Gender Representation in Saudi Tourism Social Media

Elisa Gugliotta

Université Grenoble Alpes (France), Laboratoire LIDILEM e Università degli Studi di Sassari, DiSSUF (egugliotta@uniss.it)

Abstract

This study examines the relationships between Instagram post contents and their goals on two pages promoting tourism in Saudi Arabia, targeting Arab and non-Arab tourists. It aims to understand gender representation and to analyse correlations between visual and textual content. Building on a previous study, which revealed different marketing approaches and highlighted an over-representation of women in English contents, this study uses advanced techniques such as topic modelling, keywords extraction and image analysis to identify key features and their statistical correlation with gender representation, offering insights into the strategic narratives behind the Instagram posts to promote Saudi tourism.

Parole chiave Gender representation, Saudi Arabia, Social Media.

DOI https://doi.org/10.58015/2036-2293/750

1. Introduction

The image of a country can influence the decisions of a wide audience, including investors, consumers, and, among others, tourists. In the realm of tourism, the public and media image of a nation plays a crucial role[1]. Advertising serves various purposes, one of which is to convince the target audience of the positive experience associated with a particular product. In constructing the public and media image of a destination, identifying the specific target audience is essential. Moreover, as noted by various scholars[2], there is often a gap between the public and media image of developing countries and the image proposed as a tourist destination. This gap arises because a country's image is influenced by various factors and information accumulated over time, whereas the image of a country as a touristic destination is deliberately crafted for promotional purposes. However, the tourist image can function as a tool to redeem a negative public image previously built, promoting not only tourism activities but also enhancing the overall reputation of the country[3]. Therefore, a tourism promotion campaign serves multiple purposes, including the crucial task of dispelling long-standing preconceptions. From these observations, it becomes clear how a tourism promotion campaign can be highly informative in identifying the key points a country focuses on when promoting its image. The flip side of these key points often includes negative stereotypes that the country feels are associated with it. Moreover, these campaigns can also shed light on how "Orientalist" visions of exoticism[4] are exploited by the same "oriental" countries in their promotional strategies. By using the term "orientalism" here, we are referring to postcolonial theories[5] according to which "orientalism" is a discourse created by the West to dominate and have authority over the East, representing the East in a stereotypical and reductive way, often to justify colonialism. Indeed, comparing tourism campaigns aimed at Arabic-speaking and English-speaking audiences can reveal significant differences.

In a previous study[6], we analysed two Instagram pages managed by the Saudi Ministry of Tourism: one targeting Arabic speakers and the other targeting English speakers. This analysis revealed significant differences: the Arabic contents emphasise urban events and attractions, whereas the English contents focus on natural landscapes

[1]Eli Avraham. Eran Ketter. *Perceptions, Stereotypes and Media Image of the Developing World,* in *U Tourism Marketing for Developing Countries*, Eli Avraham (ed.), 2016, pp. 9-23. Palgrave Macmillan, Houndmills.

[2] Among them e.g.: Charlotte M. Echtner. *The content of Third World tourism marketing: a 4A approach.* «International Journal of tourism research», 2002, 4.6: 413-434; Lingkun Meng, Yi Liu, Yuanlei Wang, Xiaojuan Li. *A big-data approach for investigating destination image gap in Sanya City: When will the online and the offline goes parted?*. «Regional Sustainability», 2021, 2.1, pp. 98-108; andSara C. Martinez, Maria D. Alvarez, *Country versus destination image in a developing country*. «Journal of Travel & Tourism Marketing», 2010, 27.7, pp. 748-764.

[3] *Ibidem.*

[4] Edward W. Said, *Orientalism,* New York, Pantheon Books, 1978.

[5] Robert J. C. Young, *Postcolonialism: An historical introduction, Malden,* Blackwell Publishing, 2001.

[6] Elisa Gugliotta. *Comparative Analysis of Textual and Visual Contents on Saudi Tourism Instagram Pages*. «Asia in the Mirror. Self-representations, Self-narratives, and Perception of the other», Peter Lang (forthcoming).

and sensory experiences. A particularly notable finding was the over-representation of women in the contents aimed at the English-speaking audience. This was interpreted as a strategic move to challenge and redefine international perceptions of the role of women in Saudi Arabia. Quantitative data on women perception in the media that is specific to Saudi Arabia is largely non-existent[7] and the representation of gender in marketing, particularly in regions with complex cultural dynamics, is a critical area of study. However social media allowed a glimpse into certain dynamics, and scholars noted that in Saudi Arabia, a country undergoing significant social and economic transformation, the portrayal of women in promotional content has been evolving[8]. In general, this evolution reflects broader shifts in societal norms and the strategic objectives of various sectors, including tourism[9]. Recent years have seen the Saudi tourism industry make concerted efforts to appeal to both domestic and international audiences. As part of Vision 2030, Saudi Arabia's ambitious plan to diversify its economy, there has been a notable push to promote the country as a vibrant and modern tourist destination[10]. This promotional strategy includes the creation of distinct marketing content for Arab and non-Arab tourists, aimed at catering to the different expectations and perceptions of these audiences[11].

Building on these insights, the current study delves deeper into the representation of gender in the contents of the two mentioned Instagram pages handled by the Saudi Ministry of Tourism. Our main goal is to investigate the correlation between the visual and textual elements of the posts and the gender representation they convey, aiming to uncover the nuanced ways in which Saudi Arabia is crafting its narrative to different audiences and the role that gender plays in this narrative. In this study, we combined both quantitative and qualitative data-processing techniques to examine data related to images depicting the Arabic world. For example, we manually enriched our corpus with detailed descriptions of the images. In fact, the corpus was already enriched with automatically-generated image captions, that required a rigorous manual validation because of the employed model difficulty with images describing the Arabic world[12]. At the same time, to facilitate and speed up our analyses, we employed advanced text analysis methods on textual data, which included both natural texts (published with the

[7] Aloufi, Alanoud. *Gender and national identity in Saudi Arabia*. 2017. PhD Thesis. San Francisco State University.
[8] Marianna Boero, Cristina Greco. *How advertising preserves cultural identities while communicating societal changes: A comparative study of the representation of women between Italy and Saudi Arabia,* « Lexia. Rivista di semiotica 39-40 Re-Thinking. Juri Lotman in the Twenty-First Century», 2022, pp. 331-357. Also: Aloufi, Alanoud. *Gender and national identity in Saudi Arabia*. 2017. PhD Thesis. San Francisco State University.
[9] Cristina Greco, *Food Heritage, Memory and Cultural Identity in Saudi Arabia: The Case of Jeddah,* in *Food for Thought. Numanities - Arts and Humanities in Progress*, vol 19, Stano, S., Bentley, A. (eds) Springer, 2021, pp. 55-74.
[10] Aayesha S. Khan, Ameera A. Alkohli, Samar Alnmer, Amal M. Albshri, Ismail M. H. Rushwan, Sagir A. Khan *Multimodal Discourse Analysis of Professional Tourism Campaign Titled 'Saudi by Saudis.'*. «Educational Administration: Theory and Practice», 2024, pp. 2911-2920.
[11] Anonymised for blind review purposes.
[12] See Gugliotta (forthcoming) for further details on the automatic captioning of images on our corpus. In this study, we also provided an error analysis to demonstrate the limitations of the employed captioning model with images from the Arab world.

posts) and automatically generated texts (image captions). In Section 2 we provide an overview of our methodology by outlining the exploited tools and techniques. In Section 3 we present the corpus used to perform our analyses and the data preparation steps. In Section 4, we describe and discuss our analyses.

2. Methodology

In order to perform our analyses on visual and textual data we provided our corpus with additional annotation layers, namely:

1. image features description;
2. Arabic text translation into English;
3. natural text topic classification;
4. natural text+image captions keyword extraction.

Regarding the first level of information, it was performed manually, by employing the image captions already generated for each image. All the other information levels have been produced automatically by employing techniques such as Topic Modelling, Keyword Extraction and Semantic Clustering based on Word Embeddings. Topic Modelling is a technique used in text analysis to identify main and common themes in a set of documents or texts. This methodology is particularly useful for analysing large amounts of textual data, making it possible to synthesise and organise information in an understandable way. In the case of this study, it was used to identify and understand the main themes within the texts. In particular, we employed a model previously described as effective when dealing with short, informal texts, such as social media posts[13]. This is BERTopic[14], which exploits BERT (Bidirectional Encoder Representations from Transformers) Language Models[15], and takes advantage of its capabilities to capture the contextual meaning of words using clustering techniques to identify themes. Analysing Instagram posts from the Saudi Ministry of Tourism with BERTopic helped us to confirm thematic distributions over the two Instagram pages. This allowed us to better understand how different tourism experiences are presented to Arabic and English-speaking audiences.

While Topic Modelling techniques are powerful, they do have some limitations. For instance, they may not perform well on small datasets[16], can generate a lot of outlier

[13] Roman Egger, Joanne Yu, *A Topic Modeling Comparison Between Lda, NMF, Top2Vec, and BERTopic to Demystify Twitter Posts,* «Frontiers in sociology», 7, 2022, <https://www.frontiersin.org/journals/sociology/articles/10.3389/fsoc.2022.886498/full> (Last access: 28/07/2024).

[14] Maarten Grootendorst, *BERTopic: Neural topic modeling with a class-based TF-IDF procedure,* «arXiv», 2022, <https://arxiv.org/abs/2203.05794> (Last access: 28/07/2024).

[15] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristna Toutanova, *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. «North American Chapter of the Association for Computational Linguistics: Human Language Technologies»,* vol 1, 2019, pp. 4171-4186, <https://aclanthology.org/N19-1423.pdf> (Last access: 28/07/2024).

[16] Dimosthenis Antypas, Asahi Ushio, Jose Camacho-Collados, Leonardo Neves, Vítor Silva, Francesco Barbieri, *Twitter Topic Classification. «International Conference on Computational Linguistics»,*2022, pp. 3386–3400.

topics[17], and often require existing knowledge for effective implementation[18]. Thus, to complement Topic Modelling, we employed Keyword Extraction techniques to identify significant words within the texts. Keyword Extraction is a method used to automatically identify relevant words from a body of text. For this study, we used KeyBERT[19], that leveraging BERT embeddings[20], captures the semantic meaning of words and then uses cosine similarity to find the words and phrases that are most similar to the document. This method is particularly useful for identifying keywords in short texts, such as social media posts, where traditional keyword extraction methods might struggle. These keywords were then used in combination with the previously identified topics to provide a more comprehensive analysis of the content.

To analyse the salient information extracted from the texts, we used semantic clustering techniques to group keywords based on their semantic similarity. Semantic clustering is a text analysis technique that groups words into meaningful clusters based on their semantic similarity. In our study, we used SpaCy[21] to transform keywords into word embeddings, on which the K-Means clustering algorithm was applied[22]. This allowed us to identify clusters of keywords representing similar concepts or themes[23]. Lastly, we conducted **statistical analyses** to explore correlations between the elements of the posts (images and texts). This enabled us to uncover patterns and insights within the data as discussed in Section 4.

3. Data

3.1. SAND description

The dataset employed for this study is called Saudi Arabia Self-Narrative Dataset (SAND)[24]. This corpus has been created to examine Saudi Arabia's self-representation in its current tourism promotion policy. As shown in Table 1, SAND comprises 249 Instagram posts collected over the first six months of 2024 from two official Saudi tourism Instagram accounts[25]: 122 posts from the Arabic-language page and 127 posts

[17] Raquel Silveira, Carlos G. O. Fernandes, João A. M. Neto, Vasco Furtado and José E. P. Filho, *Topic Modelling of Legal Documents via LEGAL-BERT*, «*Relations in the Legal Domain - RELATED*», 2021, pp. 64-72.

[18] Dimosthenis Antypas, Asahi Ushio, Jose Camacho-Collados, Leonardo Neves, Vítor Silva, Francesco Barbieri, *Twitter Topic Classification*. «*International Conference on Computational Linguistics*», 2022, pp. 3386–3400.

[19] Bilal Issa, Mohammed B. Jasser, Hoe N. Chua, Mohd Hamzah, *A comparative study on embedding models for keyword extraction using KeyBERT method*, «*International Conference on System Engineering and Technology (ICSET)*», 2023, pp. 40-45.

[20] Embeddings are numerical representations of words that capture their semantic relationships.

[21] Matthew Honnibal, Ines Montani, Sofie Van Landeghem, Adriane Boyd, ***S****paCy: Industrial-strength Natural Language Processing in Python*, «Zenodo», 2020.

[22] K-Means is an unsupervised clustering algorithm that divides the data into k distinct clusters, minimising the variance within each cluster.

[23] In our previous study, we tested a number of clustering algorithms on our data and K-Means outperformed the others.

[24] The corpus is available at: < https://github.com/eligugliotta/SAND/>.

[25] The Instagram pages with English and Arabic contents respectively are: <https://www.instagram.com/visitsaudi/> and <https://www.instagram.com/visitsaudi.ar> (Last access: 14/10/2024).

from the English-language page. SAND includes both textual and visual contents from these posts, with original post texts totalling 32,308 tokens (14,339 from the Arabic page and 17,969 from the English page). The dataset also incorporates images from the posts, automatically generated image captions using the ViT-GPT2 model[26], and manually validated image captions totalling 8,005 tokens (3,917 from the Arabic page and 4,088 from the English page). The Vision Transformer (ViT) combined with GPT-2 (ViT-GPT2)[27], was pre-trained on the COCO 2017 dataset[28], which includes a diverse range of images and their corresponding captions. The primary objectives of using ViT-GPT2 were to increase the quantity of data available for analysis, translate visual communication into text to facilitate comparative analyses, and test the model's efficacy on images from the Arab world. Additionally, SAND includes an error analysis of the automatic image captioning, categorising errors into types such as object recognition, attribution, action, relation, context, and specificity issues. The dataset is further enriched with pre-processed text of posts, English translations of the Arabic posts, combined pre-processed strings of natural and automatically generated texts, and cluster analysis results. SAND was created to facilitate comparative analysis of Saudi Arabia's tourism promotion strategies across different target audiences, incorporating both visual and textual data to enable comprehensive multimodal analysis of the country's self-representation in tourism marketing.

SAND Components	Arabic Page	English Page	Total
Posts	122	127	249
Original Images	122	127	249
Original Texts (words / tokens)	2,534 / 14,339	3,688 / 17,969	6,212 / 32,308
Captions (words / tokens)	1,033 / 3,917	1,103 / 4,088	2,136 / 8,005
Total (words / tokens)	**3,557 / 18,256**	**4,791 / 22,057**	**8,348 / 40,313**

Table 1: SAND numbers

The SAND dataset is available online[29]. To ensure protection of personal data, SAND has been anonymised in both the natural text column ("Natural Text") and the processed text column ("Processed Text"). This procedure was also carried out to prevent influencer names from being selected as keywords during the data processing.

[26] Ankur Kumar, *The Illustrated Image Captioning using transformers*, «Ankur | NLP Enthusiast», 2022, <https://ankur3107.github.io/> (Last access: 27/07/2024).
[27] Available on the Hugging Face transformers library at the following link: <https://huggingface.co/nlpconnect/vit-gpt2-image-captioning>.
[28] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollăr and Lawrence C. Zitnick, *Microsoft COCO: Common objects in context*, «Computer Vision–ECCV», vol 13, 2014, pp. 740-755.
[29] At < https://github.com/eligugliotta/SAND/>.

Specifically, all elements following the @ symbol were removed, with the exception of "UNESCO", given the potential usefulness of this token for Topic Modelling.

3.2. Sub-SAND description (after Extracting the Sub-Corpus with Human Subjects)

To facilitate a comprehensive analysis of gender representation within the SAND, we employed a systematic approach to isolate images containing human subjects. This process involved leveraging the automatically generated image captions as a filtering mechanism. Specifically, we implemented a keyword-based search strategy, targeting descriptors indicative of human presence such as "woman", "man", "girl", "boy", "child", "family", and related terms. This approach allowed us to create a sub-corpus that emphasizes human-centric imagery, thereby enabling a more nuanced examination of gender portrayal across the dataset. By utilizing the generated captions as a proxy for image content, we were able to efficiently identify relevant visual data. The resulting subset of SAND, which numbers are provided in Table 2, is the bases for our study.

SAND Components	Arabic Page	English Page	Total
Posts	71	66	137
Original Images	71	66	137
Original Texts (words / tokens)	1,520 / 8,511	2,007 / 9,721	3,527 / 18,232
Captions (words / tokens)	654 / 2,502	601 / 2,244	1,255 / 4,746
Total (words / tokens)	**2,174 / 11,013**	**2,608 / 11,965**	**4,782 / 22,978**

Table 2: SAND Gender Sub-Corpus

This sub-corpus allows for a quantitative comparison of gendered subjects between the Arabic and English Instagram pages. Figure 1 presents the distribution of female and male subjects across both pages, categorizing subjects into female (woman/girl) and male (man/boy) groups, while excluding subjects of unspecified gender for a focused gender analysis.

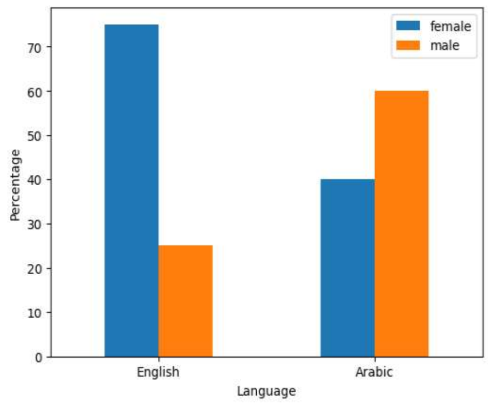

Figure 1: Gender among the Instagram Pages

The histogram reveals notable differences in gender representation between the two pages. In the Arabic-language content, male subjects appear more frequently, accounting for approximately 60% of gendered representations, while female subjects comprise about 40%. This distribution shifts dramatically in the English-language page, where female subjects are significantly more prevalent, representing 75% of gendered subjects, compared to 25% for male subjects. This marked disparity in gender representation between the two pages suggests a highly tailored approach to content creation for different target audiences. The substantial over-representation of female subjects in the English-language content appears to be a deliberate strategy to challenge international perceptions of women's roles in Saudi society. This aligns with the country's efforts to project a more progressive image to a global audience. Conversely, the Arabic-language content maintains a more traditional gender balance, with a slight male majority, possibly reflecting cultural norms or expectations within the Arab-speaking target audience. These findings reinforce the analyses performed in our previous study, highlighting the strategic differentiation in content creation for diverse audiences. In the following section, we will delve into the analyses performed on the SAND sub-corpus to better understand the underlying reasons for the observed gender distribution disparities between the English and Arabic pages.

3.3. Data Pre-processing and Annotation

As mentioned in Section 2, our methodology involves manually classification of image features to easily extract the context, subjects, and objects depicted[30]. Specifically, concerning the subjects, we focused on the activities performed in the images and for each image of the SAND sub-corpus, relying on image captions, we provided three levels of information consisting in: "image activity", "image object" and "image context". An excerpt is given in Example 1. However, since the classification

[30] While we were unable to perform inter-annotation agreement due to the absence of multiple annotators, consistency was maintained through verification steps. Given the clear distinctions in classification, we expected minimal disagreement in the results.

automatically extracted by the image captions was too granular (e.g. forty-nine classes were identified for the image-object category[31]), we decided to group the classes in order to reduce their number. Table 3 presents the image categories (image activities, objects, and contexts) and their respective final classes with the frequency in brackets.

1) Example of manual classification of image features. Between brackets we present the automatically detected image features, based on image captions, outside the brackets we present the final class assigned to the image in the example.

Image activity: (Reading) Static
Image object: (Book) Art&entertainment
Image context: (Urban) City

Category	Values
Activities	Static (27), sportive (15), gesture&emotions (15), no-action (9)
Objects	Art&entertainment (18), nature&environment (16), structures&buildings (13), miscellaneous (13), people&clothing (7).
Contexts	City (40), nature (31), culture (27).

Table 3: BERTopic results

Concerning both texts, naturally produced (post texts) and automatically generated (image captions), these were cleaned of stopwords. Example 2 shows, at its first level, the original text of an English post, and, at the second level, the same text after stopwords removal.

2) Example of text pre-processing.

1. "For the perfect family break this winter season come and explore all that Diriyah has to offer. With its cultural activities and historical buildings, it is the ideal destination from which to get to know the history of Saudi.";

[31] The classes automatically identified for image-object were: hat, palms, VR-device, cell-phone, rifle, building, traditional-clothes, tattoo, tunic, crowd-people, painting, wall, keffiyeh, room, road, village, horizon, tennis-racquet, rock, rock-formation, gazebo, costume, flower, tree, step, food, palm-tree, boat, horse, microphone, billiards, water, camera, store, sidewalk, bike, table, frame, no-object, fire, book, balls, baloon, soccer-ball, work-art, cell phone, video-game, beach. Image-activity classes were: playing, looking, descending, walking, standing, smiling, jumping, shooting, laughing, diving, sitting, holding, wearing, reading, riding, no-action, paddling. Image-context: city, nature, sport, heritage, culture, art.

2. "perfect family break winter season come explore diriyah offer cultural activity historical building ideal destination get know history saudi".

Arabic natural texts were also translated into English to perform automatic analyses on both natural texts, Arabic and English, at once. The translation was performed automatically by exploiting the Google Translate Ajax API[32]. An example of an Arabic text processing is provided below. At the first level we report the original text of the post, at the second level, the same text after stopwords removal[33], and at the third level its translation.

3) Example of Arabic text pre-processing.

1. مهرجان_قمم_الدولي للفنون الأدائية الجبلية.. فنون محلية وعالمية تعيشون فيها أجواء مميزة #
لا تُنسى

عسير | 📍

يناير 27 - 20 | 📅

2. مهرجان قمم الدولي للفنون الأدائية الجبلية فنون محلية وعالمية تعيشون أجواء مميزة تنسى عسير يناير
20 27

3. Qemam[34] International Festival for Mountain Performing Arts Local and International Arts Lived a distinctive unforgettable atmosphere Aseer[35] 20 27 January

The translation shown in Example 3, has been manually revised to correct only the translation of "Qemam", as the Google API provided a literal translation ("Summary") which was not appropriate for the context. The automatic translations from Arabic to English were not manually revised on a global scale, but only in certain cases where the text was prone to misinterpretation (as in Example 3). A full review would improve the quality of the data and represents an important step for future improvements. However, concerning Example 3, the translation result seems satisfying. We note, e.g., the fact that during the pre-processing step the Arabic negation (لا) was removed, being considered a stopword. However the negative meaning (of لا تُنسى, a passive construction that could be translated as "not to be forgotten") was retained in the translation ("unforgettable"), which is based on the pre-processed text where neither the negation (لا), nor the passive mark, consisting in the diacritic (◌ُ), are retained.

[32] <https://pypi.org/project/googletrans/> (Last access: 28/07/2024).
[33] For the Arabic stopwords removal the Python library *Arabic-Stopwords 0.4.3* was employed.
[34] The term "Qemam" in Arabic (قمم) translates to "summits" or "peaks," symbolizing the pinnacle of cultural achievements and heritage. The festival features a variety of performances and activities that showcase the cultural traditions of Saudi Arabia and other participating countries. It aims to celebrate and preserve cultural heritage.
[35] ʿAsīr is a province in Saudi Arabia.

3.4. Topic Modelling and Keywords Extraction

We applied BERTopic to the pre-processed natural texts of the English and Arabic posts. Considering the main topic of the pages, we filter out from the source texts generic words like "tourism" and "visit" to focus on more salient topics. The output included three topics. For each one we extracted the most important keyword that we report in the following table, together with the frequency of each cluster per language-context.

Language-context	Cluster	Frequencies	Keyword
Arabic	1	30	Spirit
Arabic	2	23	Saudi
Arabic	3	18	Man
English	1	34	Woman
English	2	22	Man
English	3	10	Saudi

Table 4: BERTopic results

The results from BERTopic, presented in Table 4 reveal interesting thematic differences and similarities between the Arabic and English texts. First of all, both language-contexts feature clusters focused on "man" and "saudi", but only the English-context has a cluster represented by the keyword "woman". However, the English texts have a less pronounced focus on "saudi", whereas the most frequent Arabic cluster is represented by the word "spirit". The latter suggests a deeper engagement with cultural or emotional aspects, potentially reflecting themes of heritage, pride, and identity.

Despite the encouraging results, the clustering of posts and the representation of clusters into a single theme reduced the granularity of our analysis; for this reason, we considered useful to process our texts by exploiting Keywords Extraction, with KeyBERT. For these analyses we combined natural and automatic generated texts, in order to provide as much context as possible to model and produce a complete representation of the message conveyed by the post. For instance, as shown in Example 4, KeyBERT helped us to extract keywords that highlighted important aspects such as "festival", "desert", "heritage", and "experience". The first level of the following example presents the original text of an English post; the second level shows the automatically generated caption for the image; the third level consists in the keywords beside their relevance score, computed by KeyBERT.

4) Example of Keywords Extraction.

1. "From captivating history to breath-taking entertainment, Diriyah has it all! Make sure to book your trip so you can enjoy exploring the heritage and taking in the fantastic performances over a delicious meal!";
2. "a man and a woman walking down a walkway of a desert village";

3. "diriyah: 0.51, village: 0.40, heritage: 0.40, desert: 0.35, breathtaking: 0.34".

These keywords were instrumental to generate semantic clusters based on salient features of the posts. To perform the semantic clustering, we selected the most important keywords (based on their scores) and computed their embedding, in order to apply the clustering algorithm directly on them[36]. The resulting clusters identified were four: Urban-context (19 posts), Nature (41 posts), Culture (62 posts), and Art (15 posts).

4. Analyses

The primary objective was to determine how the visual and textual representation of women and men differs between the content of the two Instagram pages. Key hypotheses include: (a) Image characteristics: Arabic and English images with female subjects will show significant differences in activities, context, and objects compared to those with male subjects. (b) Post topics: the topics of posts associated with female subjects will differ significantly from those associated with male subjects, reflecting different narratives.

4.1 Analysing the Correlations

To address these hypotheses, we performed a multinomial logistic regression with gender and languages combined into four independent categories, while the image objects, contexts, and activities were treated as dependent variables. However, as evidenced by the results reported in Table 5, our data is either insufficient in quantity or not balanced enough across classes, leading to sparsity that undermines the training of a logistic regression model capable of high accuracy. Specifically, the model for image objects achieved an overall accuracy of 40%, the model for context reached 50%, and the model for activities obtained 55%.

[36] The clustering algorithm employed was K-Means.

Category	Precision	Recall	F1-score
Image Object (Accuracy 40%)			
art&entertainment	0.50	0.83	0.62
miscellaneous	0	0	0
nature&environment	0.30	1.00	0.46
people&clothing	0	0	0
structures&buildings	0	0	0
Image Context (Accuracy: 50%)			
city	0.17	0.20	0.18
culture	0.25	0.25	0.25
nature	0.80	0.73	0.76
Image Activity (Accuracy: 55%)			
no_action	0	0	0
sportive	0	0	0
static	0.55	1.00	0.71

Table 5: Results of Multinomial Logistic Regression for Image features and Gender

Focusing on the image objects category, the classification report reveals varied performance across different classes. Unfortunately, we did not have a predefined baseline for the image classification, as our task is unique in its combination of image features and linguistic variables. Nonetheless, the performance metrics highlight the challenges posed by limited and unbalanced data, which hinder the model's ability to make accurate predictions, particularly in underrepresented categories such as the "miscellaneous" "people&clothing", and "structures&buildings" categories, all showing precision, recall, and F1-scores of 0. However, the model demonstrates a reasonable ability to identify instances within the "art&entertainment" and "nature&environment" categories, with precision scores of 0.50 and 0.30, respectively. This indicates that half of the predictions for "art&entertainment" and 30% for "nature&environment" are correct. The recall for these categories is notably higher, with "art&entertainment" at 0.83 and "nature&environment" at 1, meaning the model successfully identifies most instances in these classes. The classification report for the image contexts category also reveals a varied performance. The model demonstrates a strong ability to identify instances within the "nature" category, achieving a precision of 0.80 and a recall of 0.73, indicating that 80% of the predictions for "nature" are correct, and it successfully identifies 73% of the instances. In contrast, the performance for the "city" and "culture" categories is notably weaker, with precision and recall scores of 0.17 and 0.20 for "city" and 0.25 for both precision and recall for "culture". This suggests that the model struggles significantly with these categories. Similarly, the classification report for the image activities category highlights the model's mixed performance. The model shows high performance in identifying instances within the "static" category, achieving a precision of 0.55 and a perfect recall of 1, meaning that all instances of "static" were correctly identified. However, the model fails to correctly classify any instances of the "no_action" and "sportive" categories, with precision, recall, and F1-scores all at 0. This further underscores the challenges posed by the limited and imbalanced data in these categories. Since our data are not suitable for training a regression model, we provide statistical descriptive analyses based on frequency and performed chi-squared test (χ^2) to assess the significance of the associations observed between variables. Below we show the contingency tables for "image object", "image activity" and "image context".

Gender/Lang.	Art&entert.	Misc.	People&cloth.	Structures&build.	Nature&env.	Total
Female_ar	7	2	2	1	0	12
Female_en	2	7	0	7	11	27
Male_ar	8	2	5	3	0	18
Male_en	1	2	0	1	5	9
Total	19	13	7	12	16	66

Table 6: Image Object and Gender

Table 6 illustrates that images of women in Arabic posts (Female_ar) are predominantly associated with the categories `Art&entertainment` and `People&clothing`, with no representation of `Nature&environment`. In contrast, images of women in English posts (Female_en) exhibit a more diverse distribution, including notable occurrences in `Nature&environment` and `Structures&buildings`, which indicates a broader range of object types. Similarly, men in Arabic posts (Male_ar) are primarily represented by `Art&entertainment` and `People&clothing` objects, while men in English posts (Male_en) show a significant presence of `Nature&environment objects`. This variation highlights distinct preferences across different gender/audience-target groups. The χ^2 test results (37.23, p-value: 0.0002) confirm that these differences are statistically significant, suggesting that the observed distribution is unlikely to be due to chance.

Gender/Lang.	Gesture&emot.	No_action	Sportive	Static_position	Total
Female_ar	5	1	1	5	12
Female_en	4	2	7	14	27
Male_ar	6	5	4	3	18
Male_en	0	1	3	5	9
Total	15	9	15	27	66

Table 7: Image Activity and Gender

Table 7 presents image activity categories across different gender and language contexts. For Arabic posts featuring women, the most frequent categories are "Gesture&emotions" and "Static_position", with fewer occurrences in "No_action" and

"Sportive". In contrast, English posts featuring women show a diverse distribution, with notable counts in "Static_position" and sportive activities, indicating a wider range of activities depicted. Among men, Arabic posts predominantly feature "Gesture&emotions", followed by "No_action" and "Sportive", while English posts have fewer instances overall but include "Static_position" and sportive activities. The χ^2 test results (14.78, p-value: 0.097) indicate that the observed differences in activity categories across gender/language groups are not statistically significant. Consequently, although some variations exist in the types of activities depicted, these differences are likely due to random variation rather than a clear-cut pattern influenced by gender or language.

Gender/Lang.	City	Culture	Nature	Total
Female_ar	7	5	0	12
Female_en	3	5	19	27
Male_ar	8	7	3	18
Male_en	1	1	7	9
Total	19	18	29	66

Table 8: Image Context and Gender

Finally, Table 8 provides an overview of the distribution of image context categories across different gender and target-audience. For Arabic posts featuring women, the most common contexts are "City" and "Culture", with no occurrences of "Nature". In contrast, English posts featuring women display a prominent focus on "Nature", alongside significant counts in "City" and "Culture". Men's Arabic posts show a balanced distribution among "City", "Culture", while English posts featuring men are primarily characterised by "Nature", with minimal representation in "City" and "Culture". The χ^2 results (27.46, p-value: 0.00012) indicate that the observed distribution of image contexts across gender/language groups is highly unlikely to have occurred by chance. Therefore, there is a statistically significant association between gender/audience-target and image context categories in this dataset.

The final analyses address our second hypothesis concerning the semantic clusters derived from post texts, which categorize the thematic content of posts. Table 9 presents the contingency table depicting the distribution of semantic clusters across gender categories.

Gender/Lang.	Art	Culture	Nature	Urban	Total
Female_ar	4	5	0	3	12
Female_en	1	9	12	5	27
Male_ar	4	8	3	3	18
Male_en	3	4	2	0	9
Total	12	26	17	11	66

Table 9: Semantic Clusters and Gender (M/F)

Table 9 illustrates that the distribution of semantic clusters varies significantly among different gender and language groups. For Arabic posts featuring women, the predominant clusters are "`Culture`" and "`Urban`", with no representation in "`Nature`". Conversely, English posts featuring women exhibit a diverse range of clusters, with notable counts in "`Culture`" and "`Nature`". Arabic posts featuring men show a balanced and important distribution among "`Art`" and "`Culture`", while English posts featuring men have a small representation in "`Art`" and "`Culture`", with minimal representation in "`Nature`" and `Urban`. The χ^2 results (5.80, p-value: 0.071) suggest that while there is some evidence of a relationship between gender and semantic clusters, it is not statistically significant. This indicates that, although there are observable patterns in the distribution of semantic clusters across different genders, these patterns may not be strong enough to conclude a definitive significant association within the dataset. To improve the balance and increase the dataset's robustness, future analyses should consider including additional gender categories, including mixed category. Indeed, Table 10 provides an updated view of the previous contingency table (Table 9) including a new category for mixed gender (representing family or couples). This addition significantly impacts the analysis, offering a more nuanced view of how semantic clusters are distributed across different gender identities.

Gender/Lang.	Art	Culture	Nature	Urban	Total
Female_ar	4	5	0	3	12
Female_en	1	9	12	5	27
Male_ar	4	8	3	3	18
Male_en	3	4	2	0	9
Mixed_ar	1	22	11	7	41
Mixed_en	2	14	13	1	30
Total	15	62	41	19	137

Table 10: Semantic Clusters and Gender

From the table, it is evident that mixed categories have a substantial representation in the "Culture" cluster, with counts of 22 (Arabic posts) and 14 (English posts), respectively. This suggests that mixed gender posts, whether in Arabic or English, are strongly associated with "Culture" content. In contrast, the "Nature" cluster shows a notable presence among English posts featuring women and mixed gender categories (Arabic and English). "Urban" content is more prevalent in Arabic posts featuring women and mixed categories. The χ^2 results (31.96, p-value: 0.0065) indicate that the differences observed in the distribution of semantic clusters are statistically significant, with a strong association between gender and clusters.

4. Conclusion

In this study, we examined the relationships between the content of Instagram posts and their objectives on two pages promoting tourism in Saudi Arabia, aimed at Arab and non-Arab tourists. The aim was to understand gender representation and analyse the correlations between visual and textual content to identify key characteristics and their statistical correlation with gender representation, offering insights into the strategic narratives behind Instagram posts promoting Saudi tourism.

The results of our analysis firstly showed that the image of women on promotional tourism pages in Saudi Arabia is significantly more represented in content aimed at non-Arabic-speaking tourists. Furthermore, the content targeted at English-speaking audiences emphasises the female presence in a variety of contexts, such as natural and sporting settings, in contrast to the female figure in posts directed at an Arabic-speaking audience.

These findings can be interpreted in various ways and suggest a strategic attempt to challenge international perceptions of the role of women in the country. This communicative choice could be seen as a way for Saudi Arabia to assert its agency in the global context, seeking to control its image and counter Western representations. The tension between modernity and tradition is evident in the analysed content, where Arabic posts tend to focus on urban and cultural events, while English posts are oriented towards sensory experiences and natural landscapes. This reflects not only the different expectations of the two audiences but also a marketing strategy that seeks to present Saudi Arabia as a vibrant and modern tourist destination, in line with the economic diversification goals outlined in Vision 2030.

In analysing these results, we cannot avoid thinking that gender representation in the Instagram content of the Saudi Ministry of Tourism reflects not only the country's internal dynamics but also the complex power relations with the West. These relationships deeply influence the image Saudi Arabia wishes to project, both to the Arab world and the West. In this context, the construction of national identity emerges as a central theme. Like other Gulf states, Saudi Arabia has had to negotiate between tradition and modernity, a process reflected in its tourism campaigns. Central to Saudi Arabia's international relations with the Western world is its oil wealth, often associated with luxury, which has shaped and continues to influence the global perception of the country. In contrast, religious matters are foundational to its relations with the Arab world, with Saudi Arabia being a pilgrimage destination for Muslims.

In light of this information, our results highlight two major absent themes: luxury and religious matters, suggesting Saudi Arabia's desire to renew its public image. It is plausible, therefore, to consider that the greater representation of female figures in content aimed at non-Arabic-speaking tourists is not just a marketing issue but also a reflection of the social transformations taking place in the country. The growing visibility of women in promotional content seems indicative of an evolution in social norms and the strategic objectives of the tourism sector.

In conclusion, although Saudi Arabia was not formally colonised, its self-image can be analysed using postcolonial theory tools, since this approach allows for the consideration of the complex dynamics of power, representation, and identity that characterise its position in the contemporary world, revealing how the country is trying to assert its identity and agency in an increasingly competitive global context.