



Position Paper: **Tecnologías del**
Lenguaje aplicadas al sector
turístico

Julio 2021

Ametic
LA VOZ DE LA INDUSTRIA DIGITAL

Contenido

Contextualización	2
Tecnologías relacionadas	2
Modelos comunes de datos	¡Error! Marcador no definido.
Ontologías	¡Error! Marcador no definido.
Interacción por voz	3
Asistentes conversacionales	3
Casos de uso	¡Error! Marcador no definido.
Análisis de patologías a partir de historias clínicas	¡Error! Marcador no definido.
Diagnóstico y pronóstico temprano	¡Error! Marcador no definido.
Detección de brotes y crisis sanitarias	¡Error! Marcador no definido.
Ayuda en la salud mental	¡Error! Marcador no definido.
Retos del NLP en salud	5
Anonimización	¡Error! Marcador no definido.
Estandarización	¡Error! Marcador no definido.
Disponibilidad de datos anotados	¡Error! Marcador no definido.
Situación en España	6

Contextualización

El sector turístico es muy amplio e incluye un gran número de subsectores interrelacionados, como el transporte, las agencias de viaje, los intermediarios, la industria hotelera y la restauración, y los propios destinos turísticos. Las tecnologías de NLP aportan valor al sector turístico en tres áreas: la automatización de procesos para eliminar aquellas tareas repetitivas que consumen recursos, nuevas formas de interacción con el turista gestionando de forma eficiente aspectos como las reservas, preguntas o quejas, y el análisis de los grandes volúmenes de datos generados para los diferentes subsectores turísticos.

Tecnologías relacionadas

Las tecnologías de Procesamiento del Lenguaje Natural (NLP) permiten la interpretación del lenguaje (voz y texto) por parte de las máquinas, de modo que se facilite la interacción entre la persona y los dispositivos. La gran diversidad de idiomas y su complejidad hacen que se trate de una tarea muy retadora.

La aplicación de las tecnologías NLP está revolucionando el interfaz de usuario, ya que crea una nueva experiencia única, con una reducción en el número de clics. Por ello, existe un gran número de casos de usos potenciales para esta tecnología en el sector turístico.

Minería de textos

La minería de textos permite analizar grandes cantidades de contenidos textuales de distintas fuentes (foros de discusión online, noticias, opiniones, informes) para extraer datos y realizar resúmenes que ayuden en la toma de decisiones. En general, estas soluciones se utilizan para analizar texto digital de diferentes fuentes (motores de búsqueda, blogs y foros) y plataformas de medios sociales (Twitter y Facebook) con el objetivo de identificar patrones y tendencias en relación a la afinidad con la marca, preferencias de producto o patrones de consumo.

Existen diferentes aplicaciones básicas de las técnicas de minería de textos, aunque la mayoría de ellas se centra en la detección de topics y el análisis de sentimiento. Ambas tecnologías pueden resultar muy útiles para descubrir la opinión real de los consumidores. En primer lugar, la extracción de topics se puede utilizar para detectar los temas más relevantes para el público y seguir eventos representativos, de modo que el ecosistema turístico pueda detectar oportunidades de negocio o gestión de crisis en tiempo real. En segundo lugar, de forma complementaria, el análisis de sentimiento permite cuantificar el tipo e intensidad de la emoción expresada en el texto. No se debe olvidar que muchos de los agentes del ecosistema turístico basan su éxito en la generación de emociones, por lo que dependen de los comentarios de los clientes.

Una de las aplicaciones más interesantes para el ecosistema turístico es la inteligencia competitiva a partir de los comentarios online, ya que la minería de texto permite identificar potenciales competidores y evaluar la imagen de marca. Igualmente, la combinación del análisis de texto junto con los datos históricos de ventas de producto puede ayudar en la predicción de ventas de productos.

Finalmente, hay que indicar que la gestión de las noticias falsas (fake news) o la desinformación es una de las principales líneas emergentes para la aplicación de técnicas de minería de textos.

Así, empresas del sector turístico han comenzado a utilizar este tipo de técnicas para validar la información en el caso de quejas o incumplimientos en la cadena de valor de paquetes y servicios turísticos que conllevan varios contratos (agencias, hoteles, rentacars, líneas aéreas, etc.) de tal forma que se puedan determinar la mejor respuesta en tiempo real para la resolución de una incidencia o una prestación indemnizatoria.

Generación personalizada de textos

La generación de lenguaje natural ofrece herramientas interesantes para generar ofertas turísticas a partir de datos estructurados, ya que las empresas que se dedican a la venta de productos turísticos abordan el reto de describir cientos de productos y servicios. Por ello, las grandes empresas turísticas y algunas pequeñas cadenas hoteleras y agencias de viaje están experimentando con estas tecnologías para la creación de descripciones sobre alojamientos o la generación automática de ofertas temporales.

Por ejemplo, estas herramientas tienen en cuenta el número de habitaciones, la distancia a un recurso turístico o una característica particular de un alojamiento para crear textos de forma automática, de modo que se generen automáticamente contenidos para Web, folletos u otros medios digitales que incluyan diferentes palabras clave en función de los mercados objetivo. De esta forma, es posible personalizar los contenidos e incluso la propia forma de redacción y el tono de los textos en función de los perfiles de usuario.

Interacción por voz

Los alojamientos podrán utilizar técnicas de NLP a partir de asistentes de voz, que pueden ayudar en tareas como el ajuste de la temperatura del aire acondicionado, el ajuste de las luces de la habitación o la sintonización de las series preferidas. Igualmente, los robots con capacidad de generación de lenguaje pueden comprender las necesidades de los visitantes y ofrecerles los servicios requeridos. Por ejemplo, pueden ayudar en el servicio de habitaciones sirviendo productos o recogiendo la ropa de la lavandería.

Asistentes conversacionales

Uno de los principales usos de las tecnologías de Procesamiento del Lenguaje Natural (NLP) para el sector turístico es la provisión de información online a los clientes, principalmente a través de los medios sociales y las Apps de mensajería. Los chatbots permiten establecer conversaciones con las personas mediante interfaces de tipo chat, utilizando plataformas o buscadores ya disponibles en los dispositivos móviles de los turistas. Dichos chatbots son multilingües, ofrecen respuestas instantáneas y están disponibles 24/7, de modo que los clientes pueden obtener respuestas de forma rápida y en cualquier momento, incrementando su satisfacción y fidelización.

En relación a su aplicación en el ecosistema turístico, la aplicación de chatbots se centra actualmente en las reservas como primera etapa en la experiencia turística, lo que suele implicar elegir el medio de transporte hasta y en el destino, y el alojamiento. Se trata de ayudar en la búsqueda del vuelo más corto con el precio adecuado, o del alojamiento que cumple las necesidades del viajero de una forma más amigable, ya que las personas se sienten más cómodas manteniendo una conversación para clarificar sus dudas. Entre otros factores, la principal diferencia con una relación entre personas es la capacidad del chatbot de acceder a información previa sobre el cliente para personalizar la conversación y ofrecer diferentes productos complementarios para favorecer el cross-selling.

Igualmente, los chatbots pueden grabar sus conversaciones con los usuarios, de modo que los alojamientos y destinos pueden adquirir datos sobre la experiencia del consumidor para realizar estudios de mercado y refinar sus productos turísticos. Además, pueden incentivar la provisión de comentarios y feedback por parte de los turistas, animándoles a escribir comentarios y valorar su experiencia. En el caso de identificar una queja, el chatbot puede ofrecer canales alternativos para gestionarla, antes de recibir críticas negativas en las redes sociales.

Traducción automática

La traducción automática ofrece ventajas claves para la competitividad de las empresas en sectores como el turismo, al proveer una rápida internacionalización y aumentos de productividad. En particular, la traducción automática es uno de los componentes principales para el crecimiento de un mercado digital único europeo, donde un 43% de los ciudadanos europeos no están dispuestos a adquirir productos o servicios online que no se ofrezcan en su idioma nativo.

Las soluciones de traducción automática son más efectivas para dominios acotados (hoteles, restauración, aerolíneas), ofreciendo en estos casos la mayor calidad y usabilidad. Para desarrollar sistemas adaptados a dominios específicos, los paradigmas dominantes actuales son la traducción automática estadística y la traducción automática neuronal, que aprenden de los propios datos del dominio. La combinación de estos dos aspectos presenta una oportunidad fuerte para una integración de la tecnología de traducción automática como apoyo al desarrollo de la actividad de empresas en industrias y sectores variados.

Las tecnologías NLP ayudan en la superación de las barreras lingüísticas, ya que pueden traducir textos en señales de forma automática y ayudar en el guiado de los turistas hacia su destino.

Retos del NLP en salud

La utilización de herramientas de NLP en el ecosistema turística deberá abordar la mejora de la eficiencia de sus principales procesos. Tal y como se ha detallado anteriormente, estas tecnologías permiten una mejora en la experiencia de la reserva, reinventar la experiencia en los hoteles y en los destinos turísticos, personalizar la oferta turística y mejorar los procesos de gestión. Ahora bien, se han identificado una serie de retos que se deberán abordar en los próximos años.

- Tecnologías de voz para una gestión automática, eficiente y multilingüe de los call centres, incluyendo la generación de respuestas automáticas o la organización de las llamadas y reenvío automático a los profesionales correspondientes.
- Herramientas para la generación automática en tiempo real de contenidos textuales (narrativa, vocabulario) para los paquetes turísticos en función de los perfiles de usuarios a partir de datos estructurados, de modo que se consigan un storytelling y resúmenes personalizados.
- Herramientas tecnológicas (semi)automáticas para abordar las necesidades de traducción de textos, incluyendo la traducción de contenidos estructurados de destinos y alojamientos, o la respuesta personalizada a los comentarios de los clientes.
- Asistentes cognitivos de interacción natural por voz (chatbot) adaptada diferentes idiomas y acentos principales, que sirvan de apoyo y asistencia a turistas y profesionales.
- Algoritmos híbridos basados en AI y Procesamiento del Lenguaje Natural (NLP) para la detección automática de comentarios, rumores y noticias textuales falsas en los medios sociales (fake news), la detección de engaño, intrusión o suplantación de identidad o la comprobación automática de hechos.

Situación en España

En 2018 el Observatorio Nacional de las Telecomunicaciones y de la Sociedad de la Información (ONTSI)¹ publicó un estudio de caracterización de la industria y otros agentes de las tecnologías del lenguaje.

Las soluciones que integran la actividad del sector incluyen la traducción automática y la traducción asistida por ordenador, el procesamiento del habla y los sistemas conversacionales, el procesamiento del lenguaje natural y tecnologías semánticas basadas en análisis de texto, y los recursos para las tecnologías del lenguaje.

Los agentes del sector se distribuyen mayoritariamente entre la Comunidad de Madrid, el País Vasco, Cataluña y Comunidad Valenciana. Los agentes del sector tienen cierta madurez en la dedicación a actividades de tecnologías del lenguaje, aunque la mayoría desarrolla la actividad TL combinada con otras líneas de negocio. La mayoría de las empresas consultadas tienen menos de 49 empleados, lo que mostraría un sector conformado por microempresas y pequeñas empresas. El sector turístico supone un 45% de las ventas de tecnologías del lenguaje de las empresas.

Entre las principales barreras detectadas, se pueden mencionar la falta de datos o corpus para entrenar los algoritmos de NLP, la falta de profesionales formados, el pequeño tamaño de las empresas. El sector turístico se presenta como una oportunidad para el sector, ya que es intensivo en procesos de cara al público utilizando diferentes canales, por lo que el lenguaje como interfaz es crítico para las actividades de este sector durante toda la etapa del viaje.

Recientemente, España ha puesto en marcha un grupo de trabajo Interministerial perteneciente al Ministerio de Ciencia e Innovación enfocado a la Inteligencia Artificial. En 2019, se publicó la *Estrategia Española en I+D+I en Inteligencia Artificial*, donde se remarca el papel de las ciudades y territorios inteligentes como uno de los desarrollos más importantes a los que va a contribuir la IA, mencionando explícitamente los vehículos autónomos, las aplicaciones que mejoren y hagan más segura la movilidad, o la mejora de la conectividad del transporte. Igualmente, se menciona la aplicación de las técnicas de visión artificial como herramienta básica para mejorar la seguridad, incluyendo la monitorización de cámaras de seguridad en tiempo real en grandes eventos o centros de transporte, así como la seguridad vial.

Esta Estrategia ha sido el embrión de la recientemente aprobada *Estrategia Nacional de Inteligencia Artificial* (ENIA), que va a permitir coordinar y alinear las inversiones y políticas de estado en relación a la IA. Tal y como recoge la Estrategia, el turismo y la movilidad son sectores con un gran impacto esperado a corto y medio plazo. Entre las acciones a realizar, se mencionan la creación de repositorios de datos públicos para desarrollar nuevas aplicaciones y oportunidades que favorezcan la gestión urbana y la movilidad sostenible, o la implantación de programas de cooperación público-privada para una movilidad sostenible e inteligente.

1

<https://www.ontsi.red.es/sites/ontsi/files/Estudio%20Caracterizaci%3bn%20del%20Sector%20de%20Tecnolog%3adas%20del%20Lenguaje%20en%20Espa%3b1a%202018.pdf>