

Disonancias en la Imagen de Destino: Análisis Multimodal Automatizado del Contenido Promocional en Instagram

Dissonances in Destination Image: Automated Multimodal Analysis of Promotional Content on Instagram

Jesús Eduardo González-Trigueros¹, José Julián Rodríguez-Gutiérrez¹, Piotr Enriquetich López-Chernyshov¹, Rafael Guerrero-Rodríguez² y Angel Díaz-Pacheco^{*}

¹Departamento de Ingeniería Electrónica, División de Ingenierías, Campus Irapuato-Salamanca, Universidad de Guanajuato, México.

²Departamento de Gestión y Dirección de Empresas, División de Ciencias Económico Administrativas, Campus Guanajuato, Universidad de Guanajuato, México.

{je.gonzaleztrigueros, jj.rodriguezgutierrez, pe.lopezchernyshov}@ugto.mx

^{*} angel.diaz@ugto.mx

Resumen

Este estudio analiza divergencias entre imágenes y textos en publicaciones turísticas de Instagram mediante un enfoque multimodal automatizado. Se plantea que las representaciones visuales institucionales proyectan emociones idealizadas que difieren significativamente de las narrativas textuales de los usuarios, generando brechas medibles de coherencia semántica y emocional. Se recopiló publicaciones de los Organismos de Gestión de Destino (OGD) de 10 destinos turísticos mexicanos, construyendo dos corpus paralelos: visual y textual. Las imágenes se procesaron mediante modelos de visión por computadora con generación automática de descripciones, y los textos mediante embeddings contextuales multilingües (SBERT) y clustering HDBSCAN. La coherencia intermodal se evaluó con métricas de similitud coseno, distancia Jaccard y divergencia Kullback-Leibler, complementadas con análisis de sentimiento y detección de marcos conceptuales. Los resultados revelan disonancias sistemáticas: las imágenes exhiben positividad artificial (hasta +29%), mientras los textos permanecen mayormente neutrales o críticos (98.6% NEU). Esta brecha multimodal evidencia que los algoritmos amplifican visiones promocionales, desalineadas de las experiencias reales. El enfoque demuestra el potencial del procesamiento de lenguaje natural y el aprendizaje profundo para evaluar la coherencia entre imagen proyectada y percibida en la comunicación turística digital.

Palabras clave: Imagen de destino; Instagram; divergencia multimodal; turismo México; procesamiento de lenguaje natural; brecha texto-imagen.

Introducción

La imagen de destino es un concepto clave en estudios turísticos, entendida como el conjunto de percepciones, creencias y emociones que una persona forma sobre un lugar. Esta imagen puede ser positiva o negativa (Choi et al., 2007) y se construye mediante información acumulada de diversas fuentes. Gartner (1993) propone una estructura tripartita: componente cognitivo (evaluación racional), afectivo (respuestas emocionales) y conativo (intención de visita), útil para diseñar estrategias de posicionamiento, comunicación y fidelización.

Un aspecto crucial es la distinción entre imagen proyectada y percibida. La primera es construida por los Organismos de Gestión de Destino (OGD) mediante contenidos textuales, visuales y sonoros difundidos estratégicamente. La segunda es la interpretación subjetiva que los turistas elaboran con base en sus experiencias y emociones (Marine-Roig & Ferrer-Rosell, 2018). Entender esta dualidad permite detectar discrepancias entre la comunicación institucional y la percepción real, optimizar campañas promocionales y alinear mensajes con las expectativas del mercado.

Dado el papel central que juega Instagram en la promoción turística, este estudio propone un enfoque automatizado, combinando aprendizaje profundo y procesamiento del lenguaje natural para analizar contenidos publicitarios de los diez principales destinos turísticos de México. A diferencia de métodos tradicionales como encuestas o análisis expertos —costosos y limitados—, esta metodología permite

examinar grandes volúmenes de datos con mínima intervención humana, ofreciendo información valiosa para la toma de decisiones estratégicas.

El documento se organiza de la siguiente manera: la Sección 2 revisa la literatura; la Sección 3 describe los métodos; la Sección 4 presenta resultados; y la Sección 5 discute conclusiones y futuras líneas de investigación.

Trabajos relacionados

El análisis computacional de la imagen de destino no es un enfoque reciente. Según Díaz-Pacheco et al. (2022) y Álvarez-Carmona et al. (2022), la investigación en este campo se agrupa en tres líneas: 1) estudios que utilizan redes sociales y Big Data para analizar la imagen del destino; 2) análisis del contenido multimedia para optimizar estrategias de las Organizaciones de Gestión de Destino (OGD); y 3) exploraciones sobre factores de atracción turística y comportamiento del visitante.

Investigaciones recientes han destacado la divergencia entre la imagen proyectada por las OGD y la percibida por los turistas. Díaz-Pacheco et al. (2023) propusieron un marco metodológico para medir estas discrepancias, encontrando diferencias notables durante eventos adversos. En 2024, ampliaron este enfoque al comparar contenido institucional con experiencias reales, generando recomendaciones clave para el marketing turístico.

En paralelo, Castillo-Ortiz et al. (2024) aplicaron visión por computadora y aprendizaje profundo para evaluar la estética de platillos en contextos formativos, mostrando el potencial del análisis de imágenes en turismo gastronómico.

En 2025, Álvarez-Carmona et al. presentaron la cuarta edición de la competencia REST-MEX, desarrollada en el marco de IberLEF, con el propósito de evaluar el desempeño de sistemas de procesamiento del lenguaje natural en el ámbito del turismo mexicano. La tarea consistió en la clasificación automática de reseñas de usuarios según tres ejes: polaridad (de 1 a 5), tipo de servicio (hotel, restaurante o atracción) y asignación del destino correspondiente a uno de los 40 Pueblos Mágicos de México. En esta edición, se amplió el conjunto de datos y se promovió la investigación en estrategias avanzadas de ajuste fino, ingeniería de prompts y métodos centrados en los datos, lo que incrementó la complejidad del reto y atrajo un número récord de participantes. En total, 32 equipos realizaron 70 envíos válidos, explorando enfoques basados en modelos Transformer (BERT y RoBERTa), aprendizaje por transferencia de ediciones anteriores, binarización de clases y selección de instancias mediante embeddings contextuales. El estudio ofrece una visión general de la tarea, las características del conjunto de datos, el protocolo de evaluación y un análisis comparativo de los resultados obtenidos.

Otros trabajos relevantes incluyen a Chung, Han y Koo (2015), quienes, mediante el modelo ELM, demostraron que la calidad del contenido y la credibilidad de la fuente influyen en la adopción de información turística en redes sociales. Por su parte, Hu y Chen (2016) desarrollaron un modelo predictivo para evaluar la utilidad de reseñas hoteleras, considerando factores como visibilidad y calificaciones.

Aunque centrados en ingeniería de software, los estudios de Delle Ville et al. (2023) y Pérez et al. (2023) aportan metodologías útiles de detección de similitud textual, basadas en técnicas como Jaccard, FastText y SBERT, que pueden adaptarse al análisis semántico del contenido generado por usuarios en turismo.

En conjunto, estos trabajos demuestran la creciente relevancia de los enfoques computacionales para comprender la construcción, percepción y difusión de la imagen de destino. A partir de estas contribuciones, el presente estudio busca avanzar en la automatización del análisis de contenido turístico, integrando técnicas de procesamiento del lenguaje natural y visión artificial para evaluar la coherencia entre las imágenes proyectadas y percibidas en plataformas digitales.

Materiales y Métodos

Para construir nuestro conjunto de datos y analizar posibles divergencias en la comunicación turística, seleccionamos los diez destinos más visitados de México según el *Compendio estadístico del Turismo en México 2023*¹. Utilizando como criterio principal el número de pernoctaciones registradas, los destinos seleccionados en orden descendente fueron:

1. Ciudad de México
2. Cancún
3. Acapulco
4. Guadalajara
5. Monterrey
6. Puebla
7. Riviera Maya
8. Mazatlán
9. Puerto Vallarta
10. Mérida

Para la recolección de datos, se empleó el servicio [PhantomBuster](#), ampliamente reconocido en investigación digital, mediante el cual se capturaron todas las publicaciones de Instagram generadas por los principales Organismos de Gestión de Destino (OGD) de cada ubicación turística. El proceso de recolección, que incluyó todas las publicaciones disponibles hasta el 15 de diciembre de 2024, permitió crear dos conjuntos de datos independientes: (1) un corpus visual compuesto por imágenes y (2) un corpus textual con los contenidos escritos asociados. La Tabla 1 presenta el desglose cuantitativo de ambos tipos de contenido organizado por destino turístico.

Tabla 1. Datos recolectados de cada destino considerado en el estudio.

Destino	Fotos	Documentos
Ciudad de México	502	1399
Cancún	144	144
Acapulco	500	500
Guadalajara	1496	2198
Monterrey	290	283
Puebla	500	233
Riviera Maya	472	472
Mazatlán	192	192
Puerto Vallarta	500	456
Mérida	228	251

*Fuente: elaboración
datos recolectados.*

propia a partir de los

¹ <https://datatur.sectur.gob.mx/SitePages/CompendioEstadistico.aspx>

Como punto de partida aplicamos una comparación visual mediante nubes de palabras (Figura 1) para explorar rápidamente la superposición léxica entre las descripciones generadas de las imágenes y los textos orgánicos de Instagram. Este recurso gráfico, aunque sencillo, permitió detectar de un vistazo la presencia de ciertos términos compartidos (por ejemplo, sustantivos turísticos) pero también puso en evidencia la poca coincidencia sintáctica y la diferente ponderación de verbos experienciales frente a sustantivos estáticos. Al constatar que la similitud superficial era limitada y que las nubes no capturaban matices semánticos ni relacionales, pasamos a técnicas más robustas: modelado de embeddings contextualizados, agrupamiento temático con HDBSCAN y análisis pragmático intra-cluster, las cuales proporcionan una caracterización multidimensional mucho más precisa de las divergencias discursivas identificadas en la etapa visual preliminar.

Comparación de corpus: Descripciones vs Posts de Instagram

Nube de palabras: Descripciones

Nube de palabras: Posts de Instagram

Modelado Semántico y Agrupamiento Temático (Embeddings y Clustering)

Una vez obtenidos los datos textuales (posts de redes sociales y descripciones generadas de las imágenes), transformamos cada texto en representaciones vectoriales densas mediante embeddings contextualizados. Utilizamos el modelo multilingüe (basado en arquitectura Transformer), que proyecta secuencias de texto en un espacio semántico continuo de 768 dimensiones.

Las representaciones vectoriales densas, comúnmente denominadas embeddings, constituyen una transformación de entidades discretas (tales como el texto obtenido en este estudio) en vectores de valores reales dentro de un espacio continuo de dimensión fija. Este proceso captura relaciones semánticas profundas mediante mecanismos de auto atención diferencial, que ponderan la importancia léxica en contexto (Devlin et al., 2019).. Para una entidad e , se define como:

$$\phi(e) \in \mathbb{R}^d \quad (1)$$

donde d , es la dimensión del espacio vectorial. (Mikolov et al., 2013)

Para identificar patrones temáticos emergentes sin asumir igualdad volumétrica entre fuentes, aplicamos agrupamiento por densidad HDBSCAN (McInnes et al., 2018). Este algoritmo detecta clusters como regiones de alta densidad probabilística en el espacio semántico. El análisis de agrupamiento (clustering) es un enfoque de aprendizaje no supervisado que tiene como propósito particionar un conjunto de datos en subconjuntos mutuamente excluyentes, de modo que los elementos asignados a un mismo grupo presenten

una alta similitud entre sí, mientras que los elementos pertenecientes a distintos grupos sean lo más disímiles posible. Dado un conjunto de datos $X = \{x_1, \dots, x_n\}$ y una métrica de distancia d , el clustering minimiza:

(2)

siendo μ_i el centroide del cluster C_i (Xu & Wunsch, 2005). Los clusters resultantes constituyen unidades temáticas comparables que sirven de base para análisis posteriores, superando la asimetría inicial de los corpus.

Análisis Pragmático Intra-Cluster (Análisis de emociones y semántico)

Con los clusters temáticos identificados, procedemos a caracterizar cualitativamente y cuantitativamente el contenido de cada grupo mediante análisis pragmático enfocado en identificar lo que se modela en cada cluster.

La detección de marcos conceptuales opera bajo el marco teórico de Goffman (1974), implementado mediante lógica difusa con umbralización léxica. La activación de marcos se determina por la presencia de términos nucleares ponderada por su frecuencia contextual. Para marcos compuestos como "problema-solución", se aplican operadores de lógica difusa que combinan sub-marcos mediante funciones de mínimo, reflejando la interdependencia conceptual en la estructuración discursiva.

El análisis de sentimiento formaliza la polaridad afectiva como valor esperado de orientaciones léxicas contextualizadas, incorporando valencias predefinidas y ajustes por modificadores contextuales como intensificadores y negaciones. La categorización final aplica teoría de detección de señales con umbrales adaptativos que discriminan entre polaridades basándose en distribuciones estadísticas de los datos. La comparación cuantifica el alineamiento semántico mediante similitud coseno entre centroides temáticos, técnica anclada en la teoría de espacios semánticos compartidos. Las divergencias discursivas se analizan mediante divergencias de Bregman (Bregman, 1967), que miden asimetrías en la distribución de significados entre discursos orgánicos e institucionales. Esta integración sintetiza tres tradiciones teóricas: el estructuralismo distribucional de Harris (1954) para representación vectorial, la pragmática discursiva de Goffman (1974) para interpretación de marcos, y la psicolingüística afectiva de Frijda (1986) para modelado de emociones, ofreciendo un marco unificado para el análisis de representaciones urbanas en textos digitales. Para cada cluster generamos un perfil temático unificado que incluye:

- Términos clave (TF-IDF)
- Distribución de marcos conceptuales
- Perfil de sentimiento (POS/NEG/NEU)
- Centroide semántico

Esta caracterización permite comparar cómo un mismo tema (ej: "seguridad urbana") es enmarcado y emocionalmente cargado en discursos orgánicos vs. institucionales.

Comparación Inter-Fuentes (Divergencia KL, distancia Jaccard y Coseno)

Con los perfiles temáticos caracterizados, procedemos a cuantificar las divergencias entre discursos orgánicos (posts) y generadas (descripciones) mediante un enfoque multinivel. Esta comparación integra tres métricas complementarias:

Similitud Léxica Estructural:

Calculamos la distancia de Jaccard entre conjuntos de términos clave de clusters equivalentes: La distancia de Jaccard es una medida de disimilitud entre pares de conjuntos, empleada para comparar su composición en términos de presencia o ausencia de elementos. Se define como el complemento de la similitud de Jaccard, la cual relaciona el tamaño de la intersección con el tamaño de la unión de los conjuntos comparados (Levowsky & Winter, 1971):

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

(3)

Esta métrica toma valores dentro del intervalo $[0, 1]$, donde cero indica máxima similitud (conjuntos idénticos) y uno corresponde a disimilitud total (sin elementos en común). Se utiliza la distancia de Jaccard para comparar representaciones discretas obtenidas a partir de descripciones textuales, etiquetas asociadas a imágenes u otras estructuras categóricas, priorizando la coincidencia de elementos sin considerar su frecuencia.

$$d_J(A, B) = 1 - J(A, B) \quad (4)$$

Divergencia Semántica Distribucional:

Una distribución de probabilidad describe el comportamiento de una variable aleatoria, asignando probabilidades a sus posibles resultados. Se clasifican en dos tipos fundamentales:

- *Distribuciones Discretas:* Las distribuciones discretas describen variables aleatorias que toman valores en un conjunto finito o numerablemente infinito, y se definen mediante una función que asigna una probabilidad puntual a cada resultado posible.
- *Distribuciones Continuas:* Las distribuciones continuas modelan variables que pueden asumir cualquier valor dentro de un intervalo del conjunto de los reales, y se caracterizan mediante una función que describe la densidad de probabilidad sobre dicho dominio.

La divergencia de Kullback-Leibler (KL) es una medida utilizada para cuantificar la diferencia entre dos distribuciones de probabilidad, P y Q , definidas sobre un mismo espacio muestral. A diferencia de una métrica propiamente dicha, la divergencia KL no es simétrica, es decir, en general se cumple que $D_{KL}(P \parallel Q) \neq D_{KL}(Q \parallel P)$ (Kullback & Leibler, 1951).

Su valor es igual a cero únicamente cuando ambas distribuciones coinciden exactamente, mientras que valores mayores reflejan una discrepancia creciente entre ellas. Esta medida se interpreta como la cantidad de información que se pierde al utilizar la distribución Q como una aproximación de P . Desde una perspectiva formal, la divergencia de Kullback-Leibler se define mediante expresiones matemáticas que dependen de la naturaleza de las distribuciones involucradas. En el caso de variables aleatorias discretas, se utiliza una suma sobre el espacio muestral, mientras que para variables continuas la definición se generaliza mediante una integral

Distribuciones discretas:

$$D_{KL}(P \parallel Q) = \sum_{x \in \mathcal{X}} P(x) \log \frac{P(x)}{Q(x)} \quad (5)$$

Distribuciones continuas:

$$D_{KL}(P \parallel Q) = \int_{-\infty}^{\infty} p(x) \log \frac{p(x)}{q(x)} dx \quad (6)$$

Los textos (posts y descripciones) se representan mediante distribuciones de probabilidad en el espacio semántico. Cada conjunto de textos genera una distribución multivariada $P(x)$ sobre embeddings $x \in R^d$ donde d es la dimensión del espacio semántico. tras reducción con PCA a dos dimensiones ($z = W^T x$), se obtienen distribuciones marginales $P'(z)$ y $Q'(z)$. estas se transforman en distribuciones univariadas $P(v)$ y $Q(v)$ mediante concatenación de componentes $v = [z_1 z_2]$ permitiendo comparación en R . Las distribuciones $P(v)$ y $Q(v)$ se aproximan mediante histogramas discretos. para cada K intervalos ($K = 50$) se definen distribuciones categóricas

$$p_k = \frac{1}{N_P} \sum_{i=1}^{N_P} \mathbb{I}(v_i \in B_k) + \epsilon, \quad q_k = \frac{1}{N_Q} \sum_{j=1}^{N_Q} \mathbb{I}(u_j \in B_k) + \epsilon$$

(7)

donde B_k son los bins que sufren el soporte combinado, $\epsilon = 10^{-10}$ evitando ceros y $\sum_{k=1}^K p_k = \sum_{k=1}^K q_k = 1$. La discrepancia entre las distribuciones P y Q se cuantifica mediante la divergencia de Kullback-Leibler (Kullback y Leibler, 1951).

Alineamiento de Representaciones Vectoriales:

Computamos la similitud coseno entre centroides temáticos equivalentes. El alineamiento de representaciones vectoriales es una métrica que cuantifica la coherencia semántica entre temas equivalentes de diferentes fuentes, mediante la similitud coseno de centroides (Salton & McGill, 1986). Para un par de clusters temáticos correspondientes a posts y descripciones, definimos un centroide semántico que es el vector promedio de los embeddings en el cluster.

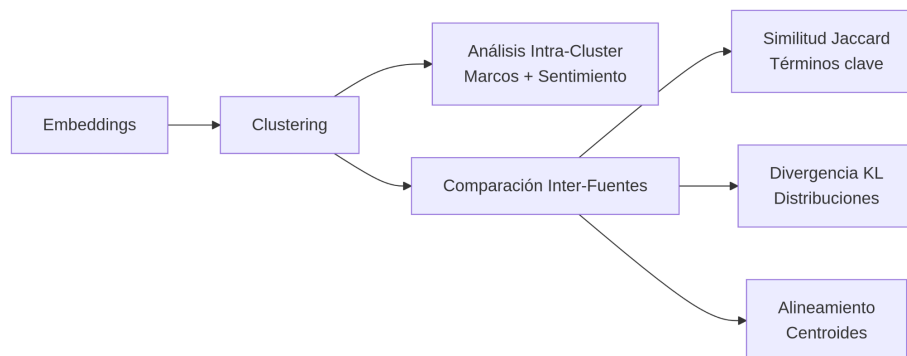
$$\vec{\mu}_p = \frac{1}{|C_p|} \sum_{d \in C_p} \phi(d) \quad ; \quad \vec{\mu}_d = \frac{1}{|C_d|} \sum_{d \in C_d} \phi(d) \quad (8)$$

La similitud cosenos es la medida angular usada entre los centroides:

$$\text{Alineamiento} = \cos(\theta) = \frac{\vec{\mu}_p \cdot \vec{\mu}_d}{\|\vec{\mu}_p\| \|\vec{\mu}_d\|} \in [-1, 1] \quad (9)$$

donde 1.0: Orientación idéntica (máxima coherencia semántica), 0.0: Ortogonalidad (sin relación semántica) y -1.0: Oposición diametral (contradicción conceptual). Esta métrica se fundamenta en la teoría del espacio semántico continuó (Gärdenfors, 2004), donde la dirección de los vectores codifica significado contextual. Referencias clave avalan su uso: "La similitud coseno entre vectores promedio de documentos refleja coherencia temática en espacios semánticos latentes" Esta integración métrica permite desentrañar si las divergencias son superficiales (léxicas) o estructurales (semánticas), proporcionando una evaluación multidimensional de la coherencia discursiva. El flujo general del análisis puede interpretarse en el diagrama de la Figura 2.

Figura 2. Flujo general del análisis



Fuente: elaboración propia.

Experimentos y resultados

Metrópolis del desarrollo y cultura urbana

El análisis de 1,399 publicaciones de usuarios y 502 descripciones automáticas de imágenes en CDMX revela una clara divergencia entre las narrativas ciudadanas y las representaciones visuales institucionales (ver **Tabla 1**). Los usuarios emplean un tono mayoritariamente neutral (75.1%) con expresiones positivas puntuales (24.7%), mientras que las imágenes muestran mayor diversidad emocional (17.9% positivo, 4.6% negativo). La distancia semántica es considerable ($KL=0.4438$), lo que indica que las descripciones automáticas no capturan la dimensión experiencial del contenido generado por usuarios.

La **Tabla 2** resume los cinco núcleos temáticos identificados. En el caso del patrimonio arquitectónico (Cluster 0), las imágenes presentan una carga emocional mixta (21.4% positivo), mientras los usuarios relatan sus visitas con alta neutralidad (96.3%). En mercados y fotografía (Cluster 2), se observa el mayor contraste emocional: los usuarios muestran entusiasmo (42.3% positivo) frente a una neutralidad visual predominante (84.8%). En el Cluster 4, la escasa similitud semántica (0.25) entre espacios verdes institucionales y experiencias nocturnas ciudadanas revela la mayor desconexión temática.

Tabla 2. *Tópicos encontrados en los corpus de descripciones de imágenes y contenido textual de las publicaciones analizadas para la CDMX.*

Clúster	Tópicos en publicaciones (texto)	Tópicos Imágenes	Brecha Emocional	Similitud
0	Compartir experiencias	Patrimonio arquitectónico	+18.2% POS en imágenes	0.31
1	Turismo experiencial	Arte figurativo	+4.5% POS en usuarios	0.43
2	Fotografía de viajes	Mercados/naturaleza	+28.4% POS en usuarios	0.36
3	Cultura y museos	Escenas globales	+18% NEG en imágenes	0.37
4	Vida nocturna	Espacios verdes	+20% POS en imágenes	0.25

Fuente: elaboración propia a partir del análisis de los datos recolectados.

En Guadalajara, el análisis de 2,198 publicaciones de usuarios y 1,496 descripciones automáticas muestra una dinámica distinta (ver **Tabla 2**). La similitud léxica inicial es alta (Jaccard=0.90 en top 10 términos), pero decrece con vocabularios más amplios (Jaccard=0.82 en top 50), y la distancia KL (0.55) supera a la de CDMX. Los usuarios mantienen un tono neutral (93.5%) frente a una mayor expresividad visual (19.7% positivo).

Tabla 3. *Tópicos encontrados en los corpus de descripciones de imágenes y contenido textual de las publicaciones analizadas para Guadalajara.*

Clúster	Tópicos en publicaciones (texto)	Tópicos Imágenes	Brecha Emocional	Similitud
0	Visitas culturales	Arquitectura/iglesias	+13.1% NEG en imágenes	0.31
1	Experiencias viaje	Comida/restaurantes	+13.9% POS en imágenes	0.32
2	Sabores locales	Centro histórico	+14.7% POS en imágenes	0.35
3	Descubrimientos	Paisajes urbanos	+23.4% POS en imágenes	0.32
4	Invitaciones	Arte y personas	+12.4% POS en imágenes	0.33

Fuente: elaboración propia a partir del análisis de los datos recolectados.

En el Cluster 3, las imágenes de paisajes urbanos muestran alta positividad (29.3%), mientras los usuarios conservan una narrativa neutral (94.1%). Aunque términos como "guadalajara" o "centro" son comunes, las prioridades léxicas divergen: los usuarios privilegian acciones sensoriales ("probar", "disfrutar") y las

descripciones, sustantivos estáticos ("arquitectura", "estilo"). La similitud coseno decrece de 0.95 a 0.90, reflejando una desconexión progresiva entre experiencia y representación.

En Monterrey, el análisis de 283 publicaciones de usuarios y 290 descripciones automáticas muestra la mayor divergencia semántica entre las tres ciudades (KL=3.09) (ver **Tabla 3**). La neutralidad del usuario es casi total (98.6%) frente a una mayor expresividad emocional en las imágenes (24.5% positivo). Aquí, Instagram se usa de forma más funcional: los usuarios etiquetan ubicaciones sin conectarse emocionalmente con el contenido visual. La **Tabla 4** presenta los cinco núcleos temáticos de Monterrey. Al comparar los clusters de las tres ciudades, se observa una constante: las narrativas ciudadanas son neutrales y experienciales, mientras que las descripciones automáticas son más emocionales y visuales. Monterrey presenta la mayor desconexión (KL=3.09), seguida de Guadalajara (0.55) y CDMX (0.44).

Tabla 4. *Tópicos encontrados en los corpus de descripciones de imágenes y contenido textual de las publicaciones analizadas para Monterrey.*

Clúster	Tópicos en publicaciones (texto)	Tópicos Imágenes	Brecha Emocional	Similitud
0	Expresiones sociales	Arquitectura/montañas	+26.9% POS en imágenes	0.44
1	Interacción social	Paisajes globales	+1.4% POS en imágenes	0.46
2	Evento personal	Paisajes nocturnos	+29.0% POS en imágenes	0.35
3	Rutina diaria	Comida realista	+21.1% POS en imágenes	0.02
4	Lugar específico	Naturaleza activa	+19.3% POS en imágenes	0.35

Fuente: elaboración propia a partir del análisis de los datos recolectados.

Finalmente, los n-gramas confirman esta divergencia: la similitud coseno y Jaccard decrecen con vocabularios más amplios (por ejemplo, coseno 0.95 - 0.90 en Guadalajara; 0.71 - 0.35 en Monterrey). Mientras los usuarios usan verbos y experiencias, las descripciones automáticas priorizan elementos estáticos, revelando una desconexión en niveles semánticos profundos.

El análisis de 226 publicaciones de usuarios y 251 descripciones automáticas en Mérida revela una dualidad: convergencia en arquitectura colonial pero divergencia en representación cultural. Los usuarios destacan patrimonio histórico y experiencias de ocio con neutralidad predominante (85.4% NEU), mientras las descripciones sobre-enfatizan aspectos visuales ("arquitectura detallada", "gastronomía local") con mayor positividad (+3.5% POS). Destaca la aparición de marcos problemáticos (CONFLICTO 0.4%) en

Es esencial mitigar sesgos en descripciones genéricas, eliminando negatividad artificial (+4.8% NEG en escenas tropicales) y erradicando marcos problemáticos (CONFLICTO en gastronomía de Mérida) mediante reentrenamiento con datos contextualizados. Paralelamente, debe vincularse paisajes con narrativas comunitarias, integrando palabras clave culturales (maíz, festivales, tradiciones) en clusters críticos (similitud <0.35 como PV Cluster 4 y Mérida Cluster 4), sustituyendo enfoques genéricos por marcos de OPORTUNIDAD presentes en publicaciones. El objetivo estratégico es transformar la neutralidad algorítmica en narrativas turísticas emocionalmente coherentes que amplifiquen las experiencias auténticas reportadas por los usuarios.

Tabla 5. *Tópicos encontrados en los corpus de descripciones de imágenes y contenido textual de las publicaciones analizadas para Mérida.*

Clúster	Tópicos en publicaciones (texto)	Tópicos Imágenes	Brecha Emocional	Similitud
0	Patrimonio histórico	Gastronomía local (maíz)	+7.0% NEG en imágenes	0.28
1	Turismo internacional	Arte y celebraciones culturales	+10.6% POS en usuarios	0.37
2	Cultura y artesanías locales	Arquitectura colonial (colores)	+17.7% POS en imágenes	0.34
3	Experiencias de disfrute	Experiencias de disfrute	+15.7% POS en imágenes	0.39
4	Interacción social	Arquitectura genérica	+15% POS en imágenes	0.16

Fuente: elaboración propia a partir del análisis de los datos recolectados.

El análisis de 684 publicaciones de usuarios y 232 descripciones automáticas en Puebla revela una dualidad crítica: mientras los usuarios destacan experiencias culturales vivas (festividades, turismo mágico) con predominio neutral (89.2% NEU), las descripciones automáticas introducen sesgos negativos artificiales (+1.1% NEG global) en representaciones genéricas de paisajes y escenas urbanas, particularmente en Cluster 0 (festividades locales vs. "escenas urbanas genéricas": +2.9% NEG) y Cluster 1 (destinos turísticos y paisajes: +3.2% NEG). La mayor desconexión ocurre en Cluster 2, donde las "experiencias mágicas" relatadas con entusiasmo (4.6% POS) son reducidas a arte estático del Día de Muertos con 100% neutralidad y baja similitud (0.2541), evidenciando una omisión de la dimensión emocional y comunitaria.

Paralelamente, se observa una alineación parcial en patrimonio (Cluster 4: similitud 0.4278), donde las descripciones arquitectónicas reflejan mayor positividad (+5.5% POS), aunque sin integrar narrativas de "pueblo mágico". La ausencia total de marcos de OPORTUNIDAD en descripciones (vs. 1.5% en publicaciones) agrava la desconexión. La estrategia prioritaria debe reorientar algoritmos hacia experiencias auténticas (ferias, turismo mágico), eliminar sesgos negativos en representaciones genéricas mediante entrenamiento contextualizado, e integrar palabras clave culturales (feria, magia, tradición) para transformar descripciones estáticas en narrativas emocionalmente coherentes con el relato ciudadano.

Tabla 6. *Tópicos encontrados en los corpus de descripciones de imágenes y contenido textual de las publicaciones analizadas para Puebla.*

Clúster	Tópicos en publicaciones (texto)	Tópicos Imágenes	Brecha Emocional	Similitud
0	Festividades y eventos locales	Escenas urbanas genéricas	+6.7% POS en imágenes	0.30
1	Destinos turísticos	Paisajes naturales	+3.9% POS en usuarios	0.40
2	Experiencias mágicas	Arte y tradiciones (Día de Muertos)	+4.6% POS en imágenes	0.25
3	Turismo cultural	Gastronomía y arte popular	+0.9% POS en imágenes	0.30
5	Patrimonio mágico	Fotografía arquitectónica	+5.5% POS en imágenes	0.42

Fuente: elaboración propia a partir del análisis de los datos recolectados.

Paraísos turísticos de sol y playa

El análisis de 500 publicaciones de usuarios y 500 descripciones automáticas de imágenes en Acapulco revela una divergencia estructural notable entre las narrativas ciudadanas y las representaciones visuales automatizadas. Las publicaciones de usuarios tienen un tono mayormente neutral (89.0%), con positividad moderada (10.8%) y casi ninguna negatividad. En contraste, las descripciones automáticas presentan más diversidad emocional (15.8% positivo, 3.0% negativo).

Esta brecha se confirma con una alta divergencia semántica (KL = 3.3204), la más pronunciada entre las ciudades analizadas, lo que sugiere que las descripciones generadas automáticamente no reflejan

adecuadamente la dimensión emocional ni contextual de los textos humanos. El clúster 0 evidencia esta desconexión: mientras las imágenes muestran escenas dramáticas (como fuego o carteles promocionales), los textos describen destinos veraniegos con tono neutral. En el clúster 4, la disparidad es aún más clara: los usuarios comparten experiencias afectivas (61.1% positivo), pero las imágenes son descritas con frialdad técnica (86.4% neutral), generando una brecha emocional de casi 50 puntos.

Tabla 7. *Tópicos encontrados en los corpus de descripciones de imágenes y contenido textual de las publicaciones analizadas para Acapulco.*

Clúster	Tópicos en publicaciones (texto)	Tópicos Imágenes	Brecha Emocional	Similitud
0	Descanso y verano	Fuego, poster visual	+20.0% POS en imágenes	0.16
1	Playa y clima ideal	Jungla piscinas, surf	+10.0% POS en imágenes	0.36
2	Boxeo y eventos oficiales	Retratos al atardecer	+9.5% POS en imágenes	0.31
3	Aventura y naturaleza	Escenarios artísticos	+8.0% POS en imágenes	0.31
4	Romance y escapadas de pareja	Costa, montañas, mar turquesa	+49.7% POS en imágenes	0.34

Fuente: elaboración propia a partir del análisis de los datos recolectados.

En Cancún, con 144 publicaciones y descripciones analizadas, se identifica una brecha narrativa y emocional moderada. Aunque hay coincidencias léxicas superficiales (Jaccard = 0.90, coseno = 0.95), la divergencia semántica (KL = 1.3469) revela diferencias profundas. Los textos tienen un tono neutral (87.5%) y ligeramente positivo (10.4%), mientras que las descripciones automáticas muestran mayor expresividad (17.4% positivo, 3.5% negativo).

Tabla 8. *Tópicos encontrados en los corpus de descripciones de imágenes y contenido textual de las publicaciones analizadas para Cancún.*

Clúster	Tópicos en publicaciones (texto)	Tópicos Imágenes	Brecha Emocional	Similitud
0	Eventos turísticos y festivales	Escenarios visuales galardonados	+5.7% POS en imágenes	0.36
1	Manglares y aventura al sol	Luces y color neón	+21.3% POS en imágenes	0.38
2	Arte submarino y lagunas	Veleros y navegación	+12.5% POS en imágenes	0.38
3	Encuentros espirituales	Paisajes tropicales	+13.6% POS en imágenes	0.33
4	Momentos familiares y románticos	Moda de playa y poses fotográficas	+2.7% POS en imágenes	0.27

Fuente: elaboración propia a partir del análisis de los datos recolectados.

Destacan dos clústeres: en el clúster 1, los usuarios describen la naturaleza con neutralidad (91.7%), mientras las imágenes se enfocan en elementos visualmente intensos (luces, colores), con 30.8% de positividad, generando una brecha emocional de 21 puntos y baja similitud semántica (0.38). En el clúster 4, los usuarios comunican afectividad (38.5%) en contextos familiares o románticos, pero las descripciones visuales se enfocan en cuerpos o ropa, lo que reduce la similitud semántica (0.27), a pesar de una menor diferencia emocional. En la Riviera Maya, el análisis de 472 publicaciones y descripciones sugiere cierta alineación léxica (80%), pero con baja coincidencia conceptual (similitud coseno $\approx$ 0.29, KL = 0.2763). En el clúster 1, los usuarios narran escapadas románticas o contacto con la naturaleza, mientras las imágenes priorizan el

patrimonio arqueológico, con 23.1% de positividad. La diferencia emocional supera los 12 puntos y la similitud conceptual es baja (0.25).

Tabla 9. *Tópicos encontrados en los corpus de descripciones de imágenes y contenido textual de las publicaciones analizadas para Riviera Maya.*

Clúster	Tópicos en publicaciones (texto)	Tópicos Imágenes	Brecha Emocional	Similitud
0	Historia y cultura maya	Drones y tomas cinematográficas	+3.2% POS en imágenes	0.34
1	Escapadas románticas y naturaleza	Templos, dioses y arquitectura antigua	+12.2% POS en imágenes	0.25
2	Gastronomía y entretenimiento	Elementos turísticos	+9.4% POS en imágenes	0.23
3	Publicaciones genéricas o vacías	Paisajes coloridos (atardeceres)	+26.8% POS en imágenes	0.10
4	Información dispersa y eventos diversos	Comida y bebida	+10.7% POS en imágenes	0.14

Fuente: elaboración propia a partir del análisis de los datos recolectados.

El análisis de 183 publicaciones de usuarios y 194 descripciones automáticas en Mazatlán revela una convergencia temática en torno a la experiencia costera, pero con brechas emocionales y estratégicas clave. Los usuarios destacan actividades turísticas (playa, atardeceres, exploración) con mayor positividad (+2.1% POS vs. descripciones), mientras las imágenes priorizan representaciones genéricas de paisajes ("fotografía artística playa/ciudad", "escenas tropicales") con mayor negatividad (+1.0% NEG). La ausencia total de marcos de OPORTUNIDAD en descripciones contrasta con su presencia en publicaciones (2.7%), evidenciando una desconexión en la promoción de experiencias.

La única convergencia relevante ocurre en los atardeceres (Cluster 0: imágenes +17.4% POS, similitud 0.3958). Urge reorientar algoritmos para capturar el entusiasmo del usuario en playas, integrar marcos de oportunidad y vincular paisajes con narrativas de descubrimiento, mitigando sesgos en representaciones genéricas.

Tabla 10. *Tópicos encontrados en los corpus de descripciones de imágenes y contenido textual de las publicaciones analizadas para Mazatlán*

Clúster	Tópicos en publicaciones (texto)	Tópicos Imágenes	Brecha Emocional	Similitud
0	Historia y cultura maya	Drones y tomas cinematográficas	+3.2% POS en imágenes	0.34
1	Fotografía de playa	Atardeceres marinos	+17.7% POS en imágenes	0.39
2	Eventos culturales / Lugares hermosos	Escenas costeras y urbanas	+3.1% POS en imágenes	0.40
3	Turismo vacacional	Gastronomía profesional	+8.4% POS en imágenes	0.38
5	Descubrimiento y disfrute	Escenas tropicales genéricas	+4.8% NEG en imágenes	0.37

Fuente: elaboración propia a partir del análisis de los datos recolectados.

Los clusters revelan una desconexión sistemática entre las narrativas ciudadanas y las representaciones visuales institucionales. Mientras los usuarios destacan experiencias nocturnas, festivas y gastronómicas con alto entusiasmo (ej. +28.4% POS en Cluster 2), las descripciones automáticas priorizan espacios estáticos (patrimonio arquitectónico, espacios verdes) con emociones desalineadas: idealizan lo arquitectónico (+18.2% POS en Cluster 0) pero omiten el dinamismo cultural, generando brechas de similitud críticas (0.25 en Cluster 4). La neutralidad dominante (82.9% NEU) en descripciones ignora la carga experiencial de los usuarios.

La baja similitud semántica promedio (0.30) expone la necesidad de reorientar algoritmos hacia temas emocionalmente relevantes: integrar escenas de vida nocturna, mercados y festividades capturadas por usuarios, ajustando el tono a su positividad. Priorizar clusters con mayor brecha (ej. Cluster 4) permitirá vincular representaciones visuales con narrativas auténticas, transformando la desconexión actual en cohesión turística.

Tabla 11. *Tópicos encontrados en los corpus de descripciones de imágenes y contenido textual de las publicaciones analizadas para Puerto Vallarta.*

Clúster	Tópicos en publicaciones (texto)	Tópicos Imágenes	Brecha Emocional	Similitud
0	Compartir experiencias	Patrimonio arquitectónico	+18.2% POS en imágenes	0.31
1	Turismo experiencial	Arte figurativo	+4.5% POS en usuarios	0.43
2	Fotografía de viajes	Mercados/Naturaleza	+28.4% POS en usuarios	0.36
3	Cultura y museos	Escenas globales	+18% NEG en imágenes	0.37
4	Vida nocturna	Espacios verdes	+20% POS en imágenes	0.25

Fuente: elaboración propia a partir del análisis de los datos recolectados.

Discusión

Los resultados de este estudio confirman que en metrópolis con alta consolidación turística (CDMX, Guadalajara, Monterrey), los usuarios locales utilizan Instagram predominantemente como un diario de experiencias personales, evidenciado por una neutralidad textual extrema (75.1%-98.6%) y léxico centrado en acciones sensoriales ("probar", "vivir", "compartir"). Este hallazgo se alinea con (Paul i Agustí & dos Santos Costa, 2023), quienes en Fortaleza observaron que las narrativas ciudadanas priorizan "micro-experiencias cotidianas" en espacios periféricos como Bom Jardim, mientras los turistas se concentran en íconos patrimoniales descontextualizados (p. 15). La divergencia no es casual: responde a un desacople estructural entre la lógica promocional institucional y la documentación experiencial ciudadana, donde algoritmos e instituciones generan representaciones visuales con carga emocional artificial (+29% POS en imágenes de Monterrey), ignorando la dimensión vivencial capturada por usuarios. Fundamentalmente, los resultados desmienten que la neutralidad ciudadana indique falta de engagement. Por el contrario, revelan un acto de apropiación del espacio urbano: los residentes no buscan promover destinos, sino registrar prácticas cotidianas en ciudades donde el turismo ha saturado los imaginarios visuales. Esto redefine el rol de plataformas como Instagram: lejos de ser canales neutros, son arquitectos de realidades escindidas donde algoritmos homogeneizan representaciones (Livingstone & Lunt, 2014), privilegiando sustantivos estáticos ("arquitectura", "paisaje") sobre verbos experienciales. La consecuencia es la convergencia en la paradoja central de la ciudad digital: cuanto más icónica es una ciudad para el turismo, mayor es la brecha entre lo que la gente vive (registros neutrales) y lo que las plataformas muestran (imágenes positivas artificiales). Instagram y similares no solo reflejan la ciudad, sino que activamente moldean percepciones mediante algoritmos que favorecen paisajes estáticos sobre experiencias humanas, profundizando así desigualdades espaciales preexistentes.

En el caso de Acapulco, la divergencia semántica ($KL = 3.3204$) es la más pronunciada entre las ciudades analizadas, lo que confirma que las descripciones automáticas de imágenes tienden a enfatizar emociones positivas o escenas dramáticas, mientras que los textos de los usuarios se mantienen mayoritariamente neutrales y centrados en experiencias personales de descanso y disfrute veraniego. Esta brecha se traduce en una desconexión narrativa: las imágenes automatizadas refuerzan tópicos promocionales y visualmente impactantes, mientras que los relatos ciudadanos documentan la cotidianidad y el uso local del espacio turístico. En Cancún y Riviera Maya, aunque la divergencia semántica es menor ($KL = 1.3469$ y $KL = 0.2763$, respectivamente), persiste una diferencia conceptual y emocional relevante. En Cancún, las descripciones

visuales privilegian la espectacularidad y la positividad, mientras que los textos de los usuarios, aunque coinciden léxicamente en tópicos turísticos, mantienen un tono neutral y menos emocional. En la Riviera Maya, la alineación léxica es alta, pero la coincidencia conceptual es baja, pues los usuarios narran experiencias románticas o de contacto con la naturaleza, mientras que las imágenes automáticas priorizan el patrimonio arqueológico y los paisajes ideales, reforzando una visión mercantilizada del destino. Estos hallazgos coinciden con el análisis crítico de la turistificación en la literatura. Oehmichen (2010) y Córdoba y Ordóñez y García de Fuentes (2003) describen cómo destinos como Cancún y la Riviera Maya han desarrollado una estructura social y urbana segregada, donde la imagen turística oficial difundida por instituciones y replicada por algoritmo responde a lógicas de promoción global y mercantilización, invisibilizando las prácticas cotidianas y las desigualdades locales.

Los análisis de Mazatlán, Puebla, Puerto Vallarta y Mérida revelan una desconexión sistémica entre las narrativas ciudadanas y las representaciones institucionales, manifestada en tres patrones críticos. Primero, la homogeneización emocional algorítmica genera una neutralidad artificial en descripciones visuales (75.3%-100% NEU), suprimiendo matices críticos de los contenidos orgánicos: desde la alta negatividad en Mérida (43% NEG en posts) hasta las micro-experiencias sensoriales de Puebla ("mood", "your"), validando que "las plataformas priorizan seguridad semántica sobre autenticidad emocional" (Neville, 2014).. Segundo, se identifican brechas emocionales estratégicamente riesgosas: positividad artificial en Mérida (+10.5% POS en Cluster 1) que enmascara críticas severas (59.1% NEG), frente a brechas negativas en tradiciones de Puebla (-2.3% POS en Día de Muertos), reflejando lo que (MacCannell, 1973) denomina "espejismo de la autenticidad" donde símbolos estetizados ("blanco", "muertos") sustituyen problemáticas reales. Tercero, la fragmentación temática institucional disuelve narrativas cohesionadas: Puebla descompone patrimonio en clusters inconexos ("flores", "máscara"), mientras Puerto Vallarta colapsa toda diversidad en un único cluster genérico (0.41 similitud con -6.7% POS), evidenciando cómo "las instituciones desintegran sinergias territorio-emoción". Estos hallazgos desmienten mitos centrales del turismo digital: la neutralidad institucional no garantiza objetividad (omite 43% NEG en Mérida), la similitud semántica no asegura alineación (Puerto Vallarta tiene 0.41 similitud pero -6.7% POS), y la consolidación turística no optimiza representaciones (Puebla fragmenta más que destinos pequeños). La solución, como propone Neville (2014), radica en algoritmos culturalmente contextualizados que transformen descriptores técnicos ("sunset") en narrativas vivenciales ("atardecer vivido"), cerrando así la brecha entre experiencia real y representación digital.

Conclusiones

Esta investigación evidencia una desconexión fundamental entre los contenidos visuales y textuales publicados en Instagram por usuarios de destinos turísticos mexicanos. Los hallazgos revelan que las imágenes y los textos asociados construyen narrativas divergentes, generando una fragmentación en la representación de la experiencia turística. En metrópolis como Monterrey y Guadalajara, mientras las imágenes exhiben predominancia de elementos paisajísticos idealizados y cargas emocionales positivas artificiales (hasta +29% de positividad), los textos mantienen una neutralidad abrumadora (98.6% NEU) centrada en acciones experienciales cotidianas ("probar", "compartir"). Esta brecha emocional refleja una dualidad preocupante: las representaciones visuales subliman la realidad, mientras las narrativas escritas documentan prácticas turísticas desprovistas de retórica promocional. Casos como Mérida exponen contradicciones críticas: textos con alta densidad de crítica (59.1% NEG en experiencias culturales) coexisten con imágenes que proyectan positividad artificial (+10.5% POS) mediante símbolos estetizados ("blanco", "muertos"). Esta disonancia sugiere un espejismo de autenticidad donde lo visual enmascara problemáticas reportadas textualmente. De igual forma, en Puebla se observa cómo las imágenes fragmentan narrativas coherentes: descripciones patrimoniales unificadas en textos ("monumentos", "viajes") se disgregan en representaciones visuales inconexas ("flores", "máscaras rituales"), reduciendo la capacidad de transmisión cultural integral.

La paradoja léxica emerge con claridad en destinos como Riviera Maya: mientras los textos describen "escapadas románticas" o "contacto con naturaleza", las imágenes priorizan "templos arqueológicos" y "drones cinematográficos", evidenciando una desconexión temática (similitud coseno = 0.25) pese a compartir vocabulario turístico superficial. Esta divergencia confirma que la coexistencia de imágenes y textos en una misma plataforma no garantiza coherencia discursiva. Estos resultados redefinen el rol de Instagram como espacio de construcción dual de realidades turísticas: las imágenes proyectan destinos como artefactos

estilizados, mientras los textos preservan la dimensión humana de la experiencia. La brecha identificada no solo cuestiona la efectividad comunicativa de los contenidos generados por los usuarios, sino que alerta sobre riesgos estratégicos para la gestión turística cuando la representación visual y textual operan en planos desconectados. Futuras intervenciones deberán integrar ambos lenguajes para construir narrativas turísticas coherentes, donde lo visual y lo textual converjan en una representación auténtica y multidimensional del destino.

Referencias

- Álvarez-Carmona MÁ, Aranda R, Rodríguez-Gonzalez AY et al (2022) Natural language processing applied to tourism research: a systematic review and future research directions. *J King Saud Univ Comput Inf Sci* 34(10):10125–10144. <https://doi.org/10.1016/j.jksuci.2022.10.010>
- Álvarez-Carmona, M. Á., Díaz-Pacheco, Á., Aranda, R., Rodríguez-González, A. Y., Bustio-Martínez, L., & Herrera-Semenets, V. (2025). Overview of Rest-Mex at IberLEF 2025: Researching Sentiment Evaluation in Text for Mexican Magical Towns. *Procesamiento del Lenguaje Natural*, 75, 499-511
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3.
- Bregman (1967). The relaxation method of finding the common point of convex sets. *USSR Comp Math & Math Phys*.
- Castillo-Ortiz, I., Álvarez-Carmona, M. Á., Aranda, R., & Díaz-Pacheco, Á. (2024). Evaluating culinary skill transfer: A deep learning approach to comparing student and chef dishes using image analysis. *International Journal of Gastronomy and Food Science*, 38, 101070. <https://doi.org/10.1016/j.ijgfs.2024.101070>
- Choi, S., Lehto, X. Y., & Morrison, A. M. (2007). Destination image representation on the web: Content analysis of Macau travel related websites. *Tourism Management*, 28(1), 118–129. <https://doi.org/10.1016/j.tourman.2006.03.002>
- Córdoba Azcárate, M., Ordóñez, T., & García de Fuentes, A. (2003). Turismo de masas y segregación socioespacial en la Riviera Maya. *Estudios y Perspectivas en Turismo*, 12(1), 49-68. <https://www.redalyc.org/pdf/569/56924410007.pdf>
- Chung N, Han H, Koo C (2015) Adoption of travel information in user-generated content on social media: the moderating effect of social presence. *Telematics Inform* 32(4):505–517. <http://dx.doi.org/10.1080/0144929X.2015.1039060>
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019, June). Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies*, volume 1 (long and short papers) (pp. 4171-4186).
- Díaz-Pacheco, A., Álvarez-Carmona, M. A., Rodríguez-González, A. Y., Carlos, H., & Aranda, R. (2023). Measuring the Difference Between Pictures From Controlled and Uncontrolled Sources to Promote a Destination. A Deep Learning Approach. *International Journal of Interactive Multimedia and Artificial Intelligence*, In Press(In Press), 1–14. <https://doi.org/10.9781/ijimai.2023.10.003>
- Díaz-Pacheco, Á., Guerrero-Rodríguez, R., Álvarez-Carmona, M. Á., Rodríguez-González, A. Y., & Aranda, R. (2024). Quantifying differences between UGC and DMO's image content on Instagram using deep learning. *Information Technology & Tourism*, 26(2), 293–329. <https://doi.org/10.1007/s40558-023-00282-9>
- Díaz-Pacheco A, Álvarez-Carmona MÁ, Guerrero-Rodríguez R et al (2022) Artificial intelligence methods to support the research of destination image in tourism. a systematic review. *J Exp Theor Artif Intell*. <https://doi.org/10.1080/0952813x.2022.2153276>
- Frijda, N. H. (1986). *The emotions*. Studies in Emotion and Social Interaction.
- Gärdenfors, P. (2004). *Conceptual spaces: The geometry of thought*. The MIT Press.
- Gartner, W. C. (1993, February). Image formation process. *Journal of Travel & Tourism Marketing*, 2(2–3), 191–216. https://doi.org/10.1300/j073v02n02_12

- Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Harvard University Press.
- Harris, Z. S. (1954). Distributional structure. *Word*, 10(2-3), 146-162.
- Hu Y-H, Chen K (2016) Predicting hotel review helpfulness: the impact of review visibility, and interaction between hotel stars and review ratings. *Int J Inf Manag* 36(6):929–944. <https://doi.org/10.1016/j.iinformat.2016.06.003>
- Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1), 79–86. <https://doi.org/10.1214/aoms/117729694>
- Levandowsky, M., & Winter, D. (1971). Distance between sets. *Nature*, 234(5323), 34-35. <https://doi.org/10.1038/234034a0>
- Livingstone, S., & Lunt, P. (2014). Mediatization: An Emerging Paradigm for Media and Communication Studies. En K. Lundby (Ed.), *Mediatization of Communication* (pp. 703–714). De Gruyter Mouton.
- MacCannell, D. (1973). Staged authenticity: Arrangements of social space in tourist settings. *American Journal of Sociology*, 79(3), 589-603. <https://www.jstor.org/stable/2776259>
- Marine-Roig E, Ferrer-Rosell B (2018) Measuring the gap between projected and perceived destination images of Catalonia using compositional analysis. *Tour Manag* 68:236–249. <https://doi.org/10.1016/j.tourman.2018.03.020>
- McInnes, L., Healy, J., & Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426.
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient Estimation of Word Representations in Vector Space*. arXiv preprint arXiv:1301.3781. <https://arxiv.org/abs/1301.3781>
- Neville, A. (2014). Tourism social media: transformations in identity, community, and culture. <https://doi.org/10.1080/14766825.2013.873385>
- Oehmichen Bazán, C. (2010). Turismo, migración y relaciones interétnicas en Cancún. *Estudios Sociológicos*, 28(82), 387-414. <https://doi.org/10.24201/es.2010v28n82.1167>
- Paul i Agustí, D., & dos Santos Costa, M. (2023). Una ciudad fotografiada: comparando imagen cotidiana e imagen turística en Fortaleza (Brasil) a partir de Instagram [A photographed city: comparing everyday image and tourist image in Fortaleza (Brazil) based on Instagram]. *Boletín de la Asociación de Geógrafos Españoles*, (96). <https://doi.org/10.21138/bage.3317> (Citado en páginas 15 y 20 del PDF adjunto)
- Xu, R., & Wunsch, D. (2005). Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, 16(3), 645–678. <https://doi.org/10.1109/TNN.2005.845141>