

Educational Study Tour Research Hotspots and Development Orientation: An LDA-based Literature Review

Feng Weihe¹ , Nur Hidayah Che Ahmat^{2*} , Yuan Xinying³ 

¹Faculty of Hotel and Tourism Management, Universiti Teknologi MARA (UiTM) Cawangan Selangor, 42300 Puncak Alam, Selangor, Malaysia;

Faculty of Cultural Creativity and Tourism, Shandong Vocational and Technical University of International Studies, Rizhao, Shandong 276826, China

Email: 2022348907@student.uitm.edu.my

²Faculty of Hotel and Tourism Management, Universiti Teknologi MARA (UiTM) Cawangan Pulau Pinang, 13500 Permatang Pauh, Pulau Pinang, Malaysia

Email: hidayah.ca@uitm.edu.my

³Faculty of Hotel and Tourism Management, Universiti Teknologi MARA (UiTM) Cawangan Selangor, 42300 Puncak Alam, Selangor, Malaysia

Email: 2022882568@student.uitm.edu.my

ABSTRACT

Educational tourism has received growing academic attention in recent years as globalization and international educational exchanges continue to expand. This study explores the major research hotspots and future development trends in educational study tourism research. Using the Web of Science (WOS) database, 12,218 publications related to “study tourism” published between 2015 and 2024 were initially collected. After screening and classification, 842 relevant publications were retained, and the 100 most relevant articles were further analysed using the Latent Dirichlet Allocation (LDA) topic modelling method. The findings show that current research mainly focuses on educational tourism, tourist experience, students, international educational tourism, rural tourism, influencing factors, and economic development. Several emerging directions were also identified, including tourism quality, awareness, behavioural intention, satisfaction, school-based programs, cultural heritage, and industry-related tourism. Overall, educational tourism research is becoming more interdisciplinary and internationally oriented. This study helps clarify the development of educational tourism research and provides references for future studies in international educational tourism.

CORRESPONDING AUTHOR (*):

Nur Hidayah Che Ahmat
(hidayah.ca@uitm.edu.my)

KEYWORDS:

Educational Study Tourism
Literature Review
Research Hotspots
LDA Analysis

CITATION:

Feng, W., Nur Hidayah, C. A., & Yuan, X. (2026). Educational Study Tour Research Hotspots and Development Orientation: An LDA-based Literature Review. *Malaysian Journal of Social Sciences and Humanities (MJSSH)*, 11(5), e004004. <https://doi.org/10.47405/mjssh.v11i5.4004>

Contribution/Originality: This study contributes to the existing literature by applying the Latent Dirichlet Allocation (LDA) topic modelling method to educational tourism research. The study systematically identifies major research hotspots and future development trends and provides a clearer understanding of the evolution of educational tourism studies and future directions for international educational tourism research.

1. Introduction

As the birthplace of modern tourism, Britain has long promoted educational study tourism, which can be traced to the “wandering study trips” of noble children in 16th–17th-century Europe. Later, the United States introduced winter camp study tours in 1885, and Japan’s Tokyo Teachers College applied study tours for students’ physical and will training in 1886. Educational study tourism emphasizes integrating education with travel. Ritchie et al. (2003) systematically discussed the concept in *Managing Educational Tourism*. With the global expansion of education tourism, scholarly interest has grown, producing increasing research outputs. This study aims to summarize current issues, predict future trends, and provide references for further studies. Using diachronic bibliometric analysis, all WOS publications on educational tourism (2015–2024) were reviewed, totalling 12,218 articles. Employing knowledge graph and topic analysis methods, this paper identifies key research themes, reveals the field’s evolution, and highlights insights for future research.

2. Literature review

2.1. Literature Review on Educational Tourism

Educational tourism involves overnight vacations and short trips, including school and adult study tours, often organized independently or formally in diverse environments (Ritchie et al., 2003). Globalization has reshaped education and increased cross-cultural exchanges (Sjøen, 2021). As a cross-border integration of education and tourism, it promotes practical learning through travel, enriching experience, and skills (Sun & Xu, 2021; Nurhuda, 2023). Its value lies in providing experiential learning, fostering cultural understanding, and supporting personal and social development. Students benefit by deepening subject knowledge, cultivating empathy, resilience, adaptability, and communication skills, while enhancing employability in the global market (Ormond & Vietti, 2022). Educational tourism also develops global citizenship and contributes economically to host destinations (Tomasi et al., 2020).

2.2. Literature Review on Latent Dirichlet Allocation (LDA)

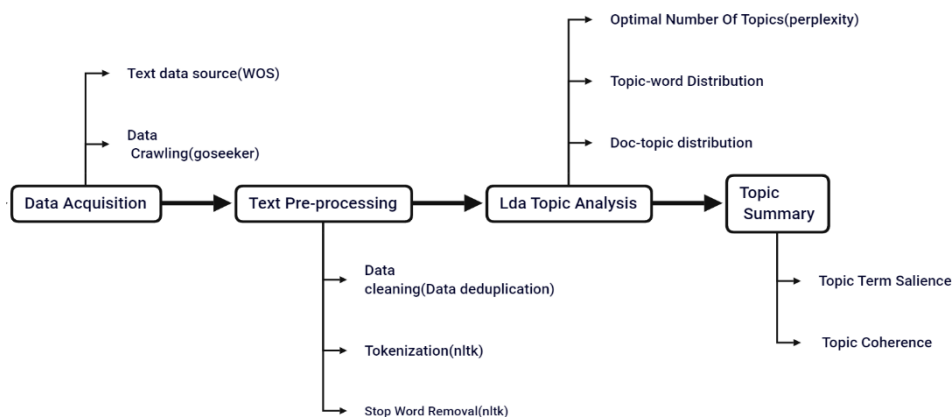
Educational study tours offer learning beyond the classroom. Vieira et al. (2022) found them effective for enriching teaching and supporting human sustainability, while Nurhuda (2023) highlighted transformative learning, improved confidence, and positive outcomes for students. Ritchie et al. (2003) classified educational tourism into school-based and adult programs, including curriculum-related travel and extracurricular field trips. In China, study tours are integrated into in- and out-of-school education, serving as innovative practical training (Sun & Xu, 2021).

Latent Dirichlet Allocation (LDA), introduced by Blei et al. (2003), is a widely used topic modelling approach for uncovering semantic structures in large corpora (Kherwa & Bansal, 2019). LDA assumes documents are mixtures of topics and has been applied across fields from psychology to blockchain, education, and tourism (Ali et al., 2022). Applications range from health message tracking to tourism trends analysis (Kurt et al., 2024). By extracting distinct research themes, LDA supports a better understanding of disciplinary development and informs future directions.

3. Research Methods

This study uses the LDA topic analysis model to examine research hotspots in study tourism. First, text data are collected and pre-processed to ensure suitability for LDA (Jelodar et al., 2019). Preprocessing includes tokenization and stop word removal (Yau et al., 2014). LDA is then applied to the cleaned data, iteratively estimating topic-word and document-topic distributions (Lancichinetti et al., 2015). The number of topics is determined using perplexity to select the optimal value (Gan & Qi, 2021). After training, results are evaluated and interpreted through topic term salience and coherence, and topic popularity and trends are analysed. As shown in Figure 1, the research design follows four main procedures: data collection, text preprocessing, LDA topic modelling, and interpretation of research hotspots and future directions.

Figure 1: Research Design



LDA is a Bayesian generative model for discovering latent topic structures in text (Gan & Qi, 2021; Jelodar et al., 2019). Each document comprises latent topics, and each topic consists of feature words. Its calculation relies on probabilistic relationships among documents, topics, and words, using Bayesian formulas and Dirichlet distributions for parameter estimation (Omar et al., 2015). Key principles and formulas are provided below:

In the LDA model, we first assume that each document d is generated by multiple topics t according to the probability distribution θ_d . θ_d is a K -dimensional vector, where k is the preset number of topics, and $\theta_d[k]$ represents the probability that document d corresponds to topic k . Secondly, assume that each topic t_i corresponds to a word distribution ϕ_t on the vocabulary. ϕ_t is also a vector, where $\phi_t[w]$ represents the probability of word w appearing under topic t . Finally, for each word w_i in document d , a topic t_i is first randomly selected according to θ_d , and then the word w_i is generated according to the word distribution $\phi_t[w_i]$ of the selected topic t_i .

For a document d and its word sequence $W_d = (w_1, w_2, \dots, w_N)$, the LDA model defines their joint probability distribution as:

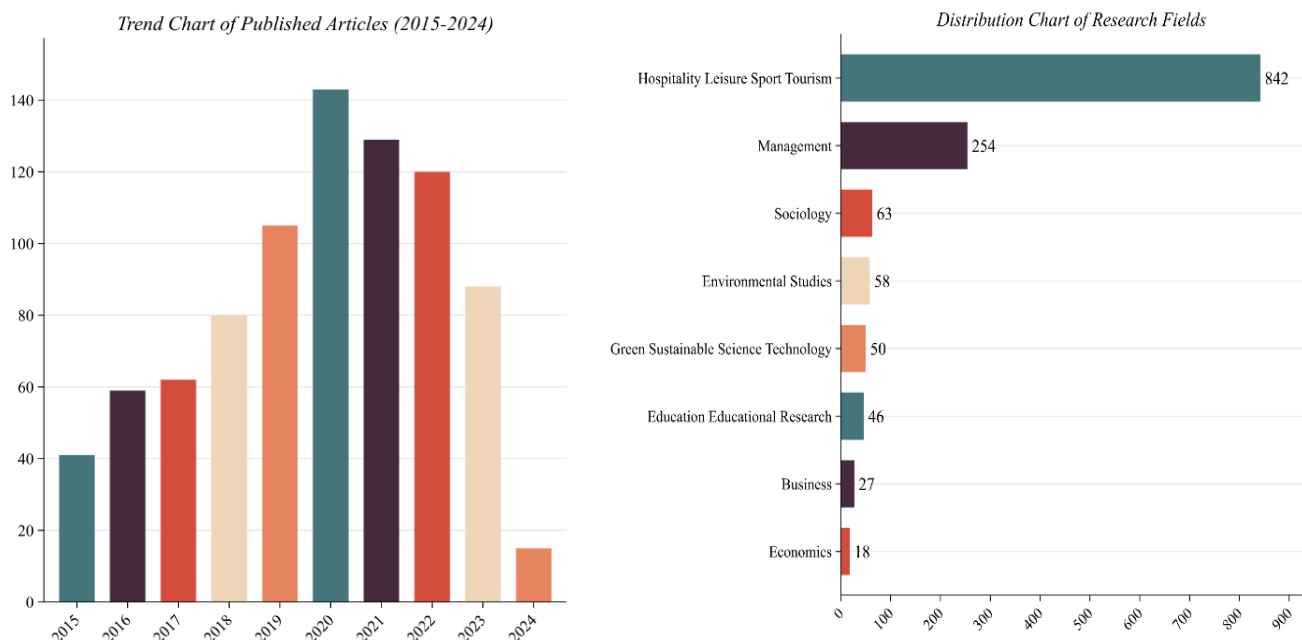
$$P(W_d, Z_d | \theta_d, \phi) = \prod_{i=1}^N P(w_i | z_i, \phi) P(z_i | \theta_d) \quad (1)$$

Among them: $Z_d = (z_1, z_2, \dots, z_N)$ is the latent topic tag sequence, and z_i represents the topic to which the i -th word belongs. $P(w_i|z_i)$ is the probability that word w_i is generated under a given topic z_i , that is, $\phi[z_i][w_i]$. $P(z_i|d)$ is the probability that the i -th word is assigned to topic z_i , that is, $\theta_d[z_i]$ (Gan & Qi, 2021; Jelodar et al., 2019; Omar et al., 2015).

LDA continuously updates the probability distribution of each word in each document and each topic through iterative calculations, and finally obtains the topic distribution of each document and the characteristic word distribution of each topic. The LDA topic analysis method has been widely used in document analysis. LDA topic analysis can quickly dig out main topics from a large number of documents and help us understand the research direction and content of the documents. In particular, it helps to understand the trends and directions of current research.

The text data for this article comes from the Web of Science (WOS) database. Using "educational study tourism" as the keyword, a total of 12,218 relevant publications were retrieved over 10 years (2015–2024), including 9,694 journal papers and other materials such as books and reviews. Many papers involved unrelated fields such as sports science, nutrition, and life sciences. Since the review focused on tourism and leisure management, irrelevant papers were excluded. After further sorting, literature was classified into Hospitality, Leisure, Sport & Tourism, Management, and Social Psychology, yielding 842 relevant publications. Figure 2 presents the annual trend and disciplinary distribution of the educational study tourism literature retrieved from the Web of Science database.

Figure 2: Literature Trend and Distribution Chart



Through a preliminary review of the literature, we have gained an understanding of the research trends and literature distribution in the field of study tourism. Finally, we selected the 100 most relevant literature for topic analysis. The top 100 academic papers with the highest relevance to study tourism in the past ten years were obtained. This article uses crawler technology to obtain the titles and abstracts of the documents.

This article uses the Goseeker crawler program to automatically obtain WOS database information through this software, and extract the titles and abstracts of relevant documents.

After obtaining the title and abstract of the literature, this article cleaned and processed the data. First, remove empty data and duplicate data. Then, format the data and unify the text format. For this text preprocessing, we will use the NLTK English word segmentation tool based on Jupyter. NLTK (Natural Language Toolkit) is a commonly used toolkit in the field of natural language processing (NLP). It provides rich text processing functions, including English word segmentation and part-of-speech tagging, named entity recognition, etc. (Chen et al., 2019).

First, we use the following to import NLTK under Jupyter, download the corpus, and use NLTK's word_tokenize function for word segmentation. This function splits the text into individual words and returns a list. Secondly, stopwords are introduced to remove stop words in the text, that is, words that have no semantic meaning, such as "of", "the", "an", etc. At the same time, NLTK's WordNetLemmatizer is used for lemmatization, for example, "learned" is restored to "learn". For the division of parts of speech, such as nouns, verbs, adjectives, etc., we use NLTK's pos_tag function for part-of-speech tagging. In order to avoid the separation of specific nouns or phrases (such as "educational tourism, international educational tourism, empirical research", etc.), we introduced a custom dictionary function to build a dictionary for this study. After the word segmentation is completed, we use the wordcloud library to generate the word cloud diagram, and use matplotlib under jupyter to display the word cloud. Figure 3 shows the keyword word cloud generated after text preprocessing. The larger keywords indicate more frequently discussed themes in educational study tourism research.

Figure 3: Study Tourism Research Keywords Word Cloud



scikit-learn library to provide the Tfidf Vectorizer class for generating TF-IDF to calculate TF-IDF (Agrawal et al. al., 2018), the results are as follows:

Table 1 presents the word frequency and TF-IDF values of the main keywords in educational study tourism research. The results show that “educational tourism”, “students”, “economic growth”, “China”, and “international students” are among the most frequently identified terms, indicating the major thematic concerns in the selected literature.

Table 1: Study tourism theme TF-IDF value table

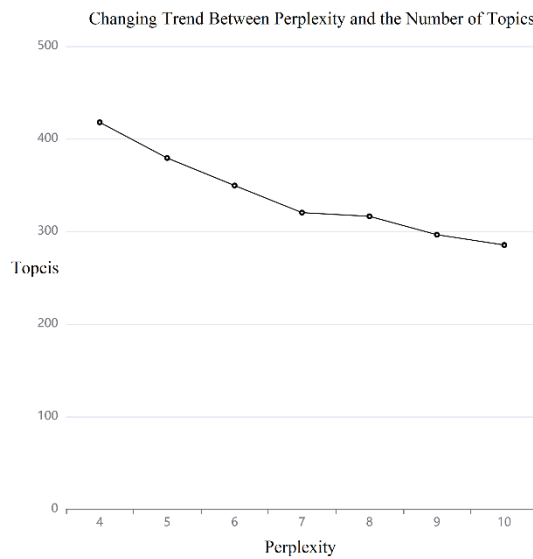
No.	Phrase	Word frequency	Number of words	TF-IDF
1	Educational Tourism	129	51	0.002954
2	Students	45	30	0.002044
3	Economic Growth	36	25	0.001911
4	China	34	22	0.001931
5	International Students	32	18	0.002135
6	Japan	27	21	0.001629
7	The Development	25	19	0.001613
8	Malaysia	24	15	0.001781
9	Economic Growth	23	16	0.001646
10	Tourism Activities	17	11	0.001475
11	Rural Areas	17	6	0.001873
12	Education	14	13	0.00112
13	Factors	14	13	0.00112
14	Perceptions	13	9	0.001231
15	Quality Education	12	11	0.001041
16	The Destination	11	8	0.001092
17	The University	10	5	0.001169
18	Educational Tourism Demand	9	6	0.000992
19	Satisfaction	8	8	0.000794
20	Interest	7	2	0.00103
21	Culture	7	5	0.000818

After word segmentation, topic analysis requires determining the optimal number of topics, a key hyperparameter in LDA. Too few topics may overlook semantic nuances, producing overly broad results, while too many can cause overfitting and redundant topics (Zhao et al., 2015). To select the optimal number, we use perplexity, an indicator of text model performance (Gan & Qi, 2021). Generally, perplexity decreases with topic growth and stabilizes, with the inflection point indicating the best choice (Cao et al., 2009).

In practice, the gensim library was applied in Jupyter to construct the LDA model and compute perplexity, with results visualized via matplotlib. Alpha and beta were adjusted to control distribution sparsity (Neishabouri & Desmarais, 2020). Analysis showed that when the number of topics reached seven, perplexity stabilized, suggesting seven as the optimal number of topics.

As shown in Figure 5, the perplexity value gradually stabilizes when the number of topics reaches seven, indicating that seven topics are appropriate for the LDA model.

Figure 5: Changing Trend Between Perplexity and the Number of Topics



After determining the optimal number of topics, we will perform LDA topic analysis. We use the LDA Model library of the gensim library to establish the number of topics (Cao et al., 2009). Under these 7 topic words, display the occurrence of each topic word. The frequency and document distribution probability are shown in the table2 and figure 6. Regarding the topic-word distribution, firstly, the 10 most popular words under each topic word were selected for identification, and secondly, the total scores of these seven hot words were sorted.

Table 2 summarizes the TF-IDF values and distribution of the seven identified topic categories generated by the LDA model. The results indicate that educational tourism, experience, students, international educational tourism, rural tourism, influencing factors, and economic development represent the main research hotspots in this field.

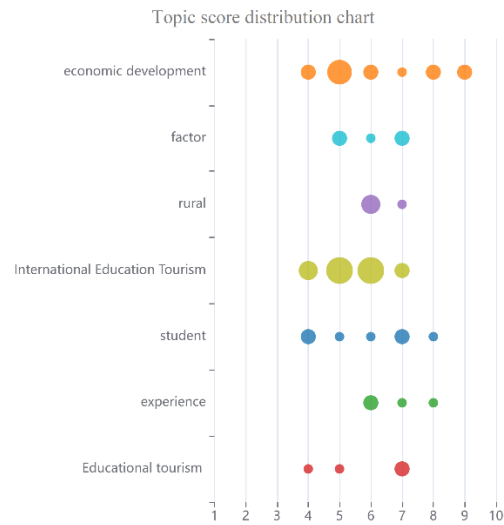
Table 2: Word TF-IDF Value Of 7 Topic Words

No	Topic	Number of items	TF-IDF	TF-IDF mean
1	Educational tourism	122(25.00%)	0.1161	0.0058
2	experience	120(22.73%)	0.1021	0.005
3	student	92(13.64%)	0.0631	0.0053
4	International Education Tourism	90(10.23%)	0.0528	0.0059
5	rural	90(10.23%)	0.0495	0.0055
6	Influencing factors	90(10.23%)	0.0457	0.0051
7	economic development	70(7.95%)	0.0388	0.0055

Figure 6 presents the topic score distribution of the seven identified topics and visually shows the relative importance of each research hotspot.

Table 3 further displays the representative hot words under each LDA topic. These keywords help explain the semantic content of each topic and provide a clearer basis for interpreting the research themes identified through the topic modelling process.

Figure 6: Topic score distribution chart



As can be seen from the above popularity chart, previous researchers focus on student-dominated educational tourism activities, internationalization, influencing factors, education, service, demand, economy, development and schools. At the same time, it can be seen from the LDA analysis table that the research hot spots in recent years are educational tourism, experience, student, International Education Tourism, rural, influencing factors, economic development.

Table 3: Topic content hot words Table

Topic	Topic content hot words									
	1	2	3	4	5	6	7	8	9	10
Economic Development	Policymaker	Income	Stakeholder	Competitor	Employ	Relationship	Causality	Demand	Country	Qualification
Educational Tourism	Product	Strategy	Destination	Accessibility	Support	Motivation	Experience	Culture	School	Museum
Experience Factor	Requirement	Awareness	Product	Marketing	Exploratory	Activity	Interest	Farm	Issue	Education
International Education	Demand	Experience	Moral	Identity	Price	Psychology	Global	Visitor	Process	Activity
Tourism	Culture	Experience	Communication	Global	Environmental	Sustainability	University	Destination	Inhabitant	Outdoor
Rural Tour	Exploratory	Farm	Agriculture	Environment	Experience	Geopark	Destination	Resident	Satisfaction	
Student	Learn	Product	Knowledge	Exploratory	Teacher	Abroad	Skill	University	Ability	Project

4. Results

4.1. Educational Tourism

Jang et al. (2022) reviewed how educational and dark tourism concepts spread in Japan, using the Fukushima Hope Tour as an example to illustrate the development of educational dark tourism. Educational tourism, combining research learning and travel, broadens tourists' horizons. Yet it remains a neglected subsector in China's tourism market (Si et al., 2019). Drawing on Britain's experience, the study suggests strategies to accelerate its development in China. Maga and Nicolau (2018) view educational tourism as both new and ancient, defining its scope, distinguishing it from other tourism and educational activities, and proposing a framework for classifying academic mobility. They identify significant growth potential through ICTs in creating value-added products and services that drive local economies. McGladdery and Lubbe (2017) argue that educational tourism remains under-researched, particularly the school tourism market. Their work offers a process-driven, results-based model, outlining future directions for school and international educational tourism.

4.2. Experience

The interrelationship between adult learning and educational tourism is discussed, offering a framework for procedural and methodological analyses of offline and online experiences adaptable to educational institutions. The economic, social, and environmental significance of educational tourism is also emphasized (McGladdery & Lubbe, 2017). Personal experiences and interests influence participants' evaluation of its value (Zatsepina et al., 2020). Tomasi et al. (2020) noted that visitors' overall experience, including leisure and tourism activities, requires further research to develop intervention models. Zaleckienė and Turčinskaitė (2019) analyzed rural educational tourism in Lithuania, identifying drivers and constraints, and highlighting work experience as a factor. Overall, experience remains a central focus in educational tourism research, both as an influencing factor and a primary research object.

4.3. Student

Nagai & Kashiwagi (2018) has shown that the number of high school students in Japan participating in overseas educational tourism programs has been increasing. The importance of educational tourism to students is mentioned, and educational tourism is a way to motivate students in the process of education, and the ability to visualize and implement knowledge under new technologies can arouse students' interest (Zatsepina et al., 2020). International students who are one of the main subjects of educational tourism can benefit themselves and the local educational tourism destinations (Tomasi et al., 2020). At the same time, it can be seen from much literature such as (Dembovska et al., 2016; Harazneh et al., 2018; Kharkovskaya et al., 2021), that the main object of educational tourism research is students. From the age range, students can be from primary school to university; from the geographical range, students can be local or domestic students or international students.

4.4. International Education Tourism

Based on Britain's successful experience, promoting internationalization can accelerate educational tourism in China (Si et al., 2019). McGladdery and Lubbe (2017) argued that international educational tourism contributes to global peace and identified future research directions through model construction. Nagai and Kashiwagi (2018) noted the growing participation of Japanese high schools in international educational tourism but highlighted the lack of overseas research. Using literature analysis, they discussed challenges by examining Japan's current trends and unique features, aiming to better understand its international educational tourism market. Tomasi et al. (2020) examined the economic impact of international educational tourism students. In addition, Bhuiyan et al. (2010) and Harazneh et al. (2018) explored aspects of national educational tourism. Clearly, international educational tourism forms a major branch of the field, attracting substantial research attention worldwide.

4.5. Rural

The development of rural educational tourism in Lithuania and its encouraging and limiting factors are analyzed (Zaleckienė & Turčinskaitė, 2019). Ohe (2018) noted that it is insufficient to view rural educational tourism only technically; it should also be studied from operators' identities. Petroman et al. (2016) argued that rural educational tourism is increasingly linked with recreational and educational leisure activities, enabling visitors to understand farmers' lives and shaping a positive image of agriculture for youth and urban consumers. Ohe (2017) and Tavares et al. (2023) also studied related aspects. Rural educational tourism has recently become an important field in educational tourism, largely due to the practicality of educational tourism and the applied nature of rural tourism projects.

4.6. Influencing Factors

McGladdery and Lubbe (2017) outlined the advantages, disadvantages, opportunities, and challenges of developing educational tourism products in Batur Global Geopark. Tang (2021) emphasized that policymakers should consider educational tourism levels, quality research output, and ICT development when formulating measures to stimulate economic growth through educational tourism. Zaleckienė & Turčinskaitė (2019) analyzed rural educational tourism in Lithuania and factors encouraging or limiting its development. Motivational factors were also conceptualized. Harazneh et al. (2018) found that cost, quality, environmental, regulatory, cultural, political, security, and social issues influence educational tourism for foreign students in Northern Cyprus. Other studies (Arcodia et al., 2021) also examined influencing factors, highlighting them as a major focus in educational tourism research.

4.7. Economic Development

The purpose of developing the tourism market is to prosper the economy, and educational tourism, as a key segment, is no exception. Research on its economic impact has become a hotspot. For instance, Nagai & Kashiwagi (2018) highlighted its significant role in Japan and the global tourism industry. Tang (2021) used a balanced panel dataset of 61 countries to examine the effect of educational tourism on economic growth. Matahir & Tang (2017) extended the Solow model to Malaysia, showing cointegration among economic growth, educational tourism, and other determinants. Results indicate

a two-way short-term causality, but a one-way long-term Granger causality from educational tourism to growth, offering insights for policy to sustain growth. Other scholars (Ohe, 2018; Tang et al., 2019; Vieira et al., 2022) also explored the reciprocal relationship between educational tourism and economic development.

5. Conclusion

The author identified seven hot topics in educational tourism research: experience, students, international educational tourism, rural tourism, influencing factors, and economic development. However, gaps remain, including research on educational tourism quality, awareness, behavioural intention, demand, satisfaction, primary and secondary school tourism, and other types beyond rural tourism, such as industry and cultural heritage. Financial support, mainly from families, and government policies, as well as social attitudes toward educational tourism, also require further study. While educational tourism impacts the economy, education should remain its primary purpose. Broadly, educational tourism covers all age groups, but primarily targets students from primary to university levels. Its development involves multiple education and tourism sectors. Post-COVID-19, domestic and international educational tourism is expected to recover rapidly, generating new phenomena. This study, despite limitations in literature selection and data capture, employs Latent Dirichlet Allocation (LDA) to identify current research hotspots, offering references and insights for future international educational tourism research.

Ethics Approval and Consent to Participate

The researchers conducted this study in accordance with academic research ethics standards of Universiti Teknologi MARA (UiTM), Malaysia. This study was based on published literature data retrieved from the Web of Science database and did not involve direct participation of human subjects.

Acknowledgement

Part of this article was developed from the doctoral research of the first author in Hotel and Tourism Management at Universiti Teknologi MARA (UiTM), Malaysia.

Funding

This work was supported by the Shandong Provincial Key Project of Art and Science (Project Title: “Resilience and Growth of Senior Educational Tourism in the Silver Economy Era: Global Perspectives and Localization Insights from Shandong, China”) [Grant No. 25ZR20021287]; and the Shandong Provincial Philosophy and Social Science Innovation and Development Research Association (Project Title: “Mechanisms for Constructing Cultural IP and High-Quality Development of Senior Educational Tourism in Shandong Province under the Silver Economy Background”) [Grant No. Z-SKY26023].

Conflict of Interest

The authors reported no conflicts of interest for this work and declare that there is no potential conflict of interest with respect to the research, authorship, or publication of this article.

References

- Agrawal, A., Fu, W., & Menzies, T. (2018). What is wrong with topic modeling? And how to fix it using search-based software engineering. *Information and Software Technology, 98*, 74–88. <https://doi.org/10.1016/j.infsof.2018.02.005>
- Ali, T., Omar, B., & Soulaïmane, K. (2022). Analyzing tourism reviews using an LDA topic-based sentiment analysis approach. *MethodsX, 9*, 101894. <https://doi.org/10.1016/j.mex.2022.101894>
- Arcodia, C., Abreu Novais, M., Cavlek, N., & Humpe, A. (2021). Educational tourism and experiential learning: Students' perceptions of field trips. *Tourism Review, 76*(1), 241–254. <https://doi.org/10.1108/TR-10-2019-0403>
- Bhuiyan, M. A. H., Islam, R., Siwar, C., & Ismail, S. M. (2010). Educational tourism and forest conservation: Diversification for child education. *Procedia - Social and Behavioral Sciences, 7*, 19–23. <https://doi.org/10.1016/j.sbspro.2010.10.004>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research, 3*, 993–1022. <https://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf>
- Cao, J., Xia, T., Li, J., Zhang, Y., & Tang, S. (2009). A density-based method for adaptive LDA model selection. *Neurocomputing, 72*(7–9), 1775–1781. <https://doi.org/10.1016/j.neucom.2008.06.011>
- Chen, Y., Zhang, H., Liu, R., Ye, Z., & Lin, J. (2019). Experimental explorations on short text topic mining between LDA and NMF based schemes. *Knowledge-Based Systems, 163*, 1–13. <https://doi.org/10.1016/j.knosys.2018.10.006>
- Ritchie, B. W., Carr, N., & Cooper, C. P. (2003). *Managing educational tourism*. Channel View Publications. <https://www.multilingual-matters.com/page/detail/managing-educational-tourism/>
- Dembovska, I., Silicka, I., & Łubkina, V. (2016). Educational tourism in the training of future tourism professionals. *Society. Integration. Education. Proceedings of the International Scientific Conference, 4*, 245–255. <https://doi.org/10.17770/sie2016vol4.1561>
- Gan, J., & Qi, Y. (2021). Selection of the optimal number of topics for LDA topic model—Taking patent policy analysis as an example. *Entropy, 23*(10), 1301. <https://doi.org/10.3390/e23101301>
- Godin, F., Slavkovikj, V., De Neve, W., Schrauwen, B., & Van de Walle, R. (2013). Using topic models for Twitter hashtag recommendation. In *Proceedings of the 22nd International Conference on World Wide Web* (pp. 593–596). Association for Computing Machinery. <https://doi.org/10.1145/2487788.2488002>
- Harazneh, I., Al-Tall, R. M., Al-Zyoud, M. F., & Abubakar, A. M. (2018). Motivational factors for educational tourism: Marketing insights. *Management & Marketing, 13*(1), 796–811. <https://doi.org/10.2478/mmcks-2018-0006>
- Jang, K., Sakamoto, K., & Funck, C. (2022). Dark tourism as educational tourism: The case of “hope tourism” in Fukushima, Japan. In D. T. Duong & P. Stone (Eds.), *Dark tourism studies* (pp. 104–115). Routledge. <https://doi.org/10.4324/9781003266723-8>
- Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: Models, applications, a survey. *Multimedia Tools and Applications, 78*, 15169–15211. <https://doi.org/10.1007/s11042-018-6894-4>
- Kharkovskaya, E. V., Posokhova, N. V., Efremova, N. V., & Miroshnichenko, E. V. (2021). Analysis of the cultural and educational tourism development in the Belgorod

- region. *Revista on line de Política e Gestão Educacional*, 25(esp.3), 3550–3567. <https://doi.org/10.22633/rpge.v25iesp.3.16116>
- Kherwa, P., & Bansal, P. (2020). Topic modeling: A comprehensive review. *EAI Endorsed Transactions on Scalable Information Systems*, 7(24), 1–16. <https://doi.org/10.4108/eai.13-7-2018.159623>
- Kurt, S., Assefa, F. Z., Tuzlukaya, S. E., & Karatepe, O. M. (2024). A 30-year journey of hospitality and tourism research: A comprehensive topic modeling analysis. *International Journal of Contemporary Hospitality Management*, 36(7), 2232–2255. <https://doi.org/10.1108/IJCHM-01-2023-0109>
- Lancichinetti, A., Siner, M. I., Wang, J. X., Acuna, D., Kording, K., & Amaral, L. A. N. (2015). High-reproducibility and high-accuracy method for automated topic classification. *Physical Review X*, 5(1), 011007. <https://doi.org/10.1103/PhysRevX.5.011007>
- Lee, J. H., Wood, J., & Kim, J. (2021). Tracing the trends in sustainability and social media research using topic modeling. *Sustainability*, 13(3), 1269. <https://doi.org/10.3390/su13031269>
- Maga, A., & Nicolau, P. (2018). Conceptualizing educational tourism and the educational tourism potential: Evidence from ASEAN countries. In *Proceedings of the International Scientific Conference “Competitive, Sustainable and Secure Development of the Regional Economy: Response to Global Challenges” (CSSDRE 2018)* (pp. 343–348). Atlantis Press. <https://doi.org/10.2991/cssdre-18.2018.72>
- Matahir, H., & Tang, C. F. (2017). Educational tourism and its implications on economic growth in Malaysia. *Asia Pacific Journal of Tourism Research*, 22(11), 1110–1123. <https://doi.org/10.1080/10941665.2017.1373684>
- McGladdery, C. A., & Lubbe, B. A. (2017). Rethinking educational tourism: Proposing a new model and future directions. *Tourism Review*, 72(3), 319–329. <https://doi.org/10.1108/TR-03-2017-0055>
- Nagai, H., & Kashiwagi, S. (2018). Japanese students on educational tourism: Current trends and challenges. In G. Richards & C. A. Arcodia (Eds.), *Asian youth travellers: Insights and implications* (pp. 117–134). Springer. https://doi.org/10.1007/978-981-10-8539-0_7
- Neishabouri, A., & Desmarais, M. C. (2020). Reliability of perplexity to find number of latent topics. In *Proceedings of the Thirty-Third International Florida Artificial Intelligence Research Society Conference* (pp. 246–251). AAAI Press. <https://cdn.aaai.org/ocs/18442/18442-79381-1-PB.pdf>
- Nurhuda, A. (2023). Benchmarking and exploring educational tourism in Malaysia. *Journal of English Language Teaching, Literature and Culture*, 2(1), 1–11. <https://doi.org/10.53682/jeltec.v2i1.6024>
- Ohe, Y. (2017). Assessing managerial efficiency of educational tourism in agriculture: Case of dairy farms in Japan. *Sustainability*, 9(11), 1931. <https://doi.org/10.3390/su9111931>
- Ohe, Y. (2018). Educational tourism in agriculture and identity of farm successors. *Tourism Economics*, 24(2), 167–184. <https://doi.org/10.1177/1354816617729021>
- Omar, M., On, B.-W., Lee, I., & Choi, G. S. (2015). LDA topics: Representation and evaluation. *Journal of Information Science*, 41(5), 662–675. <https://doi.org/10.1177/0165551515587839>
- Ormond, M., & Vietti, F. (2022). Beyond multicultural “tolerance”: Guided tours and guidebooks as transformative tools for civic learning. In A. Wearing, C. M. Hall, & T. D. Duong (Eds.), *The Routledge handbook of tourism and hospitality education* (pp. 180–193). Routledge. <https://doi.org/10.4324/9781003299578-14>

- Petroman, C., Mirea, A., Lozici, A., Constantin, E. C., Marin, D., & Merce, I. (2016). The rural educational tourism at the farm. *Procedia Economics and Finance*, 39, 88–93. [https://doi.org/10.1016/S2212-5671\(16\)30245-1](https://doi.org/10.1016/S2212-5671(16)30245-1)
- Si, Y., Zeng, Y., Wang, J., Wang, H., & Yang, F. (2019). Characteristics and implications of British educational tourism. In *Proceedings of the 2nd International Seminar on Education Research and Social Science (ISERSS 2019)* (pp. 1420–1423). Atlantis Press. <https://doi.org/10.2991/iserss-19.2019.264>
- Sjøen, M. M. (2023). From global competition to intercultural competence: What teacher-training students with cross-cultural teaching experience should be learning. *Scandinavian Journal of Educational Research*, 67(1), 140–153. <https://doi.org/10.1080/00313831.2021.1990121>
- Sun, F.-R., & Xu, T.-Z. (2021). The study tour in China: An emerging mode for practical education. *Sustainability*, 13(21), 11969. <https://doi.org/10.3390/su132111969>
- Tang, C.-F., Cheah, Y.-K., & Chua, S. Y. (2019). Does educational tourism significantly influence economic growth? Evidence from a macro-econometric modelling. *International Journal of Business and Society*, 20(3), 924–935. <https://www.ijbs.unimas.my/images/repository/pdf/Vol20-no3-paper4.pdf>
- Tang, C. F. (2021). The threshold effects of educational tourism on economic growth. *Current Issues in Tourism*, 24(1), 33–48. <https://doi.org/10.1080/13683500.2019.1694869>
- Tang, C. F., Law, S. H., & Harun, M. (2022). Demand for inbound educational tourism: A note on relative differences and non-linear relationships. *Current Issues in Tourism*, 26(3), 341–346. <https://doi.org/10.1080/13683500.2022.2026305>
- Tavares, B., Minasi, S., & Pagnussat, E. (2023). Turismo rural pedagógico em fazendas de café: Estratégias interdisciplinares no ensino médio. *Revista Rosa dos Ventos - Turismo e Hospitalidade*, 15(1), 226–243. <https://doi.org/10.18226/21789061v15i1p226>
- Tomasi, S., Paviotti, G., & Cavicchi, A. (2020). Educational tourism and local development: The role of universities. *Sustainability*, 12(17), 6766. <https://doi.org/10.3390/su12176766>
- Vieira, G., Pocinho, M., Nunes, C., & Sales Sousa, S. (2022). Engaging in educational tourism: Academic response. *Journal of Tourism, Sustainability and Well-being*, 10(2), 101–113. <https://doi.org/10.34623/as81-bk29>
- Yau, C.-K., Porter, A., Newman, N., & Suominen, A. (2014). Clustering scientific documents with topic modeling. *Scientometrics*, 100, 767–786. <https://doi.org/10.1007/s11192-014-1321-8>
- Zaleckienė, J., & Turčinskaitė, L. (2019). Development of educational tourism in Lithuanian rural areas. In *Research for Rural Development 2019: Annual 25th International Scientific Conference Proceedings* (Vol. 2, pp. 169–174). Latvia University of Life Sciences and Technologies. <https://doi.org/10.22616/rrd.25.2019.065>
- Zatsepina, M. B., Kriskovets, T. N., Vorobyev, V. K., Kolobova, L. V., Grigorieva, N. V., Fedulov, V. I., & Stolyarova, A. N. (2020). Educational tourism: Tribute to fashion or new educational opportunities? *Propósitos y Representaciones*, 8(SPE2), e642. <https://doi.org/10.20511/pyr2020.v8nSPE2.642>
- Zhao, W., Chen, J. J., Perkins, R., Liu, Z., Ge, W., Ding, Y., & Zou, W. (2015). A heuristic approach to determine an appropriate number of topics in topic modeling. *BMC Bioinformatics*, 16, (Suppl 13), S8. <https://doi.org/10.1186/1471-2105-16-S13-S8>