Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 469-479
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5294

Tourism and Travel Content Analysis for Market Segmentation using Toxicity and Sentiment Classification in Communalytic

Yerik Afrianto Singgalen

Faculty of Business Administration and Communication, Tourism Study Program, Atma Jaya Catholic University of Indonesia, Jakarta, Indonesia
Email: yerik.afrianto@atmajaya.ac.id
Correspondence Author Email: yerik.afrianto@atmajaya.ac.id
Submitted: **06/06/2024**; Accepted: **30/06/2024**; Published**: 30/06/2024**

Abstract−This research highlights the significant impact of digital content on shaping tourist perceptions and behaviors, particularly emphasizing the influence of travel vlogs. Utilizing the Tourism and Travel Content Analysis (TTCA) framework, the study analyzed 1,972 review posts out of 2,250, revealing critical insights into viewer engagement and sentiment. Toxicity score calculations indicated prevalent negative interactions, with scores ranging from 0.05542 to 0.86967 for Toxicity, 0.00536 to 0.50704 for Severe Toxicity, 0.01921 to 0.59834 for Identity Attack, 0.03305 to 0.76573 for Insult, 0.03737 to 0.78492 for Profanity, and 0.01075 to 0.48617 for Threat, underscoring the need for compelling content moderation. Sentiment analysis using VADER and TextBlob demonstrated a generally positive reception of travel vlogs, with VADER classifying 3.73% of posts as unfavorable, 19.83% as neutral, and 76.44% as positive. In comparison, TextBlob classified 2.71% of posts as unfavorable, 35.59% as neutral, and 61.69% as positive for English posts. Notably, VADER and TextBlob agreed on sentiment classification for 446 out of 587 posts (75.98%), with a Cohen's kappa statistic of 0.471, indicating moderate agreement. These findings suggest that well-regulated and thoughtfully designed digital content significantly enhances user engagement and optimizes destination marketing strategies. Future research should incorporate advanced analytical tools and comprehensive data sets to refine these insights further, supporting the development of more targeted and effective marketing efforts in the tourism sector. This study thus contributes to a deeper understanding of digital media's impact on tourism marketing, offering practical recommendations for leveraging content to foster positive and engaging tourist experiences.

Keywords: TTCA; VADER; TextBlob; Perspective; Market Segmentation

1. INTRODUCTION

Analyzing tourism and travel content from a market perspective is essential to understand the perceptions and behaviors within the tourism sector. A thorough examination of how potential travelers interact with and interpret travel-related information reveals crucial insights into decision-making processes [1]–[5]. This approach not only aids in identifying trends and preferences but also assists in tailoring marketing strategies to better align with consumer expectations [6]–[11]. Businesses develop more effective promotional tactics and enhance customer satisfaction by delving into specific market dynamics [12]. Consequently, targeted content analysis from a market-oriented viewpoint significantly contributes to tourism marketing endeavors' overall efficiency and success.

Challenges in destination marketing include aligning tourists' preferences with the availability of facilities at tourist destinations and other aspects requiring the involvement of various stakeholders. Effective marketing strategies must consider travelers' diverse needs and expectations, ensuring destinations have the appropriate amenities to meet these demands [13]–[18]. Collaboration among government entities, local businesses, and community organizations is imperative to enhance the visitor experience and maintain competitive appeal [19]–[23]. Addressing these multifaceted challenges necessitates a holistic approach integrating input from all relevant parties, ultimately leading to a more cohesive and successful marketing strategy for tourist destinations.

The digital content era stimulates significant discourse in destination marketing, particularly regarding its influence on travel intentions. Travel influencers are pivotal in shaping tourists' decisions through engaging video content shared on various media platforms [24]. These visual narratives effectively capture the essence of destinations, making them more appealing to potential visitors [25]. This phenomenon underscores the importance of incorporating influencer collaborations into marketing strategies to enhance destination visibility and attractiveness [26]. Consequently, travel influencers' integration of digital content has become indispensable in contemporary destination marketing practices.

This research analyzes public sentiment toward travel vlog videos to generate recommendations for developing tourist destinations that align with traveler preferences. Sentiment analysis of these videos provides valuable insights into visitor experiences and expectations, highlighting areas for improvement and enhancement [27]. Understanding the emotional responses elicited by travel content allows for the formulation of targeted strategies to meet the desires of potential tourists [28]. Consequently, leveraging sentiment analysis in travel vlogs significantly informs destination development, ensuring that offerings are closely aligned with the evolving preferences of the tourism market.

The urgency of this research lies in its potential to address critical gaps in understanding consumer behavior within the tourism sector. The study seeks to provide actionable insights for enhancing destination marketing strategies by examining contemporary trends and preferences [29]. In a rapidly evolving digital landscape, timely and relevant data is essential for staying competitive and meeting the dynamic needs of travelers [30]. Therefore, the findings from

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 469-479
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5294

this research are expected to contribute significantly to developing more effective, evidence-based approaches in tourism marketing, ensuring destinations remain attractive and relevant in a global market.

This research's theoretical and practical contributions are significant in advancing the field of tourism studies and improving industry practices. Theoretically, the study enhances the existing body of knowledge by providing a nuanced understanding of how digital content influences tourist behavior and destination perception [31]. Practically, the insights gained from the research offer valuable guidelines for tourism marketers to develop more targeted and effective promotional strategies. These contributions not only bridge the gap between theory and practice but also support the creation of more engaging and appealing tourist experiences, thereby fostering sustainable tourism development.

Similar research in the field of tourism marketing has explored the impact of digital media on travel behavior. Yet, this study uniquely focuses on the specific role of travel vlog videos in shaping public sentiment. Previous studies have generally emphasized broader digital marketing strategies without delving into the nuances of user-generated video content [32]. However, a limitation of this research is the potential bias in sentiment analysis due to the subjective nature of interpreting viewer comments and interactions. Despite this constraint, the study's findings are expected to provide valuable insights, though future research should aim to incorporate more diverse data sources to enhance the robustness of the conclusions.

Recommendations for further research in tourism marketing should include a deeper exploration of the interactive effects of various digital content types on tourist behavior. Investigating the combined impact of social media posts, blogs, and travel vlogs across different demographic segments could provide a more comprehensive understanding of digital influence. Additionally, employing advanced analytical methods, such as machine learning and big data analytics, could enhance the precision of sentiment analysis and reduce bias. These future studies will likely yield more granular insights, ultimately contributing to more effective and personalized marketing strategies in the tourism industry.

2. RESEARCH METHODOLOGY

2.1 Gap Analysis

Enhancing studies on tourism destination marketing is crucial to gaining insights into tourist preferences and evolving trend demands. Identifying gaps in current research reveals that data review methodologies need optimization, mainly by applying Artificial Intelligence (AI) for large-scale data processing [33]. AI's advanced analytical capabilities offer significant improvements in extracting meaningful patterns and trends from extensive datasets [34]. Therefore, integrating AI into tourism marketing analysis promises a more accurate and efficient understanding of tourist behaviors and preferences, ultimately leading to more tailored and effective marketing strategies.

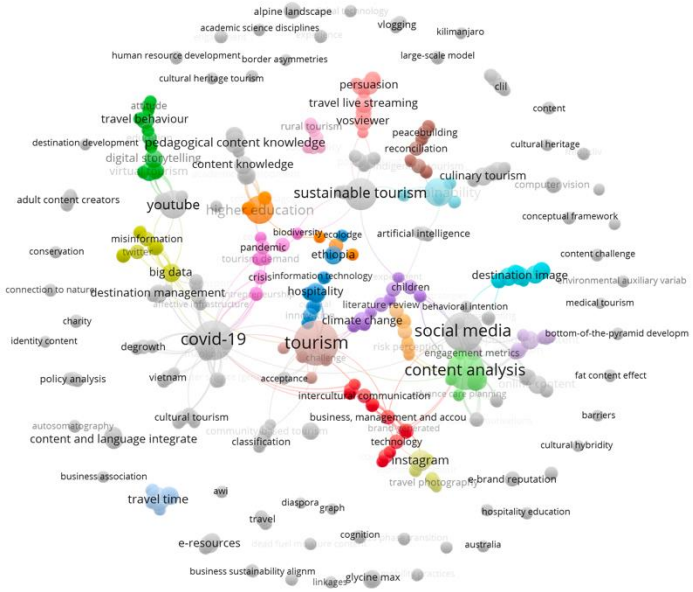

Figure 1. Network Visualization of Tourism Digital Marketing (VosViewer)

Figure 1 shows the network visualization of tourism digital marketing. The topic of tourism destination digital marketing content warrants further exploration, especially given the accessibility and motivational impact of travel vlogs on the public. Travel vlogs are potent tools for showcasing destinations, often inspiring viewers to visit through vivid, personal narratives [35]. These videos offer authentic experiences that traditional marketing methods cannot replicate, making them highly effective in influencing travel decisions [36]. Delving deeper into this subject could

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 469-479
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5294

unveil new strategies for leveraging digital content to enhance destination appeal. Consequently, a focused study on travel vlog content could significantly advance the understanding and application of digital marketing in tourism.

Based on the identified gaps in similar studies, this research employs the Tourism and Travel Content Analysis (TTCA) framework, which is pertinent to the data context and processing objectives to produce information relevant to tourism marketing. The TTCA framework facilitates a structured approach to analyzing diverse data sources, ensuring the resulting insights are comprehensive and actionable. By integrating this framework, the study aims to enhance the precision and relevance of the findings, thus contributing to more effective marketing strategies. Consequently, the TTCA framework proves essential in bridging existing research gaps and advancing the field of tourism marketing analysis.

2.2 Tourism and Travel Content Analysis (TTCA)

The Tourism and Travel Content Analysis (TTCA) framework evaluates public perceptions of digital destination content discussed through traveler vlog videos. This framework systematically assesses viewer responses and engagement, providing a detailed understanding of how such content influences audience opinions and travel intentions. By leveraging TTCA, it is possible to identify critical factors that resonate with viewers, informing more effective content creation and marketing strategies. Therefore, applying TTCA in analyzing travel vlogs is pivotal for optimizing digital tourism marketing and enhancing destination appeal based on empirical insights.

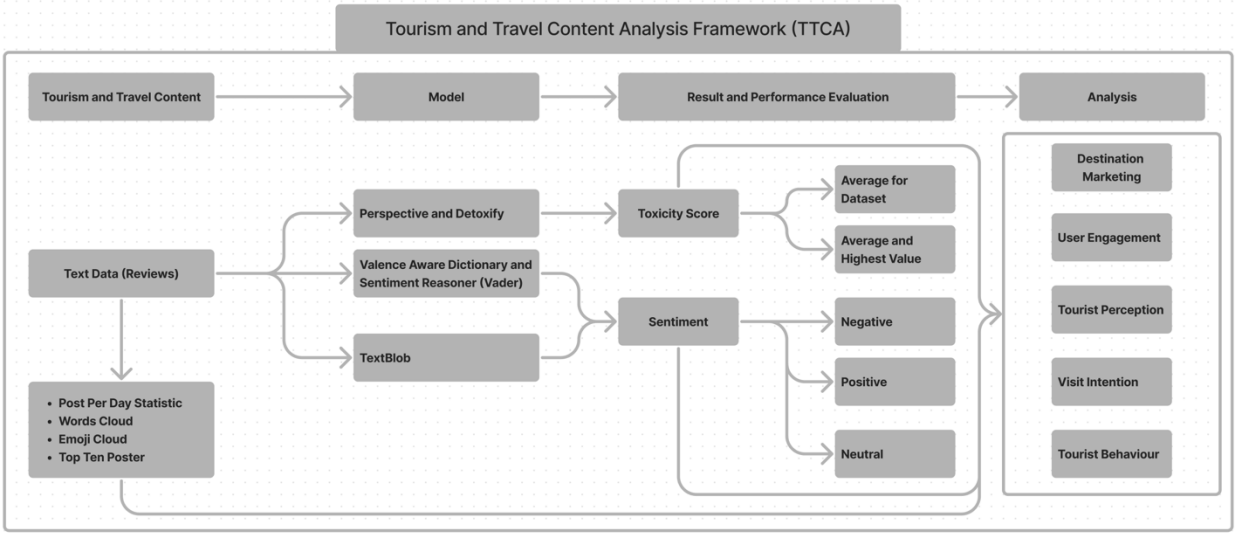

Figure 2. Tourism and Travel Content Analysis Framework

Figure 2 shows the implementation of the TTCA framework. The Tourism and Travel Content Analysis (TTCA) framework excels in transforming review data into marketing opportunities aligned with traveler preferences. This framework systematically processes and interprets diverse data sources, extracting valuable insights into tourist expectations and behaviors. By leveraging TTCA, destinations tailor marketing strategies to address specific consumer desires, enhancing appeal and competitiveness. Such targeted approaches are crucial in today's dynamic tourism landscape, where understanding and responding to nuanced traveler preferences significantly impact destination success. Consequently, TTCA is essential for optimizing tourism marketing through data-driven insights.

Considering these factors, this research utilizes the TTCA framework in processing travel vlog review data to understand viewer responses that could form the tourism market. TTCA systematically analyzes viewer comments and interactions, providing insights into preferences and expectations. This methodological approach ensures the data interpretation is comprehensive and relevant to current market trends. Employing TTCA in this context not only enhances the accuracy of the findings but also offers actionable insights for developing targeted marketing strategies. Consequently, applying TTCA is crucial in leveraging viewer responses to optimize tourism marketing efforts.

2.2.1 Text Data (Reviews)

This research focuses on processing review data from a travel vlog video uploaded to the sharing media platform YouTube with the ID EJAWYOApwJY, which has garnered 798,731 views since September 19, 2023, and received 2,255 comments. The analysis aims to extract valuable insights from these interactions, highlighting viewer preferences and sentiments towards the destination featured in the vlog. Understanding these responses is essential for developing effective tourism marketing strategies that resonate with potential travelers. Consequently, the findings from this study are expected to provide actionable recommendations for optimizing digital content in tourism promotion.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 469-479
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5294

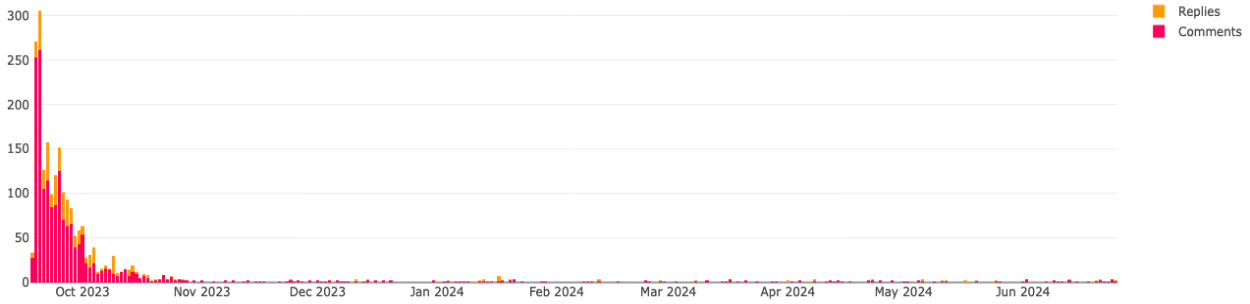

Figure 3. Post-per-day Statistic of the Video (Communalytic)

Figure 3 shows the content's post-per-day statistics. Based on the post-per-day statistics from the analyzed video, significant viewer responses are observed on specific dates: September 20, 2023, with 261 posts; September 19, 2023, with 253 posts; September 25, 2023, with 125 posts; September 22, 2023, with 115 posts, and September 21, 2023, with 105 posts. These spikes in engagement indicate particular interest and interaction from viewers during these periods. Such data points highlight critical moments of viewer engagement, suggesting heightened interest or reactions to the content. Analyzing these trends is crucial for understanding audience behavior and optimizing future content strategies. Consequently, identifying these patterns provides valuable insights for targeted marketing and content development in tourism.

Post-per-day statistical data is essential for identifying the number of viewer responses in comments after a video is published. This data provides insights into the audience's engagement levels and the timing of interactions. Analyzing these statistics helps to pinpoint specific periods of heightened activity, indicating increased viewer interest or reactions to the video content. Such information is crucial for optimizing content release strategies and enhancing viewer engagement. Consequently, leveraging post-per-day data allows a more strategic approach to content management and audience interaction in digital marketing.

Figure 4. Top Ten Poster (Communalytic)

Figure 4 shows the top ten posters. The top ten poster data reveals the most active users engaging with the video content: @adijokka and @sahrilaja9610 with 14 posts each, followed by @KristianHansen and @ahsanshop6012 with 11 posts each. Additionally, @danyipung6462, @anibaays, and @rayhannumami861 contributed 9 posts each, while @mamirmakassarmamir8015, @madistewars2576, and @ArdiArdi-hp2zh each posted 8 times. This data highlights key individuals who significantly interact with the content, indicating a higher level of engagement and interest. Understanding the behavior of these top contributors provides valuable insights for targeted marketing strategies and community building. Consequently, identifying and analyzing the activity of these top posters is crucial for optimizing engagement and fostering a loyal audience base.

Identifying the top ten posters is essential for determining the number of posts based on the accounts' identities. This analysis provides insights into which users engage most actively with the content, highlighting frequent interaction patterns. Understanding these top contributors reveals essential trends and behaviors within the audience, informing targeted engagement and community-building strategies. Marketers tailor approaches by focusing on these active users to foster loyalty and enhance viewer participation. Consequently, identifying and analyzing the top ten posters is crucial for optimizing digital content strategies and enhancing audience engagement.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 469-479
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5294

Figure 5. Frequently Used Words (Communalytic)

Figure 5 shows the frequently used words. Based on the frequently used words, it is evident that "Makassar" (326 occurrences) and "Sulawesi" (179 occurrences) are the most frequently mentioned terms, indicating a strong focus on these locations. The words "kota" (69 occurrences) and "city" (53 occurrences) further emphasize urban settings as significant topics of discussion. Additionally, the frequent mention of "Indonesia" (85 occurrences) highlights a broader geographical context within the conversation. These insights suggest that content related to these areas, particularly Makassar and Sulawesi, is highly engaging to viewers. Consequently, leveraging these critical terms in marketing strategies could enhance content relevance and viewer interest in tourism-related promotions.

A word cloud is essential for identifying the terms most viewers highlight in comments, thereby allowing for the evaluation of content and destination management for tourism marketing purposes. This analytical tool visually represents the frequency of words used, offering immediate insights into the main topics and concerns of the audience. Marketers tailor strategies to address specific interests and enhance the destination's appeal by focusing on these key terms. Consequently, using a word cloud facilitates a more targeted and practical approach to optimizing tourism content and management practices.

Figure 6. Emoji Cloud (Communalytic)

Figure 6 shows the emoji clouds. Based on the emoji cloud analysis, the most frequently used emojis include the heart emoji (❤) with 98 occurrences, the thumbs-up emoji (👍) with 41 occurrences, and the laughing emoji (😂) with 40 occurrences. Other notable emojis, such as the smiling face (😊) with 38 occurrences and the prayer hands (🙏) with 29 occurrences, also feature prominently. These emojis suggest a positive and engaging viewer response, indicating satisfaction and enjoyment. Consequently, the prevalence of these emojis highlights the importance of fostering an emotionally resonant connection with the audience in tourism marketing efforts.

2.2.2 Model

The Perspective API calcifies toxicity scores at the model stage to reduce online toxicity. Perspective utilizes machine learning models to identify abusive comments, scoring phrases based on perceived impact on conversations.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 469-479
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5294

Developers and publishers use this scoring system to provide feedback to commenters, assist moderators in reviewing comments more efficiently, or help readers filter out harmful language. The models offer scores for attributes beyond the primary Toxicity attribute, including Severe Toxicity, Insult, Profanity, Identity Attack, Threat, and Sexually Explicit content. Consequently, implementing Perspective models enhances moderation and fosters a safer online environment.

The Perspective API offers two distinct methods for analyzing comments. The first method, AnalyzeComment, allows users to request a comment to be evaluated, after which a score is returned indicating the comment's toxicity level. The second method, SuggestCommentScore, enables users to propose an improved comment score, enhancing the toxicity assessment's accuracy and reliability. These methods collectively provide a robust framework for moderating online content, ensuring that harmful language is effectively identified and managed. Consequently, the implementation of these methods supports the creation of a safer and more respectful online environment.

The Vader and TextBlob models are utilized in sentiment analysis to evaluate textual data. Vader is particularly effective for analyzing social media content due to its ability to detect the intensity of sentiments. At the same time, TextBlob provides comprehensive insights by classifying text into positive, negative, or neutral sentiments. These models offer robust methodologies for interpreting the emotional tone of large datasets. Consequently, employing Vader and TextBlob enhances the accuracy and depth of sentiment analysis, enabling more informed decisions based on the emotional responses captured in the data.

The TextBlob model was compared with the performance of the Vader model, each offering distinct advantages in sentiment analysis. TextBlob excels in providing a more granular analysis by breaking down text into subjective and objective components, thereby offering a nuanced understanding of sentiment. TextBlob's simplicity and ease of implementation make it highly accessible for various applications. While Vader is particularly adept at handling the informal language typical of social media, TextBlob's comprehensive approach to sentiment classification presents a unique strength. Consequently, leveraging TextBlob yields more detailed insights into sentiment analysis, complementing the capabilities of the Vader model.

2.2.3 Result and Performance Evaluation

The Perspective model is evaluated based on the average and highest scores obtained from the classification results across several attributes: Severe Toxicity, Insult, Profanity, Identity Attack, and Threat. This evaluation method ensures a comprehensive assessment of the model's effectiveness in identifying abusive language. Analyzing the mean and peak values gives a more accurate representation of the model's performance. Consequently, this approach highlights areas where the model excels and identifies potential improvements, thereby enhancing its utility in moderating online content and maintaining a respectful digital environment.

The TextBlob and Vader models were evaluated using Cohen's kappa statistic by the distribution of polarity values. This statistical measure assesses the agreement between the two models in sentiment classification, accounting for chance agreement. By analyzing the polarity values, Cohen's kappa provides a robust metric to compare the consistency and reliability of TextBlob and Vader. Consequently, this evaluation method offers a comprehensive understanding of each model's performance, ensuring that outputs are aligned with the expected sentiment distributions and enhancing the credibility of sentiment analysis results.

2.2.4 Analysis

At the analysis stage, sentiment and toxicity analysis data are linked to discussions related to destination marketing, user engagement, tourist perception, visit intention, and tourist behavior. This integration provides a comprehensive understanding of how various factors influence tourists' decisions and interactions with digital content. By correlating sentiment and toxicity levels with these critical aspects, valuable insights were gained into the effectiveness of marketing strategies. Consequently, this approach enhances the precision of marketing campaigns and fosters a deeper connection with the target audience, ultimately driving better engagement and satisfaction.

This research focuses on the influence of travel vlogs on visit intention and the significant impact of destination marketing through digital content. Travel vlogs provide authentic and engaging visual narratives that shape potential tourists' perceptions and motivations. The compelling nature of these videos enhances viewers' desire to visit the featured destinations. Consequently, the strategic use of digital content in destination marketing plays a crucial role in driving visit intention, making it an indispensable tool for tourism promotion and engagement. This study highlights the importance of leveraging travel vlogs to optimize marketing efforts and attract tourists.

3. RESULT AND DISCUSSION

The discussion in this research is divided into three sections: first, the process of calculating toxicity scores using the Perspective model; second, the classification of sentiment based on review data; and third, the discussion on the influence of travel vlogs on visit intention. The initial section details the methodology and effectiveness of Perspective in identifying and scoring toxic comments. The subsequent section focuses on sentiment analysis, utilizing review data to classify and understand the emotional tone of viewer responses. Finally, the third section explores how travel

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 469-479
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5294

vlogs impact potential tourists' intentions to visit destinations. These comprehensive discussions provide valuable insights into the multifaceted aspects of digital content's role in tourism marketing and visitor engagement.

3.1 Toxicity Score and Analysis

Based on the toxicity score calculations using the Perspective model, the results are as follows: Toxicity scores range from 0.05542 to 0.86967, Severe Toxicity from 0.00536 to 0.50704, Identity Attack from 0.01921 to 0.59834, Insult from 0.03305 to 0.76573, Profanity from 0.03737 to 0.78492, and Threat from 0.01075 to 0.48617. These scores indicate varying levels of abusive language detected in the analyzed comments, with Profanity and Insult having the highest maximum scores, suggesting these types of toxic behavior are more prevalent. Consequently, these findings underscore the importance of effective moderation strategies to manage and mitigate online toxicity, particularly in digital content related to travel and tourism.

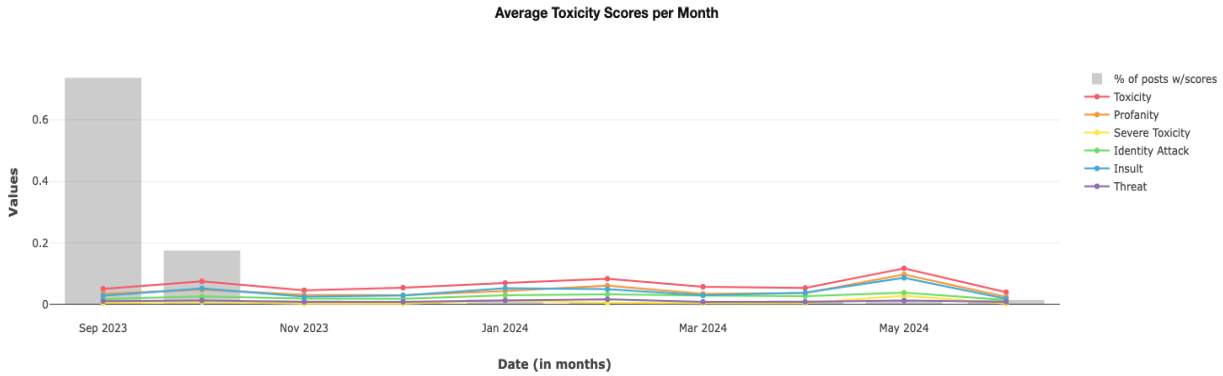

Figure 7. Toxicity Score (Communalytic)

Figure 7 shows the toxicity score based on the score per month. These figures indicate a spectrum of toxic behaviors in the comments, with Profanity and Insult exhibiting the highest maximum scores. This suggests that such forms of abusive language are more prevalent among the analyzed comments. Consequently, these insights highlight the necessity for robust moderation techniques to address and reduce online toxicity, thereby ensuring a healthier digital environment for discussions related to travel and tourism.

Ensuring a healthier digital environment for discussions related to travel and tourism is paramount for fostering positive and constructive interactions. Implementing advanced moderation tools, such as machine learning models like Perspective, effectively identifies and mitigates toxic comments, enhancing the overall quality of online conversations. Promoting community guidelines and encouraging respectful and inclusive behavior are crucial to a welcoming atmosphere. By combining technological solutions with proactive community management, digital platforms create a safer and more engaging space for users to share experiences and information about travel and tourism, ultimately benefiting both the audience and the industry.

Based on Communalytic's analysis of 1,972 posts out of 2,250, the toxicity score calculations indicate the necessity for well-regulated and thoughtfully designed content to foster a positive digital interaction environment, particularly concerning travel vlog content and tourism destination marketing. Effective content moderation strategies are essential to minimize the presence of toxic comments and enhance user engagement. Developing clear community guidelines and using advanced sentiment analysis tools contribute significantly to maintaining a constructive online atmosphere. Consequently, these measures are vital for ensuring digital platforms are beneficial spaces for promoting travel and tourism-related discussions.

3.2 Sentiment Classification

Based on the analysis of 616 out of 2,250 posts, the performance evaluation of the Vader and TextBlob sentiment classification models demonstrated optimal results. These models effectively identified and categorized the sentiments expressed in the posts, showcasing reliability and accuracy. The comparative analysis highlighted respective strengths in handling different types of textual data, contributing to a comprehensive understanding of audience sentiment. Consequently, the successful application of these models underscores the utility of sentiment analysis for improving content strategies and enhancing user engagement in digital marketing efforts.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 469-479
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5294

	# of Posts	Negative Sentiment [-1..-0.05]	Neutral Sentiment (-0.05..0.05)	Positive Sentiment [0.05..1]
VADER (English/EN)	590	22 (3.73%)	117 (19.83%)	451 (76.44%)
TextBlob (English/EN)	590	16 (2.71%)	210 (35.59%)	364 (61.69%)
TextBlob (French/FR)	2	0 (0.00%)	2 (100.00%)	0 (0.00%)
TextBlob (German/DE)	24	1 (4.17%)	21 (87.50%)	2 (8.33%)

Figure 8. Vader and TextBlob Performance in Sentiment Classification

Figure 8 shows the performance of the Vader and TextBlob models. The sentiment analysis of 590 posts using the VADER model in English revealed distinct sentiment distributions: 22 posts (3.73%) exhibited negative sentiment, 117 posts (19.83%) were categorized as neutral, and 451 posts (76.44%) displayed positive sentiment. These results indicate a predominantly positive response from the audience, with most posts reflecting favorable opinions. The relatively low percentage of negative sentiment underscores the overall positive engagement of users. Consequently, these findings validate the effectiveness of the VADER model in accurately classifying sentiments, providing valuable insights for enhancing content strategies in digital marketing.

The sentiment analysis conducted using TextBlob in different languages revealed the following distributions: for English (EN), out of 590 posts, 16 posts (2.71%) were negative, 210 posts (35.59%) were neutral, and 364 posts (61.69%) were positive. For French (FR), out of 2 posts, both (100%) were neutral. For German (DE), out of 24 posts, one post (4.17%) was negative, 21 posts (87.50%) were neutral, and two posts (8.33%) were positive. These findings highlight the effectiveness of TextBlob in classifying sentiments across multiple languages, with a notable predominance of positive and neutral sentiments, particularly in the English and German datasets. Consequently, these results underscore the utility of TextBlob for multilingual sentiment analysis in digital marketing.

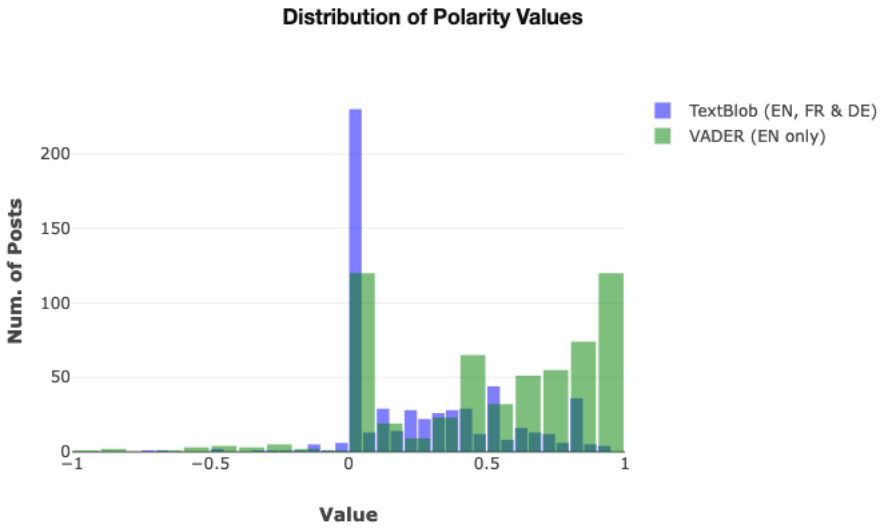

Figure 9. Distribution of Polarity Values

Figure 9 shows the distribution of polarity values. The VADER and TextBlob concur on categorizing 446 (75.98%) out of 587 English language posts, indicating a moderate agreement level (Cohen's kappa statistic = 0.471). Expressly, both sentiment analysis libraries agree on seven posts (1.57%) with negative sentiments (polarity scores ≤ -0.05), 96 posts (21.52%) with neutral sentiments (polarity scores between -0.05 and 0.05), and 343 posts (76.91%) with positive sentiments (polarity scores ≥ 0.05). This concurrence demonstrates a significant level of reliability in sentiment classification between the two models, reinforcing the validity of the sentiment analysis outcomes. Consequently, the alignment between VADER and TextBlob enhances confidence in an application for sentiment evaluation in digital content analysis.

Based on the analysis of the dataset, discrepancies between VADER and TextBlob sentiment assignments were observed. Specifically, there are seven posts where VADER assigned positive polarity scores while TextBlob assigned negative scores. Additionally, five posts were identified where VADER assigned negative polarity scores, contrasting with TextBlob's favorable assignments. These inconsistencies highlight the differences in sentiment detection algorithms and underscore the need for careful interpretation when using multiple models. Consequently, understanding these variations is crucial for improving sentiment analysis accuracy and ensuring reliable data-driven insights.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 469-479
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5294

3.3 Discussion: Travel Vlog Content and Visit Intention

The influence of travel vlog content on visit intention is significant, as these videos provide potential tourists with immersive and authentic experiences of destinations. Through engaging narratives and visual appeal, travel vlogs effectively capture the essence of a location, making it more attractive to viewers [37]. This persuasive power often translates into a heightened desire to visit the featured destinations, as viewers feel a personal connection to the experiences depicted. Consequently, leveraging travel vlog content as part of destination marketing strategies substantially enhances visit intention, demonstrating the critical role of digital media in contemporary tourism promotion.

The impact of toxicity and sentiment analysis on tourists' perceptions of a destination is profound, as these analyses provide insights into the overall discourse surrounding the destination. High toxicity levels deter potential visitors by highlighting negative interactions and hostile environments within the community. Conversely, positive sentiment analysis reflects favorable opinions and experiences, enhancing the destination's attractiveness [38], [39]. Understanding these dynamics allows for more effective management of online content and reputation. Consequently, leveraging sentiment and toxicity analysis is crucial for shaping tourists' perceptions and fostering a positive image of the destination, ultimately influencing the decision-making process.

Effective content management significantly enhances user engagement and optimizes the performance of destination marketing. Well-curated content that resonates with the audience's interests and preferences encourages more interaction and sharing, thereby increasing visibility [40]. Additionally, strategically designed content incorporating engaging visuals, compelling narratives, and informative elements attract a broader audience and sustain interest. Consequently, a robust content management strategy is crucial for maximizing the impact of marketing efforts, ensuring that the destination remains appealing and relevant in a competitive tourism market.

The limitations of this research primarily stem from the scope and methodology employed. The dataset analyzed may not be comprehensive enough to represent the entire spectrum of tourist perceptions and behaviors. Additionally, reliance on sentiment and toxicity analysis tools, such as VADER and TextBlob, introduces potential biases and inaccuracies inherent in automated text analysis. While compelling, these tools may not fully capture the nuances of human emotion and contextual subtleties. Consequently, these limitations suggest further research incorporating more diverse datasets and advanced analytical methods to understand the subject matter better and more holistically.

Further research is recommended to expand the scope and depth of the current study's findings. Future studies should incorporate a more extensive and diverse dataset to ensure a comprehensive understanding of tourist perceptions and behaviors. Additionally, employing advanced sentiment analysis techniques, such as deep learning and natural language processing, could enhance the accuracy and reliability of the results. These improvements would provide more nuanced insights into the complex dynamics of digital content and its impact on tourism marketing. Consequently, these recommendations aim to refine and build upon the existing research, contributing to more effective and informed strategies in the tourism industry.

4. CONCLUSION

In conclusion, this research underscores the significant role of digital content in shaping tourist perceptions and behaviors, with a particular focus on the influence of travel vlogs. Utilizing the Tourism and Travel Content Analysis (TTCA) framework, the study effectively processed and analyzed 1,972 review posts out of 2,250, providing valuable insights into viewer engagement and sentiment. The toxicity score calculations revealed the prevalence of negative interactions within the analyzed content, with scores ranging from 0.05542 to 0.86967 for Toxicity, 0.00536 to 0.50704 for Severe Toxicity, 0.01921 to 0.59834 for Identity Attack, 0.03305 to 0.76573 for Insult, 0.03737 to 0.78492 for Profanity, and 0.01075 to 0.48617 for Threat. These findings highlight the need for compelling content moderation strategies to maintain a positive digital environment. Sentiment analysis using models such as VADER and TextBlob demonstrated a positive reception of travel vlogs. Specifically, VADER classified 3.73% of posts as unfavorable, 19.83% as neutral, and 76.44% as positive out of 590 analyzed posts. TextBlob, on the other hand, classified 2.71% of posts as unfavorable, 35.59% as neutral, and 61.69% as positive for English posts, with similar trends observed for French and German posts. The analysis revealed that VADER and TextBlob agreed on the sentiment classification for 446 out of 587 posts (75.98%), with a Cohen’s kappa statistic of 0.471, indicating moderate agreement. There were seven posts where VADER assigned positive polarity scores, TextBlob assigned negative, and five posts with the opposite classification. The findings indicate that well-regulated and thoughtfully designed digital content enhances user engagement and optimizes destination marketing strategies' performance. Integrating advanced analytical tools and comprehensive data sets in future research is recommended to refine these insights further and support the development of more targeted and effective marketing efforts in the tourism sector. Consequently, this study contributes to a deeper understanding of the impact of digital media on tourism marketing, offering practical recommendations for leveraging content to foster a positive and engaging tourist experience.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 469-479
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5294

ACKNOWLEDGMENT

Thanks to the Tourism Study Program, Faculty of Business Administration and Communication, Atma Jaya Catholic University of Indonesia, LPPM, and Center of Digital Transformation and Tourism Development for the support and invaluable contributions to this research publication.

REFERENCES

[1] H. Xu, R. Law, J. Lovett, J. M. Luo, and L. Liu, "Tourist acceptance of ChatGPT in travel services: the mediating role of parasocial interaction," *J. Travel Tour. Mark.*, vol. 41, no. 7, pp. 955–972, 2024, doi: 10.1080/10548408.2024.2364336.
[2] R. Luekveerawattana, "Enhancing innovation in cultural heritage tourism: navigating external factors," *Cogent Soc. Sci.*, vol. 10, no. 1, p., 2024, doi: 10.1080/23311886.2024.2301813.
[3] E. Surucu-Balci and G. Balci, "Building social capital in cruise travel via social network sites," *Curr. Issues Tour.*, vol. 26, no. 7, pp. 1096–1111, 2023, doi: 10.1080/13683500.2022.2047904.
[4] S. Chen and J. M. Luo, "Assessing barriers to the development of convention tourism in Macau," *Cogent Soc. Sci.*, vol. 7, no. 1, 2021, doi: 10.1080/23311886.2021.1928978.
[5] P. Kumar, B. Aggarwal, V. Kumar, and H. Saini, "Sustainable tourism progress: a 10-year bibliometric analysis," *Cogent Soc. Sci.*, vol. 10, no. 1, p., 2024, doi: 10.1080/23311886.2023.2299614.
[6] K. Berhanu and S. Raj, "The role of social media marketing in Ethiopian tourism and hospitality organizations: Applying the unified theory of acceptance and use of technology model," *Cogent Soc. Sci.*, vol. 10, no. 1, p., 2024, doi: 10.1080/23311886.2024.2318866.
[7] S. Nieuwland, "Urban tourism transitions: doughnut economics applied to sustainable tourism development," *Tour. Geogr.*, vol. 26, no. 2, pp. 255–273, 2024, doi: 10.1080/14616688.2023.2290009.
[8] F. P. Leite, N. Pontes, and B. Schivinski, "Influencer marketing effectiveness: giving competence, receiving credibility," *J. Travel Tour. Mark.*, vol. 41, no. 3, pp. 307–321, 2024, doi: 10.1080/10548408.2024.2317748.
[9] R. Zengeya, P. W. Mamimine, and M. C. Mwando, "Diaspora based tourism marketing conceptual paper: A conceptual analysis of the potential of harnessing the diaspora to improve tourism traffic in Zimbabwe," *Cogent Soc. Sci.*, vol. 9, no. 1, 2023, doi: 10.1080/23311886.2023.2164994.
[10] O. A. George and C. M. Q. Ramos, "Sentiment analysis applied to tourism: exploring tourist-generated content in the case of a wellness tourism destination," *Int. J. Spa Wellness*, no. May, pp. 1–23, 2024, doi: 10.1080/24721735.2024.2352979.
[11] G. M. Ji, J. H. Cheah, M. Sigala, S. I. Ng, and W. C. Choo, "Tell me about your culture, to predict your tourism activity preferences and evaluations: cross-country evidence based on user-generated content," *Asia Pacific J. Tour. Res.*, vol. 28, no. 10, pp. 1052–1070, 2023, doi: 10.1080/10941665.2023.2283599.
[12] A. Gerou, "Monitoring emotions throughout the onboard customer journey–evidence from the travel and tourism industry," *Cogent Soc. Sci.*, vol. 10, no. 1, p., 2024, doi: 10.1080/23311886.2024.2318883.
[13] C. Shah and A. Trupp, "Trends in consumer behaviour and accommodation choice: perspectives from India," *Anatolia*, vol. 31, no. 2, pp. 244–259, 2020, doi: 10.1080/13032917.2020.1747213.
[14] W. Bahadur and A. Ali, "Investigating the effect of service quality dimensions on travellers' satisfaction with couchsurfing accommodation and subjective well-being in a sharing economy," *Econ. Res. Istraz.*, vol. 36, no. 3, pp. 1–18, 2023, doi: 10.1080/1331677X.2023.2217892.
[15] İ. A. Özen and E. Özgül Katlav, "Aspect-based sentiment analysis on online customer reviews: a case study of technology-supported hotels," *J. Hosp. Tour. Technol.*, vol. 14, no. 2, pp. 102–120, 2023, doi: 10.1108/JHTT-12-2020-0319.
[16] S. Ding, R. Zhang, Y. Liu, P. Lu, and M. Liu, "Visitor crowding at World Heritage Sites based on tourist spatial-temporal distribution: a case study of the Master-of-Nets Garden, China," *J. Herit. Tour.*, vol. 18, no. 5, pp. 632–657, 2023, doi: 10.1080/1743873X.2023.2214680.
[17] W. S. Lei and K. I. Loi, "Exploring onshore excursion preferences among cruisers: application of the travel career pattern in the Hong Kong cruise tourism context," *Tour. Recreat. Res.*, vol. 0, no. 0, pp. 1–14, 2024, doi: 10.1080/02508281.2024.2319563.
[18] X. Zhang, P. Tavitiyaman, and W. Y. Tsang, "Preferences of Technology Amenities, Satisfaction and Behavioral Intention: The Perspective of Hotel Guests in Hong Kong," *J. Qual. Assur. Hosp. Tour.*, vol. 24, no. 5, pp. 545–575, 2023, doi: 10.1080/1528008X.2022.2070817.
[19] M. Koščak *et al.*, "Exploring the neglected voices of children in sustainable tourism development: A comparative study in six European tourist destinations," *J. Sustain. Tour.*, vol. 31, no. 2, pp. 561–580, 2023, doi: 10.1080/09669582.2021.1898623.
[20] F. Bazzanella, M. Schnitzer, M. Peters, and B. F. Bichler, "The role of sports events in developing tourism destinations: a systematized review and future research agenda," *J. Sport Tour.*, vol. 27, no. 2, pp. 77–109, 2023, doi: 10.1080/14775085.2023.2186925.
[21] E. Happ, "Tourism destination competitiveness with a particular focus on sport: the current state and a glance into the future–a systematic literature analysis," *J. Sport Tour.*, vol. 25, no. 1, pp. 66–82, 2021, doi: 10.1080/14775085.2021.1888775.
[22] A. Bavik, Y. K. P. Wan, and F. Okumus, "Not Just Good, but Fair Service: The Mediating Role of Psychological Contract Breach between Service Quality and Behavioral Outcomes," *J. China Tour. Res.*, vol. 19, no. 3, pp. 654–676, 2023, doi: 10.1080/19388160.2022.2112353.
[23] H. Liu, H. Alli, and I. S. M. Yusoff, "The impact of user preference and perceived value on customer satisfaction and marketability at traditional handicraft product," *Cogent Bus. Manag.*, vol. 11, no. 1, p., 2024, doi: 10.1080/23311975.2024.2327476.
[24] E. Ketter and E. Avraham, "#StayHome today so we can #TravelTomorrow: tourism destinations' digital marketing strategies during the Covid-19 pandemic," *J. Travel Tour. Mark.*, vol. 38, no. 8, pp. 819–832, 2021, doi: 10.1080/10548408.2021.1921670.
[25] M. U. Nazir, I. Yasin, H. H. Tat, M. Khalique, and S. A. Mehmood, "The Influence of International Tourists' Destination

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 469-479
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5294

Image of Pakistan on Behavioral Intention: The Roles of Travel Experience and Media Exposure," *Int. J. Hosp. Tour. Adm.*, vol. 23, no. 6, pp. 1266–1290, 2022, doi: 10.1080/15256480.2021.1938782.
[26] S. S. Matusse, X. Xi, and I. M. Joaquim, "Assessment of strategies to enhance the online presence of the Mozambican government website on tourism destination marketing," *Int. Trade, Polit. Dev.*, vol. 7, no. 1, pp. 16–35, Jan. 2023, doi: 10.1108/ITPD-07-2022-0012.
[27] M. Gharzouli, A. K. Hamama, and Z. Khattabi, "Topic-based sentiment analysis of hotel reviews," *Curr. Issues Tour.*, vol. 0, no. 0, pp. 1–8, 2021, doi: 10.1080/13683500.2021.1940107.
[28] G. Le *et al.*, "Entotourism potential in Sabah, Malaysia: A Tourists' perspective.," *Cogent Soc. Sci.*, vol. 7, no. 1, 2021, doi: 10.1080/23311886.2021.1914950.
[29] Á. Dias, M. R. González-Rodríguez, and M. Patuleia, "Creative tourism destination competitiveness: an integrative model and agenda for future research," *Creat. Ind. J.*, vol. 16, no. 2, pp. 180–203, 2023, doi: 10.1080/17510694.2021.1980672.
[30] K. K. Eluwole, F. V. Bekun, and T. T. Lasisi, "Fresh insights into tourism-led economic growth nexus: a systematic literature network analysis approach," *Asia Pacific J. Tour. Res.*, vol. 27, no. 4, pp. 374–410, 2022, doi: 10.1080/10941665.2022.2075775.
[31] M. B. Gillette and E. Boyd, "Mining for tourists in China: a digital ethnography of user-generated content from coal mining heritage parks," *J. Herit. Tour.*, vol. 19, no. 1, pp. 1–19, 2024, doi: 10.1080/1743873X.2023.2255691.
[32] B. Vita, T. Deitiana, and W. Ruswidiono, "The online marketing of Indonesian street food in Jakarta," *Cogent Bus. Manag.*, vol. 8, no. 1, 2021, doi: 10.1080/23311975.2021.1996215.
[33] G. M. Majid, I. Tussyadiah, Y. R. Kim, and A. Pal, "Intelligent automation for sustainable tourism: a systematic review," *J. Sustain. Tour.*, vol. 31, no. 11, pp. 2421–2440, 2023, doi: 10.1080/09669582.2023.2246681.
[34] W. Wei, I. Önder, and M. Uysal, "Smart tourism destination (STD): developing and validating an impact scale using residents' overall life satisfaction," *Curr. Issues Tour.*, pp. 1–24, 2023, doi: 10.1080/13683500.2023.2296587.
[35] M. Boyden, "From effluvia to chemicals: techniques of self and somatic ethics in tropical health travel narratives," *Stud. Neophilol.*, vol. 93, no. 2, pp. 155–172, 2021, doi: 10.1080/00393274.2021.1916991.
[36] N. Basaraba, "The rise of paranormal investigations as virtual dark tourism on YouTube," *J. Herit. Tour.*, vol. 19, no. 2, pp. 287–309, 2024, doi: 10.1080/1743873X.2023.2268746.
[37] C. Ryan, "Mud, sweat and cameras–Irish trail and mountain running vlogging," *Sport Soc.*, vol. 27, no. 6, pp. 965–982, 2024, doi: 10.1080/17430437.2024.2334600.
[38] X. Guo and J. A. Pesonen, "The role of online travel reviews in evolving tourists' perceived destination image," *Scand. J. Hosp. Tour.*, vol. 22, no. 4–5, pp. 372–392, 2022, doi: 10.1080/15022250.2022.2112414.
[39] E. Mele and L. Egberts, "Exploring travel blogs on tourism and landscape heritage: representations of the Swiss Alps," *J. Herit. Tour.*, vol. 18, no. 6, pp. 785–806, 2023, doi: 10.1080/1743873X.2023.2237617.
[40] A. Agrawal, D. Sharma, M. K. Mishra, M. A. Shah, M. Alam, and V. Khandelwal, "How to cook a meme: exploring content strategies in brand and user-generated memes on Instagram," *Cogent Bus. Manag.*, vol. 11, no. 1, p., 2024, doi: 10.1080/23311975.2024.2366001.