Insights*, 2024, 7(3): 1615–1635.

[20] Rahmadian E, Feitosa D, Zwitter A. A systematic literature review on the use of big data for sustainable tourism [J]. *Current Issues in Tourism*, 2022, 25(11): 1711–1730. doi: 10.1080/13683500.2021.1974358.

[21] Verma S, Warrier L, Bolia B. Past, present, and future of virtual tourism-a literature review [J]. *Int J Inf Manage Data Insights*, 2022, 2(2): 100085.

[22] Luther W, Baloian N, Biella D. Digital twins and enabling technologies in museums and cultural heritage: An overview [J]. *Sensors*, 2023, 23(3): 1583.

[23] Madzík P, Falát L, Copuš L. Digital transformation in tourism: bibliometric literature review based on machine learning approach [J]. *Eur J Innovation Manage*, 2023, 26(7): 177–205. doi: 10.1108/EJIM-09-2022-0531.

[24] Augello A, Infantino I, Pilato G. Site experience enhancement and perspective in cultural heritage fruition—a survey on new technologies and methodologies based on a "four-pillars" approach [J]. *Future Internet*, 2021; 13(4): 92.

[25] Khan M, Israr S, S Almogren A. Using augmented reality and deep learning to enhance taxila museum experience [J]. *J Real-Time Image Proc*, 2021; 18(2): 321–332. doi: 10.1007/s11554-020-01038-y.

[26] Marques C, Pedro J, Araújo I. A systematic literature review of gamification in/for cultural heritage: leveling up, going beyond [J]. *Heritage*, 2023, 6(8): 5935–5951.

[27] Kumar S, Kumar V, Kumari Bhatt I. Digital transformation in tourism sector: Trends and future perspectives from a bibliometric-content analysis [J]. *J Hosp Tour Insights*, 2024, 7(3): 1553–1576.

[28] Le T, Arcadia C, Novais M. Proposing a systematic approach for integrating traditional research methods into machine learning in text analytics in tourism and hospitality [J]. *Current Issues in Tourism*, 2021; 24(12): 1640–1655. doi: 10.1080/13683500.2020.1829568.

[29] Baker J, Nam K, Dutt CS. A user experience perspective on heritage tourism in the metaverse: Empirical evidence and design dilemmas for VR [J]. *Inf Technol Tour*, 2023, 25(3): 265–306. doi: 10.1007/s40558-023-00256-x.

[30] Barrientos F, Martin J, De Luca C. Computational methods and rural cultural & natural heritage: A review [J]. *Journal of Cultural Heritage*, 2021; 49: 250–259.

[31] Siddiqui M, Syed T, Nadeem A. Virtual tourism and digital heritage: an analysis of VR/AR technologies and applications [J]. *Int J Adv Comput Sci Appl*, 2022, 13(7). doi: 10.14569/IJACSA.2022.0130739.

[32] Boboc R, Băutu E, Gîrbacia F. Augmented reality in cultural heritage: an overview of the last decade of applications [J]. *Applied Sciences*, 2022, 12(19): 9859.

[33] Li J, Wider W, Ochiai Y. A bibliometric analysis of immersive technology in museum exhibitions: exploring user experience [J]. *Front Virtual Reality*, 2023, 4: 1240562.

[34] Michalakis K, Caridakis G. Context awareness in cultural heritage applications: A survey [J]. *J Comput Cult Herit*, 2022, 15(2): 1–31. doi: 10.1145/3480953.

[35] Casillo M, De Santo M, Mosca R. An ontology-based chatbot to enhance experiential learning in a cultural heritage scenario [J]. *Frontiers in artificial intelligence*, 2022, 5: 808281.

[36] Rosário A, Dias J. Exploring the landscape of smart tourism: a systematic bibliometric review of the literature of the Internet of things [J]. *Administrative Sciences*, 2024, 14(2): 22.

[37] Chong H, Lim C, Rafi A. Comprehensive systematic review on virtual reality for cultural heritage practices: coherent taxonomy and motivations [J]. *Multimedia Syst*, 2022, 28(3): 711–726. doi: 10.1007/s00530-021-00869-4.

[38] Goodarzi P, Ansari M, Rahimian F. Incorporating sparse model machine learning in designing cultural heritage landscapes [J]. *Automation in Construction*, 2023, 155: 105058.

[39] García-Madurga M, Grilló-Méndez A. Artificial intelligence in the tourism industry: An overview of reviews [J]. *Administrative Sciences*, 2023, 13(8): 172.

[40] De Masi F, Larosa F, Porrini D. Cultural heritage and disasters risk: a machine-human coupled analysis [J]. *International Journal of Disaster Risk Reduction*, 2021; 59: 102251.

[41] Cuomo M, Tortora D, Foroudi P. Digital transformation and tourist experience co-design: big social data for planning cultural tourism [J]. *Technol Forecasting Social Change*, 2021; 162: 120345. doi: 10.1016/j.techfore.2020.120345.