

IMAGEN DEL DESTINO NATAL, BRASIL

Un análisis a partir del contenido generado por los usuarios en Tripadvisor

Marlene Medeiros*

Diego Silveira Costa Nascimento**

Instituto Federal de Educación, Ciencia y

Tecnología de Rio Grande do Norte, Natal - Brasil

Lissa Valéria Fernandes Ferreira***

Aline Soares Dantas****

Universidad Federal de Rio

Grande do Norte, Natal, Brasil

Resumen: El presente artículo contribuye con un análisis de la imagen del destino Natal/RN-Brasil, a partir del contenido generado por los usuarios en TripAdvisor. El análisis y la codificación textual buscaron identificar los componentes cognitivos, afectivos y conativos presentes en los comentarios de los turistas. Se aplicaron técnicas computacionales avanzadas de Procesamiento de Lenguaje Natural y Aprendizaje de Máquinas para el tratamiento de 31.779 comentarios y la clasificación de modelos. En general, los resultados alcanzados tanto por el análisis descriptivo como por el modelo automático, indican una alta probabilidad de los aspectos cognitivos de la ciudad. Al mismo tiempo que alerta sobre la baja probabilidad de los aspectos afectivos y conativos en los comentarios, lo que refleja un debilitamiento de la recomendación del destino. Los resultados también indican el fortalecimiento del itinerario de playas, mientras señala escasa visibilidad del itinerario cultural e histórico.

PALABRAS CLAVE: Contenido generado por el usuario, imagen de destino, minería web.

Abstract: Image of Destination Natal (Brazil) from The User-Generated Content by The User on Tripadvisor. This article contributes to an analysis of the image of the destination Natal / RN-Brasil, from the user-generated content on TripAdvisor. The analysis and textual codification sought to identify the cognitive, affective and connative components present in the comments of the tourists. In general, the results achieved by both the descriptive analysis and the automatic model, point to a high probability of the cognitive aspects of the city. At the same time it is a warning for the low probability of affective and conative in the comments, which reflects a weakening of the recommendation of the destination. The results also indicate the strengthening of the itinerary of beaches, while indicating little visibility of the cultural and historical route.

KEY WORDS: User-generated content, destination image, mining web.

* Doctoranda en Administración por la Universidad Federal de Rio Grande do Norte (UFRN), Natal, Brasil. Graduación y Maestría en Administración por la UFRN. Profesora del Instituto Federal de Educación, Ciencia y Tecnología de Rio Grande do Norte (IFRN), Natal, Brasil. Dirección: Rua Brusque, 2926, Conj. Santa Catarina, Potengi, Natal/RN - CEP: 59112-490 Brasil. E-mail: lenedeiros76@gmail.com

** Doctor en Ciencias de la Computación por la Universidad Federal de Rio Grande do Norte (UFRN), Natal, Brasil. Maestría en Informática Aplicada por la Universidad de Fortaleza (UNIFOR), Fortaleza, Brasil. Bachiller en Informática por el Centro de Estudios Superiores de Maceió (CESMAC), Maceió, Brasil. Profesor del Instituto Federal de Educación, Ciencia y Tecnología de Rio Grande do Norte (IFRN), Natal, Brasil. Dirección: Rua Brusque, 2926, Conj. Santa Catarina, Potengi, Natal/RN - CEP: 59112-490 E-mail: diego.nascimento@ifrn.edu.br

*** Doctora en Administración por la Universidad Federal de Rio Grande do Norte (UFRN), Natal, Brasil. Profesora adjunta en la UFRN. Miembro permanente del Programa de Posgrado en Turismo (PPGTUR) de la UFRN. Investigadora del Grupo de Estudios Gestión del Turismo (GESTUR) de la UFRN. Dirección: Avenida Governador Sílvio Pedroza, 176 – Apto. 1101, Areia Preta Natal/RN -59014-100 Brasil. E-mail: lissafferreira.iadb@yahoo.es

**** Doctoranda en Administración por la Universidad Federal de Rio Grande do Norte (UFRN), Natal, Brasil. Graduación y Maestría en Ingeniería de Producción por la UFRN. E-mail: eng.alinedantas@gmail.com

INTRODUCCIÓN

El punto de partida de un itinerario de viajes es la búsqueda de información acerca del destino. Este proceso, antes de la propagación de la internet interactiva (*WEB 2.0*), se limitaba a la información ofrecida por los gestores turísticos: agencias de viaje, red hotelera, gobierno, etc. (Hsu, Lin, & Ho, 2012; Li *et al.*, 2015)

Actualmente, las organizaciones de marketing ya no son las únicas fuentes a las que recurren los turistas en busca de información sobre el destino. Los medios sociales de libre acceso han sido el canal más procurado por los turistas potenciales (Johnson *et al.*, 2012).

Entre la información disponible en los medios, los comentarios de los turistas que ya pasaron sus vacaciones en el lugar tienen un papel fundamental en la planificación del viaje y en la formación de la imagen del destino para quienes están planificando su viaje (Filieri, 2015; Gretzel, 2011; Pantano, 2013).

Estos comentarios online son conocidos como contenido generado por el usuario (CGU) y reúnen comentarios, evaluaciones, recomendaciones, información e interacciones producidas por los usuarios (Baka, 2016; Wong & Qi, 2017)

Los CGUs preocupan a los gestores turísticos dada la influencia que éstos tienen en la toma de decisiones de turistas potenciales, pudiendo estimular o no su motivación para la elección de un destino turístico (Ayeh, Au, & Law, 2013a; 2013b; Pantano *et al.*, 2017; Volo, 2010).

Además de ser un canal de información sobre el destino, el CGU coloca al turista que vivió la experiencia como actor principal en el proceso de reputación del producto turístico, compitiendo directamente con la imagen del destino transmitida por las instituciones formales (Akehurst, 2009; Baka, 2016). Asimismo, los relatos de las experiencias de los turistas, ricos en detalles y libres de intereses, han ganado fuerza y la confianza de los turistas potenciales (Baka, 2016; Dickinger, 2010; Munar, 2011).

El CGU del turismo se presenta como una herramienta de análisis capaz de proveer información estratégica a los stakeholders acerca del destino, sus puntos fuertes y sus fragilidades (Korfiatis, García-Bariocanal & Sánchez-Alonso, 2012; Nguyen & Cao, 2015). El gerenciamiento de esta información por parte de los gestores turísticos es necesario para entender al destino desde el punto de vista de su atractivo, de modo de potenciar su imagen y desarrollar acciones específicas para cada segmento de mercado (Carson, 2008; Pantano *et al.*, 2017; Wong & Qi, 2017).

A pesar de la cantidad de información disponible vía CGU, el número de investigaciones que tratan de sus efectos en la imagen del destino aún es limitado (Ghazali & Cai, 2014; Tamajón & Valiente,

2015). Este hecho puede ser explicado, en parte, en función de la necesidad del uso de técnicas y sistemas de captura y codificación que posibiliten el tratamiento de datos en masa.

Ante la importancia de la temática, el presente estudio tiene por objetivo analizar la imagen del destino Natal/RN-Brasil, bajo la óptica del turista que lo visita, recurriendo a técnicas computacionales avanzadas de procesamiento del lenguaje natural y el aprendizaje de máquinas. Las técnicas fueron utilizadas para captar y tratar 31.779 evaluaciones de usuarios disponibles en *TripAdvisor* en la sección “qué hacer”, posteadas por los turistas en el período de 01/01/2016 hasta 26/06/2017.

Considerando los ocho agentes apuntados por Gartner (1994) que influyen en la formación de la imagen del destino, este estudio analiza la imagen del destino Natal/RN desde la óptica del agente orgánico, o sea el turista que vivió la experiencia en el destino. El análisis y la codificación textual hacen uso de técnicas auxiliares de la minería web (mining web), buscando identificar los componentes cognitivos, afectivos y conativos presentes en los comentarios de estos turistas.

La relevancia del estudio se da por su contribución teórica, visto que en las bases de datos *Scopus*, *Web of Science* y *Science Direct*, no se encontró ningún artículo sobre la imagen del destino Natal/RN utilizando técnicas de minería de datos. Además, el estudio tiene relevancia práctica ya que sus resultados sirven de orientación para los gestores turísticos en cuanto a los aspectos y atributos del destino que son más relevantes y a la forma en que se difunde la imagen, pudiendo a partir de estos resultados desarrollar estrategias adecuadas para cada segmento. A continuación se expone la fundamentación teórica del estudio, la metodología, los resultados, el análisis de los mismos y finalmente las consideraciones finales.

CONTENIDO GENERADO POR EL USUARIO

Los datos disponibles online pueden ayudar a predecir, categorizar y seleccionar los diferentes elementos (productos, lugares, servicios, clientes, etc.), constituyendo una base de información para subsidiar los procesos decisorios de los usuarios y empresas (Berezina *et al.*, 2015).

El CGU disponible en las plataformas abiertas permite la recolección de una cantidad de información, sea actual o de series temporales, y en una velocidad que no sería posible con técnicas tradicionales como *survey*, generando así una reducción de costos en las investigaciones de mercado, opinión, académicas, etc. (Soler & Germán, 2016).

La disponibilidad del CGU, aliada a la facilidad de acceso a través de diversos medios, la ascendente tendencia a que los usuarios dispongan sus opiniones y la alta visibilidad, ha impactado directamente en las organizaciones públicas y privadas que precisan adaptarse a esta nueva realidad (Baka, 2016; Marine-Roig & Anton Clavé, 2015; 2016; Pantano *et al.*, 2017).

En el contexto del turismo el impacto del CGU fue aún mayor a través de las evaluaciones y comentarios online en torno a los productos turísticos (Bing *et al.*, 2011; Melián-González, Bulchand-Gidumal & López-Valcárcel, 2013). Para los turistas las evaluaciones funcionan como principal fuente de información y feedback. Para la industria del turismo sirven como termómetro de los servicios, productos y destinos ofrecidos (Pantano, 2013; Pantano *et al.*, 2017).

El turista potencial utiliza la experiencia de otros turistas relatada en los comentarios online y en base a ellos saca sus conclusiones sobre el destino. El CGU funciona como un reductor de incertidumbre en cuanto a la elección del destino y lo que se puede esperar del mismo (Ayeh, Au & Law, 2016; Xiang, Magnini & Fesenmaier, 2015).

Si bien por un lado la industria del turismo pone a disposición en tiempo real y en forma gratuita información importante que refleja cómo está siendo percibido y difundido su producto, este factor es preocupante pues parte de estas experiencias compartidas no son positivas y pueden desgastar la imagen y la reputación del producto (Baka, 2016; Berezina *et al.*, 2015; Pantano *et al.*, 2017). De este modo, el CGU ejerce una fuerte influencia en la formación y transformación de la imagen del destino (Ring, Tkaczynski, & Dolnicar, 2016; Tamajón & Valiente, 2015; Volo, 2010).

Entre los diversos sitios y plataformas en el contexto del turismo que disponen CGUs, *TripAdvisor* resulta el sitio de mayor impacto dado el volumen de información generada y la gama de destinos, productos y servicios turísticos disponibles y su visibilidad (Baka, 2016).

Han crecido los estudios que utilizan los comentarios de *TripAdvisor* y dicho sitio ha sido utilizado como herramienta de recolección de CGUs. Cenni & Goethals (2017) analizaron 300 evaluaciones negativas en *TripAdvisor* respecto de la red hotelera buscando identificar características semejantes o divergentes entre usuarios en idioma inglés, holandés e italiano. Pantano *et al.*, (2017) desarrollaron un modelo para identificar la previsión de las preferencias futuras de los turistas en base a las evaluaciones de los usuarios de *TripAdvisor*. Soler & Germán (2016) desarrollaron un modelo de marca para hoteles españoles con datos de *TripAdvisor*. Lu & Stepchenkova (2012) evaluaron las experiencias de ecoturismo relatadas en línea por viajeros de Estados Unidos en Costa Rica a partir de 373 evaluaciones de usuarios de la plataforma. Berezina *et al.* (2015) compararon las evaluaciones de los usuarios satisfechos e insatisfechos entre 2.510 huéspedes de hoteles de Sarasota, Florida. Mientras que Baka (2016) exploró la influencia de las evaluaciones de *TripAdvisor* en la reputación de las empresas del sector de viajes.

LA INFLUENCIA DEL CGU EN LA IMAGEN DEL DESTINO

La imagen del destino es un conjunto de impresiones, atributos, características, ideas y expectativas acerca de un lugar (Crompton, 1979; Datzira-Masip & Poluzzi, 2014; Kim & Richardson, 2003). Este conjunto de percepciones genera una identidad del lugar que pasa a ser su referencia,

afectando las decisiones de los turistas en cuanto a la elección del destino y en el destino (Çakmak & Isaac, 2012).

A pesar de que no hay consenso sobre la formación de la imagen de un destino (Gallarza, Saura, & García, 2002), la obra de Gunn (1972) es una de las principales referencias (Ghazali & Cai, 2014; Tamajón & Valiente, 2015). Gunn (1972) afirmó que la imagen del destino puede generarse por la inducción de empresas, organizaciones de gerenciamiento del destino e instituciones, o de forma orgánica como fruto de la experiencia comentada por los turistas.

Expandiendo la visión de Gunn (1972), Gartner (1994) indicó que la imagen del destino se compone por tres elementos interrelacionados: cognitivo, afectivo y conativo. El componente cognitivo se refiere a los atributos mensurables del destino. El componente afectivo se refiere a los sentimientos y sensaciones de los turistas en relación al destino. El componente conativo se refiere a las recomendaciones hechas por los turistas en función de la vivencia en los dos componentes anteriores (Gartner, 1994; Ghazali & Cai, 2014). Gartner (1994) enumeró ocho agentes que influyen en la imagen del destino mediante los componentes cognitivos, afectivos y conativos.

Cuadro 1: Síntesis de los elementos de la imagen del destino

Agente de Influencia	Forma	Posición	Agente	Elementos de Influencia	Imagen Formada del Destino
Overt Induced I	Inducida Controlable	Oferta	Agencias de Propaganda e Publicidad	Medios visuales, con evocaciones emocionales	Imágenes positivas
Overt Induced II	Inducida Controlable	Oferta	Agentes de Viaje	Interactiva con el cliente	Atractivas, favorables, irrealistas.
Covert Induced I	Inducida Controlable	Oferta	Personas Influyentes, Celebrities	Programas de TV, Shows.	Atractiva
Covert Induced II	Inducida Controlable	Oferta	Personas Influyentes, Celebrities	Interactiva con el público, por medio de relatos de experiencias	Atractiva
Autónomo	Orgánica Semi-controlable	Independiente	Informes, films	Documentos, films, artículos de noticias	Independiente
Orgánico Solicitado	Orgánica no controlable	Demanda	Información de fuentes conocedoras	Conversaciones	Independiente
Orgánico No Solicitado	Orgánica no controlable	Independiente	Información no solicitada de fuentes desconocidas	Conversaciones	Independiente
Orgánico	Orgánica no controlable	Demanda	Basado en la experiencia vivida	Conversaciones, testimonios, evaluaciones	Independiente

Fuente: Basado en Gunn (1972), Gartner (1994) y Tasci & Gartner (2007)

Tasci & Gartner (2007) a partir de lo propuesto por Gartner (1994), desarrollaron un modelo agregando una clasificación en cuanto a la posición de quien forma la imagen: del lado de la oferta, referente a los empresarios de turismo; del lado de la demanda, quienes serían los consumidores; y la parte independiente en relación a la imagen generada por actores que no están directamente involucrados con la actividad turística. El Cuadro 1 sintetiza los principales puntos discutidos en los modelos de Gunn (1972), Gartner (1994) y Tasci & Gartner (2007).

En relación al nivel de control, los actores orgánicos generan información más detalladas. Sus comentarios que forman la imagen del destino son incontrolables y han ganado credibilidad y aceptación por mostrar una percepción independiente y libre de intereses (Baka, 2016; Gartner, 1994; Ghazali & Cai, 2014; Munar, 2011; Tasci & Gartner, 2007). Por este motivo tienen una fuerte influencia en las decisiones de los turistas potenciales que confían en estos testimonios (Ayeh *et al.*, 2013a; Marine-Roig & Anton Clavé, 2016; Volo, 2010).

En la internet interactiva los contenidos generados por usuarios (CGU) online representan el mayor canal de formación de las imágenes orgánicas de los destinos y el mayor fragmento de los marcadores privados no comerciales de un destino (Wong & Qi, 2017).

Mientras los segmentos formales de la industria del turismo invierten mucho dinero y tiempo intentando construir una imagen del destino (Crompton, 1979; Ghazali & Cai, 2014), el CGU de las evaluaciones online es reconocido como uno de los elementos más ágiles en la formación y transformación de la imagen de un destino (Wong & Qi, 2017).

METODOLOGIA

Para este estudio se utilizó como fuente de recolección de datos los comentarios posteados por los turistas entre el 01/01/2016 y el 26/06/2017 en la sección “qué hacer” del portal *TripAdvisor*. El destino elegido para investigar fue la ciudad de Natal, localizada en la Región Nordeste de Brasil, en el Estado de Rio Grande do Norte (RN). Natal es una ciudad del litoral con fuerte presencia de bellezas naturales. Esta localización fue seleccionada por su expresividad en el escenario del turismo brasileño. En un estudio realizado por el Ministerio de Turismo de Brasil (BRASIL, 2017), la región Nordeste de Brasil fue reconocida como la de mayor potencial turístico y el premio “*Travelers’ Choice™ Destinos de Brasil*” de *TripAdvisor* reconoció a Natal/RN como la séptima ciudad más visitada por los turistas brasileños. Natal también fue una de las capitales brasileñas elegidas para ser sede de la Copa del Mundo 2014, lo que proyectó aún más el potencial turístico de la ciudad. Colaborando con esto, la actividad turística ejerce una fuerte influencia en la economía de la ciudad (Chagas, Sampaio, & Santos, 2013).

El portal elegido para este estudio fue *TripAdvisor* debido al amplio acceso a los comentarios y a su reconocimiento como archivo de información turística y confiable (Dickinger, 2010). La sección “qué

hacer” fue la elegida porque contiene varias subcarpetas sobre los atractivos de la ciudad que destacan los turistas. La definición del recorte temporal se dio para destacar la imagen más actual del destino.

La sección “qué hacer” estaba compuesta por 266 subcarpetas de atracciones de la ciudad. El primer análisis implicó la clasificación de estas carpetas, eliminando las que no tenían comentarios. Luego se verificó la existencia de más de una carpeta remetida a un mismo lugar (por ejemplo, existían dos carpetas relacionadas con la Praia de Pitangui).

Las carpetas de contenidos similares fueron condensadas en una sola, incorporando sus comentarios. Así, quedaron 72 carpetas. Éstas fueron clasificadas por tipo de atractivo, formando 4 clases de atractivos por afinidad: “playas y lagunas”, “edificaciones y monumentos históricos”, “plazas y parques” y “compras”. En la minería se utilizaron 31.779 comentarios. Fueron recolectados mediante la técnica conocida como *Web Scraping* (WS), que es utilizada en trabajos que involucran minería web. Para extraer los datos brutos se desarrolló un programa en el lenguaje de programación *Python* utilizando una biblioteca específica de WS llamada *BeautifulSoup*, que lee todos los comentarios directamente del sitio de *TripAdvisor*, posibilitando guardarlos en archivos con formato texto.

Una vez generados estos archivos se seleccionaron comentarios o parte de ellos aleatoriamente y sin reposición, rotulados con la clase que mejor representaba los aspectos del destino (cognitivo, afectivo o conativo), para la construcción de la Base de Entrenamiento (BTR). La BTR está formada por líneas y columnas. Las columnas están formadas por un vector donde los valores de las posiciones son representados por cada una de las palabras clave que aparecen en la BTR, siendo la última posición reservada para la clase. Cada línea es una representación binaria de las columnas, siendo el valor representado por “1” cuando la palabra clave está presente en el comentario o “0” en el caso contrario. La última posición del vector recibe la descripción de la clase que representa el comentario. La BTR posee un total de 422 modelos no balanceados distribuidos entre las clases: cognitivo (58,06%), afectivo (18,72%) y conativo (23,22%). La distribución sigue la proporción de las palabras clave que surgen en los comentarios.

Tanto la BTR como la base de tests (BTE), representados por los comentarios de los cuatro grupos de afinidad, pasaron por un pre-procesamiento de texto realizado en tres etapas utilizando la biblioteca de procesamiento del lenguaje natural NLTK, también desarrollada en Python (Guo & Morris, 2017): (i) *tokenización* (ii) remoción de la palabra vacía y (iii) *stemización*. En la etapa (i) cada parte del texto es desmembrada, tanto las palabras como los signos de puntuación, formando una lista de *tokens*. En la etapa (ii) son removidos de la lista generada anteriormente los *tokens* que no son significativos para la clasificación futura del texto (artículos, preposiciones, numerales o símbolos). En la etapa (iii) se aplica la técnica de stemización que busca reducir todas las palabras a su radical, evitando las variaciones lingüísticas.

Para la clasificación de los 31.779 comentarios no rotulados en relación a los aspectos cognitivo, afectivo y conativo, se utilizó el algoritmo del aprendizaje de máquinas o automático, basado en la estadística bayesiana *Naïve Bayes* (NB) (Guo & Morris, 2017; John & Langley, 1995). El algoritmo NB fue entrenado en base al entrenamiento descripto anteriormente. El NB clasificó cada comentario de acuerdo con la distribución de probabilidad para cada clase. La próxima sección describe los resultados del análisis.

ANÁLISIS DE LOS RESULTADOS

En este estudio la imagen que se analiza es una imagen orgánica de destino (Gartner, 1994; Gunn, 1972), visto que los comentarios parten de turistas que visitaron Natal/RN. La Tabla 1 muestra las categorías generadas y los atractivos asociados a éstas. Se presentan la cantidad de comentarios para cada atractivo, el subtotal de éstos por categoría y el porcentaje individual y por categoría en relación al total general de comentarios.

Tabla 1: Categorías y atractivos del destino Natal-RN (Brasil) comentados en TripAdvisor

LUGAR	CATEGORIA	QT. COMENTARIOS	%
SHOPPING MIDWAY MALL	COMPRAS	461	1,45%
CENTRO DE ARTESANADO DE PONTA NEGRA		280	0,88%
MANOS DE ARTE SHOPPING DE ARTESANÍAS		20	0,06%
MERCADO DE PONTA NEGRA		13	0,04%
CANTO DO MANGUE		6	0,02%
SUBTOTAL CATEGORIA "COMPRAS"		780	2%
FUERTE DOS REIS MAGOS	EDIFICACIONES Y MONUMENTOS HISTÓRICOS	1.511	4,75%
PUENTE NEWTON NAVARRO		1.190	3,74%
CENTRO DE LANZAMIENTO BARREIRA DO INFERNO		760	2,39%
ARENA DAS DUNNAS		505	1,59%
CATEDRAL METROPOLITANA		109	0,34%
CASA/MUSEO/ MEMORIAL CÂMARA CASCUDO		80	0,25%
IGLESIA MATRIZ		52	0,16%
IGLESIA DE ST.ANTÔNIO Y MUSEO DE ARTE SACRO		51	0,16%
FAROL DE MÃE LUIZA		36	0,11%
PALÁCIO POTENGI (PINACOTECA POTIGUAR / ESTADO)		29	0,09%
SOLAR BELA VISTA		18	0,06%
TEATRO ALBERTO MARANHÃO		16	0,05%
CENTRO DE CONVENCIONÕES		12	0,04%
MUSEO DE MINÉRIOS DO RN		9	0,03%
MUSEO DO BRINQUEDO POPULAR		8	0,03%
PIEDRA DO ROSÁRIO		8	0,03%
CENTRO CULTURAL CASA DA RIBEIRA		7	0,02%
IGLESIA NOSSA SEN. DOS NAVEGANTES		5	0,02%
COLUMNA CAPITOLINA		4	0,01%
IGLESIA DO ROSÁRIO N.SRA. DOS PRETOS		4	0,01%
MUSEO CAFÉ FILHO		4	0,01%
MEMORIAL DE NATAL		3	0,01%
BAOBÁ DO POETA		2	0,01%
INSTITUTO HISTÓRICO Y GEOGRÁFICO		2	0,01%
MEMORIAL DO IFRN		2	0,01%
MUSEO DE CIENCIAS MORFOLÓGICAS		2	0,01%
MUSEO ARY PARREIRAS		1	0,00%
MUSEO DE LA CULTURA POPULAR		1	0,00%
MUSEO DEL PUERTO DE NATAL		1	0,00%
MUSEO DR. SÓLON DE MIRANDA GALVÃO		1	0,00%
MUSEO PÂMPANO		1	0,00%
MEMORIAL ALUÍZIO ALVES		1	0,00%

SUBTOTAL CATEGORIA "EDIFICAC. Y MONUM. HISTÓR."		4.435	14%
PARQUE DAS DUNNAS	PLAZAS Y PARQUES	1.494	4,70%
PARQUE DE LA CIUDAD DOM NIVALDO MONTE		66	0,21%
DEL RELÓJ DO SOL		17	0,05%
PEDRO VELHO		9	0,03%
AUGUSTO SEVERO		6	0,02%
ÁRBOL DE MIRASSOL		6	0,02%
PLAZA AUGUSTO LEITE		3	0,01%
DE LAS MÃDRES MARTHA SALEM		1	0,00%
PLAZA 7 DE SETEMBRO		1	0,00%
SUBTOTAL CATEGORIA "PLAZAS Y PARQUES"		1.603	5%
GENIPABU / DUNNAS DE GENIPABU	PLAYAS Y LAGUNAS	7.811	24,58%
PONTA NEGRA/ PRAIA DO MORRO DO CARECA		7.592	23,89%
PARRACHOS DE MARACAJÁ		2.212	6,96%
LAGOA DE PITANGUI		1.316	4,14%
CAMURUPIM		1.081	3,40%
DE LOS ARTISTAS / DEL FUERTE / DEL MEDIO		817	2,57%
BARRA DE PUNAÚ		682	2,15%
LAGOA DE JACUMÃ		632	1,99%
BARRA DE CUNHAÚ		566	1,78%
PIRANGI DO SUL / NORTE		559	1,76%
BARRA DE TABATINGA		358	1,13%
PITANGUI		302	0,95%
BÚZIOS		264	0,83%
REDINHA		246	0,77%
JACUMÃ		160	0,50%
MURIÚ		84	0,26%
SANTA RITA		47	0,15%
PIUM		42	0,13%
PORTO MIRIM		41	0,13%
BARRA DO RIO		38	0,12%
GRAÇANDU		38	0,12%
AREIA PRETA / MIAMI		29	0,09%
PASSEIOS DE BUGGY		20	0,06%
PASSEIO DE DROMEDÁRIO		13	0,04%
MIRANTE DA PONTA DO MOCEGO		9	0,03%
BARREIRA D'ÁGUA		2	0,01%
SUBTOTAL CATEGORIA "PLAYAS Y LAGUNAS"		24.961	79%
TOTAL GENERAL DE COMENTARIOS		31.779	100%

Fuente: Datos de la investigación (2017)

Al analizar los comentarios se percibe que la imagen del destino Natal-RN/Brasil tiene una fuerte asociación con sus playas. De los 31.779 comentarios disponibles en *TripAdvisor*, hasta la fecha de la recolección de los datos, el 79% se refería a las playas y/o lagunas de la capital potiguar, haciendo referencia a la identidad de la ciudad (Çakmak & Isaac, 2012). Esta información demuestra que las playas son el principal atractivo del destino Natal, es lo que más destacan los turistas que pasan por la ciudad y lo que recibe mayor número de comentarios.

Estos comentarios generan mayor influencia en la elección de futuros turistas que pretenden visitar la ciudad y buscan en los CGUs lo más recomendado (Ayeh *et al.*, 2013b; Marine-Roig & Anton Clavé, 2016; Volo, 2010). Esto coincide con la imagen inducida del destino (Gartner, 1994; Gunn, 1972). A lo largo de los años la Secretaría de Turismo (SETUR) y las agencias de turismo han fortalecido la imagen del destino Natal como "ciudad del sol" y "ciudad del sol y del mar".

Otro factor para destacar es que el 48% del total general de comentarios está centrado en sólo dos playas: Praia de Ponta Negra y Praia de Genipabu. Estas playas son muy difundidas, incluso con fotos,

tanto en el portal de la SETUR y en los medios de los agentes turísticos como a través de tarjetas postales de la ciudad.

La segunda categoría más comentada es la de edificaciones y monumentos históricos, que incluye museos, teatros, iglesias, puntos históricos y edificios. A pesar de ser la categoría con mayor cantidad de atractivos (32), éstos representan sólo el 14% del total de los comentarios. Entre los atractivos de la categoría se destacan el Fuerte de los Reis Magos, el Puente Newton Navarro y el Centro de Lanzamiento de Cohetes de la Barreira do Inferno. Estos tres atractivos suman casi el 11% de los comentarios, representando el 78% del total de la categoría.

En este estudio los comentarios relacionados con las iglesias de la ciudad corresponden al 0,7% del total de CGUs. En coincidencia con la realidad indicada por el modelo generado, la Secretaría Municipal de Turismo (SETUR) publicó una nota oficial en el sitio de la prefectura de la ciudad (NATAL, 2017) el 25 de abril de 2017 informando del lanzamiento del “Itinerario Natal Sagrado e Histórico” de modo de incrementar el turismo religioso.

La categoría plazas y parques presenta una baja representatividad en el destino Natal, recibiendo sólo el 5% de todos los comentarios generados. Se destaca el Parque das Dunas (4,7%), pero al analizar la nube de palabras de este atractivo (Figura 1) se percibe que muchos de los comentarios realizados en esta carpeta no se refieren al parque en sí. Los aspectos cognitivos destacados en esta categoría (“dunas” y “buggy”) y los aspectos afectivos (“emoción” y “aventura”) se refieren al paseo en las Dunas de la Praia de Genipabu.

Figura 1: Nube de palabras de la categoría plazas y parques

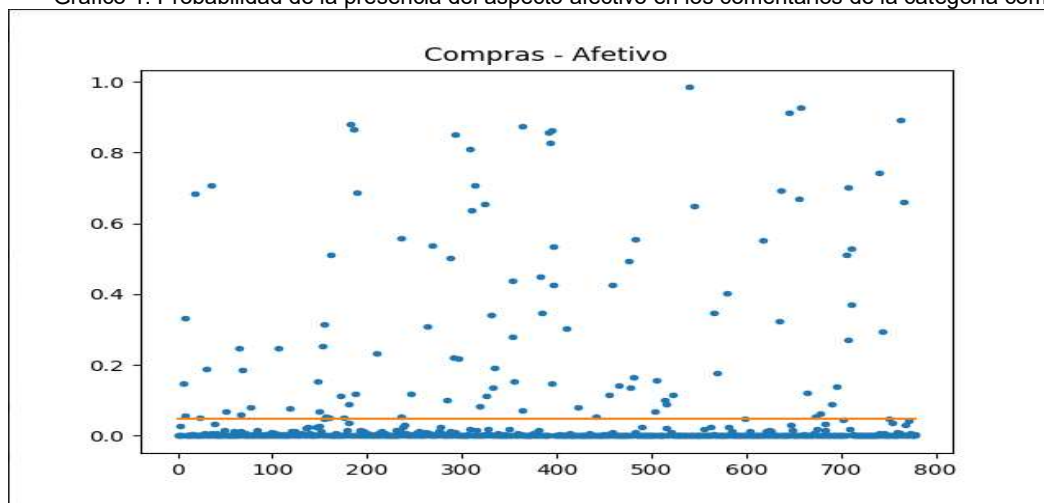


Fuente: Datos de la investigación (2017)

Este dato revela que el nombre del parque confunde a los turistas y genera una imagen errónea del atractivo para quien, al leer los comentarios, intenta tener una idea del lugar. Finalmente, la categoría compras fue la menos expresiva, destacándose uno de los shoppings de la ciudad y el Centro de Artesanos del barrio Ponta Negra.

El modelo desarrollado permitió identificar la probabilidad de que sean destacados los aspectos afectivo, conativo y cognitivo en la imagen del destino. En todas las categorías se obtuvo una fuerte presencia de los aspectos cognitivos en los comentarios y una baja presencia de los aspectos afectivos y conativos. Para algunas categorías ya se esperaba una baja presencia de aspectos afectivos y el modelo lo confirmó. El Gráfico 1 presenta el resultado de la categoría compras y muestra que la probabilidad de que los comentarios de dicha categoría tengan aspectos afectivos es próxima a 0%.

Gráfico 1: Probabilidad de la presencia del aspecto afectivo en los comentarios de la categoría compras



Fuente: Datos de la investigación (2017)

Reforzando lo expuesto puede decirse que el algoritmo NB presentó un resultado muy próximo al análisis descriptivo. La concentración de los puntos de la probabilidad es mayor cerca de la línea horizontal de la media. La justificación de los puntos dispersos en el gráfico se centra en que en los comentarios se describe más de un aspecto de la imagen del destino. El algoritmo NB también se aproxima en la Figura 2, donde las palabras más destacadas no hacen referencia a los aspectos afectivos, asumiendo una técnica plausible de utilizar el aprendizaje de máquinas para esta naturaleza de problema.

La nube de palabras de la categoría compras (Figura 2) muestra los términos más citados en los 780 comentarios de dicha categoría, como “shopping”, “tiendas”, “estacionamiento”, “artesanías” y “alimentación”; los cuales hacen referencia a los aspectos cognitivos presentes en la misma. Los aspectos afectivos no aparecen entre los términos más comentados, reforzando los resultados del Gráfico 1.

Figura 2: Nube de palabrras de la categoría compras



Fuente: Datos de la investigación (2017)

La Figura 3 muestra la nube de palabras de los términos más comentados en los 31.779 comentarios. En base a éstos es posible identificar que la imagen del destino Natal generada a partir de estos CGUs, conforme afirman Crompton (1979), Datzira-masip & Poluzzi (2014) y Kim & Richardson (2003), tiene como principales atributos, impresiones y características cognitivas las playas, paseos, dunas, paseos en *buggy*, lagunas, el Morro do Careca, el Forte dos Reis Magos y los restaurantes. En relación al aspecto afectivo la ciudad es vista como un lugar tranquilo, agradable y con atractivos que provocan emoción. En el aspecto conativo, formado a partir de las percepciones cognitivas y conativas, Natal es un destino excelente, de lindos paisajes, con atracciones recomendadas para la familia y los niños, que vale la pena conocer.

Figura 3: Nube de palabras de todos los comentarios



Fuente: Datos de la investigación (2017)

La nube de palabras de la Figura 3 también refleja de qué manera los turistas potenciales, que están planeando su viaje, generan a priori la imagen del destino Natal a partir de las experiencias de otros turistas (Filieri, 2015; Gretzel, 2011; Pantano, 2013). Los aspectos positivos de la ciudad son destacados en los comentarios, que de forma directa o indirecta recomiendan el destino a otros turistas (Santos & Cruz, 2013).

Este estudio confirma los hallazgos de Baka (2016), Marine-Roig & Anton Clave (2015; 2016) y Pantano *et al.* (2017), quienes afirman que es creciente la tendencia de los turistas a utilizar los medios sociales para describir sus experiencias de viaje.

A diferencia de la imagen inducida por los medios formales, los comentarios online son detallados y aproximan el destino (lo que se puede esperar de él) a quien aún no lo conoce (Dickinger, 2010; Munar, 2011; Pantano *et al.*, 2017). Otro aspecto presente en los comentarios online es la descripción de la experiencia en el destino, la cual se convierte en el producto fundamental del turismo (Molina, Moreno & Moreno, 2010). Estas experiencias relatan lo que se vivió en el destino, positivamente e negativamente, dando mayor credibilidad y realidad a la información (Wong & Qi, 2017).

Algunos ejemplos extraídos de los CGUs analizados ilustran el aspecto citado anteriormente, como es el caso del siguiente comentario de la Praia de Genipabu: *“La playa es tranquila, durante la mañana había corrientes que tiraban hacia los laterales. El agua es salada, pero muy limpia tanto dentro como fuera del mar. Hay sombrillas y sillas en la arena y se cobra 10,00 para quien las quiera utilizar”*. En un comentario sobre el Centro de lanzamiento de la Barreira do Inferno se puede apreciar la riqueza de los detalles: *“es un lugar que tiene mucha historia, fue la primera base área de América del Sur para lanzamiento de cohetes. En el lugar se exponen fotos, un avión, un misil y un cañón. Posee una historia muy interesante, vale la pena visitarlo. Queda camino a las playas del litoral sur”*.

Tabla 2: Desempeño del algoritmo *Naïve Bayes* para todas las categorías de comentarios

Categorías	Aspectos					
	Afectivo		Cognitivo		Conativo	
	Madia	Desvío Patrón	Madia	Desvío Patrón	Madia	Desvío Patrón
Playas y Lagunas	0.0697	0.1859	0.8256	0.2813	0.1047	0.2105
Edificaciones y Monumentos históricos	0.0598	0.1756	0.7884	0.3078	0.1518	0.2624
Plazas y Parques	0.0770	0.1976	0.7814	0.3117	0.1415	0.2512
Compras	0.0486	0.1560	0.7279	0.3568	0.2235	0.3353

Fuente: Datos de la investigación (2017)

Una visión general del desempeño del algoritmo NB para todas las previsiones de los aspectos del destino (afectivo, cognitivo y conativo), en relación a las cuatro categorías establecidas de comentarios (playas y lagunas, edificaciones y monumentos históricos, plazas y parques y compras), se presenta en la Tabla 2. Los valores corresponden a la media y desvío patrón de las probabilidades de los aspectos/categorías del destino. Los valores presentados en la Tabla 2 validan el abordaje automático utilizado, reforzando que el aspecto cognitivo es el que se presenta de manera más fuerte en los comentarios mediante las medias más altas para este tipo de aspecto y un bajo desvío patrón.

El resultado preocupa porque el aspecto conativo es muy bajo, lo que representa que la recomendación del destino es poco verbalizada, a pesar de que sus atractivos son bastante evidentes. Esta constatación señala para los portales turísticos la necesidad de insertar una evaluación de la recomendación del atractivo. En algunos comentarios no queda claro si el visitante recomienda o no el atractivo comentado, dado que indica aspectos positivos y negativos.

CONSIDERACIONES FINALES

El estudio contribuye a entender qué atributos son reconocidos por el turista que visita Natal/RN-Brasil. Los datos del CGU de *TripAdvisor* acerca del destino Natal-RN/Brasil remiten a algunas conclusiones y recomendaciones detalladas a continuación.

Las técnicas de procesamiento del lenguaje natural y el aprendizaje de máquinas, permiten afirmar que el aspecto cognitivo es el más fuerte presentado en los comentarios de los turistas. Estas técnicas ayudan a identificar los atributos más relevantes de la imagen del destino analizado, pudiendo ayudar a los organismos y empresas responsables del turismo de la región a tomar decisiones estratégicas de desarrollo e incentivo al turismo.

Las playas, lagunas, paseos, dunas y restaurantes, son algunos de los atractivos más destacados del destino. Los aspectos afectivos y conativos se presentan en una proporción inferior. En consonancia con la imagen difundida por la oferta, la imagen de la demanda, generada de forma orgánica a partir del CGU (turistas que visitan la ciudad), también fortalece el destino Natal/RN como itinerario de sol y playa. No obstante, un destino que ha fortalecido los atractivos de una única categoría termina por perder competitividad ante los destinos que ofrecen opciones en diferentes categorías.

Se entiende que los ítems más evaluados en *TripAdvisor* sobre el destino Natal influyen en la llegada de turistas potenciales hacia este destino. Para que este ciclo no afecte negativamente a otros itinerarios turísticos de la ciudad, menos comentados, es necesaria una política de incentivo por parte del gobierno y la iniciativa privada. De esta manera, se recomienda el desarrollo de políticas estratégicas para fomentar los atractivos que ya existen, pero que no tienen visibilidad en el destino, como es el caso de los puntos históricos y culturales de la ciudad; así como la prospección de nuevos productos turísticos, que tornen al destino más atractivo y competitivo

La relevancia del estudio se da desde el punto de vista práctico ya que desarrolla un método para captar la imagen del destino a partir del inmenso contenido de datos disponibles online. El uso de la minería de datos permitió analizar un volumen de 31.779 comentarios. La sistematización y análisis de este contenido permite una reducción del tiempo, esfuerzo y dinero gastado en la formación de la imagen por parte de los gestores turísticos.

La imagen del destino que se forma por medio del CGU tiene un alcance mayor y una alta velocidad de propagación, dada la visibilidad y la posibilidad de compartir la información. Estos datos deben ser utilizados como punto de partida para políticas específicas de los gestores turísticos, que sabiendo cómo es percibida la imagen pueden invertir sus recursos en enfatizar determinados segmentos, mejorar otros e invertir en puntos de la ciudad con potenciales turísticos no explotados.

REFERENCIAS BIBLIOGRÁFICAS

- Akehurst, G.** (2009) "User generated content: The use of blogs for tourism organisations and tourism consumers." *Service Business* 3(1): 51–61
- Ayeh, J. K.; Au, N. & Law, R.** (2013a) "Do we believe in TripAdvisor? Examining credibility perceptions and online travelers: Attitude toward using user-generated content." *Journal of Travel Research* 52(4): 437–452
- Ayeh, J. K.; Au, N. & Law, R.** (2013b) "Predicting the intention to use consumer-generated media for travel planning." *Tourism Management* 35: 132–143
- Ayeh, J. K.; Au, N. & Law, R.** (2016) "Investigating cross-national heterogeneity in the adoption of online hotel reviews." *International Journal of Hospitality Management* 55: 142–153
- Baka, V.** (2016) "The becoming of user-generated reviews: Looking at the past to understand the future of managing reputation in the travel sector." *Tourism Management* 53: 148–162
- Berezina, K.; Bilgihan, A.; Cobanoglu, C. & Okumus, F.** (2015) "Understanding satisfied and dissatisfied hotel customers: text mining of online hotel reviews." *Journal of Hospitality Marketing & Management* 25(1): 1–24
- Bing, P.; Zheng, X.; Law, R. & Fesenmaier, D. R.** (2011) "The dynamics of search engine marketing for tourist destinations." *Journal of Travel Research* 50(4): 365–377
- Brasil** (2017) <http://www.brasil.gov.br/turismo/2017/08/nordeste-e-a-regiao-que-mais-aproveita-potencial-turistico-diz-pesquisa>. Acesso em 05.07.2017
- Çakmak, E. & Isaac, R. K.** (2012) "What destination marketers can learn from their visitors' blogs: An image analysis of Bethlehem, Palestine." *Journal of Destination Marketing & Management* 1(2): 124–133
- Carson, D.** (2008) "The 'blogosphere' as a market research tool for tourism destinations: A case study of Australia's Northern Territory." *Journal of Vacation Marketing* 14(2): 111–119
- Cenni, I. & Goethals, P.** (2017) "Negative hotel reviews on TripAdvisor: A cross-linguistic analysis." *Discourse, Context & Media* 16: 22–30
- Chagas, M. M. das; Sampaio, L. M. B. & Santos, K. E. B. S.** (2013) "Análise do processo de formação

da imagem de destinos turísticos de sol e praia: um estudo em Canoa Quebrada/CE.” *Revista Brasileira de Pesquisa Em Turismo* 7(2): 296–316

Crompton, J. L. (1979) “An assessment of the image of Mexico as a vacation destination and the influence of geographical location upon that image.” *Journal of Travel Research* 17(4): 18–23

Datzira-Masip, J. & Poluzzi, A. (2014) “Brand architecture management: The case of four tourist destinations in Catalonia.” *Journal of Destination Marketing and Management* 3(1): 48–58

Dickinger, A. (2010) “The trustworthiness of online channels for experience - and goal- directed search tasks.” *Journal of Travel Research* 50(4): 378-391

Filieri, R. (2015) “Why do travelers trust TripAdvisor? Antecedents of trust towards consumer-generated media and its influence on recommendation adoption and word of mouth.” *Tourism Management* 51: 174–185

Gallarza, M. G.; Saura, I. G. & García, H. C. (2002) “Destination image.” *Annals of Tourism Research* 29(1): 56–78

Gartner, W. C. (1994) “Journal of travel & image formation process.” *Journal of Travel & Tourism Marketing* 2(2–3): 191–216

Ghazali, R. M. & Cai, L. (2014) “Social media sites in destination image formation.” *Journal of Tourism and Cultural Change* 12(1): 73–86

Gretzel, U. (2011) “Intelligent systems in tourism.” *Annals of Tourism Research* 38(3): 757–779

Gunn, C. (1972) “Designing tourist regions.” Von Nostrand Reinhold, New York

Guo, C. & Morris, S. A. (2017) “Engineering cell identity: establishing new gene regulatory and chromatin landscapes”. *Current Opinion in Genetics and Development* 46: 50-57

Hsu, F.-M.; Lin, Y.-T. & Ho, T.-K. (2012) “Design and implementation of an intelligent recommendation system for tourist attractions: The integration of EBM model, Bayesian network and Google Maps.” *Expert Systems with Applications* 39(3): 3257–3264

John, G. H. G. & Langley, P. (1995) “Estimating continuous distributions in bayesian classifiers.” In proceedings of the eleventh conference on uncertainty in artificial intelligence. Montreal, Quebec, 1: 338--345

Johnson, P. A.; Sieber, R. E.; Magnien, N. & Ariwi, J. (2012) “Automated web harvesting to collect and analyse user-generated content for tourism”. *Current Issues in Tourism* 15(3): 293-299

Kim, H. & Richardson, S. L. (2003) “Motion picture impacts on destination images.” *Annals of Tourism Research* 30(1): 216–237

Korfiatis, N., García-Bariocanal, E., & Sánchez-Alonso, S. (2012) “Evaluating content quality and helpfulness of online product reviews: The interplay of review helpfulness vs. review content.” *Electronic Commerce Research and Applications* 11(3): 205–217

Li, G.; Law, R.; Vu, H. Q.; Rong, J. & Zhao, X. (2015) “Identifying emerging hotel preferences using emerging pattern mining technique.” *Tourism Management* 46: 311–321

Lu, W. & Stepchenkova, S. (2012) “Ecotourism experiences reported online: Classification of satisfaction attributes.” *Tourism Management* 33(3): 702–712

Marine-Roig, E. & Anton Clavé, S. (2015) “Tourism analytics with massive user-generated content: A case study of Barcelona.” *Journal of Destination Marketing and Management* 4(3): 162–172

- Marine-Roig, E. & Anton Clavé, S.** (2016) "Perceived image specialisation in multiscale tourism destinations." *Journal of Destination Marketing and Management* 5(3): 202–213
- Melián-González, S.; Bulchand-Gidumal, J. & López-Valcárcel, B. G.** (2013) "Online customer reviews of hotels." *Cornell Hospitality Quarterly* 54(3): 274–283
- Molina, C. M.; Moreno, M. R. & Moreno, R. R.** (2010) "Imagen de los destinos turísticos urbanos y lealtad del turista: ¿Actitud o comportamiento?" *Estudios y Perspectivas en Turismo* 19(2): 279–298
- Munar, A. M.** (2011) "Tourist-created content: rethinking destination branding." *International Journal of Culture, Tourism and Hospitality Research* 5(3): 291–305
- Natal** (2017a) <http://www.natal.rn.gov.br/noticia/ntc-25933.html>. Acesso em 10.08.2017
- Nguyen, H. T. H. & Cao, J.** (2015) "Trustworthy answers for top-k queries on uncertain Big Data in decision making." *Information Sciences* 318: 73–90
- Pantano, E.** (2013) "From e-tourism to f-tourism: emerging issues from negative tourists' online reviews." *Journal of Hospitality and Tourism Technology* 4(3): 211–227
- Pantano, E.; Priporas, C. V. & Stylos, N.** (2017) "'You will like it!' using open data to predict tourists' response to a tourist attraction." *Tourism Management* 60: 430–438
- Ring, A.; Tkaczynski, A. & Dolnicar, S.** (2016) "Word-of-mouth segments." *Journal of Travel Research* 55(4): 481–492
- Santos, E. D. S. & Cruz, G. da** (2013) "Imágen turística de Ilhéus (BA- Brasil) según la percepción de los visitantes." *Estudios y Perspectivas en Turismo* 22(4): 607–629
- Soler, I. P. & Germán, G.** (2016) "Brand equity research using online customer ratings of Spanish hotels." *International Journal of Tourism Research* 113(November): 101–113
- Tamajón, G. L. & Valiente, C. G.** (2015) "Barcelona seen through the eyes of TripAdvisor: actors, typologies and components of destination image in social media platforms." *Current Issues in Tourism* (Dezember): 1–5 DOI <http://dx.doi.org/10.1080/13683500.2015.1073229>
- Tasci, A. D. & Gartner, W. C.** (2007) "Destination image and its functional relationships." *Journal of Travel Research* 45(4): 413–425
- Volo, S.** (2010) "Bloggers' reported tourist experiences: Their utility as a tourism data source and their effect on prospective tourists." *Journal of Vacation Marketing* 16(4): 297–311
- Wong, C. U. I. & Qi, S.** (2017) "Tracking the evolution of a destination's image by text-mining online reviews - the case of Macau." *Tourism Management Perspectives* 23: 19–29
- Xiang, Z.; Magnini, V. P. & Fesenmaier, D. R.** (2015) "Information technology and consumer behavior in travel and tourism: Insights from travel planning using the internet." *Journal of Retailing and Consumer Services* 22: 244–249

Recibido el 20 de agosto de 2017

Reenviado el 06 de septiembre de 2017

Aceptado el 09 de septiembre

Arbitrado anónimamente

Traducido del portugués