



UNIVERSIDAD  
DE CANTABRIA

**ESCUELA DE DOCTORADO DE LA UNIVERSIDAD DE CANTABRIA  
PROGRAMA DE DOCTORADO EN CIENCIAS JURÍDICAS Y  
EMPRESARIALES**

TESIS DOCTORAL

**APLICACIÓN DE LA CIENCIA DE DATOS AL ESTUDIO DE LA  
SOSTENIBILIDAD TURÍSTICA: ENFOQUE TEÓRICO Y ANÁLISIS  
EMPÍRICO DE EMPRESAS Y TURISTAS EN ESPAÑA**

**AUTORA**

MARÍA FERNANDA BERNAL SALAZAR

**DIRECTORES**

ELISA BARAIBAR-DIEZ  
JESÚS COLLADO AGUDO

Santander, 2025



## AGRADECIMIENTOS

Al finalizar esta Tesis Doctoral, me doy cuenta de que sin las personas que me han acompañado en este proceso, este esfuerzo y logro no tendría el mismo significado. Este trabajo es el resultado de un enorme crecimiento académico y personal, así como el apoyo incondicional y la colaboración de muchas personas que han sido fundamentales para su culminación.

En primer lugar, quiero agradecer a mis directores de Tesis, la Dra. Elisa Baraibar-Diez y el Dr. Jesús Collado Agudo, por su guía, enseñanza y constante motivación. Son profesionales admirables y personas excepcionales. A lo largo de estos tres años, su apoyo, amabilidad y confianza han sido primordiales para el desarrollo y finalización de este trabajo. Sinceramente, ha sido un privilegio contar con dos directores que destacan tanto por su excelencia profesional como por su calidad humana.

En segundo lugar, agradezco al grupo de investigación del Área de Comercialización e Investigación de Mercados de la Universidad de Cantabria: Marimar, Patri, Héctor, Ángel e Inna, por su cálida acogida y por brindarme siempre un aliento de motivación. Gracias por su colaboración y por su disposición en cada etapa del proceso. Ha sido un placer conocerlos y compartir tantas experiencias que reflejan su profesionalismo y valores personales.

Mi agradecimiento se extiende a todas las personas del Departamento de Administración de Empresas de la Universidad de Cantabria. Su apoyo, su disposición para escucharme y sus palabras de aliento me han impulsado a seguir adelante. En particular, quiero agradecer a Estefanía Palazuelos, Javier Montoya y Paula San Martín, por sus constantes palabras de aliento, sus sonrisas y buenos días, que siempre fueron una motivación para comenzar con energía el día. También a Carlos López Gutiérrez, por su amabilidad, disposición y colaboración.

Quiero agradecer y dedicar esta Tesis Doctoral a mi madre, hermana y abuela, quienes han sido mi motor constante y mi inspiración en la vida. Gracias por su apoyo y amor incondicional, por sus palabras de aliento y por siempre estar a mi lado. También agradezco a mi padre, mi tía Lucero, mis tíos Pedro y Marta, mis primos y, en general, a toda mi familia, por su amor y respaldo, incluso desde la distancia. Asimismo, agradezco profundamente y desde un lugar especial de mi corazón a Diego, por motivarme a tomar este camino, por su amor, apoyo constante, colaboración y paciencia. Asimismo, quiero dar un agradecimiento especial a todas aquellas personas que llegaron a mi vida para enriquecerla y motivarme durante el proceso de esta Tesis Doctoral.

Por último, no puedo concluir estos agradecimientos sin nombrar el proyecto de investigación que ha sido esencial para la realización y culminación de este trabajo: “TED2021-131314B-I00 SOSTENIBILIDAD CORPORATIVA Y TURISMO INTELIGENTE EN COMUNIDADES RURALES: INFLUENCIA EN EL DESARROLLO ECONÓMICO Y SOCIAL DEL TERRITORIO”, financiado por MICIU/AEI /10.13039/501100011033 y por la Unión Europea NextGenerationEU/ PRTR.

## **ÍNDICE DE CONTENIDOS**



## ÍNDICE DE CONTENIDOS

PRESENTACIÓN DE LA TESIS DOCTORAL .....	25
<b>1. Justificación de la investigación .....</b>	<b>25</b>
1.1. Ámbito de la investigación .....	29
<b>2. Objetivos del estudio .....</b>	<b>31</b>
<b>3. Metodología y estructura de la Tesis Doctoral .....</b>	<b>33</b>
CAPÍTULO I. MARCO CONCEPTUAL GENERAL .....	39
<b>1. Sostenibilidad y turismo .....</b>	<b>39</b>
1.1. Diferencia entre el concepto de desarrollo sostenible y sostenibilidad .....	39
1.2. El concepto de turismo sostenible .....	40
1.3. La sostenibilidad en el contexto turístico.....	44
1.4. Dimensiones de la sostenibilidad en el turismo.....	45
1.5. Destino turístico y sostenibilidad .....	48
1.6. Actores involucrados en la sostenibilidad turística .....	49
1.6.1. Residentes.....	50
1.6.2. Turistas.....	50
1.6.3. Empresas.....	51
1.6.4. Gobierno .....	51
1.7. Turismo regenerativo.....	52
<b>2. Nuevos enfoques para analizar la sostenibilidad turística: Ciencia de datos y <i>big data</i> .....</b>	<b>54</b>
2.1. Ciencia de datos en la investigación.....	54
2.2. Inteligencia artificial y ciencia de datos .....	56
2.3. Tipos de análisis en ciencia de datos.....	57
2.4. Fases del proceso de ciencia de datos .....	59
2.5. Análisis de grandes volúmenes de datos: <i>Big data</i> .....	61
2.6. Ciencia de datos y <i>big data</i> en la investigación turística.....	64
2.7. Fuentes de datos para análisis de <i>big data</i> en turismo.....	66
<b>3. Conclusiones .....</b>	<b>67</b>

CAPÍTULO II. RED DEL CONOCIMIENTO SOBRE SOSTENIBILIDAD TURÍSTICA A TRAVÉS  
DEL ANÁLISIS BIBLIOMÉTRICO Y UNA REVISIÓN COMPUTACIONAL DE LA LITERATURA 71

<b>1. Introducción</b> .....	<b>71</b>
<b>2. Revisión de la literatura</b> .....	<b>73</b>
2.1. Análisis bibliométricos previos sobre sostenibilidad turística.....	73
2.2. Métodos computacionales en la revisión de la literatura sobre sostenibilidad turística.....	75
<b>3. Metodología de investigación</b> .....	<b>76</b>
3.1. Definición del objetivo conceptual.....	76
3.2. Recopilación de literatura .....	76
3.3. Análisis bibliométrico .....	77
3.4. Revisión computacional de la literatura.....	78
3.4.1. Modelado de temas BERTopic .....	78
3.5. Análisis de distribución y tendencias .....	79
<b>4. Resultados</b> .....	<b>79</b>
4.1. Distribución de la producción científica.....	79
4.2. Cooperación científica.....	81
4.3. Impacto e influencia científica .....	82
4.4. Estructura del conocimiento y publicaciones clave .....	84
4.5. Temáticas y áreas de investigación .....	85
4.6. Tendencias de las temáticas de investigación.....	90
<b>5. Conclusiones</b> .....	<b>93</b>

CAPÍTULO III. EVOLUCIÓN TEÓRICA Y PRÁCTICA DESDE LA RESPONSABILIDAD  
SOCIAL CORPORATIVA HACIA LA SOSTENIBILIDAD CORPORATIVA: UNA APLICACIÓN DE  
LA CIENCIA DE DATOS EN LAS EMPRESAS TURÍSTICAS DE ESPAÑA ..... 99

<b>1. Introducción</b> .....	<b>99</b>
<b>2. Revisión de la literatura</b> .....	<b>102</b>
2.1. Responsabilidad Social Corporativa y Sostenibilidad Corporativa: ¿Cambio de paradigma?102	
2.2. RSC y SC en el sector turístico .....	107
2.3. Aplicación de la ciencia de datos en informes de RSC y SC.....	108

<b>3. Preguntas e hipótesis de investigación .....</b>	<b>110</b>
<b>4. Metodología de la investigación .....</b>	<b>112</b>
4.1. Metodología aplicable al enfoque teórico de la revisión computacional de la literatura ....	112
4.2. Metodología aplicable al enfoque empírico de la revisión de las memorias corporativas de las empresas turísticas españolas .....	113
4.2.1. Recopilación de datos .....	114
4.2.2. Métodos y técnicas de investigación .....	115
4.2.3. Codificación y categorización.....	115
4.2.4. Preprocesamiento.....	116
4.2.5. Modelado de temas basado en el enfoque BERTopic .....	117
4.2.6. Análisis estadístico de los datos.....	117
4.2.7. Clasificación bajo los criterios ESG.....	118
<b>5. Resultados .....</b>	<b>119</b>
5.1. Resultados del enfoque teórico: Revisión computacional de la literatura .....	119
5.2. Resultados del enfoque empírico: Análisis de los informes corporativos de las empresas turísticas.....	125
5.2.1. Análisis de títulos y palabras frecuentes.....	125
5.2.2. Modelado de temas con el enfoque BERTopic .....	128
5.2.3. Áreas claves de gestión en RSC y SC .....	129
5.2.4. Análisis estadístico.....	133
5.2.5. Pruebas de normalidad y <i>Bootstrapping</i> .....	134
5.2.6. Importancia de los criterios ESG en las empresas turísticas de España .....	135
<b>6. Conclusiones .....</b>	<b>139</b>
6.1. Conclusiones del enfoque teórico: Revisión computacional de la literatura .....	139
6.2. Conclusiones del enfoque empírico: Revisión de los informes corporativos de las empresas turísticas de España .....	140
6.3. Limitaciones y futuras líneas de investigación .....	142
<b>CAPÍTULO IV. DISCURSO DE LOS TURISTAS SOBRE SOSTENIBILIDAD: INSIGHTS A PARTIR DE RESEÑAS EN LÍNEA DE ATRACCIONES TURÍSTICAS RURALES Y URBANAS DE ESPAÑA .....</b>	<b>145</b>
<b>1. Introducción .....</b>	<b>145</b>
<b>2. Revisión de la literatura .....</b>	<b>147</b>

2.1. Sostenibilidad percibida y satisfacción de los turistas: Enfoque de <i>big data</i> y UGC.....	147
2.2. Diferencias de sostenibilidad en zonas turísticas rurales y urbanas .....	152
<b>3. Metodología de investigación .....</b>	<b>153</b>
3.1. Recopilación de datos .....	154
3.2. Clasificación de los comentarios con Inteligencia Artificial.....	157
3.3. Modelos Ologit y Gologit.....	159
<b>4. Resultados .....</b>	<b>162</b>
4.1. Pruebas estadísticas .....	162
4.1.1. Influencia del discurso sostenible de los turistas sobre la calificación .....	163
4.1.2. Diferencias de percepción sostenible del turista en zonas rurales y urbanas .....	168
<b>5. Conclusiones .....</b>	<b>173</b>
<b>CAPÍTULO V. CONCLUSIONES DE LA TESIS DOCTORAL.....</b>	<b>179</b>
<b>1. Conclusiones teóricas.....</b>	<b>180</b>
1.1. Conclusiones relativas al concepto e interpretación de la sostenibilidad turística .....	180
1.2. Conclusiones relativas a la ciencia de datos en la investigación turística .....	182
1.3. Conclusiones relativas al análisis de la red de conocimiento sobre sostenibilidad y turismo .....	183
1.4. Conclusiones relativas a la evolución teórica de la RSC a la SC en el turismo.....	185
<b>2. Conclusiones empíricas .....</b>	<b>186</b>
2.1. Implicaciones para la gestión empresarial turística .....	187
2.2. Implicaciones para la gestión de atractivos turísticos.....	189
<b>3. Limitaciones de la investigación .....</b>	<b>192</b>
3.1. Limitaciones asociadas a los estudios teóricos .....	192
3.2. Limitaciones asociadas a los estudios empíricos.....	192
<b>4. Futuras líneas de investigación.....</b>	<b>193</b>
<b>REFERENCIAS.....</b>	<b>199</b>

## **ÍNDICE DE TABLAS**



## ÍNDICE DE TABLAS

TABLA 1. CONSOLIDADO DE ANEXOS ASOCIADOS A TÉCNICAS Y HERRAMIENTAS METODOLÓGICAS APLICADAS EN LA TESIS DOCTORAL.....	34
TABLA 2. PRINCIPALES NOCIONES DEL CONCEPTO DE TURISMO SOSTENIBLE. ....	41
TABLA 3. DIMENSIONES DE LA SOSTENIBILIDAD EN EL TURISMO. ....	47
TABLA 4. DEFINICIONES DE LA CIENCIA DE DATOS POR PERSPECTIVAS. ....	56
TABLA 5. ANÁLISIS DE VOSVIEWER UTILIZADOS. ....	77
TABLA 6. CATEGORÍAS TEMÁTICAS DE LA INVESTIGACIÓN SOBRE RSC Y SC EN TURISMO.....	121
TABLA 7. RESULTADOS ESTADÍSTICOS DE LA REGRESIÓN LINEAL DE RSC Y SC ENTRE 2016 Y 2023.....	124
TABLA 8. RESULTADOS ESTADÍSTICOS DE LA REGRESIÓN LINEAL DE RSC Y SC ENTRE 2018 Y 2023.....	124
TABLA 9. DETALLE DE LA MUESTRA DE ATRACCIONES TURÍSTICAS DE ESPAÑA.....	156
TABLA 10. EJEMPLOS DE COMENTARIOS CON SU CLASIFICACIÓN Y PUNTUACIÓN. ....	158
TABLA 11. ECUACIONES DE LOS MODELOS OLOGIT Y GOLOGIT. ....	160
TABLA 12. NÚMERO DE RESEÑAS POR CATEGORÍA DE LA VARIABLE DEPENDIENTE.....	161
TABLA 13. VARIABLES INCLUIDAS EN LOS ANÁLISIS DE REGRESIÓN. ....	161
TABLA 14. RESULTADOS DE LOS MODELOS GOLOGIT.....	164
TABLA 15. MODELOS GOLOGIT POR ZONAS RURALES Y URBANAS DE M1 Y M2.....	170
TABLA 16. MODELOS GOLOGIT POR ZONAS RURALES Y URBANAS DE M3 Y M4.....	172



## ÍNDICE DE ILUSTRACIONES



## ÍNDICE DE ILUSTRACIONES

ILUSTRACIÓN 1. TOP 10 DE ECONOMÍAS LÍDERES EN FOMENTAR EL DESARROLLO DE VIAJES Y TURISMO. .....	30
ILUSTRACIÓN 2. OBJETIVOS ESPECÍFICOS, HIPÓTESIS Y/O PREGUNTAS DE INVESTIGACIÓN. ....	33
ILUSTRACIÓN 3. MARCO METODOLÓGICO GENERAL DE LA TESIS DOCTORAL. ....	35
ILUSTRACIÓN 4. ESTRUCTURA DE LA TESIS DOCTORAL. ....	36
ILUSTRACIÓN 5. ACTORES CLAVES DE LA SOSTENIBILIDAD EN TURISMO. ....	49
ILUSTRACIÓN 6. CONCEPTOS CLAVE DE INTELIGENCIA ARTIFICIAL Y CIENCIA DE DATOS. ....	57
ILUSTRACIÓN 7. TIPOS DE ANÁLISIS EN CIENCIA DE DATOS. ....	58
ILUSTRACIÓN 8. FASES DE LA CIENCIA DE DATOS. ....	61
ILUSTRACIÓN 9. TIPOS DE DATOS. ....	63
ILUSTRACIÓN 10. DIAGRAMA DEL MARCO METODOLÓGICO DE LA INVESTIGACIÓN. ....	76
ILUSTRACIÓN 11. EVOLUCIÓN DE LOS TIPOS DE PUBLICACIONES POR AÑOS. ....	80
ILUSTRACIÓN 12. DISTRIBUCIÓN DE PUBLICACIONES POR TIPO DE DOCUMENTO. ....	80
ILUSTRACIÓN 13. ANÁLISIS DE COAUTORÍA DE AUTORES. ....	81
ILUSTRACIÓN 14. ANÁLISIS DE COAUTORÍA DE PAÍSES. ....	82
ILUSTRACIÓN 15. ANÁLISIS DE CITACIÓN POR DOCUMENTOS (MÁS DE 250 CITAS). ....	83
ILUSTRACIÓN 16. ANÁLISIS DE CITACIÓN POR FUENTES. ....	83
ILUSTRACIÓN 17. ANÁLISIS DE CO-CITACIÓN DE REFERENCIAS CITADAS. ....	85
ILUSTRACIÓN 18. ANÁLISIS DE ACOPLAMIENTO BIBLIOGRÁFICO POR AUTORES. ....	85
ILUSTRACIÓN 19. ANÁLISIS DE CO-OCURRENCIA DE PALABRAS CLAVE. ....	86
ILUSTRACIÓN 20. TEMAS DE INVESTIGACIÓN SOBRE SOSTENIBILIDAD Y TURISMO. ....	88
ILUSTRACIÓN 21. POPULARIDAD DE LOS TEMAS Y ÁREAS DE INVESTIGACIÓN SOBRE SOSTENIBILIDAD Y TURISMO. ....	89
ILUSTRACIÓN 22. EVOLUCIÓN DEL VOLUMEN DE PUBLICACIONES SOBRE SOSTENIBILIDAD Y TURISMO. ....	91
ILUSTRACIÓN 23. TENDENCIAS TEMÁTICAS EN LA LITERATURA SOBRE SOSTENIBILIDAD Y TURISMO (AUGE Y DECLIVE). ....	93
ILUSTRACIÓN 24. MARCO METODOLÓGICO DE LA INVESTIGACIÓN EMPÍRICA. ....	114
ILUSTRACIÓN 25. CATEGORIZACIÓN DE RSC Y SC. ....	116
ILUSTRACIÓN 26. TEMAS DE INVESTIGACIÓN ASOCIADOS A LA RSC Y LA SC EN TURISMO. ....	120
ILUSTRACIÓN 27. EVOLUCIÓN DEL VOLUMEN DE PUBLICACIONES CIENTÍFICAS SOBRE RSC Y SC EN TURISMO. ....	122
ILUSTRACIÓN 28. VOLUMEN DE PUBLICACIONES CIENTÍFICAS EN LA ÚLTIMA DÉCADA SOBRE RSC, SC Y PANDEMIA DE COVID-19. ....	123
ILUSTRACIÓN 29. ANÁLISIS DE REGRESIÓN LINEAL DE RSC Y SC ENTRE 2016 Y 2023. ....	123
ILUSTRACIÓN 30. ANÁLISIS DE REGRESIÓN LINEAL DE RSC Y SC ENTRE 2018 Y 2023. ....	124

ILUSTRACIÓN 31. TOP 5 - TÍTULOS DE LOS INFORMES CORPORATIVOS EN P1 (2010-2015). .....	125
ILUSTRACIÓN 32. TOP 5 - TÍTULOS DE LOS INFORMES CORPORATIVOS EN P2 (2016-2021). .....	126
ILUSTRACIÓN 33. PALABRAS MÁS FRECUENTES DE RSC EN P1 (2010-2015). .....	126
ILUSTRACIÓN 34. PALABRAS MÁS FRECUENTES DE RSC EN P2 (2016-2021). .....	127
ILUSTRACIÓN 35. PALABRAS MÁS FRECUENTES DE SC EN P1 (2010-2015). .....	127
ILUSTRACIÓN 36. PALABRAS MÁS FRECUENTES DE SC EN P2 (2016-2021). .....	128
ILUSTRACIÓN 37. MATRIZ DE SIMILITUD DE TEMAS PARA RSC Y SC EN P1 (2010-2015). .....	130
ILUSTRACIÓN 38. MATRIZ DE SIMILITUD DE TEMAS PARA RSC Y SC EN P2 (2016-2021). .....	130
ILUSTRACIÓN 39. TEMAS RELACIONADOS CON LA RSC EN P1 (2010-2015). .....	131
ILUSTRACIÓN 40. TEMAS RELACIONADOS CON LA RSC EN P2 (2016-2021). .....	131
ILUSTRACIÓN 41. TEMAS RELACIONADOS CON LA SC EN P1 (2010-2015). .....	132
ILUSTRACIÓN 42. TEMAS RELACIONADOS CON LA SC EN P2 (2016-2021). .....	132
ILUSTRACIÓN 43. DIAGRAMAS DE CAJA DE LAS FRECUENCIAS RELATIVAS DE LOS INFORMES CORPORATIVOS. ....	133
ILUSTRACIÓN 44. DIAGRAMAS DE CAJA DE LAS FRECUENCIAS RELATIVAS DE LOS INFORMES CORPORATIVOS. ....	134
ILUSTRACIÓN 45. DISTRIBUCIONES DE <i>BOOTSTRAPPING</i> . .....	134
ILUSTRACIÓN 46. DISTRIBUCIÓN DE LAS CATEGORÍAS ESG (MEDIA) DE LOS INFORMES CORPORATIVOS. .....	136
ILUSTRACIÓN 47. EVOLUCIÓN DE LAS CATEGORÍAS ESG A LO LARGO DEL TIEMPO. ....	136
ILUSTRACIÓN 48. PROMEDIO DE LAS CATEGORÍAS ESG POR SUBSECTORES TURÍSTICOS. ....	137
ILUSTRACIÓN 49. EVOLUCIÓN DE CADA CATEGORÍA ESG A LO LARGO DEL TIEMPO. ....	138
ILUSTRACIÓN 50. METODOLOGÍA DE INVESTIGACIÓN. ....	154
ILUSTRACIÓN 51. COEFICIENTES DEL MODELO M2. ....	167
ILUSTRACIÓN 52. COEFICIENTES DEL MODELO M2 PARA ESPAÑA, ZONAS RURALES Y URBANAS. ....	173

## **ÍNDICE DE ANEXOS**



## ÍNDICE DE ANEXOS

ANEXO 1. TEMAS IDENTIFICADOS EN LOS INFORMES DE LAS EMPRESAS TURÍSTICAS DE 2010 A 2015 (DEL 0 AL 24). .....	251
ANEXO 2. TEMAS IDENTIFICADOS EN LOS INFORMES DE LAS EMPRESAS TURÍSTICAS DE 2010 A 2015 (DEL 25 AL 49). .....	252
ANEXO 3. TEMAS IDENTIFICADOS EN LOS INFORMES DE LAS EMPRESAS TURÍSTICAS DE 2016 A 2021 (DEL 0 AL 24). .....	253
ANEXO 4. TEMAS IDENTIFICADOS EN LOS INFORMES DE LAS EMPRESAS TURÍSTICAS DE 2016 A 2021 (DEL 25 AL 49). .....	254
ANEXO 5. SOFTWARE, HERRAMIENTAS, METODOLOGÍAS Y LENGUAJES DE PROGRAMACIÓN UTILIZADOS EN LA TESIS DOCTORAL. ....	255
ANEXO 6. PRINCIPALES TÉCNICAS DE CIENCIA DE DATOS APLICADAS EN LA TESIS DOCTORAL. ....	256
ANEXO 7. MODELOS DE PROCESAMIENTO DE LENGUAJE NATURAL (PLN) APLICADOS EN LA TESIS DOCTORAL. ....	257
ANEXO 8. PRINCIPALES BIBLIOTECAS DE PYTHON EMPLEADAS EN EL DESARROLLO DE LA TESIS DOCTORAL. ....	258



## **PRESENTACIÓN DE LA TESIS DOCTORAL**



## PRESENTACIÓN DE LA TESIS DOCTORAL

Este capítulo introductorio tiene como objetivo ofrecer una visión general de la Tesis Doctoral. Para ello, se organiza en tres apartados principales. La primera sección presenta la justificación de la investigación, detallando las razones que motivaron la elección de los dos temas de investigación, respaldados por la literatura existente. Igualmente, se explora el contexto en el que se desarrolla el estudio empírico, enfocándose en el sector turístico de España. En la segunda sección, se plantean los objetivos de la investigación que se abordarán a lo largo de la Tesis Doctoral, junto con sus respectivas hipótesis o preguntas de investigación. Por último, en la tercera sección, se detalla el enfoque metodológico empleado y se describe la estructura completa de la Tesis Doctoral.

### 1. Justificación de la investigación

En las últimas décadas, el crecimiento económico y el aumento poblacional han generado un impacto negativo significativo sobre el planeta y sus recursos naturales (Serra et al., 2009). Esta situación ha desencadenado una creciente preocupación por el bienestar de las futuras generaciones, dado que el agotamiento de recursos tanto renovables como no renovables a corto plazo amenaza el desarrollo económico y la calidad de vida de las nuevas generaciones (Schilling & Chiang, 2011). En respuesta a esta problemática, ha surgido una necesidad ética de garantizar condiciones de vida óptimas a largo plazo, considerando los límites planetarios y la disponibilidad de los recursos naturales (Sudhir & Amartya, 2000).

El paradigma del desarrollo sostenible, establecido en el Informe Brundtland de 1987 (WCED, 1987), marcó un hito al posicionar la sostenibilidad como uno de los pilares fundamentales del desarrollo global. Este enfoque busca abordar de manera integral los desafíos ambientales, sociales y económicos (Chaminade, 2020; Fernandes et al., 2022). Desde entonces, el concepto de “desarrollo sostenible” se entiende como la satisfacción de las necesidades del presente sin comprometer la capacidad de satisfacción de futuras generaciones (WCED, 1987). Por consiguiente, la sostenibilidad se ha convertido en una estrategia esencial para el desarrollo futuro, impulsando nuevos modelos de crecimiento económico que tienen como objetivo disminuir los impactos negativos en un plazo temporal más amplio.

Con el tiempo, este paradigma se ha expandido y ha ido involucrando a diversos actores de la sociedad (gobiernos, empresas, individuos, etc.), generando nuevas iniciativas y conceptos que conducen al desarrollo sostenible buscado. Uno de estos es el paradigma del desarrollo turístico sostenible, que hace referencia a la actividad turística que contempla los impactos

actuales y futuros, satisfaciendo las necesidades de los turistas, la industria y el medio ambiente (S. W. Lee & Xue, 2020). Además, la sostenibilidad adquiere una relevancia particular en el turismo, debido a los significativos impactos tanto positivos como negativos que esta industria genera en el planeta y la sociedad. Por un lado, el turismo genera empleo y contribuye a mejorar la calidad de vida; por otro lado, implica un elevado consumo de recursos naturales, una considerable producción de residuos y altos niveles de contaminación (Henderson, 2007).

La investigación turística en sostenibilidad ha evolucionado desde el desarrollo de principios teóricos hacia la implementación de prácticas que involucren activamente a los distintos *stakeholders* (Bramwell et al., 2017; Ruhanen et al., 2015; Sharpley, 2020). Un enfoque holístico del turismo sostenible identifica cuatro grupos de interés fundamentales: residentes, turistas, empresas y gobierno (McIntosh et al., 1995; Salee et al., 2022). Estos grupos de interés, con sus diversas perspectivas, necesidades e influencias, conforman un ecosistema integral y dinámico que resulta crucial para la transición hacia un modelo turístico más responsable y equilibrado. El análisis de estos actores clave y sus interacciones puede proporcionar una comprensión más profunda y detallada de los desafíos y oportunidades en la implementación de prácticas sostenibles en el sector turístico. Asimismo, investigaciones recientes como la de Alonso-Muñoz et al. (2023), señalan que la investigación sobre sostenibilidad y turismo continúa siendo fragmentada, puesto que la mayoría de los estudios se orientan a subsectores específicos, destinos particulares o problemas concretos de sostenibilidad. Por lo tanto, la literatura actual no ofrece una visión integral y actualizada del campo.

La presente Tesis Doctoral aborda la sostenibilidad en turismo desde tres enfoques: el desarrollo teórico, el análisis de empresas y la percepción de los turistas. Para ello, se lleva a cabo un análisis detallado de la literatura y una investigación empírica en el sector turístico de España, focalizándose específicamente en dos actores: las empresas y los turistas. La elección de estos grupos responde a su influencia directa en la oferta y demanda turística, así como en la implementación de prácticas sostenibles. De acuerdo con la investigación de Font et al. (2021), el valor en sostenibilidad se co-crea a través de la interacción directa entre quienes proporcionan el servicio (las empresas) y quienes lo consumen (los turistas). Es decir, las empresas hacen propuestas de valor sostenibles, pero es a través del uso y la experiencia de los turistas como se define el verdadero valor de esas prácticas. Esta interacción es fundamental para comprender cómo se genera el valor compartido de la sostenibilidad en el contexto turístico.

Las empresas y los turistas son actores clave que tienen un impacto directo y significativo en la sostenibilidad del sector (Font & McCabe, 2017). Font et al. (2016) afirman que las empresas turísticas tienen la capacidad de fomentar en sus clientes comportamientos más sostenibles y respetuosos con el medioambiente. Esto se puede lograr mediante una adecuada comunicación de la Responsabilidad Social Corporativa (RSC) y una eficaz estrategia de marketing verde (por ejemplo, compartiendo conocimiento ambiental que refuerce los valores ecológicos de los clientes) (Carvalho & Alves, 2023; Tölkes, 2018). Por otro lado, los comportamientos irrespetuosos de los turistas afectan la sostenibilidad de los destinos, amenazando su viabilidad a largo plazo y disminuyendo la calidad de la experiencia turística (Im & Kim, 2023). Por tanto, es importante atraer turistas respetuosos con el medio ambiente, puesto que sus comportamientos y decisiones de consumo minimizan la huella ecológica durante sus viajes, fomentando el turismo sostenible (Dolnicar et al., 2008). De acuerdo con Martínez García de Leaniz et al. (2018), los clientes con mayor conciencia ambiental tienen una intención más fuerte de recomendar la empresa y pagar un precio más elevado. El análisis separado pero complementario de estos dos *stakeholders* puede ofrecer una perspectiva sobre la brecha entre las prácticas de sostenibilidad de las empresas y las percepciones de los turistas sobre sostenibilidad.

De igual modo, el avance científico está impulsado cada vez más por capacidades avanzadas de computación que permiten a los investigadores transformar y analizar grandes volúmenes de datos (Hey et al., 2009). Pennington et al. (2020) enfatizan en la necesidad urgente de formar a futuros científicos con conocimientos tanto en sus campos específicos como en ciencia de datos. Los autores puntualizan que ninguna disciplina por sí sola puede abordar todos los desafíos actuales asociados con la grave situación de insostenibilidad del planeta. En consecuencia, se requiere un enfoque integrador, donde el conocimiento de diferentes disciplinas se combine para encontrar soluciones más sostenibles y efectivas. No obstante, Boyd & Crawford (2012) destacan que, aunque los enfoques de ciencia de datos y *big data* ofrecen nuevas oportunidades para el análisis y la investigación, su uso viene acompañado de importantes desafíos éticos, sociales y metodológicos que deben ser abordados críticamente.

La ciencia de datos ha emergido como una herramienta clave para el análisis y comprensión de grandes volúmenes de información en el sector turístico, especialmente en lo relacionado con la sostenibilidad. Sin embargo, pese a los evidentes beneficios que ofrece el análisis de *big data*, los académicos del ámbito turístico aún no han integrado por completo estas herramientas

y metodologías en sus investigaciones (Köseoglu et al., 2020; van der Zee et al., 2020). Existen múltiples fuentes de datos relevantes para el análisis de *big data* en turismo, como los datos provenientes de dispositivos, transacciones y contenido generado por los usuarios (*User Generated Content*, *UGC*, en sus siglas en inglés) (Li et al., 2018). Cada una de las anteriores ha proporcionado perspectivas valiosas sobre diferentes aspectos del comportamiento y las decisiones en el turismo sostenible.

Esta investigación analiza y procesa una gran cantidad de datos, tanto en el análisis teórico como en el empírico. Por un lado, se cuenta con un gran número de publicaciones científicas sobre el campo de estudio, dado que la investigación en sostenibilidad turística en las últimas décadas ha experimentado un crecimiento constante e incluye una amplia gama de temas (Bramwell et al., 2017; Demeter et al., 2023; Mihalic, 2020; Prerana et al., 2024). Por otro lado, la información generada por las empresas a través de informes de sostenibilidad y RSC y UGC son fundamentales para comprender cómo se implementan y perciben las estrategias de sostenibilidad dentro del sector turístico de España. Las empresas, a lo largo del tiempo, han incrementado la divulgación de informes corporativos y memorias de sostenibilidad. Esto fortalece sus relaciones con *stakeholders* y refuerza su legitimidad y transparencia (Crane & Glozer, 2016; Ning et al., 2021; Schuler & Cording, 2006). El UGC ofrece una visión directa del grado de conciencia y percepción de los consumidores sobre las prácticas sostenibles de las empresas turísticas (D'Acunto et al., 2020a). Asimismo, la literatura en turismo señala un desarrollo emergente de estudios que aplican técnicas de *big data* al UGC para identificar y extraer información útil sobre temas específicos, como puede ser la sostenibilidad, la privacidad o el uso de robots de servicios (Cai et al., 2024).

Por tanto, en esta Tesis Doctoral se adopta un enfoque multidisciplinario innovador que integra la sostenibilidad y la ciencia de datos en el sector turístico, dos campos de estudio en constante evolución que demandan investigación integrada y análisis robustos. Con una perspectiva particular en empresas y turistas, la investigación enriquece la literatura existente y proporciona información valiosa para los actores clave en el desarrollo del turismo sostenible. Las empresas juegan un papel crucial en el desarrollo sostenible turístico, ya que sus actividades tienen un impacto significativo en los individuos, la sociedad y el planeta.

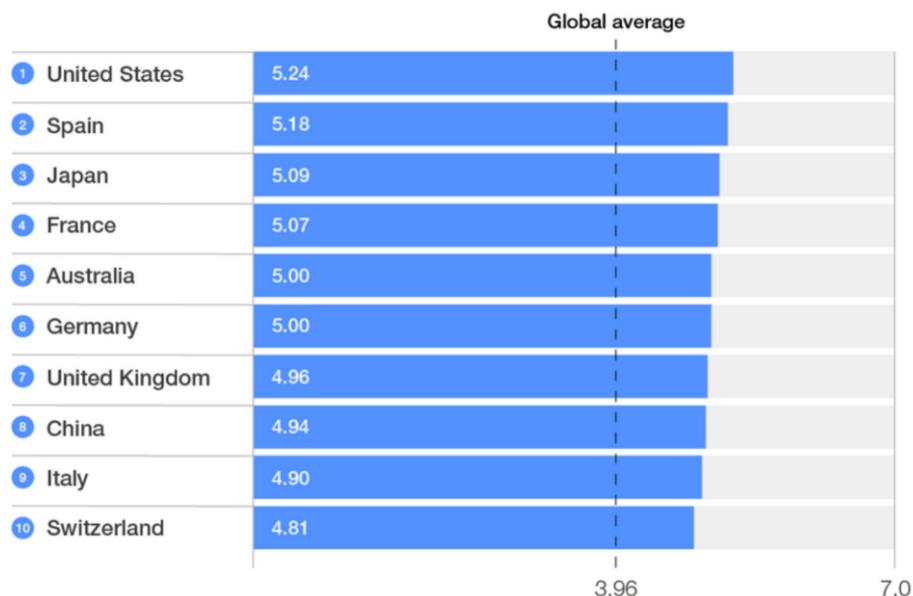
En este contexto, las compañías se enfrentan al desafío de integrar la sostenibilidad en su gestión empresarial, impulsadas por una creciente regulación, la presión del mercado y la intensa competitividad. Para mantenerse a la vanguardia, las empresas deben adaptarse

rápidamente para cumplir y superar las expectativas de consumidores cada vez más conscientes del medio ambiente y socialmente responsables. Por su parte, los turistas también asumen un papel activo, reconociendo su responsabilidad hacia los destinos que visitan y contribuyendo a la sostenibilidad a través de sus decisiones de compra, actitudes y comportamientos.

En resumen, este estudio permite comprender mejor cómo las empresas implementan estrategias sostenibles y cómo los turistas perciben la sostenibilidad, promoviendo así un turismo más responsable y duradero. Mediante la aplicación de técnicas avanzadas de ciencia de datos, esta Tesis Doctoral pretende mejorar la capacidad de ambos actores para adaptarse a las crecientes demandas de la sostenibilidad, contribuyendo al desarrollo de un sector turístico más sostenible y consciente. Igualmente, proporciona una visión actualizada, completa y detallada de la red de conocimiento sobre sostenibilidad y turismo, identificando temáticas, áreas y tendencias en la investigación científica.

### **1.1. Ámbito de la investigación**

La industria turística es un motor fundamental para el desarrollo social y económico a nivel global (Henderson, 2007). El estudio empírico de esta Tesis Doctoral se centra en el destino turístico de España, dado que es un país reconocido como una de las principales potencias turísticas a nivel mundial. El Índice de Desarrollo de Viajes y Turismo 2024 del Foro Económico Mundial (World Economic Forum, 2024) posiciona a España en el segundo lugar a escala global, solo por detrás de Estados Unidos, reflejando la competitividad de su sector turístico (Ilustración 1). Con una puntuación media de 5,18 puntos, y una mejora del 0,9% desde 2019, España demuestra su capacidad de adaptación y crecimiento incluso frente a desafíos globales como la pandemia de covid-19. El índice destaca a España como la mejor en la región del Sur de Europa gracias a su infraestructura avanzada, así como a sus recursos culturales y naturales.

**Ilustración 1. Top 10 de economías líderes en fomentar el desarrollo de viajes y turismo.**

**Fuente:** (World Economic Forum, 2024).

De acuerdo con el informe “Viajeros y destinos NextGen: nuestra visión sobre la transformación del sector turístico” (Google & Deloitte, 2024), las perspectivas para España como destino turístico son positivas. Se espera que para 2040 España continúe siendo un destino líder, con una proyección de 110 millones de viajeros internacionales, consolidando su lugar en el top 5 de destinos turísticos a nivel global. Además, España forma parte del clúster turístico del Mediterráneo, una de las cuatro áreas clave que atraerán a casi el 45% de los viajeros internacionales para 2040. Este clúster incluye a otros países como Italia, Francia y Turquía, y se caracteriza por su clima templado, riqueza cultural y patrimonio histórico.

El turismo es, por lo tanto, un pilar fundamental de la economía española. En 2023, la actividad turística alcanzó los 186.596 millones de euros, representando el 12,8% del PIB (Producto Interno Bruto) español, el mayor valor registrado en la serie histórica. De igual modo, el turismo creó 95.224 empleos más que en 2022, lo que representó el 17% del total de empleos generados en todo el país. Además, se prevé que el PIB turístico supere los 200.000 millones de euros por primera vez en la historia, alcanzando el 13,4% del PIB español e impulsando el 41,4% del crecimiento económico de España en 2024 (Exceltur, 2024).

Otro aspecto por destacar es la posición de España como destino turístico en términos de sostenibilidad. El informe “Expectativas y comportamiento de los viajeros europeos en materia de sostenibilidad”, elaborado por Mastercard y Turespaña, puntualiza que un 38% de los viajeros que se consideran sostenibles o muy sostenibles posicionan a España favorablemente en

comparación con otros destinos. El 75% de los turistas están satisfechos con la sostenibilidad del país, y el 97% quedó satisfecho o muy satisfecho con su experiencia de viaje en general. Por lo tanto, España es reconocido como uno de los destinos turísticos más atractivos para los viajeros comprometidos con la sostenibilidad (Mastercard & Turespaña, 2023).

En definitiva, el análisis de la sostenibilidad de un destino turístico se convierte en un tema fundamental para asegurar el crecimiento y la competitividad a largo plazo (Mazanec et al., 2007). Específicamente en España, y ante las proyecciones de crecimiento turístico en aumento, surge la necesidad de abordar de manera equilibrada los beneficios económicos del sector junto con sus impactos ambientales y sociales. Baggio (2020) manifiesta que la gestión de un destino orientada hacia la sostenibilidad requiere una estrategia integral que considere los desafíos del entorno cambiante y complejo, las tensiones de la competencia interna, las estimaciones de la demanda y los avances tecnológicos.

De esta forma, resulta inevitable considerar el impacto del cambio climático en la demanda turística. Zhou et al. (2024) basándose en estimaciones de elasticidad climática y cambios de temperatura proyectados para diferentes zonas climáticas y países, proyectan un cambio global en la demanda turística para 2050. Sus resultados indican que las regiones situadas en latitudes medias y altas (países como Canadá, Rusia, Finlandia, Noruega y Suecia) se beneficiarán del cambio climático. Por el contrario, las regiones tropicales, áridas y mediterráneas sufrirán una disminución en el número de turistas (como el Sur de Europa, partes de África o las islas del Caribe). Esto implicaría que España, en particular sus áreas más afectadas por el calor, podría experimentar un descenso en la cantidad de turistas en comparación con otras regiones más frías de Europa. Por lo tanto, España deberá adaptarse rápidamente a esta redistribución de la demanda turística, gestionando de manera integral sus destinos costeros, urbanos y rurales para asegurar su competitividad y seguir siendo un referente en el turismo global.

## **2. Objetivos del estudio**

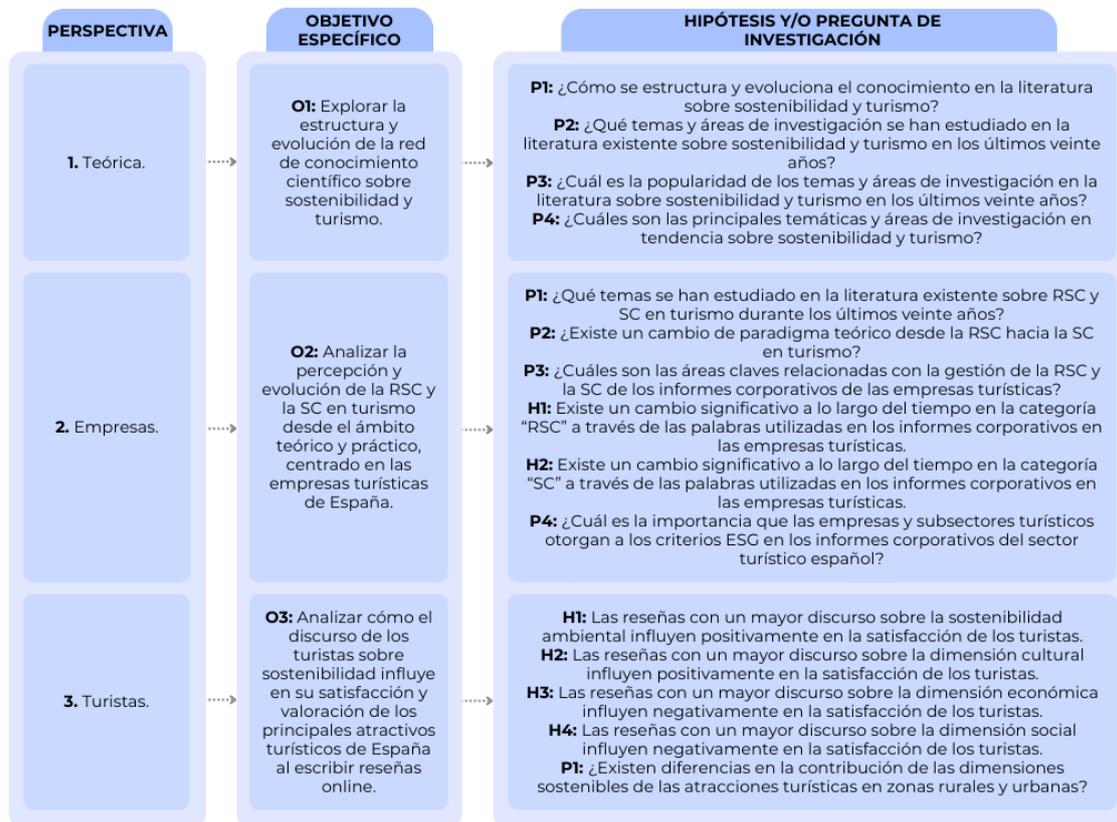
El objetivo principal de esta Tesis Doctoral es analizar la sostenibilidad turística mediante la aplicación de la ciencia de datos, desde tres perspectivas: teórica, empresarial y de turistas. Para ello, se lleva a cabo un análisis teórico y una investigación empírica, centrada en las empresas y turistas de España. A través de la aplicación de técnicas de ciencia de datos y el análisis de grandes volúmenes de datos (*big data*), se logra una mayor comprensión del paradigma del desarrollo turístico sostenible, permitiendo identificar patrones y tendencias que serían difíciles

de detectar mediante métodos tradicionales. En la Ilustración 2 se detallan los objetivos específicos, junto con sus respectivas hipótesis y/o preguntas de investigación propuestas.

- El primer objetivo (O1) explora la estructura y evolución de la red de conocimiento científico sobre sostenibilidad y turismo. Se formulan preguntas de investigación que analizan los autores y revistas más influyentes, la popularidad de áreas y temáticas de investigación, las tendencias en auge y declive. Los hallazgos pueden servir como guía para futuras investigaciones y apoyar la toma de decisiones informadas tanto en el ámbito empresarial como gubernamental, puesto que se sintetiza la literatura previa y se reflejan las áreas que están ganando mayor o menor atención en la comunidad científica.
- El segundo objetivo (O2) se centra en analizar la percepción teórica y práctica de la Responsabilidad Social Corporativa (RSC) y la Sostenibilidad Corporativa (SC). Para el análisis empírico se emplea la técnica de minería de datos para analizar los informes corporativos de las empresas turísticas de España. Las preguntas e hipótesis de investigación pretenden establecer estadísticamente la evolución desde la RSC hacia la SC a nivel teórico y práctico. Además, se aplica el modelado de temas para detectar los temas de investigación y las áreas clave que guían la gestión sostenible y responsable en las empresas turísticas de España. Por último, se cuantifica la importancia que estas empresas otorgan a los criterios ESG (*Environmental, Social and Governance*, en sus siglas en inglés).
- El tercer objetivo (O3) analiza cómo el discurso de los turistas sobre sostenibilidad influye en su satisfacción y valoración de los atractivos turísticos españoles. Las hipótesis exploran cómo las reseñas de turistas, enfocadas en distintas dimensiones de la sostenibilidad (ambiental, cultural, económica y social), impactan de manera positiva o negativa en su satisfacción, y si existen diferencias en la valoración de los destinos rurales y urbanos de España.

Esta estructura proporciona un enfoque integral para estudiar la sostenibilidad en el turismo, abarcando tanto las percepciones teóricas, como las de empresas y turistas en el destino de España.

**Ilustración 2. Objetivos específicos, hipótesis y/o preguntas de investigación.**



Fuente: Elaboración propia.

### 3. Metodología y estructura de la Tesis Doctoral

La ciencia de datos se emplea en esta investigación como herramienta metodológica para llevar a cabo los análisis. Este enfoque permite procesar y analizar grandes volúmenes de datos estructurados y no estructurados, provenientes de diversas fuentes, tales como bases de datos científicas, informes corporativos, y comentarios online de turistas. Igualmente, se aplican análisis descriptivos y predictivos, utilizando técnicas como la minería de datos y algoritmos de aprendizaje automático (*machine learning*), que permiten identificar patrones y relaciones subyacentes no evidentes a través de métodos convencionales. Estas técnicas se complementan con análisis estadísticos que sirven para validar las hipótesis planteadas y proporcionar respuestas a las preguntas de investigación.

El lenguaje de programación usado para el manejo, preprocesamiento y análisis de los datos es Python, aprovechando su amplia gama de bibliotecas para procesamiento de datos, aprendizaje automático y análisis de texto. También se emplean herramientas de *web scraping* para extraer información relevante de plataformas en línea y reseñas de turistas de forma automatizada y ordenada. Además, se utiliza el software estadístico Stata y la herramienta de software VOSviewer, lo que garantiza un enfoque riguroso y complementario en el tratamiento de los datos. La Tabla 1 detalla los anexos relacionados con los aspectos técnicos, metodologías y enfoques de ciencia de datos empleados en el desarrollo de esta investigación.

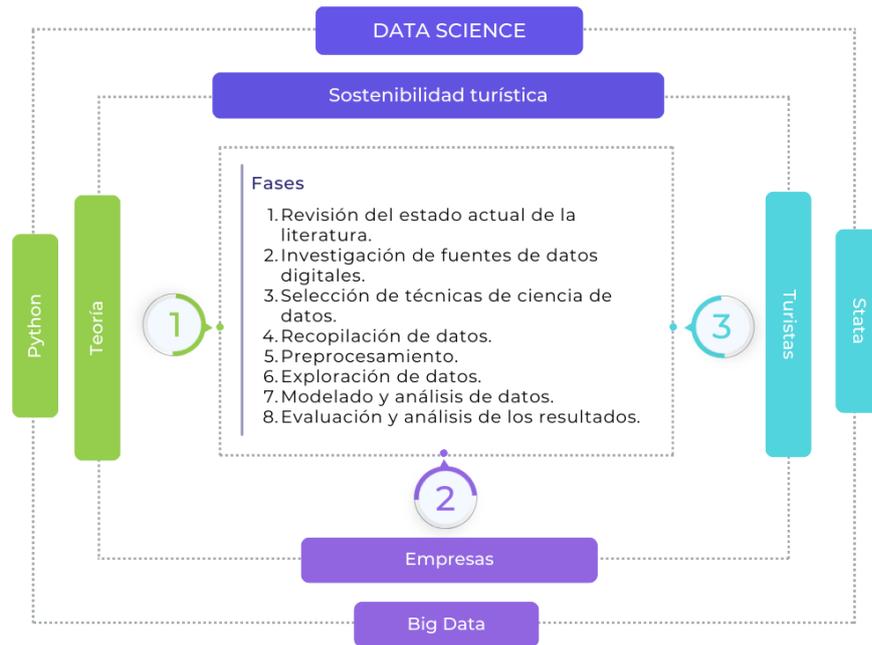
**Tabla 1. Consolidado de anexos asociados a técnicas y herramientas metodológicas aplicadas en la Tesis Doctoral.**

No.	Descripción
<b>Anexo 5</b>	Software, herramientas, metodologías y lenguajes de programación utilizados en la Tesis Doctoral.
<b>Anexo 6</b>	Principales técnicas de ciencia de datos aplicadas en la Tesis Doctoral.
<b>Anexo 7</b>	Modelos de Procesamiento de Lenguaje Natural (PLN) aplicados en la Tesis Doctoral.
<b>Anexo 8</b>	Principales bibliotecas de Python empleadas en el desarrollo de la Tesis Doctoral.

**Fuente: Elaboración propia.**

Por otro lado, en la Ilustración 3 se muestra el marco metodológico general de la Tesis Doctoral, el cual sirve como referencia integral para todo el proceso de investigación. Cada objetivo específico propuesto se aborda con su propio marco metodológico particular, adaptado a las necesidades y características de cada análisis, garantizando así un enfoque personalizado y riguroso para cada una de las perspectivas del estudio. Se inicia con una revisión del estado actual de la literatura, estableciendo así una base teórica sólida. Posteriormente, se procede a la investigación de fuentes de datos digitales y la selección de técnicas apropiadas de ciencia de datos. Las fases siguientes se centran en la recopilación de datos, el tratamiento y análisis de estos datos. Por último, se realiza una evaluación e interpretación de los resultados obtenidos.

Ilustración 3. Marco metodológico general de la Tesis Doctoral.



Fuente: Elaboración propia.

La Tesis Doctoral está organizada en cinco capítulos (ver Ilustración 4). El primer capítulo establece el marco conceptual fundamental de esta investigación, ofreciendo una exploración de los conceptos centrales que sustentan la Tesis Doctoral. Se presentan las nociones esenciales como el turismo sostenible, la sostenibilidad turística, y las dimensiones sostenibles, entre otros. También se introduce la ciencia de datos como herramienta metodológica innovadora en la investigación turística, así como los tipos de análisis y datos que existen. El segundo capítulo se centra en el análisis del conocimiento científico existente sobre sostenibilidad turística, mediante el análisis bibliométrico y la metodología *Computational Literature Review* (CLR) basada en el modelado de temas. Este término podría traducirse al español como “Revisión computacional de la literatura”, aunque su uso en este idioma es aún menos común. El propósito de este capítulo es ofrecer una base sólida que respalde el desarrollo conceptual y práctico de las siguientes dos líneas de investigación.

El tercer capítulo aborda la sostenibilidad desde la perspectiva de las empresas del sector turístico en España. Este capítulo está dedicado al análisis de las prácticas de Responsabilidad Social Corporativa (RSC) y Sostenibilidad Corporativa (SC) a partir de los informes corporativos. Se utilizan técnicas de minería de datos y análisis textual para identificar patrones en los informes corporativos, además de evaluar la evolución de estas prácticas a lo largo del tiempo. El cuarto capítulo estudia la sostenibilidad desde el punto de vista de los turistas, evaluando sus

perspectivas con la sostenibilidad. Aquí se explora cómo las percepciones de los turistas en relación con la sostenibilidad (económica, social, ambiental y cultural) influyen en su satisfacción y valoración de los atractivos turísticos de España. Se emplean técnicas de procesamiento del lenguaje natural (PLN) para analizar las reseñas online de turistas y se aplican modelos predictivos para explorar la influencia de diferentes dimensiones de la sostenibilidad en la experiencia del turista. Finalmente, el quinto capítulo presenta las conclusiones generales de la investigación, resumiendo los hallazgos más relevantes en cada una de las tres perspectivas: teórica, empresarial y de turistas. Asimismo, se ofrecen recomendaciones prácticas para mejorar la sostenibilidad en el sector turístico y se destacan futuras líneas de investigación.

**Ilustración 4. Estructura de la Tesis Doctoral.**



**Fuente: Elaboración propia.**

## **CAPÍTULO I. MARCO CONCEPTUAL GENERAL**



## CAPÍTULO I. MARCO CONCEPTUAL GENERAL

En este capítulo se exploran los dos pilares fundamentales de la Tesis Doctoral: la sostenibilidad y la ciencia de datos, proporcionando una base teórica para entender cómo se relacionan e integran en el ámbito del turismo. Para ello, el presente capítulo se ha estructurado en tres apartados y subapartados. En primer lugar, se brinda una introducción a la sostenibilidad y su relación con el turismo, abordando conceptos clave como el desarrollo sostenible, la sostenibilidad turística, sus dimensiones y los actores fundamentales implicados en el sector. En segundo lugar, se presentan las principales nociones de la ciencia de datos, su diferencia con la inteligencia artificial, los tipos de análisis y las fases del proceso. Además, se exploran las aplicaciones del *big data* y la ciencia de datos en el sector turístico. Finalmente, se exponen las conclusiones clave que servirán de orientación para los estudios teóricos y empíricos de la investigación.

### 1. Sostenibilidad y turismo

#### 1.1. Diferencia entre el concepto de desarrollo sostenible y sostenibilidad

La creciente preocupación por los límites planetarios y la escasez de los recursos naturales (Sudhir & Amartya, 2000) ha impulsado a lo largo de las últimas décadas, una evolución y expansión de los términos “sostenibilidad” y “desarrollo sostenible”. Ambos términos reflejan un cambio de paradigma que sustituye el enfoque sociopolítico del desarrollo económico neoliberal por uno sostenible (Mihalic, 2024). La transición de un modelo de crecimiento ilimitado a uno de desarrollo sostenible ha generado cambios tanto en políticas y prácticas empresariales, como nuevas iniciativas y conceptos en busca de un equilibrio entre aspectos económicos, ambientales y sociales.

La definición de desarrollo sostenible acuñada por la Comisión Mundial sobre el Medio Ambiente y Desarrollo de las Naciones Unidas en 1987 en el “Informe Brundtland” o “Nuestro Futuro Común” es ampliamente reconocida y con gran aceptación internacional. La Comisión lo define como aquel que satisface las necesidades del presente sin comprometer la capacidad de satisfacción de las futuras generaciones (WCED, 1987). Sin embargo, la ambigüedad inherente de la definición permite que una amplia gama de intereses la respalde, pero también facilita que su aplicación práctica sea evitada o distorsionada, generando desafíos en su implementación efectiva (Butler, 2018). Diesendorf (2000) también critica parcialmente la definición de la Comisión y propone otra en su lugar: “El desarrollo sostenible comprende tipos de desarrollo económico y social que protegen y mejoran el entorno natural y la equidad social” (p. 3). En esta

definición, el autor incluye explícitamente como aspectos primarios, tanto la protección y mejora del entorno natural, como la equidad social.

En la literatura y en diversos sectores de la sociedad, es común observar una tendencia generalizada del uso indiscriminado de los conceptos “sostenibilidad” y “desarrollo sostenible”. Sin embargo, algunos autores afirman que la confusión entre los términos puede derivarse en una falta de claridad en el discurso, así como la formulación e implementación de políticas o prácticas ineficaces (Diesendorf, 2000; Hector et al., 2014; Ruggerio, 2021). Hector et al. (2014) exploran los conceptos desde sus fundamentos filosóficos, estableciendo dos posiciones fundamentales. La primera posición de preservación ambiental, asociada con la sostenibilidad, valora la naturaleza por su propio derecho, defendiendo la protección de los ecosistemas y la biodiversidad para reconocer la interdependencia de todas las especies. La segunda posición de conservación prudencial, referente al desarrollo sostenible, tiende a valorar la naturaleza principalmente en términos de su utilidad para los seres humanos, destacando la utilidad de los recursos naturales y considerando la relación con la naturaleza importante solo si beneficia a los humanos.

Esto sugiere que la sostenibilidad se centra en un estado ideal de equilibrio entre humanos y naturaleza, integrando componentes sociales, económicos, políticos, culturales y ecológicos, y sus interrelaciones (Diesendorf, 2000; Hector et al., 2014; Ruggerio, 2021). Mientras que el desarrollo sostenible se refiere a los procesos y estrategias que se implementan para alcanzar ese estado u objetivo de la sostenibilidad, a menudo desde una perspectiva más centrada en los intereses humanos (Diesendorf, 2000; Hector et al., 2014).

## **1.2. El concepto de turismo sostenible**

En el ámbito del turismo, la aplicación de los principios del desarrollo sostenible dio origen al paradigma del desarrollo turístico sostenible o turismo sostenible (Butler, 2018). Sin embargo, en la literatura turística previa al uso oficial del término desarrollo sostenible introducida por el “Informe Brundtland” ya se reflejaba una preocupación por los principios que lo definen. Investigaciones como las de Rosenow & Pulsipher (1979) y los modelos teóricos de Butler (1980), Stankey (1973) y Tivy (1973) revelan que el concepto de turismo sostenible estaba emergiendo, enfocándose en la preservación, la capacidad de carga y la educación de los turistas. En la Tabla 2 se presentan algunas nociones de la definición de turismo sostenible.

Tabla 2. Principales nociones del concepto de turismo sostenible.

Autor	Nociones del concepto de turismo sostenible
Eber (1992)	Turismo e infraestructuras asociadas que tanto en el presente como en el futuro operan respetando la capacidad natural de regeneración y mantienen la productividad futura de los recursos naturales; reconocen y valoran la aportación de las personas, comunidades, costumbres y estilos de vida locales a la experiencia turística; aseguran que estas comunidades locales participen equitativamente en los beneficios económicos generados (p. 3).
World Tourism Organization (1993)	Actividad turística que satisface las necesidades de los turistas actuales y las regiones anfitrionas mientras se protege y mejoran las oportunidades para el futuro (p. 7).
Butler et al. (1993)	El turismo que se desarrolla y mantiene en un área específica (comunidad o un entorno natural), de una manera y a una escala que garantizan su viabilidad, durante un período indefinido y no degrada ni altera el medio ambiente (humano y físico). Al hacerlo, se asegura que el turismo no obstaculice el desarrollo exitoso y el bienestar de otras actividades y procesos que también dependen de ese entorno (p. 99).
Payne (1993)	El turismo sostenible debe ser capaz de aumentar las opciones económicas disponibles para las personas, sin afectar de forma negativa la estructura de la actividad económica existente. No debe interferir con las formas actuales de organización social y debe respetar los límites establecidos por las comunidades ecológicas (p. 154-155).
Woodley (1993)	El turismo sostenible en los parques (y otras zonas) debe definirse principalmente en términos de ecosistemas sostenibles (p. 94).
Countryside Commission (1995)	Turismo que es capaz de sostener las economías locales sin perjudicar el medio ambiente del que depende (p. 2).
Bramwell et al. (1996)	Aquel que se desarrolla rápidamente, contemplando la capacidad actual de alojamiento, la población local y el medio ambiente. Es la actividad turística que respeta el medio ambiente y no contribuye a

Autor	Nociones del concepto de turismo sostenible
	su propia desaparición, especialmente en zonas saturadas. Además, el turismo sostenible es un turismo responsable (p. 10-11).
Hunter (1997)	El turismo sostenible debe ser entendido como un paradigma adaptativo con un enfoque flexible que puede ajustarse a las circunstancias específicas de cada destino, las necesidades de las comunidades locales y debe integrar las preocupaciones ambientales, sociales y económicas de manera holística.
Butler (1999)	El turismo sostenible requiere un equilibrio entre los tres pilares fundamentales de la sostenibilidad: ambiental, social y económico.
Bramwell et al. (2017)	Se considera una orientación normativa que busca redirigir los sistemas sociales y comportamientos hacia un desarrollo sostenible integral, incluyendo aspectos sociales, culturales, económicos, políticos y ambientales.
World Tourism Organization (2019)	Turismo que tiene en cuenta las consecuencias actuales y futuras, económicas, sociales y medioambientales para satisfacer las necesidades de visitantes, la industria, el entorno y las comunidades anfitrionas (p. 72).

**Fuente: Elaboración propia basada en los autores.**

Como se observa en la Tabla 2, el concepto de turismo sostenible ha evolucionado a lo largo del tiempo para adoptar un enfoque más holístico y equilibrado entre los aspectos sociales, ambientales y económicos. Inicialmente, el concepto comenzó enfocándose en el respeto a la capacidad natural del entorno y la participación de las comunidades locales (Eber, 1992). La Organización Mundial del Turismo (1993) introdujo la idea de satisfacer las necesidades de los turistas y las regiones anfitrionas, mientras que autores como Butler et al. (1993) añadieron la importancia de la viabilidad a largo plazo y el equilibrio con otras actividades económicas. Payne (1993) profundizó en la dimensión económica, centrándose en la necesidad de aumentar las oportunidades para las personas, sin perjudicar la estructura existente. Sin embargo, Woodley (1993) extendió y aplicó la definición a áreas específicas como parques naturales. La *Countryside Commission* (1995) y Bramwell et al. (1996) aportaron perspectivas adicionales, destacando la importancia de sostener las economías locales y la necesidad de un desarrollo rápido pero responsable, especialmente en zonas saturadas.

A finales de los 90, autores como Hunter (1997) y Butler (1999) argumentaron que el turismo sostenible requería un equilibrio entre los tres pilares de la sostenibilidad (ambiental, social y económico). Hunter (1997) abogó por un enfoque adaptativo y flexible, en el que el turismo sostenible se ajusta tanto al contexto y circunstancias de cada destino, como a las necesidades de las comunidades locales. Bramwell et al. (2017) señala que el turismo sostenible se puede considerar como una orientación normativa hacia un desarrollo sostenible integral, contemplando aspectos sociales, culturales, económicos, políticos y ambientales. La definición más reciente de la Organización Mundial del Turismo (2019) consolida la evolución conceptual del turismo sostenible, integrando las dimensiones económicas, sociales y medioambientales, y extendiendo su alcance temporal para considerar las consecuencias actuales y futuras. Esta definición contempla varios actores de la actividad turística, buscando satisfacer las necesidades de turistas, la industria, el medio ambiente y las comunidades locales.

Esta evolución refleja un concepto cada vez más completo y multidimensional del turismo sostenible, resaltando la creciente preocupación por los impactos económicos, sociales y ambientales. Sin embargo, diversos académicos han debatido ampliamente las limitaciones de la conceptualización del término (Butler, 1999; Espiner et al., 2017; C. Hunter, 1997; Miller & Torres-Delgado, 2023; Weaver, 2007; Wheeler, 2007; S. Zhang & Chan, 2020) señalando la necesidad de una definición más precisa, así como la falta de claridad en los impactos locales versus globales, la duración de la sostenibilidad, los beneficiarios específicos, los objetivos de sostenibilidad, las medidas e indicadores concretos para evaluar su desarrollo. Esto sugiere que el concepto de turismo sostenible es dinámico y se encuentra en constante evolución, reflejando la naturaleza compleja y cambiante de los desafíos globales que enfrenta la actividad turística. Por ende, es necesario una revisión y actualización continua de los conocimientos teóricos y las aplicaciones prácticas del turismo sostenible. De acuerdo con Miller & Twining-Ward (2005), una comprensión conceptual sólida y un marco teórico conciso pueden promover cambios coherentes, reduciendo la ambigüedad tanto en debates académicos como en la formulación de políticas y prácticas empresariales.

Tras explorar el concepto de turismo sostenible como una aplicación práctica del desarrollo sostenible, es importante analizar los enfoques e interpretaciones tanto del turismo sostenible, como de la sostenibilidad en turismo y destinos, dado que en la literatura es común que los términos sostenibilidad y turismo sostenible se utilicen indistintamente, como si fueran

sinónimos. Ambos términos han ganado un protagonismo en las discusiones académicas y estrategias de desarrollo turístico a nivel mundial (Serrano et al., 2019).

### **1.3. La sostenibilidad en el contexto turístico**

En el sector del turismo, la sostenibilidad adquiere una relevancia significativa, puesto que se derivan impactos positivos y negativos en el planeta y la sociedad. Por un lado, se crea empleo y se incrementa la calidad de vida de las personas, mientras que por otro, se genera un alto uso de recursos naturales, una gran acumulación de residuos y elevados niveles de contaminación (Henderson, 2007). Estudios previos (Bramwell et al., 2017; Ruhanen et al., 2015; Sharpley, 2020) evidencian una transición en la literatura turística sobre sostenibilidad. Inicialmente, los académicos se dedicaban principalmente a desarrollar y comprender los principios teóricos. Sin embargo, el enfoque actual destaca la importancia de aplicar esos principios teóricos a la práctica, con la participación de los *stakeholders*. Es decir, la literatura en turismo sobre sostenibilidad ha evolucionado desde una etapa de desarrollo teórico hacia una fase donde se busca aplicar esos conocimientos de manera práctica, involucrando a diversas partes interesadas para asegurar un enfoque integral y colaborativo.

Como se discutió previamente, a pesar de la evolución del turismo sostenible a lo largo del tiempo, este concepto continúa presentando limitaciones que fomentan la discusión académica y diversas interpretaciones. De acuerdo con Coccossis (1996) el turismo en el contexto del desarrollo sostenible puede interpretarse de cuatro formas: 1) Perspectiva sectorial, enfocándose en la sostenibilidad económica del turismo; 2) Perspectiva ecológica, que resalta la necesidad de un turismo respetuoso con el medio ambiente; 3) Perspectiva de viabilidad a largo plazo, reconociendo la importancia de la competitividad de los destinos turísticos; 4) Perspectiva estratégica, que integra el turismo como parte de una estrategia más amplia de desarrollo sostenible, abarcando tanto el entorno físico como el humano. Estas perspectivas sugieren que el sector turístico fomenta el desarrollo económico y social, sin dañar el medio ambiente. Además, resaltan la importancia de la viabilidad a largo plazo de los destinos turísticos y su competitividad. Las cuatro perspectivas no son excluyentes, por el contrario, son complementarias y sinérgicas. El desafío futuro radica en implementar políticas y prácticas que armonicen estas perspectivas, creando un modelo de turismo que sea económicamente rentable y ecológicamente responsable, y que a su vez, también contribuya al bienestar a largo plazo de las comunidades locales y la sociedad en general.

Por su parte, Butler (1999) argumenta que la diferencia entre turismo sostenible y turismo desarrollado con principios de desarrollo sostenible radica en el enfoque y la aplicación de cada concepto. Para el autor, el turismo sostenible es un tipo de turismo que se mantiene viable en un área durante un período indefinido sin degradar el entorno. En cambio, el concepto de turismo desarrollado con principios de desarrollo sostenible es más amplio y se refiere a la planificación y desarrollo del turismo que considera la sostenibilidad ambiental, social y económica. La distinción que realiza Butler (1999) permite identificar dos perspectivas. Una sectorial, asociada al turismo sostenible, enfocándose principalmente en la sostenibilidad de la actividad turística en sí misma. Otra, holística, de desarrollo sostenible, en la que se integran las tres dimensiones de la sostenibilidad en la planificación y desarrollo de cualquier actividad, incluido el turismo. En otras palabras, se manifiesta un enfoque holístico de la sostenibilidad en el contexto turístico, mediante el proceso de desarrollo sostenible que incluye la gestión y planificación de la actividad turística.

Por consiguiente, a efectos de esta Tesis Doctoral se explora la sostenibilidad del turismo desde un enfoque holístico y centrado en grupos de interés. La sostenibilidad turística se puede entender como el equilibrio entre los beneficios obtenidos del turismo y los costes socioculturales, ambientales y económicos que implica su desarrollo. Es decir, los beneficios deben superar los impactos negativos en la sociedad, la cultura, el medio ambiente y la economía local (Butler, 2005; Butowski, 2019). Ng et al. (2017) afirman que la sostenibilidad (objetivo final) se alcanza cuando todas las partes interesadas realizan contribuciones positivas en los ámbitos social, ambiental y económico. Igualmente, se reconoce la importancia de la sostenibilidad de un destino turístico para mejorar su competitividad y viabilidad a largo plazo, como argumentan varios académicos (Butowski, 2019; Coccossis, 1996; Rahmadian et al., 2022). Rahmadian et al. (2022) afirman que el objetivo clave de la sostenibilidad en un destino turístico es la capacidad de fomentar la revisita y al mismo tiempo fomentar la preservación de las características sociales, culturales y económicas del lugar, protegiendo el medio ambiente.

#### **1.4. Dimensiones de la sostenibilidad en el turismo**

La sostenibilidad es percibida como un constructo multidimensional. Como se ha mencionado previamente, en turismo se destacan tres dimensiones clave interrelacionadas: ambiental, social (o sociocultural) y económica. Esto significa que un cambio en una puede afectar a las otras. Por ejemplo, un impacto ambiental negativo puede afectar la cultura local y

la economía, y viceversa. Por esta razón, es esencial que el turismo busque un equilibrio entre estos aspectos para lograr una sostenibilidad efectiva.

Sin embargo, en la literatura se han explorado diferentes enfoques que integran diversas dimensiones (ver Tabla 3). McCool (1995) especificó tres dimensiones sostenibles en el contexto turístico. La dimensión ambiental refleja la necesidad de que los recursos naturales y su entorno en los destinos turísticos no cambien significativamente debido al turismo. Es decir, los impactos ambientales deben ser controlados para mantener la integridad del ecosistema. La dimensión social se enfoca en mantener la estabilidad cultural y social de las comunidades locales, protegiendo las tradiciones, orden social y estructuras de las comunidades existentes. Por su parte, la dimensión económica garantiza la igualdad de oportunidades laborales, genera empleos de calidad, y fomenta precios de bienes y servicios razonables, beneficiando la estabilidad económica de la comunidad.

Algunos autores han demostrado que la sostenibilidad puede ser una fuente de ventaja competitiva para empresas y destinos turísticos. Por ejemplo, Martínez & Rodríguez del Bosque (2014) confirman que las dimensiones económica, social y ambiental contribuyen a una imagen corporativa positiva y crean un efecto mediador que fortalece la reputación de las empresas hoteleras de España. Iniesta-Bonillo et al. (2016) analizan cómo la sostenibilidad percibida por los visitantes en un destino turístico influye en su valoración y satisfacción con el viaje. Los resultados confirman que la sostenibilidad, compuesta por las tres dimensiones económicas, socioculturales y ambientales, es clave para desarrollar destinos competitivos y mejorar la satisfacción de los turistas.

Otro enfoque sugiere incluir una cuarta dimensión política, dado que podría disminuir la brecha entre la comprensión conceptual del turismo sostenible y su aplicación práctica. Ritchie & Crouch (2003) y Mihalič et al. (2016) argumentan que un entorno político bien desarrollado y una gobernanza efectiva de destino, son esenciales para implementar prácticas de turismo sostenible y mejorar la satisfacción de los residentes con el desarrollo turístico. Esta dimensión puede abarcar aspectos como el poder, la autoridad y la toma de decisiones en el ámbito del turismo.

De manera similar al enfoque anterior, aunque no necesariamente idéntico, otros académicos se han centrado en cuatro dimensiones de la sostenibilidad: ambiental, económica, sociocultural e institucional (Cottrell et al., 2013; Cottrell & Vaske, 2006; Huayhuaca et al., 2010; Jasrotia et al., 2023). Desde este enfoque, la sostenibilidad institucional se puede entender

como la capacidad del gobierno o las instituciones para gestionar recursos, establecer marcos legales, aumentar la seguridad, y fomentar la responsabilidad tanto de la comunidad local como de los turistas, contribuyendo al desarrollo integral y al manejo adecuado del destino (Jasrotia et al., 2023). Autores como Cottrell & Vaske (2006), Huayhuaca et al. (2010) y Cottrell et al. (2013) revelan que las cuatro dimensiones predicen significativamente la satisfacción de los residentes con el desarrollo del turismo sostenible y que deberían considerarse en la planificación y evaluación de la actividad turística. Por otro lado, Jasrotia et al. (2023) examinan cómo las cuatro dimensiones afectan a la satisfacción de los turistas, demostrando que las dimensiones ambiental, sociocultural e institucional tienen una relación positiva con ella.

Con el objetivo de abarcar una visión más completa, varios autores han considerado hasta siete dimensiones de sostenibilidad (Agyeiwaah et al., 2017; Bramwell et al., 1996). Agyeiwaah et al. (2017) a través de un análisis exhaustivo de la literatura existente sobre sostenibilidad en el turismo, identifican cuatro dimensiones centrales (económica, ambiental, social y cultural) y tres periféricas (política, gerencial y tecnológica). Las centrales, se alinean con la teoría del “*Triple Bottom Line*” (TBL) (Elkington & Rowlands, 1999), donde los aspectos sociales se dividen en impactos en la sociedad y la cultura, promoviendo un desarrollo turístico sostenible. Las dimensiones periféricas son relevantes, pero pueden generar dos inconvenientes. Primero, desviar la atención de los elementos esenciales de la sostenibilidad (económica, ambiental, social y cultural). Segundo, fomentar la falta de responsabilidad personal, justificando la inacción de algunos actores involucrados.

**Tabla 3. Dimensiones de la sostenibilidad en el turismo.**

Dimensiones		Autores
Tres	Económica, Sociocultural y Ambiental.	McCool (1995); Donohoe (2011); Martínez & Rodríguez del Bosque (2014); Stylos & Vassiliadis (2015); Iniesta-Bonillo et al. (2016); Mihalic (2020); Santos et al. (2022)
Cuatro	Económica, Sociocultural, Ambiental y Política.	Ritchie & Crouch (2003); Mihalič et al. (2016)
Cuatro	Económica, Sociocultural, Ambiental e Institucional.	Cottrell & Vaske (2006); Shen & Cottrell (2008); Cottrell et al. (2013); Huayhuaca et al. (2010); Jasrotia et al. (2023); Trišić et al. (2024); Ristić et al. (2024)

Dimensiones		Autores
Siete	Ambiental, Cultural, Política, Económica, Social, Gerencial y Gubernamental.	Bramwell et al. (1996)
Siete	Económica, Social, Ambiental, Cultural, Política, Gerencial (o Institucional) y Tecnológica.	Agyeiwaah et al. (2017)

**Fuente: Elaboración propia basada en los autores.**

En definitiva, los anteriores enfoques multidimensionales reflejan la importancia de considerar diferentes aspectos para comprender y abordar la sostenibilidad en el turismo de manera integral. Las dimensiones se interconectan y deben ser consideradas en conjunto para lograr un turismo verdaderamente sostenible, ya que cambios en una dimensión pueden influir en las otras. En el cuarto capítulo de la Tesis Doctoral, “Discurso de los turistas sobre sostenibilidad: *insights* a partir de reseñas en línea de atracciones turísticas rurales y urbanas de España” (página 145), se examina la sostenibilidad desde sus cuatro dimensiones centrales (económica, social, ambiental y cultural) para fomentar el desarrollo turístico sostenible (Agyeiwaah et al., 2017). Además, este enfoque permite destacar la dimensión cultural de manera individual, subrayando su importancia para el turismo.

### 1.5. Destino turístico y sostenibilidad

En ocasiones, la actividad turística puede ser la principal fuente de desarrollo económico para muchos destinos. Baggio (2020b) expone que el destino turístico se considera una unidad esencial para entender y gestionar el fenómeno complejo y dinámico del turismo. Sin embargo, si el desarrollo turístico no se gestiona correctamente, puede generar grandes impactos negativos en el medio ambiente y en la comunidad local (Iniesta-Bonillo et al., 2016). La gestión de un destino para lograr el progreso social, económico y ambiental equilibrado requiere una estrategia que contemple principalmente los desafíos del entorno cambiante, tensiones internas de la competencia entre los actores involucrados, aumento continuo de visitantes y nuevos avances tecnológicos para la mejora continua (Baggio, 2020). Por consiguiente, la sostenibilidad es fundamental para asegurar el crecimiento y la competitividad de un destino turístico en el largo plazo (Mazanec et al., 2007).

Jovicic (2017) examina la evolución del concepto de destino turístico, resaltando que inicialmente, un destino se concebía como un área geográfica particular y debía cumplir con determinados criterios, tales como poseer atracciones turísticas, alojamientos y una red de transporte que permitiera el movimiento de las personas hacia, desde y dentro del destino (Howie, 2003). Sin embargo, este enfoque clásico fue criticado por no contemplar la cooperación entre los actores del destino ni el papel activo de los turistas, ignorando que los cambios en la demanda pueden conducir a cambios en la estructura del destino (Jovicic, 2017). Por ende, autores como Leiper (2000) y Baggio & Sainaghi (2011) han propuesto considerar un destino turístico como un sistema abierto y adaptable, en el que interactúan continuamente diferentes actores, como empresas, residentes, autoridades locales y turistas. Este nuevo enfoque permite captar mejor la complejidad y la interconexión de los actores que forman parte del destino turístico, así como la manera en que estos influyen y responden a los cambios internos y externos (Jovicic, 2017).

### 1.6. Actores involucrados en la sostenibilidad turística

Desde la perspectiva holística de la sostenibilidad en el turismo, autores como McIntosh et al. (1995) y Salee et al. (2022) identifican cuatro grupos de interés claves (ver Ilustración 5): 1) los residentes, que perciben el turismo como una fuente de beneficios (oportunidades de empleo) y desventajas (impactos negativos en la sociedad y medio ambiente); 2) los turistas, que buscan experiencias memorables y únicas; 3) las empresas, que proveen servicios y productos turísticos; y 4) el gobierno, que considera el turismo como una oportunidad para impulsar la economía local mediante la generación de mayores ingresos y empleos.

**Ilustración 5. Actores claves de la sostenibilidad en turismo.**



**Fuente:** Elaboración propia a partir de McIntosh et al. (1995) y Salee et al. (2022).

### 1.6.1. Residentes

El conjunto de actitudes, comportamientos y valores promovidos por una comunidad local o destino turístico puede ayudar a fortalecer la sostenibilidad ambiental y sociocultural. La literatura ha identificado este enfoque como cultura proturística (C. Y. Hsu et al., 2020; Rasoolimanesh et al., 2019). La participación activa de los residentes locales en la actividad turística puede influir de manera positiva y contribuir al desarrollo de un destino atractivo y acogedor para los turistas. En la literatura existen múltiples investigaciones que analizan las percepciones positivas y negativas de los residentes, sus reacciones al desarrollo turístico y los impactos generados (Farmaki et al., 2019). Asimismo, las estrategias actuales para planificar y desarrollar el turismo han considerado la perspectiva de este grupo de interés (H. S. C. Choi & Sirakaya, 2005; G. Kim et al., 2020).

Se espera que los residentes tengan una visión más positiva de los impactos del turismo cuando perciben que los beneficios obtenidos son mayores que los costes asociados al desarrollo turístico (Ap, 1992) pero, cuando sucede lo contrario, los residentes tienden a percibir los impactos del turismo de manera negativa. Es decir, los costes asociados al desarrollo turístico superan los beneficios, lo que genera insatisfacción y oposición entre la comunidad local (Rockey & Lankford, 2012).

### 1.6.2. Turistas

Al momento de elegir un destino turístico, los viajeros buscan experiencias únicas y memorables a través de productos/servicios tangibles e intangibles (Torres-Delgado et al., 2023). En consecuencia, su satisfacción se convierte en un aspecto muy relevante para la economía y la industria del turismo, puesto que aumenta la lealtad entre los turistas y abre nuevas oportunidades para el marketing de productos y servicios innovadores (Campón-Cerro et al., 2017). La satisfacción de los turistas depende de varios aspectos, incluyendo el paisaje, los sonidos del entorno, el ecosistema, la ubicación geológica y la actitud acogedora de los habitantes y las instituciones locales (Jiang et al., 2018).

Gracias a la creciente concienciación sobre el cambio climático y los límites planetarios, cada vez más personas se preocupan por el turismo sostenible. Este aumento en la sensibilización ha llevado a un mayor interés en prácticas turísticas que minimicen el impacto ambiental, preserven los recursos naturales y apoyen a las comunidades locales. Por ejemplo, Berezan et al. (2013) demostraron que la sostenibilidad tiene un impacto positivo en la satisfacción de los huéspedes y su intención de regresar. Los resultados también mostraron que los huéspedes

apoyan las prácticas sostenibles de los hoteles (como el reciclaje) y no son percibidas como inconvenientes que pueden convertirse en factores de insatisfacción.

Es relevante comprender la perspectiva sostenible de los turistas con el fin de fomentar comportamientos responsables y sensibilizar sobre las dimensiones económica, ambiental, social y cultural. Esto contribuye a minimizar los impactos negativos del turismo, atraer más turistas y mejorar la competitividad de los destinos (Palacios-Florencio et al., 2021).

#### 1.6.3. Empresas

Desde el ámbito empresarial, la Responsabilidad Social Corporativa (RSC) ha representado históricamente una estrategia empresarial para gestionar aspectos sociales, económicos y ambientales (Montiel, 2008; Van Marrewijk, 2003). Sin embargo, en los últimos años se ha generado un “efecto halo” que no le permite desarrollar todo su potencial, dado que se asocia en mayor proporción con la filantropía o acciones sociales que buscan mejorar la imagen y la reputación corporativa (Andreu Pinillos et al., 2018). En su lugar, surge con fuerza el concepto de Sostenibilidad Corporativa (SC), dado que se considera una estrategia empresarial más útil (Andreu Pinillos et al., 2018). La SC va en línea con las expectativas y exigencias del mercado (Chaminade, 2020; Fernandes et al., 2022), debido a que se considera la aplicación práctica del concepto de “desarrollo sostenible” al ámbito empresarial.

Las empresas turísticas demuestran una participación activa en la sostenibilidad, puesto que están interesadas por implementar avances tecnológicos que aporten un impacto positivo al medio ambiente y las comunidades locales. No obstante, se enfrentan a nuevos desafíos para cambiar el entorno y fomentar comportamientos más sostenibles en los turistas (Streimikiene et al., 2021). A pesar de que la integración de la RSC y SC en la gestión empresarial es de carácter voluntario, se consideran aspectos fundamentales para el desarrollo del negocio, imagen y proyección a largo plazo. Ambas aportan a las organizaciones múltiples beneficios, mejorando el rendimiento económico; la satisfacción de clientes; la reputación y legitimidad; la retención de talento; también minimizan los costes y los riesgos laborales o medioambientales (Berezan et al., 2013; González-Rodríguez et al., 2019; Martínez & Rodríguez del Bosque, 2014; Wut et al., 2023).

#### 1.6.4. Gobierno

La sostenibilidad de un destino depende de la participación y colaboración entre los diversos *stakeholders*, incluidos los organismos públicos (Mihalič et al., 2016b). El gobierno se

debe encargarse de brindar las condiciones adecuadas para garantizar la sostenibilidad de un destino. Autores como Ferrer-Roca et al. (2020), Kapera (2018) y Vieira et al. (2016) señalan que los gobiernos pueden optar por el apoyo financiero y asesoramiento de empresas turísticas; promover el empleo de calidad en el sector; mejorar el vínculo entre el turismo y las políticas comerciales para fomentar el crecimiento local, nacional e internacional de las empresas turísticas; desarrollar directrices, políticas y procesos hacia la sostenibilidad; y exigir el cumplimiento de la legislación referente a sostenibilidad. Para lograr un desarrollo turístico sostenible es esencial que el gobierno desempeñe un liderazgo político, mediante el cual se implementen políticas y estrategias efectivas, contemplando la diversidad de destinos, actividades turísticas y distintos actores involucrados (Vieira et al., 2016).

En la investigación de Ruhanen (2013) se detalla el rol intervencionista de los gobiernos en el turismo a lo largo del tiempo. En un principio, y motivados por el potencial económico de la actividad turística, los gobiernos invirtieron en infraestructura y promoción. Sin embargo, a medida que los efectos negativos del turismo comenzaron a ser más evidentes y se intensificaron las críticas contra la actividad turística, los gobiernos adoptaron un rol más activo en la planificación y regulación para controlar el desarrollo turístico y mitigar los impactos adversos. Específicamente, el rol de los gobiernos evolucionó de aspectos centrados únicamente en beneficios económicos hacia una gestión más holística, incluyendo impactos físicos y sociales. Esto está en línea con la expansión y adopción generalizada del concepto de desarrollo sostenible, fomentando prácticas que armonizan el crecimiento económico, la protección del medio ambiente y recursos naturales, así como el bienestar de la sociedad y los individuos.

### **1.7. Turismo regenerativo**

El debate actual en torno a la sostenibilidad del turismo se ha desplazado hacia propuestas más innovadoras, como el turismo regenerativo. Aunque la sostenibilidad es esencial, ya no es suficiente para garantizar el bienestar humano a largo plazo, lo que hace imprescindible adoptar prácticas regenerativas que restauren la capacidad de los sistemas para satisfacer tanto las necesidades actuales como las futuras (Konietzko et al., 2023). Muñoz & Branzei (2021) conceptualizan la regeneración como aquella “capacidad de volver a dar existencia a algo” (p. 509). Este proceso va más allá de la conservación o protección, enfocándose en una restauración activa que permite a un sistema restablecer su capacidad de autorregulación y evolución constante (Dias, 2019). Ateljevic (2020) subraya la urgencia de un cambio transformador en el sector turístico, orientado hacia la adopción de prácticas regenerativas para mitigar el cambio

climático y restaurar los ecosistemas. Estas prácticas pueden incluir la agricultura regenerativa, el uso de productos locales y la creación de experiencias turísticas transformadoras que conecten a los turistas con la naturaleza.

Según Bellato & Pollock (2023), el turismo regenerativo tiene el potencial de transformar la industria, pero enfrenta obstáculos como la simplificación del concepto y la exclusión de conocimientos y prácticas locales debido al predominio de enfoques occidentales. Los enfoques neoliberales y extractivos predominan en la investigación turística, priorizando el crecimiento económico sobre el bienestar social y ambiental, y promoviendo la explotación de culturas, lugares y recursos en lugar de su cuidado o regeneración (Bellato et al., 2024). Por el contrario, el turismo regenerativo se basa en la teoría de sistemas vivos, la cual considera a las comunidades y empresas como componentes integrados dentro de un sistema socioecológico mayor (Dredge, 2022; Major & Clarke, 2022). En consecuencia, es fundamental reemplazar los enfoques occidentales por un paradigma que no solo integre diversas perspectivas y conocimientos locales e indígenas, sino que también promueva una relación armónica y de reciprocidad entre humanos y naturaleza (Bellato et al., 2024; Bellato & Pollock, 2023).

A pesar de los grandes desafíos, el turismo regenerativo está despertando un interés creciente en la literatura académica. Según Corral-Gonzalez et al. (2023), este enfoque ha ganado atención, especialmente tras la pandemia de COVID-19, registrando un notable incremento en investigaciones desde 2020, con el 81% de las publicaciones realizadas a partir de ese año. Los autores identifican varios temas clave en este campo emergente: el “diseño regenerativo” como tema motor; “turismo sostenible” y “sostenibilidad” como temas básicos; y “viajes regenerativos” y “turismo esperanzador” como temas emergentes. Además, destacan otros temas de nicho como el “cambio climático” y el “turismo lento”. En definitiva, el turismo regenerativo es un campo de investigación emergente en constante crecimiento.

Otro estudio relevante en este campo es el de Bellato et al. (2022), ya que desarrolla un marco conceptual para el turismo regenerativo, diferenciándolo del turismo sostenible al priorizar la generación de efectos netos positivos en los sistemas sociales y ecológicos. Los autores presentan un modelo estructurado en cinco dimensiones clave: mentalidad regenerativa (alinear el turismo con una visión ecológica), potencial inherente (potenciar las capacidades únicas de los lugares), capacidad sistémica (fortalecer los sistemas sociales y ecológicos), efectos del sistema (definir objetivos compartidos) y tareas y recursos (implementar herramientas colaborativas). Además, presentan siete principios prácticos, como incorporar una

visión ecológica, adoptar un pensamiento de sistemas vivos, potenciar el potencial único de los lugares, catalizar transformaciones sistémicas, promover la justicia cultural, crear comunidades regenerativas y fomentar la colaboración evolutiva. Este estudio ofrece una guía integral y práctica para avanzar en la implementación del turismo regenerativo.

También es relevante destacar que Bellato et al. (2024) identifican dos campos clave para impulsar el turismo regenerativo: la ciencia de la sostenibilidad, que proporciona herramientas metodológicas robustas para avanzar hacia futuros regenerativos, y la investigación basada en conocimientos indígenas, que aporta enfoques alineados con la visión ecológica y facilita el cambio. Estos enfoques complementarios promueven un paradigma regenerativo inclusivo y sostenible. Por lo tanto, estudios que exploren las posibles contribuciones de estas metodologías serían fundamentales para el avance del turismo regenerativo. En este sentido, el análisis de la sostenibilidad turística desarrollado en esta Tesis Doctoral resulta relevante y contribuye en esta línea.

## **2. Nuevos enfoques para analizar la sostenibilidad turística: Ciencia de datos y *big data***

### **2.1. Ciencia de datos en la investigación**

En la actual era digital, el volumen de datos crece exponencialmente, lo que hace imprescindible su procesamiento, análisis y transformación en información útil para la toma de decisiones (George et al., 2014; Janssen et al., 2017). La gestión de los datos se presenta como un reto que empresas, universidades y todo tipo de organizaciones deben enfrentar y asumir, dado que estos se consideran la “materia prima del siglo XXI” (Alonso-Arévalo, 2019). En este contexto, la ciencia de datos ha emergido como una oportunidad disruptiva en la investigación científica. Donoho (2017) analiza su evolución a lo largo de medio siglo, iniciando con la visión pionera de John Tukey, quien en su artículo “*The Future of Data Analysis*” de 1962 ya señalaba la necesidad de un enfoque más amplio en el análisis de datos. Específicamente, Tukey (1962) identificó cuatro fuerzas impulsoras de la ciencia de datos: 1) el desarrollo de teorías estadísticas, 2) los avances en computadoras y dispositivos de visualización, 3) el aumento de la cantidad de datos disponibles, y 4) el enfoque cada vez mayor en la cuantificación en diversas disciplinas, mediante el uso de métodos cuantitativos para abordar problemas y tomar decisiones basadas en datos. Esta visión revolucionaria en su momento marcó un cambio radical al redefinir la estadística, pasando de ser una disciplina abstracta dentro de las matemáticas a convertirse en una herramienta esencial para el análisis aplicado y la toma de decisiones basada en datos en distintas áreas de conocimiento.

Otro estudio que constituye un excelente punto de partida y que resulta particularmente relevante para comprender el proceso de la ciencia de datos es el de Fayyad et al. (1996). Este trabajo es clave, ya que el descubrimiento de conocimiento en bases de datos (*Knowledge Discovery in Databases, KDD*, en sus siglas en inglés) puede considerarse el precursor de lo que hoy se conoce como ciencia de datos. Los autores precisan que el KDD es el proceso de descubrir conocimiento útil a partir de datos, el cual se caracteriza por su naturaleza interdisciplinaria y constante evolución. A pesar de que el estudio de Fayyad et al. (1996) fue publicado antes de que el término “ciencia de datos” se popularizara, se destaca por su importante contribución a la literatura al puntualizar definiciones claves, describir métodos y establecer las fases del proceso KDD.

Las fases del proceso KDD siguen siendo fundamentales y han sido adaptadas y expandidas en la literatura más reciente sobre ciencia de datos. La primera es la selección, que implica identificar y recolectar los datos relevantes que serán objeto de análisis. La segunda es el preprocesamiento, en la que se realiza una limpieza de los datos, eliminando patrones o elementos no informativos (símbolos, números, etc.), manejando valores faltantes y corrigiendo inconsistencias, con el objetivo de garantizar la calidad de la información. En la tercera fase de transformación, se optimizan los datos mediante la reducción de su dimensionalidad o la generación de representaciones adecuadas para el análisis. La cuarta fase de minería de datos es donde se aplican algoritmos específicos para descubrir patrones significativos, como clasificaciones, agrupamientos o asociaciones. En la última fase de interpretación y evaluación, se analizan y validan los patrones extraídos, asegurando que los resultados sean coherentes con los objetivos del estudio y generen información valiosa.

El proceso KDD es interactivo e iterativo, puesto que en cada una de las fases el usuario interactúa y toma decisiones según el contexto u objetivos de análisis (qué datos seleccionar, qué algoritmos aplicar, cómo interpretar los resultados, etc.). En algunas ocasiones las fases del proceso no se llevan a cabo de forma secuencial, por lo que se requiere volver a pasos anteriores para ajustar o mejorar los análisis. Por ejemplo, después de aplicar un algoritmo de minería de datos, el usuario puede descubrir que es necesario limpiar mejor los datos o seleccionar otras variables más relevantes, lo que lo lleva a repetir esas fases.

La evolución del KDD y otros enfoques clave han sido fundamentales en el desarrollo de la ciencia de datos, donde los estadísticos han desempeñado un papel crucial en la creación y difusión del término “ciencia de datos” (Cao, 2017). George et al. (2016) definen la ciencia de

datos como aquel campo interdisciplinario que integra la estadística, la minería de datos, el aprendizaje automático (o *machine learning*) y el análisis de información con el objetivo de extraer patrones, comprender mejor los datos y generar nuevos conocimientos a partir de grandes volúmenes de datos. Por su parte, Cao (2017) presenta tres definiciones de la ciencia de datos como resultado de su revisión minuciosa del campo, abarcando su concepto general o abstracto, su naturaleza interdisciplinaria y sus resultados tangibles en forma de productos de datos (Tabla 4).

**Tabla 4. Definiciones de la ciencia de datos por perspectivas.**

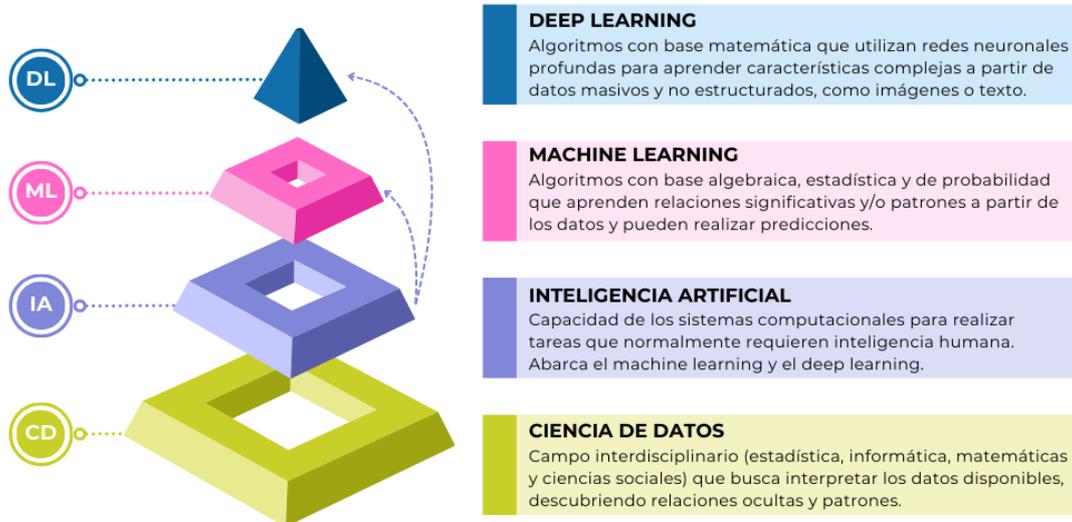
<b>Perspectiva</b>	<b>Definición de la ciencia de datos</b>
Definición de alto nivel (o general)	“La ciencia de datos es la ciencia de los datos” o “la ciencia de datos es el estudio de los datos”.
Definición disciplinaria	La ciencia de datos es un campo interdisciplinario que combina elementos de estadística, informática, matemáticas, comunicación y ciencias sociales para analizar, interpretar y extraer valor de los datos en diversos contextos. Su objetivo es convertir los datos en conocimiento útil para la toma de decisiones.
Definición de producto de datos	Un producto de datos es un resultado derivado de los datos, o que está impulsado por estos, y puede ser un descubrimiento, predicción, servicio, recomendación, conocimiento para la toma de decisiones, modelo, herramienta o sistema. Los productos de datos de valor definitivos son el conocimiento, la inteligencia, la sabiduría y la decisión.

Fuente: Cao (2017).

## 2.2. Inteligencia artificial y ciencia de datos

Para lograr una comprensión clara del campo de la ciencia de datos, es fundamental distinguirlo de la Inteligencia Artificial (IA) y sus subdisciplinas. Por tanto, en la Ilustración 6 se presentan los conceptos claves (Bishop, 2006; Cao, 2017; Goodfellow et al., 2016; Janiesch et al., 2021; Kok et al., 2009; Russell & Norvig, 2021).

Ilustración 6. Conceptos clave de inteligencia artificial y ciencia de datos.



**Fuente:** Elaboración propia basada en los autores: (Bishop, 2006; Cao, 2017; Goodfellow et al., 2016; Janiesch et al., 2021; Kok et al., 2009; Russell & Norvig, 2021).

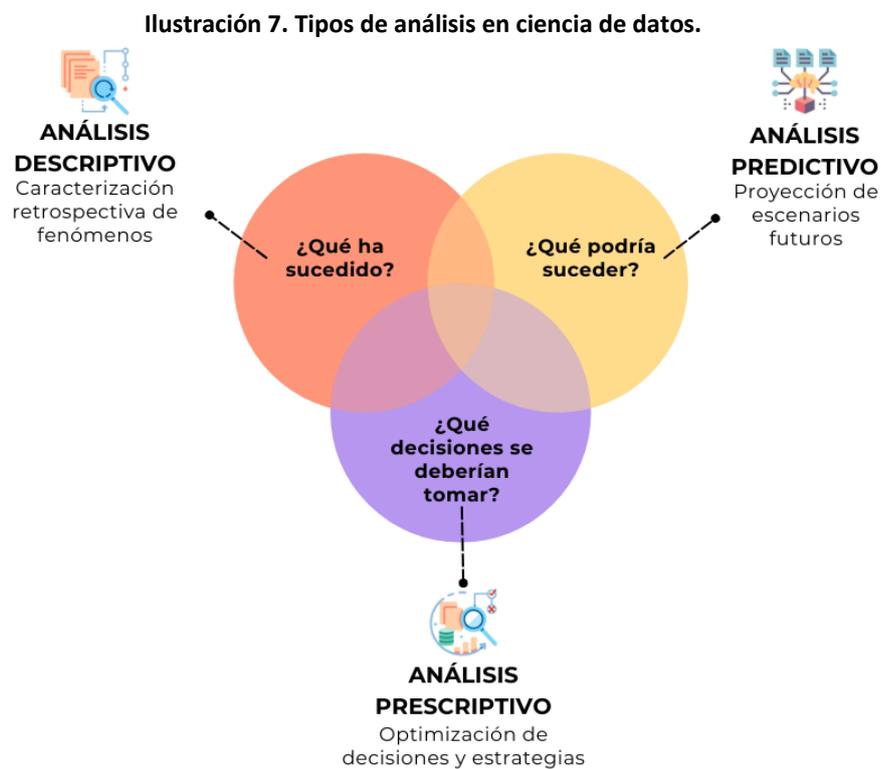
Como se mencionó anteriormente, la ciencia de datos es un campo interdisciplinario cuyo objetivo es interpretar los datos disponibles y descubrir relaciones ocultas y patrones. Para esto, emplea una variedad de herramientas y enfoques, uno de ellos es la IA, que hace referencia a la capacidad de los sistemas para resolver problemas complejos (Janiesch et al., 2021). Kok et al. (2009) puntualizan que la IA se enfoca en el desarrollo de sistemas computacionales que puedan llevar a cabo procesos de pensamiento comparables a los humanos, tales como aprender, razonar y autocorrección. Es decir, la IA abarca cualquier técnica que permite a las computadoras imitar el comportamiento humano, resolviendo tareas complejas de forma autónoma o con mínima intervención humana (Russell & Norvig, 2021).

Dentro de la IA, se encuentra el *machine learning* (aprendizaje automático) y el *deep learning* (aprendizaje profundo). El primero, se enfoca en identificar y aprender, de forma automática, relaciones y patrones útiles a partir de los datos y experiencias previas (Bishop, 2006). El segundo, suele superar a los modelos de *machine learning* más simples y a los enfoques tradicionales de análisis de datos en muchas aplicaciones. El *deep learning* utiliza redes neuronales artificiales profundas para aprender características complejas a partir de grandes conjuntos de datos, con mínima intervención humana (Janiesch et al., 2021).

### 2.3. Tipos de análisis en ciencia de datos

Dentro de la evolución del campo de la ciencia de datos, Cao (2017) identifica la consolidación de tres tipos de análisis avanzados que representan un cambio en el tratamiento

de datos (Ilustración 7). En primer lugar, la analítica descriptiva se centra en descubrir conocimiento oculto de datos correspondientes a eventos pasados o actuales, mediante herramientas y algoritmos preconfigurados. En segundo lugar, surge la analítica predictiva, que emplea algoritmos de aprendizaje automático y modelos estadísticos avanzados para pronosticar eventos futuros. Finalmente, la analítica prescriptiva representa el estado más avanzado, proporcionando recomendaciones óptimas para la toma de decisiones. Se desarrollan algoritmos y herramientas personalizadas para comprender profundamente los datos y las necesidades específicas.



**Fuente: Elaboración propia a partir de (Cao, 2017).**

Esta progresión desde lo descriptivo hasta lo prescriptivo refleja un claro aumento en la complejidad y el valor añadido de los análisis de ciencia de datos, así como un cambio de enfoque desde la comprensión del pasado hacia la orientación de acciones futuras. Cada uno de los análisis puede contribuir al ámbito académico y empresarial para una mejor toma de decisiones basada en datos. De acuerdo con George et al. (2016) el uso de la ciencia de datos en la investigación puede proporcionar respuestas más precisas a preguntas teóricas existentes y permite la formulación de nuevas preguntas con un alcance más amplio a partir de grandes volúmenes de datos.

Concretamente, esta investigación adopta un enfoque de ciencia de datos, combinando análisis descriptivos y predictivos tanto para el estudio teórico (capítulo 2, página 71) como para el empírico (capítulos 3 y 4, páginas 99 y 145, respectivamente). El análisis descriptivo se centra en caracterizar y resumir los patrones y tendencias existentes en los datos recopilados, mientras que el análisis predictivo emplea técnicas avanzadas de modelado estadístico y aprendizaje automático para descubrir relaciones ocultas entre variables.

#### **2.4. Fases del proceso de ciencia de datos**

El proceso de ciencia de datos se compone de varias fases clave que van desde la recolección de datos hasta su análisis o interpretación. El estudio de Ramasamy et al. (2023) sintetiza diversas actividades del flujo de trabajo de ciencia de datos propuestas en estudios previos, que pueden aplicarse en diferentes contextos, incluidas las empresas. En la primera fase de obtención y preparación de los datos, se recopilan y organizan para que sean aptos para su uso. Luego, en la fase de preprocesamiento, los datos se limpian y transforman para corregir problemas como valores faltantes o inconsistentes. La ingeniería de características es la tercera fase, en la cual se crean o ajustan variables para optimizar el rendimiento de los modelos. Posteriormente, en la cuarta fase de modelado, se construyen y ajustan algoritmos para analizar los datos y hacer predicciones. La siguiente fase de visualización permite que los resultados del modelo se interpreten de forma correcta. Finalmente, en la fase de despliegue el modelo puede ser implementado o aplicado para la toma de decisiones o la automatización de tareas. Cabe resaltar que, al igual que en el proceso KDD, estas fases no necesariamente siguen una secuencia lineal, dado que pueden formar parte de un proceso iterativo y cíclico.

Garijo et al. (2014) identifican un conjunto similar de actividades vinculadas a los datos dentro de los flujos de trabajo científico. Estos flujos constituyen una secuencia estructurada de pasos o tareas diseñadas para realizar un experimento computacional o un análisis científico de manera automatizada y replicable. Las fases claves detalladas son: 1) La preparación de datos, que incluye combinar, filtrar y transformar la información para posteriores análisis. 2) El análisis de los datos para obtener resultados científicos significativos. 3) La limpieza de datos para corregir errores (o inconsistencias) y mejorar la calidad de los datos. 4) El movimiento, donde los datos son transferidos a herramientas externas y recuperados para continuar el análisis. 5) La recuperación de datos involucra obtener información de bases de datos externas, integrando múltiples fuentes en el análisis. 6) La visualización implica generar visualizaciones, como gráficos

y tablas, para facilitar la interpretación de los resultados y comunicar los hallazgos de manera efectiva.

En la misma línea, George et al. (2016) proponen una guía con desafíos y soluciones a los que se pueden enfrentar los investigadores, abordando cinco aspectos clave de la ciencia de datos: recolección, almacenamiento, procesamiento, análisis y visualización de datos. Asimismo, Labrinidis & Jagadish (2012) puntualizan que el proceso de extracción de información a partir de *big data*, se divide en dos subprocesos clave: gestión de datos y análisis. La gestión de datos encierra actividades como la recolección, almacenamiento, organización y preparación de los datos, con el objetivo de asegurar que estén en las condiciones adecuadas para ser analizados. Luego, el subproceso de análisis emplea técnicas avanzadas, como algoritmos de minería de datos y métodos estadísticos, para extraer patrones e información relevante a partir de los datos procesados.

Tomando en cuenta estudios previos sobre ciencia de datos y *big data* (Fayyad et al., 1996; Gandomi & Haider, 2015; Garijo et al., 2014; George et al., 2016; Kwon, 2023; Labrinidis & Jagadish, 2012; Manning et al., 2009; Nikolay et al., 2011; Ramasamy et al., 2023; Varian, 2014), se pueden establecer varias fases fundamentales en el proceso de la ciencia de datos como metodología de investigación científica, las cuales guiarán el desarrollo de cada uno de los objetivos específicos de la presente Tesis Doctoral (Ilustración 8).

Ilustración 8. Fases de la ciencia de datos.



**Fuente:** Elaboración propia basada en los autores: (Fayyad et al., 1996; Gandomi & Haider, 2015; Garijo et al., 2014; George et al., 2016; Kwon, 2023; Labrinidis & Jagadish, 2012; Manning et al., 2009; Nikolay et al., 2011; Ramasamy et al., 2023; Varian, 2014).

Primero, se recopilan los datos de diversas fuentes, incluyendo APIs, bases de datos y *web scraping*. Luego, en el preprocesamiento, estos datos se limpian y organizan, eliminando errores y preparándolos para su uso. En el tercer paso de exploración de datos se lleva a cabo un análisis preliminar detallado, utilizando visualizaciones para comprender distribuciones y relaciones entre variables. Después, se usan técnicas avanzadas como minería de texto, análisis de regresión y aprendizaje no supervisado, permitiendo extraer patrones y conocimientos de los datos. Por último, la fase de interpretación de resultados se centra en evaluar el rendimiento del modelo mediante métricas apropiadas, visualizar los resultados de manera efectiva y validar la capacidad de generalización del modelo.

## 2.5. Análisis de grandes volúmenes de datos: *Big data*

En ocasiones, los términos “ciencia de datos” y “*big data*” pueden ser utilizados de manera indistinta, lo cual puede llevar a confusión. Sin embargo, algunos autores como George et al. (2016) y Cao (2017) enfatizan en la diferencia entre ambos conceptos, debido a que están interrelacionados, pero son distintos. El *big data* hace referencia a grandes y complejos

conjuntos de datos generados por tecnologías modernas, que desafían las capacidades de procesamiento tradicionales. Mientras que la ciencia de datos abarca un campo más amplio que incluye el manejo de grandes volúmenes de datos, su análisis, interpretación y aplicación para obtener conocimientos. Es decir, la ciencia de datos desarrolla modelos para capturar, visualizar y analizar los patrones subyacentes en el *big data*.

Laney (2001) fue uno de los primeros en introducir el marco de las “tres V” (volumen, variedad y velocidad) para describir las características del *big data*. A continuación, se presentan cada una de estas, tomando como referencia el trabajo de Gandomi & Haider (2015):

- El *volumen* hace referencia a la magnitud de los datos o cantidades de datos que superan la capacidad de las herramientas de software tradicionales. Esta característica es relativa porque varía en función del momento y del tipo de datos. Lo que actualmente se considera *big data* podría no serlo en el futuro debido al constante avance en las capacidades de almacenamiento y procesamiento. Asimismo, el tipo de datos influye en cómo se percibe el volumen, debido a que dos conjuntos de datos con el mismo tamaño pueden requerir tecnologías de gestión diferentes (por ejemplo, tablas o videos).
- La *variedad* abarca los diferentes tipos y formatos en los que los datos se presentan (Ilustración 9). Estos pueden ser datos estructurados, como los que se encuentran en bases de datos tradicionales, o no estructurados, como texto, imágenes, videos y datos de redes sociales. También existen datos semiestructurados, como los archivos XML o JSON, que no se ajustan completamente a un formato rígido, pero contienen alguna organización. Esta heterogeneidad exige diferentes enfoques y tecnologías para almacenar, preprocesar y analizar los datos, ya que cada tipo de información puede requerir condiciones específicas de procesamiento.
- La *velocidad* describe la rapidez con la que los datos se generan, transmiten y procesan. Hoy en día los datos crecen de manera exponencial, gracias a las nuevas tecnologías como sensores, redes sociales y teléfonos inteligentes. Esta característica no solo implica la capacidad de capturar datos de manera continua, sino también de analizarlos en tiempo real o casi en tiempo real, para tomar decisiones rápidas y basadas en datos actualizados.

Ilustración 9. Tipos de datos.



Fuente: Elaboración propia.

Más allá de las conocidas tres V, existen otras características clave del *big data* que también merecen ser mencionadas. IBM introdujo la *veracidad*, que hace referencia a la incertidumbre y fiabilidad de los datos, puesto que no todos los datos son precisos o confiables. Por ejemplo, la información de redes sociales puede ser subjetiva o sesgada, lo que hace necesario el uso de herramientas o algoritmos de preprocesamiento adecuados. SAS añadió la *variabilidad* y la *complejidad*, donde la variabilidad se refiere a las fluctuaciones en la generación de datos. Por ejemplo, la cantidad de datos generados por una red social puede variar considerablemente dependiendo del momento del día o de eventos particulares. La complejidad se asocia con el desafío de integrar y procesar datos provenientes de diversas fuentes. Por último, Oracle, destacó el *valor*, puntualizando que no todos los datos tienen valor por sí mismos, pero al analizar grandes cantidades de datos es posible extraer información valiosa para la toma de decisiones (Gandomi & Haider, 2015; M. Khan et al., 2014).

Es evidente que el concepto y las características del *big data* continúan evolucionando y expandiéndose para incluir aspectos adicionales que reflejan la complejidad y los desafíos del manejo de datos a gran escala. Actualmente, es posible identificar una amplia variedad de V's en *big data*. Panimalar et al. (2017) proponen un total de 17 V's, incluyendo características como Viralidad (*virality*), que mide la rapidez con la que los datos se propagan en redes, Viscosidad

(*viscosity*), que trata el retraso entre la ocurrencia de un evento y su procesamiento, y Vaguedad (*vagueness*), que se refiere a la incertidumbre en la interpretación de los datos.

Por otro lado, Owais & Hussein (2016) agrupan 9 V's en cinco categorías CPIVW: Recopilación, Procesamiento, Integridad, Visualización y Valor de los Datos (*Collecting, Processing, Integrity, Visualization, Worth*, en sus siglas en inglés) para mejorar la gestión de los datos, integrando V's como Veracidad (*veracity*), que mide la fiabilidad de los datos, y Volatilidad (*volatility*), que describe la duración de la utilidad de los datos. Saeed & Husamaldin (2021) analizan la importancia y aplicación del *big data* en diversos sectores industriales e identifican 10 V's. Estas incluyen la Vulnerabilidad (*vulnerability*), que aborda la seguridad de los datos, y Visualización (*visualization*), que se refiere a la representación gráfica de grandes conjuntos de datos para facilitar su interpretación.

En esta investigación, se utiliza una combinación de datos estructurados y no estructurados para obtener una visión integral del tema en estudio. Los datos estructurados provienen de las bases de datos científicas, proporcionando información cuantitativa y sistemáticamente organizada. Por otro lado, los datos no estructurados se obtienen de fuentes como comentarios en línea, capturando percepciones y opiniones de los turistas, así como los informes corporativos. Estos últimos se pueden considerar como datos no estructurados, puesto que generalmente consisten en documentos de texto, memorias o presentaciones que no siguen un formato o estructura de datos predefinida.

## **2.6. Ciencia de datos y *big data* en la investigación turística**

El avance tecnológico e internet han permitido la creación y el almacenamiento de grandes volúmenes de datos, transformando la investigación en el sector turístico (Li et al., 2018; X. Li & Law, 2020; Mariani et al., 2018). Egger (2022) destaca el potencial de la ciencia de datos en el turismo, reconociendo tanto sus oportunidades como sus desafíos. Las técnicas empíricas tradicionales (como encuestas, análisis de contenidos, etc.) se están enriqueciendo o mejorando con técnicas de ciencia de datos como el análisis de grandes volúmenes de datos (o *big data*), aprendizaje automático (*machine learning*), procesamiento del lenguaje natural (PLN), o modelos predictivos, entre otros. De esta manera, se amplían las capacidades de análisis y se mejora la calidad de los resultados, detectando patrones, haciendo predicciones más precisas y generando información útil para la toma de decisiones. El análisis de *big data* mejora la investigación turística, puesto que permite superar las limitaciones de tamaño de muestra que

suelen enfrentar los estudios que utilizan datos de encuestas (Yang et al., 2015), aportan nuevos conocimientos sobre variables previamente estudiadas en la literatura hotelera y ofrecen una mejor comprensión del comportamiento de los turistas al proporcionar datos abundantes y libres de sesgos de muestreo (X. Li et al., 2017; Xiang, Schwartz, Gerdes, et al., 2015).

Durante la última década, se evidencia un crecimiento de las investigaciones en turismo que emplean técnicas de ciencia de datos y *big data* (António & Rita, 2023; Boyd & Crawford, 2012; Cai et al., 2024; Mariani & Baggio, 2022). No obstante, Mariani & Baggio (2022) sostienen que la mayor parte de los estudios publicados no provienen de revistas especializadas en hospitalidad y turismo, sino de campos como la ciencia de la computación y el transporte, lo que refleja una posible falta de adopción de estas tecnologías en turismo. Por tanto, puede que los investigadores del campo tengan una cierta resistencia a usar enfoques cuantitativos avanzados y algoritmos computacionales, debido a la falta de experiencia en lenguajes de programación y a la escasez de recursos de *hardware* y *software*, en comparación con otras disciplinas académicas (Mariani et al., 2018). Además, la naturaleza caótica de los datos (no estructurados y en grandes volúmenes) que pueden generar incertidumbre y confusión (Egger, 2022).

El estudio bibliométrico de Egger (2022) identifica cuatro áreas de investigación clave, en las que se aplica la ciencia de datos al turismo. Primero, el *análisis de sentimientos* se destaca por su importancia en la comprensión de percepciones y emociones de los turistas en redes sociales, reseñas y otros textos. En segundo lugar, el *big data* proporciona la base para el manejo de grandes volúmenes de información turística. Tercero, el *aprendizaje automático (machine learning)* se presenta como una herramienta central y versátil, con aplicaciones como la predicción de la demanda, comportamiento turístico y personalización de servicios. Por último, el *pronóstico (forecasting)* relacionado con el uso de técnicas estadísticas y de *machine learning* para predecir patrones turísticos, como la demanda y el comportamiento de los turistas. La investigación de Cai et al. (2024), más reciente y en la misma línea, confirma que dentro de los temas actuales de análisis de *big data* en turismo se encuentra el seguimiento de patrones de movimiento de turistas; la evaluación de aspectos psicográficos (sentimientos, percepciones, actitudes); marketing en redes sociales; la evaluación del desempeño y sus determinantes; la predicción de demanda; y el desarrollo y validación empírica de nuevos métodos y algoritmos (sistemas de recomendación o personalización).

A pesar del creciente interés en ciencia de datos y *big data*, la integración de estas metodologías en la investigación turística no ha alcanzado su máximo potencial (Egger, 2022;

Mariani et al., 2018; Mariani & Baggio, 2022). En consecuencia, se hace evidente la necesidad de incorporar metodologías innovadoras en la investigación turística que permitan aprovechar al máximo las nuevas formas de datos no estructurados, en combinación con los conjuntos de datos y técnicas tradicionales.

## **2.7. Fuentes de datos para análisis de *big data* en turismo**

Para tener una mayor comprensión de las posibles fuentes de datos susceptibles de ser analizadas mediante *big data* en turismo, es preciso tomar como referencia el estudio de Li et al. (2018), en el que se realiza una clasificación en tres grandes grupos: datos del dispositivo, datos de transacciones, y datos UGC o contenido generado por el usuario (*User Generated Content*, *UGC*, en sus siglas en inglés). En primer lugar, el internet de las cosas (*IoT*) ha promovido la creación y uso de dispositivos sensores para monitorear los movimientos de los turistas y el entorno, generando datos de GPS, móviles y Bluetooth (datos del dispositivo) (Shoval & Ahas, 2016). En segundo lugar, los datos de transacciones u operaciones se generan de las acciones que los turistas realizan para su viaje, como pueden ser búsquedas en la web, reservar o comprar en línea (Li et al., 2018).

En tercer lugar, el crecimiento de los datos UGC se ha visto impulsado por el mayor uso de internet y los canales en social media (foros, blogs, wikis, sitios de opinión, redes sociales, etc.), puesto que ha permitido que los turistas compartan sus experiencias y opiniones (Xiang et al., 2017). Los motivos por los cuales los turistas comparten sus vivencias y percepciones pueden ser: 1) Preocupación altruista por otros viajeros o el deseo de ayudar a los demás, 2) Interacción social como forma de participación en la comunidad virtual, 3) Reconocimiento de los demás como expertos en turismo, (superación personal o estatus social), 4) Intención de ayudar a las empresas del sector o generar presión (manifestando su grado de satisfacción o insatisfacción) (Y. Chen et al., 2011; Tian, 2013).

El auge de las redes sociales y el UGC proporciona un enorme volumen de información que permite comprender las vivencias, opiniones y sentimientos de los turistas o consumidores (Marine-Roig & Clave, 2015; Xiang, Schwartz, & Uysal, 2015). Asimismo, el UGC puede ser útil para estudiar la imagen online de un destino turístico, apoyando la toma de decisiones y la mejora continua (Lalicic et al., 2021; Marine-Roig & Anton Clavé, 2015). Los métodos de análisis vinculados al UGC están alcanzando niveles de sofisticación y mayores avances. Se ha pasado de extraer únicamente los contenidos (modelado de temas, análisis de sentimientos) a análisis

sofisticados como distribución de temas, puntuaciones de sentimientos, relaciones causales e identificar determinantes (Cai et al., 2024).

Estas nuevas fuentes de información y el análisis de *big data* brindan a los investigadores la oportunidad de emplear metodologías innovadoras. En consecuencia, los académicos del ámbito turístico empiezan a aplicar técnicas de Procesamiento del Lenguaje Natural (PLN) (Guerrero-Rodriguez et al., 2023; Sann & Lai, 2020), algoritmos de *machine learning* (Cervera et al., 2024; X. Li et al., 2021; Mor et al., 2023) y redes neuronales (Esteban Curiel et al., 2018; Höpken et al., 2021; N. Li et al., 2023; Liang et al., 2024). De igual modo, se emplean modelos de lenguaje basados en aprendizaje profundo, dentro de los cuales sobresale BERT (*Bidirectional Encoder Representations from Transformers*)<sup>1</sup>. Por ejemplo, Rey-Moreno et al. (2023) aplican modelado de temas BERTopic y el análisis de “*Zero-shot text classification*” para clasificar las reseñas de los usuarios de *Airbnb* con etiquetas relacionadas con su satisfacción y confianza. Chai et al. (2021) utilizan el modelo BERT para analizar las reseñas de turistas internacionales sobre el Monte Huangshan en China. H. Zhang et al. (2023) emplean BERT para comprender qué hace que las experiencias turísticas de *Airbnb* sean únicas y valiosas para los viajeros. Viñán-Ludeña & de Campos (2022) aportan un estudio novedoso sobre la aplicación del análisis de sentimientos en turismo, para lo cual entrenan un modelo utilizando el modelo BETO (versión de BERT en español de Cañete et al. (2020)). El modelo final clasifica sentimientos junto con un proceso para localizar lugares mediante hashtags, destacando los aspectos negativos clave de cada uno.

### 3. Conclusiones

El desarrollo del marco conceptual general de esta Tesis Doctoral ha permitido establecer un fundamento sólido que destaca la relevancia de la sostenibilidad en el ámbito turístico y posiciona la ciencia de datos como un marco metodológico innovador y pertinente en el contexto de la era digital actual. Un enfoque multidisciplinar que integra la sostenibilidad y la ciencia de datos enriquece la literatura turística existente, resalta la importancia de estas temáticas en el sector y proporciona información valiosa para la toma de decisiones. Además,

---

<sup>1</sup> BERT es un modelo de lenguaje desarrollado por Google que utiliza la arquitectura de *Transformers* (Devlin et al., 2018). Los modelos *Transformers* son redes neuronales que comprenden el significado contextual al examinar cómo se relacionan entre sí las palabras en una secuencia. Utilizan técnicas matemáticas como la atención o atención propia para identificar estas conexiones y capturar el contexto de la información (Vaswani et al., 2017).

el análisis de grandes volúmenes de datos (o *big data*) permite llevar a cabo estudios más robustos y precisos, superando las limitaciones del procesamiento y análisis manual.

En primer lugar, la investigación ha puesto de manifiesto que la sostenibilidad en el turismo es un aspecto clave debido a los significativos impactos que esta industria genera en el planeta y la sociedad. La implementación de la sostenibilidad en el sector es cada vez más urgente, puesto que se considera que muchos de los problemas ambientales más graves tienen su origen en las actividades y el desarrollo turístico (Trang et al., 2019; S. Wang et al., 2020). En este sentido, la sostenibilidad turística busca un equilibrio entre los beneficios y costes de la actividad. Por tanto, los beneficios obtenidos del turismo deberán ser mayores a los costes de su desarrollo a nivel ambiental, sociocultural y económico (Butler, 2005; Butowski, 2019). No obstante, para que la sostenibilidad sea alcanzable, es imprescindible que todas las partes involucradas (residentes, turistas, empresas y gobiernos) se comprometan y trabajen en conjunto hacia ese objetivo (Ng et al., 2017).

En segundo lugar, la enorme cantidad de datos generados diariamente en todos los sectores, incluido el turismo, ha cambiado la forma en que se analiza y gestiona la información. La ciencia de datos emerge como un campo interdisciplinario imprescindible para transformar datos en conocimiento útil, proporcionando nuevas oportunidades en el ámbito académico y empresarial, especialmente en la sostenibilidad. La capacidad de analizar grandes volúmenes de datos permite detectar patrones y tendencias que antes pasaban desapercibidos, facilitando la toma de decisiones basadas en datos. A pesar de los beneficios que ofrece el *big data*, se observa que los académicos del ámbito turístico no han adoptado completamente estas herramientas y metodologías (Egger, 2022; Mariani et al., 2018; Mariani & Baggio, 2022). Por tanto, se evidencia un *gap* en la investigación, dado que la comunidad científica no ha aprovechado completamente el potencial que ofrecen estas metodologías.

**CAPÍTULO II. RED DEL CONOCIMIENTO SOBRE SOSTENIBILIDAD  
TURÍSTICA A TRAVÉS DEL ANÁLISIS BIBLIOMÉTRICO Y UNA  
REVISIÓN COMPUTACIONAL DE LA LITERATURA**



## **CAPÍTULO II. RED DEL CONOCIMIENTO SOBRE SOSTENIBILIDAD TURÍSTICA A TRAVÉS DEL ANÁLISIS BIBLIOMÉTRICO Y UNA REVISIÓN COMPUTACIONAL DE LA LITERATURA**

En este segundo capítulo se desarrolla la perspectiva teórica, alineada con el primer objetivo específico de la Tesis Doctoral: Explorar la estructura y evolución de la red de conocimiento científico sobre sostenibilidad y turismo. Para ello, se lleva a cabo una revisión detallada de la literatura existente, utilizando un enfoque innovador que combina el análisis bibliométrico con la metodología *Computational Literature Review* (CLR) (traducida al español como “Revisión computacional de la literatura”) con el modelado de temas. Se analizan 9.449 publicaciones científicas, identificando las temáticas y áreas de investigación claves y predominantes. Además, se destacan las tendencias emergentes y en declive, respaldadas con análisis estadísticos. Este capítulo está organizado en diversas secciones, incluyendo la introducción que justifica el estudio, la revisión de la literatura, el marco metodológico desarrollado, los resultados y las conclusiones, junto con las limitaciones y las futuras líneas de investigación.

### **1. Introducción**

El vínculo entre sostenibilidad y turismo tiende a reforzarse cada vez más por los impactos positivos y negativos que este sector genera en la economía, la sociedad y el medio ambiente. La investigación en sostenibilidad turística ha experimentado un crecimiento en las últimas décadas, abarcando una amplia cantidad de temas que incluyen la conservación del medio ambiente, gobernanza y políticas, gestión de recursos naturales, participación de *stakeholders* y el relevante papel de los residentes, entre otros temas de interés (Bramwell et al., 2017; Demeter et al., 2023; Mihalic, 2020; Prerana et al., 2024). A medida que la literatura continúa ampliándose, surge la necesidad de sintetizar e integrar el conocimiento para proporcionar tanto una visión completa del desarrollo y estado actual del campo, como para identificar tendencias claras (Nunkoo et al., 2023; Rasoolimanesh et al., 2023; Sharma et al., 2021). En este contexto, se han llevado a cabo varias revisiones de la literatura sobre sostenibilidad y turismo, generalmente haciendo uso del análisis bibliométrico (Alonso-Muñoz et al., 2023; P. Kumar et al., 2024; Niñerola et al., 2019; Serrano et al., 2019; Yoopetch & Nimsai, 2019). Los análisis bibliométricos son útiles para revelar tendencias emergentes y brechas en la investigación, identificar autores y revistas influyentes, y mapear la estructura y evolución del campo de estudio.

Algunos estudios bibliométricos anteriores se han enfocado en un número limitado de investigaciones sobre sostenibilidad turística, analizando menos de 1.000 documentos (Alonso-Muñoz et al., 2023; P. Kumar et al., 2024; Moyle et al., 2020; Ruhanen et al., 2015). Otros han ampliado el conjunto de estudios, abarcando entre 1.000 y casi 5.000 documentos (Niñerola et al., 2019; Serrano et al., 2019; Yoopetch & Nimsai, 2019). Sin embargo, investigaciones recientes adoptan dos enfoques, algunas complementan los análisis bibliométricos con distintos métodos tradicionales o estadísticos, mientras que otras, aunque en menor medida, implementan técnicas novedosas. Por ejemplo, Alonso-Muñoz et al. (2023) aplican un análisis bibliométrico a un conjunto de artículos desde 1998 hasta 2021, seguido de una revisión exhaustiva de los artículos más citados en los años 2019 y 2020 para identificar las tendencias de investigación. Por su parte, Agarwal et al. (2024) emplean una técnica avanzada de modelado de temas estructurales para analizar 3.289 artículos de investigación sobre turismo sostenible publicados entre 1978 y 2022. Por consiguiente, es evidente que se requieren tanto análisis bibliométricos más exhaustivos que incluyan una mayor cantidad de datos, como métodos innovadores que complementen y enriquezcan los enfoques tradicionales. A menudo las investigaciones no integran métodos computacionales avanzados que podrían proporcionar una visión más profunda de las temáticas subyacentes.

El presente estudio tiene como propósito cubrir esta brecha en la investigación mediante la aplicación de un análisis bibliométrico y una revisión computacional de la literatura que emplea algoritmos de aprendizaje automático para analizar grandes volúmenes de literatura científica (Antons et al., 2021). Se ha optado por un enfoque de modelado de temas BERTopic, puesto que permite detectar las temáticas inmersas en un gran volumen de textos y es capaz de capturar el contexto semántico (Grootendorst, 2021). Igualmente, se utilizan pruebas estadísticas, incluyendo regresiones lineales, para identificar las tendencias emergentes en los temas de investigación. En consecuencia, el objetivo principal es proporcionar una visión exhaustiva y actualizada de la investigación sobre sostenibilidad en el turismo, a partir del análisis de 9.449 documentos de la base de datos Scopus. Esto incluye la identificación de temáticas predominantes, emergentes y en declive. Para esto se plantean las siguientes cuatro preguntas de investigación:

**P1:** ¿Cómo se estructura y evoluciona el conocimiento en la literatura sobre sostenibilidad y el turismo?

**P2:** ¿Qué temas y áreas de investigación se han estudiado en la literatura existente sobre sostenibilidad y turismo en los últimos veinte años?

**P3:** ¿Cuál es la popularidad de los temas y áreas de investigación en la literatura sobre sostenibilidad y turismo en los últimos veinte años?

**P4:** ¿Cuáles son las principales temáticas y áreas de investigación en tendencia sobre sostenibilidad y turismo?

Las contribuciones de esta investigación son múltiples. En primer lugar, se brinda una visión general y actualizada de la investigación sobre sostenibilidad en turismo. Al sintetizar y consolidar el conocimiento existente, se proporciona a investigadores, profesionales y responsables de políticas una comprensión más profunda de la sostenibilidad en la industria turística, lo cual puede ser útil para la toma de decisiones informadas y el desarrollo de estrategias efectivas. En segundo lugar, se amplían los marcos metodológicos para revisiones de literatura sobre sostenibilidad turística mediante la incorporación de grandes volúmenes de datos y la aplicación de métodos innovadores. Específicamente, se combina el análisis bibliométrico con una revisión computacional de la literatura a través del modelado de temas para 9.449 documentos. Estos métodos permiten realizar una evaluación más exhaustiva y detallada de la literatura existente. En tercer lugar, se identifican las tendencias temáticas de la investigación sobre sostenibilidad y turismo, manifestando las áreas que están ganando mayor atención en la comunidad científica, así como aquellos temas de investigación que han alcanzado un alto grado de madurez y aquellos que tienden a estar en declive.

## **2. Revisión de la literatura**

### **2.1. Análisis bibliométricos previos sobre sostenibilidad turística**

Los análisis bibliométricos han ganado gran reconocimiento y aprobación en la investigación académica y empresarial, puesto que, a diferencia de otras técnicas, como el metaanálisis y las revisiones sistemáticas de literatura, permiten manejar grandes volúmenes de datos bibliométricos para mapear y analizar el conocimiento existente (Donthu et al., 2021). Estudios previos han intentado sintetizar el conocimiento sobre sostenibilidad y turismo mediante el análisis bibliométrico y con múltiples enfoques. Por ejemplo, Moyle et al. (2020) evalúan 839 artículos sobre turismo sostenible publicados entre 1987 y 2017, centrándose en la relación con los Objetivos de Desarrollo Sostenible (ODS) de la ONU. Della Corte et al. (2019) hacen énfasis en el contexto de la innovación abierta y las tecnologías digitales, confirmando que esa

combinación es crucial para el turismo sostenible, puesto que no solo facilita la colaboración y creación de productos turísticos sostenibles, sino que ayuda a reducir los impactos negativos ambientales. Garrigos-Simon et al. (2018) proporcionan una visión general del turismo sostenible y su relación con el crecimiento de ingresos y empleo a partir de 2.279 referencias. Prerana et al. (2024) revisan los artículos publicados durante 25 años (1997 a 2021), identificando cinco temáticas referentes al turismo sostenible: reconceptualización y crítica, el rol de los residentes, ecoetiquetado y participación de *stakeholders*, turismo basado en la comunidad y gobernanza e indicadores de sostenibilidad.

Según Alonso-Muñoz et al. (2023), la investigación sobre sostenibilidad y turismo continúa siendo fragmentada y no proporciona una visión completa y actualizada del campo. Esto se debe a que muchos estudios son limitados en su alcance y la mayoría se enfocan en subsectores específicos, destinos particulares o problemas de sostenibilidad concretos. No obstante, diversas investigaciones han optado por un ámbito más general (P. Kumar et al., 2024; Niñerola et al., 2019; Prerana et al., 2024), brindando una base sólida para comprender la evolución y estructura de la sostenibilidad en el turismo. Por lo tanto, es evidente que son escasos los estudios en esta línea que emplean un volumen significativo de datos y técnicas avanzadas para complementar los análisis tradicionales. La presente investigación enriquece la literatura y amplía los marcos metodológicos para análisis de literatura con un enfoque innovador. Esto responde a la era digital actual y el abrumador volumen de información y conocimiento científico disponible en internet y, asimismo, superan las limitaciones de lectura y procesamiento del ser humano.

Investigaciones recientes de análisis de literatura como la de Agarwal et al. (2024) sugieren que se deben integrar métodos computacionales avanzados. La minería de textos y los algoritmos de inteligencia artificial podrían revelar las bases intelectuales y las estructuras conceptuales del turismo sostenible (o sostenibilidad turística). La literatura previa muestra un cambio en las metodologías y enfoques de investigaciones que buscan mapear el conocimiento y tendencias. En primer lugar, se evidencia un incremento en la cantidad de documentos analizados (Moyle et al., 2020; Ruhanen et al., 2015; Serrano et al., 2019; Yoopetch & Nimsai, 2019). En segundo lugar, se adopta un enfoque hacia la fusión de análisis bibliométricos con otras técnicas de análisis (por ejemplo, revisiones sistemáticas, pruebas estadísticas) (Alonso-Muñoz et al., 2023). En tercer lugar, se observa la implementación de métodos novedosos como el modelado estructural de temas (STM) y el enfoque 3M (Agarwal et al., 2024).

## 2.2. Métodos computacionales en la revisión de la literatura sobre sostenibilidad turística

El uso de métodos computacionales avanzados en la revisión de la literatura podría derivarse del aumento constante de publicaciones académicas y la limitada capacidad humana para procesar y analizar la información. En este contexto, la revisión computacional de la literatura ha surgido como una metodología que emplea técnicas de minería de texto y algoritmos para analizar y sintetizar grandes volúmenes de literatura científica de un campo específico (Antons et al., 2021). Además, puede enriquecer y complementar los resultados obtenidos de un análisis bibliométrico, puesto que analiza el contenido de los estudios mediante algoritmos de aprendizaje automático (Antons et al., 2021). Este enfoque permite extraer contenido temático detallado (Braun & Clarke, 2006) de los estudios a partir de sus resúmenes o textos completos.

Estudios preliminares han empleado satisfactoriamente la revisión computacional de la literatura en múltiples ámbitos como la innovación (Caloffi et al., 2023); informática y sistemas (Guler et al., 2024; Mortenson & Vidgen, 2016); competitividad de destinos turísticos (Xia et al., 2024); economía verde (Polewsky et al., 2024) o responsabilidad social corporativa (J. Xie & Jain, 2024). Sin embargo, no ha sido posible encontrar ningún estudio con esta misma metodología que analice la literatura sobre sostenibilidad y turismo. La aplicación de esta técnica brinda un panorama más detallado y completo del conocimiento existente en el campo, lo que a su vez permite identificar temáticas, áreas de investigación y tendencias en auge o declive.

La revisión computacional de la literatura sigue un proceso de seis pasos (Antons et al., 2021). En el primer paso se definen los objetivos conceptuales como explicar, relacionar, explorar o debatir. El segundo paso es la operacionalización de la revisión, en el que se especifica el dominio del conocimiento que se revisará, se compila el corpus de textos y se definen tanto las reglas de alcance, como el marco temporal y las bases de datos. En el tercer paso se seleccionan las técnicas computacionales, es decir, los algoritmos de inteligencia artificial, considerando tanto los objetivos establecidos, como las capacidades técnicas y las restricciones de recursos. En el cuarto paso se llevan a cabo los análisis de contenido o aplicación de los algoritmos. Esto puede incluir la preparación y limpieza de datos (dependiendo de la técnica elegida) y la verificación de la precisión de los resultados. En el quinto paso se consolidan y sintetizan los resultados, verificando nuevamente la validez de los resultados obtenidos. Por último, el sexto paso es la presentación de los hallazgos de manera clara y estructurada, asegurando que la revisión sea replicable y accesible para otros investigadores.

### 3. Metodología de investigación

Esta investigación emplea un marco metodológico que integra el análisis bibliométrico y la revisión computacional de la literatura. Los detalles de cada paso se muestran en la Ilustración 10.

Ilustración 10. Diagrama del marco metodológico de la investigación.



Fuente: Elaboración propia.

#### 3.1. Definición del objetivo conceptual

El objetivo principal de esta investigación es explorar la estructura y evolución de la red de conocimiento científico sobre sostenibilidad y turismo. Este enfoque exploratorio permite mapear el estado actual del conocimiento, identificar temas y áreas de investigación, detectar tendencias en auge y declive, proporcionando una visión integral y actualizada.

#### 3.2. Recopilación de literatura

En esta segunda fase se descarga la información de la base de datos Scopus, conocida por su alta calidad (Mortenson & Vidgen, 2016) y amplia adopción en la realización de análisis bibliométricos. La búsqueda se realiza en inglés utilizando términos clave y comillas para limitar los resultados a aquellos que incluyeran exactamente las expresiones seleccionadas. Además, se emplea el operador lógico *OR* para incluir variaciones de los conceptos. Así, la búsqueda queda de la siguiente forma sin restricción de año: *“sustainable tourism development” OR*

*“responsible tourism development” OR “sustainable tourism” OR “responsible tourism” OR “sustainable destination” OR “sustainable tourism destination” OR “responsible destination” OR “tourism sustainability”*. Se obtienen 9.449 documentos compuestos por artículos de revistas, actas de congresos, libros y capítulos de libros, abarcando el periodo de 1989 a julio de 2024 (mes de descarga de la información). Los metadatos de los documentos seleccionados y descargados incluyeron los resúmenes, autores, títulos, fuentes y año de publicación, palabras clave y número de citas.

### 3.3. Análisis bibliométrico

Para la tercera fase se utilizó el software bibliométrico VOSviewer (van Eck & Waltman, 2010). Se analizaron las tendencias temporales en la producción científica, las redes de colaboración entre autores y países, así como las revistas y temáticas más influyentes. También se mapearon y visualizaron las relaciones entre los términos más frecuentes en los documentos. La Tabla 5 presenta los análisis específicos realizados con VOSviewer para cada uno de estos aspectos.

**Tabla 5. Análisis de VOSviewer utilizados.**

Aspecto	Tipo de análisis	Descripción
Cooperación científica	Análisis de coautoría de autores	Muestra las relaciones de colaboración entre dos o más autores en la redacción de un artículo científico.
	Análisis de coautoría de países	Visualiza la colaboración internacional en investigación, conectando países cuyos investigadores han publicado juntos.
Impacto e influencia científica	Análisis de citación por documentos	Representa las relaciones de citación entre documentos altamente citados (en este caso, con más de 250 citas).
	Análisis de citación por fuentes	Ilustra las relaciones de citación entre revistas u otras fuentes de publicación.
Estructura del conocimiento y publicaciones clave	Análisis de co-citación de referencias citadas	Muestra qué referencias tienden a ser citadas juntas.
	Análisis de acoplamiento bibliográfico por autores	Analiza la relación entre investigadores individuales basándose en los trabajos que citan en común.

Aspecto	Tipo de análisis	Descripción
Temáticas de investigación	Análisis de co-ocurrencia de palabras clave	Visualiza qué términos clave aparecen juntos con frecuencia en publicaciones.

Fuente: Elaboración propia.

### 3.4. Revisión computacional de la literatura

En esta fase se delimita el marco temporal de los documentos a los últimos veinte años, excluyendo los documentos de los primeros años debido a la baja cantidad de publicaciones, puesto que podrían sesgar el análisis. Además, se decide no incluir el año 2024, dado que, al estar en curso, los datos incompletos podrían generar tendencias erróneas o incompletas. De esta manera, en la revisión computacional de la literatura se exploran 8.023 documentos del año 2004 al año 2023. Se define que la técnica computacional a emplear es el modelado de temas, debido a que se considera una herramienta útil para la revisión de literatura en áreas con un alto volumen de publicaciones, identificando de forma automatizada y precisa de los temas clave a partir de resúmenes (Mortenson & Vidgen, 2016). Cabe señalar que para el preprocesamiento y análisis posteriores de los datos se utiliza el lenguaje de programación Python.

#### 3.4.1. Modelado de temas BERTopic

Se selecciona la técnica de *machine learning* de modelado de temas BERTopic porque es capaz de capturar el contexto semántico (Grootendorst, 2021). Este enfoque se basa en los modelos *Transformers* y el modelo de lenguaje BERT (*Bidirectional Encoder Representations from Transformers*) desarrollado por Google en 2018. Para aplicar esta técnica fue necesario realizar un preprocesamiento de los datos, que incluyó la lematización y limpieza de resúmenes, transformándolos en minúsculas y eliminando elementos no informativos. BERTopic asume que cada documento aborda un tema específico y utiliza un proceso de agrupamiento para identificar y representar cada tema mediante las palabras clave más relevantes. De este modo, para cada tema se extraen las palabras más representativas y su puntuación c-TF-IDF, que indica su relevancia dentro del grupo. Cuanto mayor es el puntaje, más representativa es la palabra en ese tema específico, por lo que los puntajes varían entre los diferentes temas. Tras el modelado de temas, se realizó una categorización manual de los temas identificados, basada en las palabras representativas de cada uno. De igual modo, se llevó a cabo una agrupación de los temas en áreas de investigación más amplias. Este proceso permitió sintetizar los temas en

categorías generales coherentes con las principales líneas de investigación en el campo de sostenibilidad y turismo.

### 3.5. Análisis de distribución y tendencias

En la quinta fase y con el fin de comprender la relevancia de cada tema y área de investigación en el corpus de literatura, se realizó un análisis de distribución. Este análisis cuantitativo permitió identificar las áreas de investigación predominantes en sostenibilidad y turismo. Se calculó la frecuencia relativa de cada tema y área de investigación (o categorías generales), lo cual brindó una visión clara de la estructura temática del campo y las áreas de mayor concentración.

Por otro lado, para explorar la evolución y tendencia de los temas investigados, se aplicó el enfoque descrito por Xia et al. (2024). Se calculó la regresión de mínimos cuadrados ordinarios (Ordinary Least Squares, OLS en sus siglas en inglés) para cada tema identificado, lo cual permite capturar los cambios longitudinales en la popularidad de los temas. Se establece como variable dependiente el porcentaje de documentos relacionados con un tema específico en cada año. Este porcentaje se puede expresar como  $P(T/D)$ , donde  $T$  representa la cantidad de documentos sobre un tema en particular y  $D$  el total de publicaciones de ese año. Específicamente, se modela la relación entre el tiempo (años) y la popularidad de los temas (porcentaje de documentos relacionados con un tema en cada año). De esta forma, la pendiente de la regresión indica si un tema está en auge o en declive, y el valor  $p$  determina la significancia estadística de la tendencia.

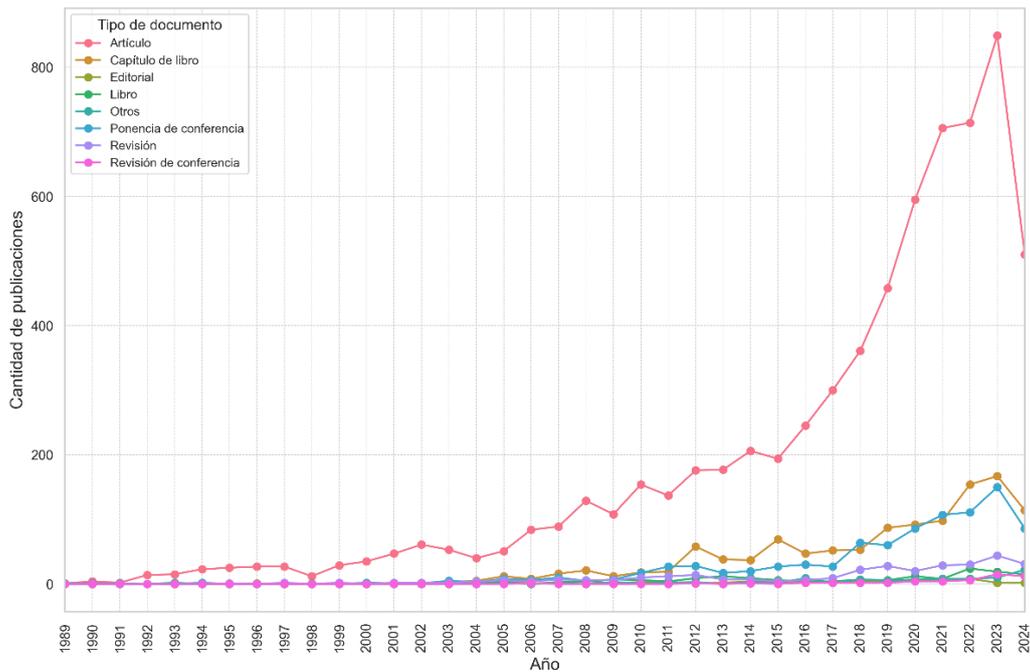
## 4. Resultados

### 4.1. Distribución de la producción científica

La Ilustración 11 muestra la evolución temporal de las publicaciones sobre sostenibilidad y turismo para cada tipo de documento desde 1989 hasta julio de 2024 (fecha de descarga de la información). Al igual que en Kumar et al. (2024), se evidencia un aumento en la producción científica, especialmente a partir de 2016, lo cual podría atribuirse a la aprobación de la Agenda 2030 de los Objetivos de Desarrollo Sostenible de ese año (Agarwal et al., 2024). Los artículos científicos han experimentado el mayor crecimiento, destacando los años entre 2016 y 2020. Los capítulos de libros y las ponencias de conferencias también muestran un aumento, aunque más moderado, indicando su continua relevancia. Por otro lado, al analizar al detalle la distribución por tipo de documentos (Ilustración 12) de la muestra total de 9.449 publicaciones, se observa que los artículos científicos son el tipo de documento predominante, representando

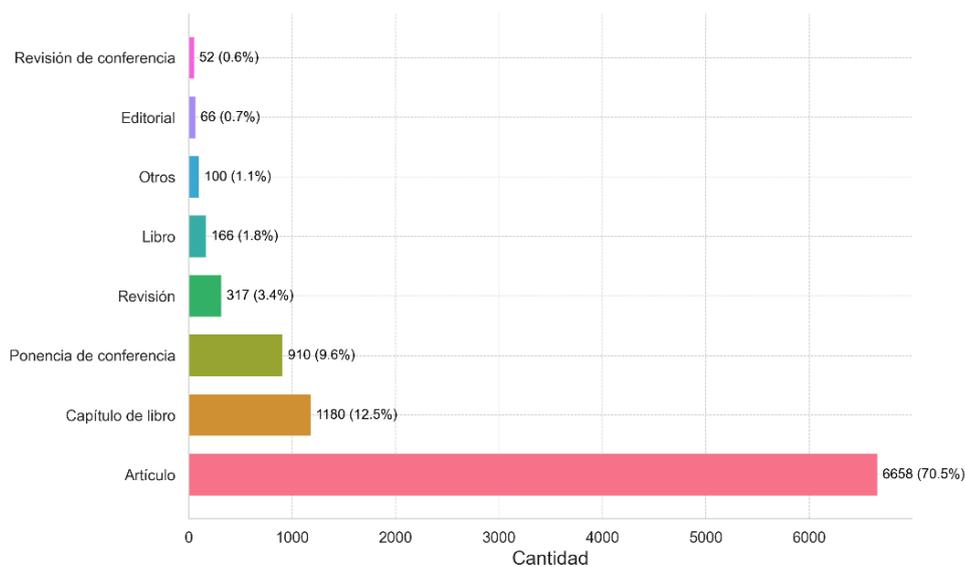
el 70,5% del total con 6.658 publicaciones. Los capítulos de libros constituyen el 12,5% (1.180 publicaciones), seguido por las ponencias de conferencias con un 9,6% y 910 publicaciones, reflejando la importancia de estas en la presentación de resultados preliminares y la discusión académica. La categoría “Otros” incluye documentos como notas, encuestas breves, cartas, representando el 1,1%.

**Ilustración 11. Evolución de los tipos de publicaciones por años.**



Fuente: Elaboración propia.

**Ilustración 12. Distribución de publicaciones por tipo de documento.**

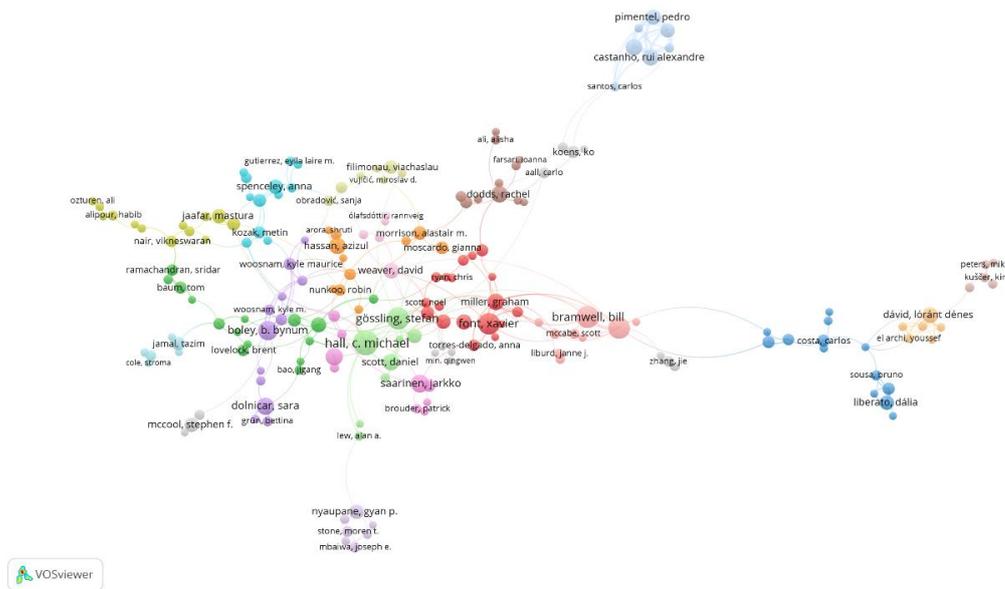


Fuente: Elaboración propia.

## 4.2. Cooperación científica

El análisis de la cooperación científica en sostenibilidad y turismo, tanto entre autores como entre países, revela una red global e interconectada. La red de coautoría de autores (Ilustración 13) destaca a Hall C. Michael y Gössling Stefan como nodos centrales, con 48 y 38 documentos respectivamente y múltiples colaboraciones fuertes. Por su parte, el análisis de coautoría de países (Ilustración 14) muestra la posición dominante de Reino Unido y Estados Unidos, ambos con 880 documentos publicados y una alta conectividad global (*link strength*<sup>2</sup> de 809 y 699, respectivamente). Otros países con una importante capacidad de colaboración internacional son China (730 documentos, 613 enlaces) y Australia (666 documentos, 537 enlaces). Los resultados son consistentes con los de Niñerola et al. (2019), en los que se concluye que los países con mayor producción son Estados Unidos, Reino Unido y Australia.

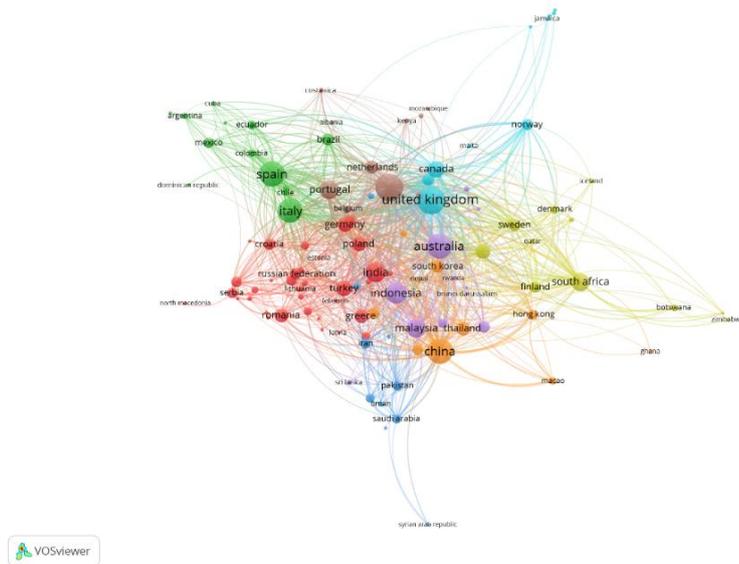
Ilustración 13. Análisis de coautoría de autores.



Fuente: Elaboración propia/procesamiento en VOSviewer.

<sup>2</sup> El *“link strength”* o *“fuerza de enlace”* indica la intensidad de la colaboración entre países en términos de publicaciones científicas conjuntas.

Ilustración 14. Análisis de coautoría de países.



Fuente: Elaboración propia/procesamiento en VOSviewer.

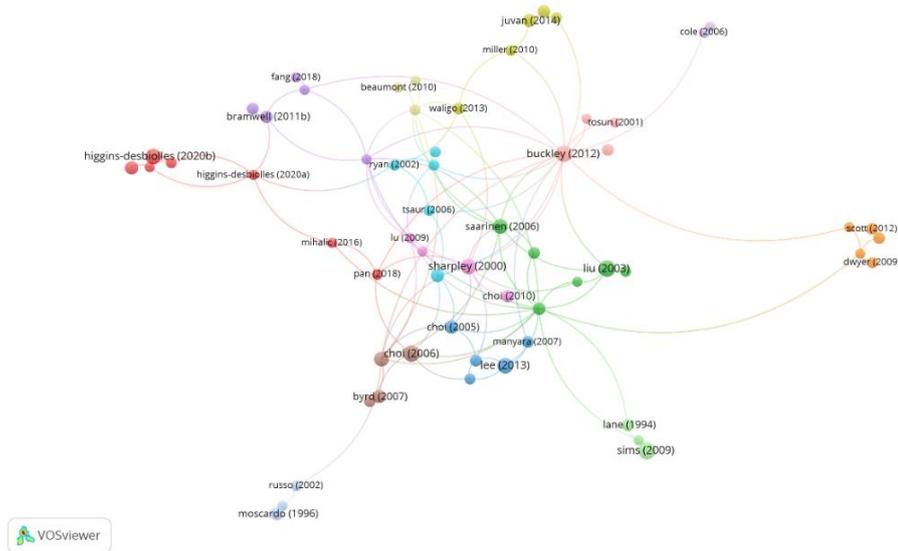
### 4.3. Impacto e influencia científica

A través del análisis de citación se identifican las publicaciones y fuentes más influyentes en sostenibilidad y turismo. El análisis de citación por documentos (Ilustración 15) destaca que la publicación más citada es la de Sims (2009) con 853 citas. Este autor estudia cómo la comida local puede contribuir a una experiencia turística sostenible al proporcionar la autenticidad que los visitantes buscan. La segunda publicación con más citas (838) es la de Gössling (2002), en la que se investiga cómo las actividades turísticas afectan el medio ambiente a escala global, a pesar de que estos efectos a menudo se consideran en un contexto más local. Otras investigaciones han sido cruciales en el desarrollo conceptual de la temática de estudio, por ejemplo, Butler (1999) con 830 citas o Liu (2003) con 719 citas.

Al analizar las conexiones entre documentos (enlaces) se observa que el artículo *“Sustainable tourism: Research and reality”* de Buckley (2012), aunque tiene 709 citas, posee 18 enlaces, indicando que es altamente citado y muy central en la red de citación. Esto sugiere que es un documento que conecta diferentes estudios y es fundamental para la estructura del conocimiento en sostenibilidad y turismo. Lo mismo sucede con las publicaciones de autores como Bramwell & Lane (1993), con 393 citas y 16 enlaces; Sharpley (2000), con 648 citas y 11 enlaces; Hunter (1997), con 622 citas y 10 enlaces. En contraste, la publicación de Gössling (2002) con 838 citas y 0 enlaces, demuestra que tiene un alto índice de citación, pero está más especializado o aislado en términos de su interconexión con otros estudios dentro del campo.

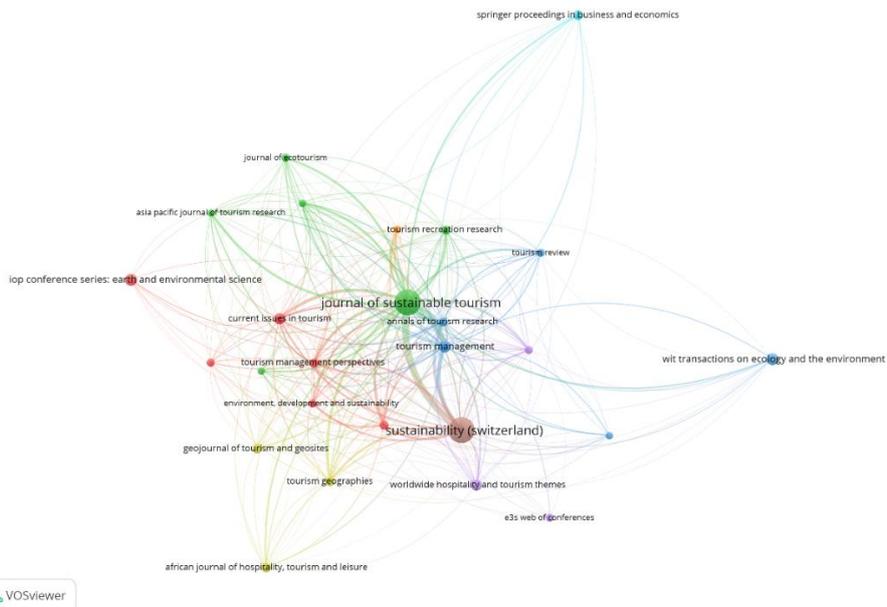
Con respecto al análisis de fuentes (Ilustración 16), “*Journal of Sustainable Tourism*” lidera con 734 documentos y 47.420 citas, seguida de “*Sustainability (Switzerland)*” con 737 documentos y 15.292 citas. Otras revistas importantes incluyen “*Tourism Management*” con 141 documentos y 15.663 citas, y “*Annals of Tourism Research*” con 82 documentos y 8.789 citas.

Ilustración 15. Análisis de citación por documentos (más de 250 citas).



Fuente: Elaboración propia/procesamiento en VOSviewer.

Ilustración 16. Análisis de citación por fuentes.



Fuente: Elaboración propia/procesamiento en VOSviewer.

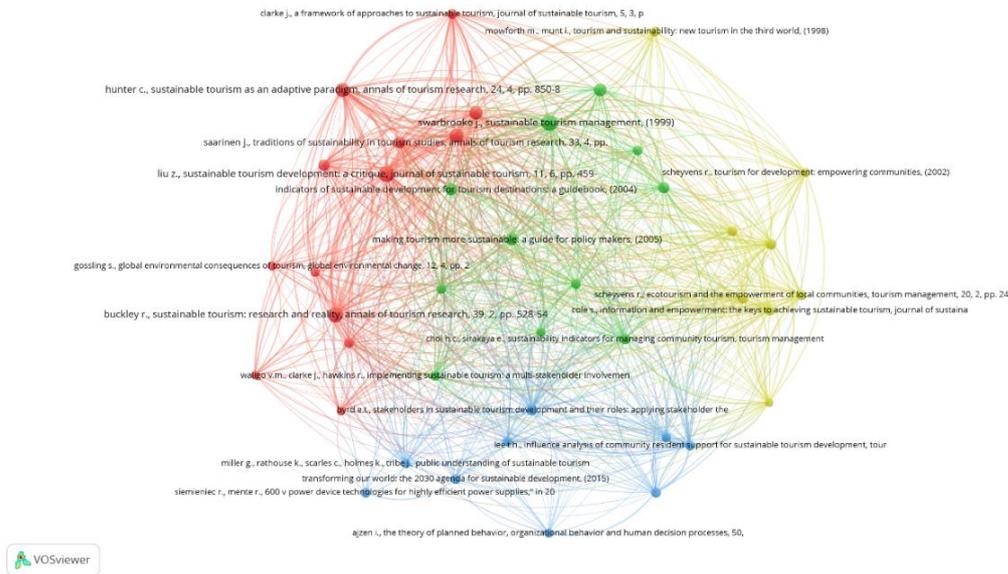
#### 4.4. Estructura del conocimiento y publicaciones clave

A continuación, se provee una visión integral de cómo se estructura el conocimiento sobre sostenibilidad y turismo, detectando los trabajos y autores que han sido fundamentales en la configuración y evolución de la literatura. El análisis de co-citación de referencias citadas de la Ilustración 17, destaca los estudios de Liu (2003), Hunter (1997) y Buckley (2012) como los más citados, con 224, 181 y 205 citas respectivamente. Estas publicaciones también presentan una cantidad elevada de enlaces (*links*) asociados (524, 432 y 416, respectivamente), lo que indica su centralidad en la red de conocimiento y su capacidad para conectar diferentes áreas de investigación.

Como se visualiza en la Ilustración 17, las publicaciones se pueden agrupar en clústeres temáticos que reflejan distintas áreas de investigación. Los trabajos incluidos en el clúster rojo brindan una base teórica sólida para la literatura (R. Buckley, 2012; Clarke, 1997; C. Hunter, 1997; Z. Liu, 2003; Sharpley, 2000). El clúster verde se enfoca en la gestión sostenible del turismo (H. S. C. Choi & Sirakaya, 2006; Swarbrooke, 1999; UNEP, 2005). La importancia del rol de las comunidades locales dentro del turismo sobresale en el clúster amarillo con las publicaciones de Scheyvens (1999), Scheyvens (2002) y Choi & Murray (2010). Por su parte, el clúster azul hace referencia a la comprensión del comportamiento humano y un enfoque centrado en *stakeholders* (Ajzen, 1991; Byrd, 2007).

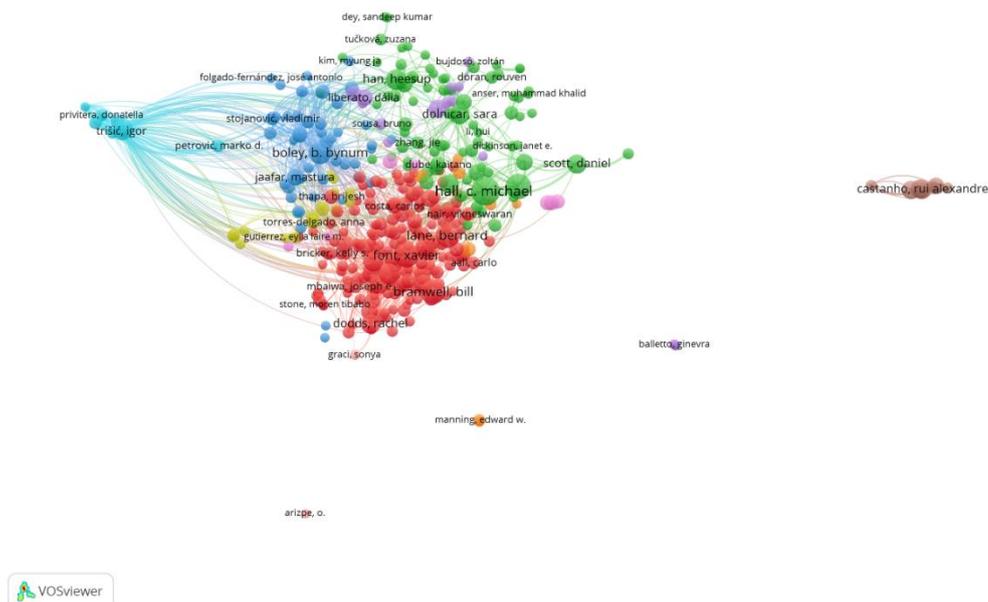
Complementando este análisis, la Ilustración 18 de acoplamiento bibliográfico entre autores resalta a Trisić Igor y Štetić Snežana por la elevada cantidad de enlaces (38.769), lo que refleja un notable solapamiento en las referencias que citan. A pesar de tener un menor número de citas y documentos comparados con autores como Hall C. Michael o Gössling Stefan se puede considerar que sus investigaciones tienen una alta influencia al conectar diferentes estudios en la literatura sobre sostenibilidad y turismo.

**Ilustración 17. Análisis de co-citación de referencias citadas.**



Fuente: Elaboración propia/procesamiento en VOSviewer.

**Ilustración 18. Análisis de acoplamiento bibliográfico por autores.**



Fuente: Elaboración propia/procesamiento en VOSviewer.

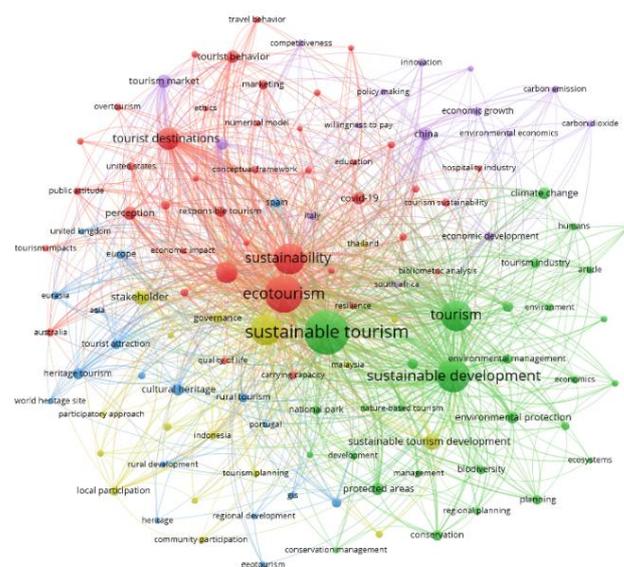
#### 4.5. Temáticas y áreas de investigación

A continuación, se presentan los temas clave en la literatura sobre sostenibilidad y turismo, identificados mediante el análisis bibliométrico y la revisión computacional de la literatura. En primer lugar, el análisis de co-ocurrencia de palabras clave realizado con el software VOSviewer

evidencia una estructura temática compleja y multidimensional (Ilustración 19). El primer clúster (rojo) se centra en la sostenibilidad, el ecoturismo y el destino turístico, destacando términos como “percepción”, “comportamiento del turista” y “marketing” (nótese que se mencionan con su traducción al español), lo que puede reflejar la gestión sostenible y holística que deben desempeñar los destinos. El segundo clúster (verde) aborda el desarrollo sostenible y la conservación de los ecosistemas naturales, con palabras clave como “conservación”, “gestión ambiental” y “áreas protegidas”. El tercer clúster (azul) resalta el patrimonio cultural y el turismo rural, incluyendo términos como “turismo cultural”, “patrimonio cultural” y “turismo rural”. El cuarto clúster (amarillo) indica la importancia de los *stakeholders*, haciendo énfasis en la comunidad local, así como en la gobernanza y la planificación de la actividad turística. Por último, el quinto clúster (morado) señala el interés en el turismo como motor económico, utilizando términos como “crecimiento económico”, “mercado turístico” y “competitividad”.

Los resultados están alineados con los estudios previos de Molina-Collado et al. (2022) y Prerana et al. (2024), quienes han identificado temas clave en el ámbito del turismo y la sostenibilidad mediante análisis bibliométricos. Estos académicos identifican como temas centrales la conservación de la biodiversidad, las áreas protegidas, la gestión ambiental, el papel de las partes interesadas y los residentes, destacando la importancia de la dimensión ambiental y el enfoque de *stakeholders* dentro del turismo.

Ilustración 19. Análisis de co-ocurrencia de palabras clave.

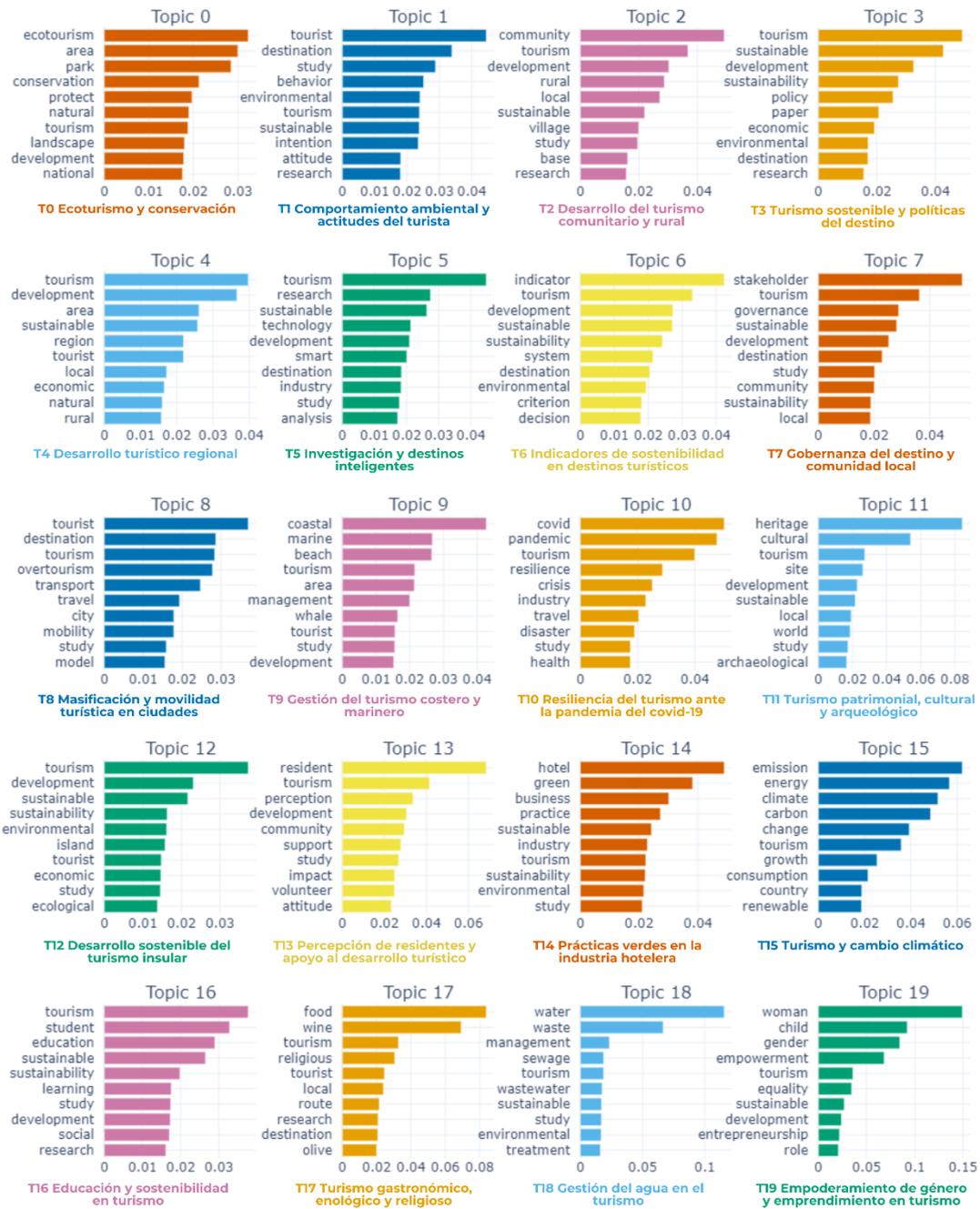


Fuente: Elaboración propia/procesamiento en VOSviewer.

Por otro lado, la revisión computacional de la literatura mediante modelado de temas ofrece un análisis más detallado de las temáticas y áreas de investigación de los últimos veinte años (2004 al 2023). La Ilustración 20 muestra los veinte temas identificados por BERTopic (0 al 19), junto con las palabras claves que los representan. Como se observa, existe una diversidad de enfoques temáticos en la investigación sobre sostenibilidad y turismo. Los principales temas, como el ecoturismo, destino turístico, desarrollo comunitario y rural, coinciden con los hallazgos del análisis bibliométrico. También destacan la política económica, gobernanza, impacto del covid-19, hotelería sostenible, entre otros.

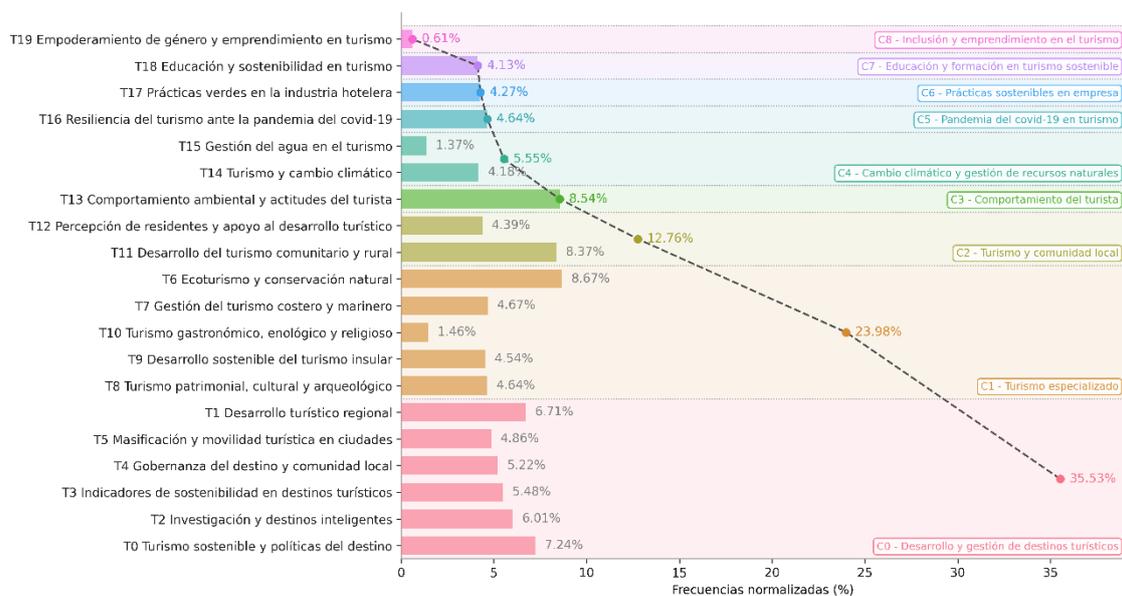
Para facilitar la comprensión de las principales temáticas y áreas de investigación, cada tema fue categorizado y nombrado de manera específica (ver Ilustración 20). Además, se agruparon los temas interrelacionados dentro de categorías generales que abarcan grandes áreas de investigación. En la Ilustración 21 se presentan las nueve áreas de investigación identificadas con sus respectivos temas y frecuencias relativas en porcentajes. Cada barra representa un tema específico y está coloreada según la categoría a la que pertenece. La línea negra punteada conecta las frecuencias de cada categoría, proporcionando una visión general del peso relativo de cada una en comparación con las demás.

Ilustración 20. Temas de investigación sobre sostenibilidad y turismo.



Fuente: Elaboración propia/procesamiento con BERTopic.

**Ilustración 21. Popularidad de los temas y áreas de investigación sobre sostenibilidad y turismo.**



Fuente: Elaboración propia.

Los hallazgos reflejan que la categoría de “Desarrollo y gestión de destinos turísticos” es la más predominante sobre sostenibilidad y turismo con un 35,53% de la literatura analizada. Sobresalen temas como “Turismo sostenible y políticas del destino” (7,24%), “Desarrollo turístico regional” (6,71%) e “Investigación y destinos inteligentes” (6,01%). Esto subraya la importancia de la planificación y gestión sostenible de los destinos turísticos a través de políticas integrales y procesos de gestión sostenibles (Bramwell & Lane, 2011; R. Buckley, 2012; Henry & Jackson, 1996).

El “Turismo especializado” se posiciona como la segunda categoría más relevante (23,98%). Dentro de esta, “Ecoturismo y conservación natural” sobresale con un 8,67%, reflejando la importancia de implementar formas de turismo sostenible que contribuyen a la conservación del medio ambiente y el desarrollo socioeconómico en áreas naturales poco alteradas (Ceballos-Lascurain, 1996; Tsaur et al., 2006). Además, emergen tipos de turismo especializados como el patrimonial, cultural, gastronómico o religioso que reflejan los múltiples intereses y motivaciones de los turistas. Por ejemplo, el turismo patrimonial atrae a turistas interesados en experimentar tradiciones culturales e históricas (Chhabra et al., 2003) y las motivaciones para realizar turismo cultural pueden asociarse con el aprendizaje, la identidad y la experiencia transformadora (Richards, 2018). El turismo religioso está estrechamente vinculado con la sostenibilidad, dado que existen cinco dimensiones clave que influyen en la experiencia del visitante en sitios religiosos: factores religiosos y culturales, ambientales, económicos, servicios

e instalaciones, y humanos (Pande & Shi, 2023). Por su parte, la gastronomía local se considera un producto auténtico que representa la cultura del destino y brindan una experiencia turística sostenible (Sims, 2009).

La categoría “Turismo y comunidad local” (12,76%) destaca la importancia de la participación comunitaria con énfasis en el “Desarrollo del turismo comunitario y rural” (8,37%). Esto concuerda con el estudio de Hardy et al. (2002) sobre la importancia de la participación de los residentes en el turismo sostenible. El comportamiento del turista se manifiesta como una categoría independiente con un 8,54%, mostrando la relevancia de la sostenibilidad ambiental y el rol de los turistas. Según, Han (2021) el comportamiento del consumidor sostenible es fundamental para proteger el medio ambiente y beneficiar a la sociedad a largo plazo.

Por otro lado, se identifican áreas de investigación relacionadas con desafíos contemporáneos, como “Cambio climático y gestión de recursos naturales” y “Pandemia del covid-19 en turismo”. Estas áreas reflejan la adaptabilidad del turismo a crisis globales, impactando la sostenibilidad de la propia actividad turística y sus *stakeholders*. De acuerdo con Scott et al. (2012), el cambio climático afecta al turismo y, a su vez, el turismo contribuye al cambio climático. Por ende, es necesario integrar el cambio climático en la planificación y gestión del turismo. Igualmente, Gössling et al. (2020) argumentan que la crisis de covid-19 se presenta como una oportunidad para transformar el sistema turístico hacia un modelo más sostenible y resiliente.

La literatura también se ha enfocado en áreas como “Prácticas sostenibles en empresa”, “Educación y formación en turismo sostenible” e “Inclusión y emprendimiento en el turismo”, subrayando el auge de la implementación práctica de la sostenibilidad en el sector y la formación de profesionales, respectivamente. Estos hallazgos se alinean con las investigaciones de Kim et al. (2017) y Yu et al. (2017) sobre prácticas sostenibles en hoteles, así como los estudios de Villena et al. (2021) sobre la educación en turismo para enfrentar los desafíos del desarrollo sostenible.

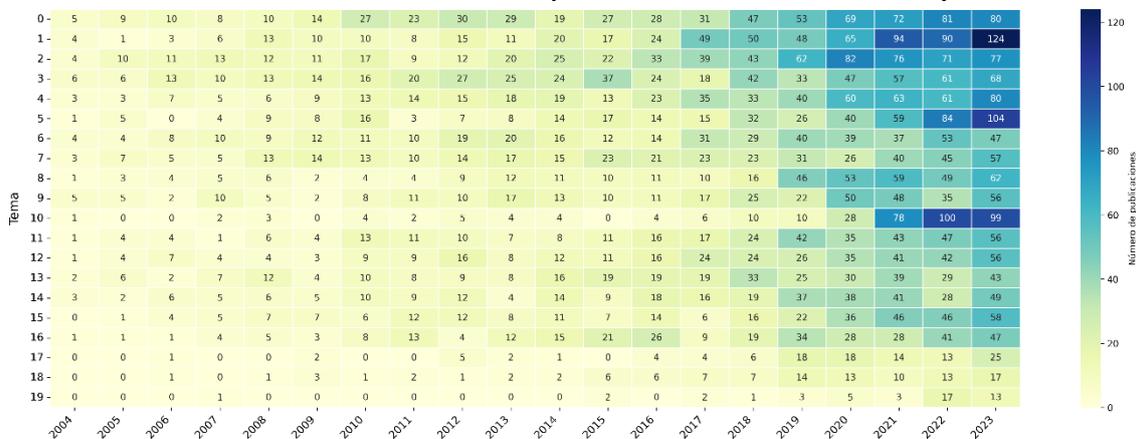
#### **4.6. Tendencias de las temáticas de investigación**

Tal como se especificó en la metodología, para explorar la evolución y tendencia de los temas investigados se analizó el volumen de publicaciones por año y se calcularon regresiones de mínimos cuadrados ordinarios (OLS) para cada tema específico. Siguiendo el enfoque de Xia et al. (2024), se modela la relación entre el tiempo (años) y la popularidad de los temas

(porcentaje de documentos relacionados con un tema específico en cada año). Es decir, la variable dependiente corresponde a la popularidad de los temas, mientras que la variable independiente es el tiempo (años).

En primer lugar, la Ilustración 22 detalla la cantidad de publicaciones por tema a lo largo del tiempo, permitiendo visualizar la evolución del interés en determinados temas y su concentración anual en la producción científica sobre sostenibilidad y turismo. Se evidencia un aumento en el número de publicaciones en la mayoría de los temas a partir de 2016, alcanzando su punto máximo en los años más recientes (2021-2023). Este incremento predomina más en los temas 1 (Comportamiento ambiental y actitudes del turista), 5 (Investigación y destinos inteligentes) y 10 (Resiliencia del turismo ante la pandemia del covid-19), indicando un creciente interés y desarrollo de estos. Por el contrario, los temas 17 (Turismo gastronómico, enológico y religioso), 18 (Gestión del agua en el turismo) y 19 (Empoderamiento de género y emprendimiento en turismo) presentan el menor número de publicaciones durante todo el periodo. Esto podría sugerir que son temáticas emergentes, menos exploradas o con menor relevancia en la literatura actual.

**Ilustración 22. Evolución del volumen de publicaciones sobre sostenibilidad y turismo.**



Fuente: Elaboración propia.

En segundo lugar, la Ilustración 23 muestra los resultados de las regresiones OLS con las pendientes ( $\beta$ ) y los  $p$ -value asociados a cada tema. Esto permite comprobar estadísticamente los cambios longitudinales en la popularidad de los temas para determinar si se encuentran en auge (barras verdes: pendiente positiva y estadísticamente significativo con un  $p$ -value  $\leq 0,05$ ) o declive (barras rojas: pendiente negativa y estadísticamente significativo con un  $p$ -value  $\leq 0,05$ ). Las tendencias no estadísticamente significativas se representan en barras grises.

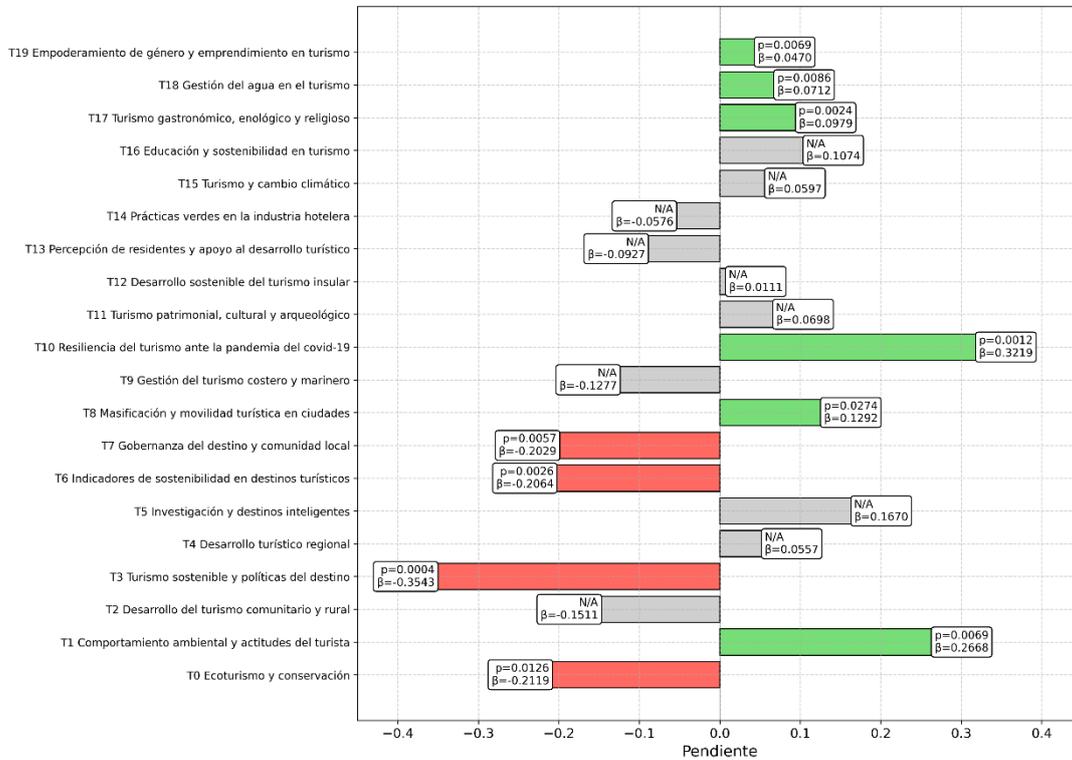
Los temas que han ganado atención en la investigación reciente son T1 (Comportamiento ambiental y actitudes del turista), T8 (Masificación y movilidad turística en ciudades), T10 (Resiliencia del turismo ante la pandemia del covid-19), T17 (Turismo gastronómico, enológico y religioso), T18 (Gestión del agua en el turismo) y T19 (Empoderamiento de género y emprendimiento en turismo). Los hallazgos son congruentes con estudio previos. Por ejemplo, Yooetch & Nimsai (2019) destacan el interés de explorar los comportamientos turísticos. Gössling et al. (2020) enfatizan en la necesidad de investigar y comprender los cambios provocados por la pandemia covid-19 para avanzar hacia un modelo más sostenible. Estos resultados también confirman que la gestión del agua dentro del sector es un tema emergente con gran relevancia, puesto que se considera un recurso crítico en el desarrollo turístico (Gössling et al., 2012). Otro aspecto polémico y relevante es el fenómeno del sobreturismo o masificación, convirtiéndose en una temática emergente de gran interés. Desarrollar investigaciones científicas en este ámbito es esencial, ya que pueden ofrecer soluciones valiosas para mitigar sus efectos negativos y contribuir a mejorar la calidad de vida de los residentes. Según Butler (1999), el principal desafío del desarrollo sostenible en turismo no radica en fomentar continuamente formas de turismo a pequeña escala que sean respetuosas con el medio ambiente y la cultura local, sino en cómo lograr que los actuales desarrollos de turismo masivo sean lo más sostenibles posible.

Por otro lado, los temas T0 (Ecoturismo y conservación), T3 (Turismo sostenible y políticas del destino), T6 (Indicadores de sostenibilidad en destinos turísticos) y T7 (Gobernanza del destino y comunidad local) presentan una tendencia negativa significativa, indicando un posible declive en la atención académica. La pendiente negativa y los *p-value* bajos sugieren que estos temas, aunque centrales en la literatura sobre sostenibilidad y turismo, pueden estar alcanzando un punto de saturación o desplazándose a nuevas áreas de estudio, manifestando la madurez de estos.

Finalmente, algunas tendencias no se pueden comprobar estadísticamente, puesto que su *p-value* es mayor a 0,05. Por tanto, no es posible asegurar que la tendencia observada (ya sea positiva o negativa) sea real y no se derive de una variabilidad aleatoria en los datos. Esto podría deberse a que son tendencias emergentes, aún en fases iniciales, o a que se dispone de datos insuficientes para alcanzar un nivel de confianza estadísticamente significativo en esos temas específicos. Entre las tendencias en crecimiento, se identifican temas como “Educación y sostenibilidad en turismo” (T16), “Turismo y cambio climático” (T15) y “Turismo patrimonial,

cultural y arqueológico” (T11). Por el contrario, las tendencias decrecientes incluyen temas como “Desarrollo del turismo comunitario y rural” (T2), “Gestión del turismo costero y marinerero” (T9) y “Prácticas verdes en la industria hotelera” (T14).

**Ilustración 23. Tendencias temáticas en la literatura sobre sostenibilidad y turismo (auge y declive).**



Fuente: Elaboración propia.

## 5. Conclusiones

Esta investigación realiza tres contribuciones principales. En primer lugar, se brinda una visión general y actualizada del conocimiento científico sobre sostenibilidad y turismo, a partir de un análisis exhaustivo y respaldado por un gran corpus de literatura (9.449 documentos). El enfoque metodológico mixto utilizado combina el análisis bibliométrico y la revisión computacional de literatura mediante modelado de temas, lo que permite generar una comprensión detallada de la producción, cooperación, impacto e influencia científica en el campo. Los hallazgos también revelan la estructura del conocimiento destacando las temáticas clave, las áreas de investigación, así como las tendencias emergentes y aquellas que están en declive. Esta información puede ser útil para investigadores, profesionales y formuladores de políticas, debido a que brinda una comprensión más profunda de las dinámicas de sostenibilidad en turismo, destaca conexiones clave que pueden orientar futuras investigaciones e identifica tendencias emergentes.

Al igual que Agarwal et al. (2024), se ratifica que la aprobación de la Agenda 2030 en el año 2015 marcó un hito en el ámbito académico puesto que se observó un incremento elevado en la producción científica sobre sostenibilidad y turismo desde 2016. Sin embargo, conforme a lo señalado por Moyle et al. (2020), los Objetivos de Desarrollo Sostenible (ODS) han sido reconocidos como un marco importante, pero aún se evidencia una brecha en la integración de estos en la investigación sobre turismo sostenible. Es decir, existe una necesidad por desarrollar investigaciones teóricas y prácticas centradas en los ODS y el turismo sostenible. Algunos de los temas asociados con los ODS en la literatura previa son las alianzas (ODS 17), la vida submarina (ODS 14), el agua limpia y saneamiento (ODS 6) y la acción climática (ODS 13) (Moyle et al., 2020). Los hallazgos de este estudio confirman estadísticamente que la gestión del agua en el turismo se ha convertido en un tema de creciente relevancia, atrayendo un notable interés dentro de la comunidad académica. Por el contrario, y aunque la tendencia alcista en la temática de cambio climático no puede ser comprobada estadísticamente, los resultados sugieren que es un tema emergente.

También, se evidencia que existe una red global e interconectada de cooperación en el campo, con una destacada participación de académicos como Hall C. Michael y Gössling Stefan. Los países con mayor producción científica son Reino Unido y Estados Unidos, lo cual coincide con los resultados del estudio de Niñerola et al. (2019). Entre los indicadores de impacto, se resaltan cuatro publicaciones con un alto número de citas y enlaces, dado que actúan como pilares fundamentales en la construcción del marco teórico y la estructura del conocimiento (Bramwell & Lane, 1993; R. Buckley, 2012; C. Hunter, 1997; Sharpley, 2000). La revista científica más influyente es *Journal of Sustainable Tourism* con 734 documentos y 47.420 citas. Además, mediante el análisis de co-citación de referencias citadas, se identifican cuatro grandes áreas de investigación que estructuran la literatura sobre sostenibilidad y turismo. Este hallazgo sugiere la madurez y consolidación de estas áreas, abarcando el marco teórico, la gestión sostenible del turismo, la importancia de las comunidades locales, la comprensión del comportamiento humano y *stakeholders*. Estos enfoques han sido fundamentales para el desarrollo teórico y práctico de la sostenibilidad turística.

En segundo lugar, se amplían los marcos metodológicos para revisiones de literatura. Se adopta una metodología con técnicas avanzadas que complementan las revisiones de literatura existentes (P. Kumar et al., 2024; Niñerola et al., 2019; Prerana et al., 2024), especialmente en la detección de temas y tendencias clave. A diferencia de los enfoques previos, este estudio

innovador permite identificar de forma detallada y precisa los temas y áreas de investigación, y a su vez, comprueba estadísticamente las tendencias en auge o declive de las temáticas. De acuerdo con Agarwal et al. (2024), los métodos avanzados de análisis y minería de textos son una oportunidad para revelar las bases intelectuales y las estructuras conceptuales de la literatura.

En tercer lugar, se detectan tanto los temas y áreas de investigación, como las tendencias que pueden servir de guía para futuras investigaciones. Este análisis proporciona una base sólida para orientar a los investigadores o profesionales en la exploración de nuevas direcciones que fomenten el desarrollo de estudios que contribuyan al desarrollo turístico sostenible. Los hallazgos demuestran que el análisis bibliométrico brinda una visión valiosa de los temas de investigación, pero, su alcance es limitado en cuanto a la profundidad con la que puede explorar las complejidades de la literatura existente. Por ende, la revisión computacional de la literatura con modelado de temas puede complementar, validar y profundizar en las temáticas investigadas. Asimismo, ofrece a los investigadores en turismo nuevas oportunidades para sintetizar grandes corpus de literatura.

Los resultados obtenidos en esta investigación son congruentes con la literatura previa, ya que, en términos generales, se resalta el enfoque de la sostenibilidad en el turismo centrada en los *stakeholders*. Al principio, los estudios sobre sostenibilidad y turismo se enfocaban en el desarrollo de marcos teóricos, sin embargo, estudios recientes indican que esos principios teóricos se deben llevar a la práctica con la participación de los *stakeholders* para asegurar un enfoque integral y colaborativo (Bramwell et al., 2017; Ruhanen et al., 2015; Sharpley, 2020). Dentro de las áreas de investigación detectadas mediante el modelado de temas sobresalen el “Desarrollo y gestión de destinos turísticos”, “Turismo y comunidad local”, “Comportamiento del turista” y “Prácticas sostenibles en empresas”, las cuales se pueden asociar con los cuatro grupos de interés clave en el turismo identificados por McIntosh et al. (1995) y Salee et al. (2022): gobiernos, residentes, turistas y empresas, respectivamente. Igualmente, los temas y áreas de investigación reflejan de manera implícita las cuatro dimensiones centrales de la sostenibilidad - económica, ambiental, social y cultural - descritas por Agyeiwaah et al. (2017). Estas dimensiones se alinean con la teoría del “*Triple Bottom Line*” (TBL) (Elkington & Rowlands, 1999), donde los aspectos sociales se dividen en impactos en la sociedad y la cultura, promoviendo un desarrollo turístico sostenible. De esta forma, se puede concluir que la literatura sobre sostenibilidad y turismo ha evolucionado con el tiempo para adoptar un enfoque

holístico que considera tanto los intereses y expectativas de los actores involucrados como los cuatro pilares de la sostenibilidad: ambiental, social, cultural y económico.

Esta investigación no está exenta de limitaciones, las cuales pueden abrir futuras líneas de investigación. En primer lugar, y a pesar de que se utilizó un extenso corpus de literatura, este se limita a una sola base de datos. Por tanto, próximos estudios podrían utilizar la misma metodología e incluir múltiples bases de datos para mejorar la representatividad de los hallazgos y validar las tendencias observadas. En segundo lugar, aunque se intentó destacar cada temática y área de investigación identificada, el enfoque de este estudio fue un mapeo y exploración general de la literatura, lo que implica que no se exploraron en profundidad todos los aspectos relevantes. Sería interesante que futuras investigaciones analicen al detalle estas temáticas específicas, así como los conceptos asociados. Esto permitiría identificar nuevas tendencias en ámbitos especializados, por ejemplo, en la planificación de destinos o gestión empresarial, ya que ambos requieren la implementación de estrategias concretas como la Responsabilidad Social del Destino (RSD), la Responsabilidad Social Corporativa (RSC) o la Sostenibilidad Corporativa (SC).

**CAPÍTULO III. EVOLUCIÓN TEÓRICA Y PRÁCTICA DESDE LA  
RESPONSABILIDAD SOCIAL CORPORATIVA HACIA LA  
SOSTENIBILIDAD CORPORATIVA: UNA APLICACIÓN DE LA  
CIENCIA DE DATOS EN LAS EMPRESAS TURÍSTICAS DE ESPAÑA**



### **CAPÍTULO III. EVOLUCIÓN TEÓRICA Y PRÁCTICA DESDE LA RESPONSABILIDAD SOCIAL CORPORATIVA HACIA LA SOSTENIBILIDAD CORPORATIVA: UNA APLICACIÓN DE LA CIENCIA DE DATOS EN LAS EMPRESAS TURÍSTICAS DE ESPAÑA**

En el marco de este tercer capítulo se desarrolla la investigación empírica referente al segundo objetivo específico de la Tesis Doctoral: Analizar la percepción y evolución de la Responsabilidad Social Corporativa (RSC) y la Sostenibilidad Corporativa (SC) en turismo desde el ámbito teórico y práctico, centrado en las empresas turísticas de España. Esta investigación pretende validar estadísticamente si existe un cambio de paradigma de RSC a SC a nivel teórico y práctico. En primer lugar, se aplica la metodología *Computational Literature Review* (CLR) (traducida al español como “Revisión computacional de la literatura”) a 1.505 documentos científicos mediante el modelado de temas, en un marco temporal de dos décadas. En segundo lugar, se desarrolla un estudio empírico en las empresas turísticas de España entre 2010 y 2021, a través de la aplicación de técnicas de ciencia de datos a 364 informes corporativos. Se aplican dos modelos de *machine learning*, BERTopic para detectar las áreas clave relacionadas con la gestión de la RSC y SC, y FinBERT para cuantificar la importancia que las empresas y subsectores turísticos otorgan a los criterios ESG (*Environmental, Social and Governance*, en sus siglas en inglés). Este capítulo adopta la estructura de un artículo científico, iniciando con la introducción para dar contexto a las temáticas de estudio, la revisión de la literatura establece el fundamento teórico de los conceptos examinados. Posteriormente, se exponen la metodología empleada, los hallazgos obtenidos y las conclusiones.

#### **1. Introducción**

En ocasiones, los términos Responsabilidad Social Corporativa (RSC) y Sostenibilidad Corporativa (SC) suelen emplearse de manera indistinta, si bien es fundamental diferenciarlos para aplicarlos y analizarlos correctamente (Bansal & Song, 2016; Montiel, 2008; Van Marrewijk, 2003). Van Marrewijk (2003), basado en el estudio de Panapanaan et al. (2003), sostiene que existe una relación jerárquica entre la RSC y SC, donde la RSC actúa como un paso intermedio hacia la SC (el objetivo final). Sheehy & Farneti (2021) subrayan que la RSC se centra en las obligaciones sociales y éticas de las empresas hacia sus *stakeholders* (Sheehy, 2014), mientras que la SC se enfoca más en el medio ambiente y en la implementación de políticas ambientales dentro de las operaciones empresariales, orientándose hacia la integración de la sostenibilidad en la estrategia empresarial (Montiel, 2008).

Aunque los orígenes de la RSC y la SC tuvieron enfoques distintos (Damtoft et al., 2024), con el tiempo, ambas teorías han llegado a un punto en común: buscar un equilibrio entre las responsabilidades económicas, sociales y ambientales de las empresas (Montiel, 2008). Es evidente que esto puede llevar a que tanto empresas como académicos utilicen los términos de manera intercambiable, lo que puede generar confusión y dificultar una comprensión de sus diferencias conceptuales. Burbano et al. (2023) precisa que se debe replantear el foco de los estudios sobre SC para abordar los desafíos ambientales y sociales actuales, desarrollando nuevas metodologías que midan los efectos de las prácticas sostenibles de las empresas.

En el contexto turístico, la RSC y SC son fundamentales ya que permiten que las empresas alcancen sus objetivos económicos y respondan de manera efectiva a las demandas sociales y ambientales, contribuyendo al desarrollo sostenible del sector (Font et al., 2012; Font & Lynes, 2018; S. Lee et al., 2022). De acuerdo con Back (2024), la RSC en la industria turística se ha transformado en prácticas estándar de gestión sostenible bajo los criterios ESG (*Environmental, Social and Governance*, en sus siglas en inglés). Este enfoque es una de las formas más comunes de implementar y materializar la SC. Sin embargo, autores como Legendre et al. (2024) sostienen que, aunque los criterios ESG han ido ganando popularidad, la transición de RSC a ESG aún no se ha completado, y ambos términos a menudo se usan como sinónimos. En consecuencia, se puede decir que la literatura evidencia un cambio de paradigma de RSC a SC. Esta transición refleja un avance relevante en el gobierno corporativo, enfatizando la importancia de un enfoque integral de la sostenibilidad (David et al., 2024). Asimismo, Damtoft et al. (2024) puntualizan que, aunque la RSC es comúnmente utilizada, en el ámbito empresarial muchas empresas se centran más en la SC, como lo reflejan recientes marcos y directivas de la Unión Europea sobre finanzas sostenibles y reportes de sostenibilidad.

Igualmente, las partes interesadas no solo exigen que las empresas del sector turístico operen de manera responsable y sostenible, sino que también esperan que comuniquen de manera clara sus resultados económicos, sociales y ambientales (Grosbois, 2012; Medrado & Jackson, 2015). Hooghiemstra (2000) afirma que la divulgación de informes de RSC se utiliza principalmente para gestionar la percepción pública y garantizar (o restaurar) la legitimidad de una empresa. Este autor también destaca la importancia de la comunicación corporativa como una herramienta para influir en la opinión pública y construir una imagen positiva. Por otro lado, Uyar et al. (2020) concluyen que, a pesar de que los informes de sostenibilidad en el sector

turístico han mejorado con el tiempo, aún existen áreas que necesitan mayor atención, como la energía renovable, la biodiversidad, la preservación del patrimonio cultural e histórico.

Diwan & Amarayil Sreeraman (2023) señalan que los informes de sostenibilidad han evolucionado hacia informes ESG, y actualmente se presentan como informes integrados, donde se combina la información financiera con la no financiera, abordando estos aspectos en conjunto bajo el enfoque del *Triple Bottom Line (TBL)*, que evalúa el desempeño económico, social y ambiental (Elkington & Rowlands, 1999). Como resultado de lo anterior, se ha experimentado un creciente interés en los informes de RSC, donde tanto empresas como académicos han tratado de entender mejor cómo las empresas comunican sus compromisos con la sociedad y la importancia de asegurar que estos informes reflejen las prácticas sostenibles y responsables que se llevan a cabo (Medrado & Jackson, 2015). Esto sugiere que los informes de RSC y SC son una fuente clave para evaluar cómo las empresas cumplen sus compromisos con las comunidades y adoptan prácticas sostenibles, garantizando la transparencia ante sus diversos *stakeholders*.

Ante la confusión y superposición que puede existir entre los conceptos de Responsabilidad Social Corporativa (RSC) y Sostenibilidad Corporativa (SC), resulta imprescindible llevar a cabo más estudios que ofrezcan una visión actualizada y detallada sobre estas prácticas en el sector turístico. Esto puede ser útil para evidenciar si realmente está ocurriendo un cambio de paradigma de RSC a SC. Asimismo, las memorias corporativas son una fuente clave para analizar dicha evolución y las prácticas empresariales implementadas a nivel social, ambiental y económico. De esta forma, las memorias sintetizan y consolidan información valiosa de las empresas que manifiesta la transformación hacia un enfoque más integral de SC.

Esta investigación tiene como objetivo principal analizar la percepción y evolución de la Responsabilidad Social Corporativa (RSC) y la Sostenibilidad Corporativa (SC) en turismo en el ámbito teórico y práctico. Desde el punto de vista teórico, se ofrece una visión detallada y actualizada de la investigación sobre ambas temáticas mediante una revisión computacional de literatura del 2004 al 2023. Esta revisión emplea el modelado de temas para identificar las principales corrientes de investigación y su evolución temporal. Adicionalmente, se valida estadísticamente la evolución de los temas asociados con la RSC y SC para confirmar si existe un cambio teórico entre RSC y SC. Desde una perspectiva práctica, se realiza un estudio empírico en el sector turístico de España, analizando 364 informes corporativos publicados entre 2010 y 2021. Se aplica el modelado de temas BERTopic para detectar las áreas clave relacionadas con la gestión de la RSC y SC en el sector, y la minería de texto para examinar la frecuencia de uso

de términos asociados con cada temática, permitiendo comprender su presencia y relevancia en los informes a lo largo del tiempo. Finalmente, se analiza el contenido textual de las memorias corporativas con el modelo de procesamiento del lenguaje natural (PLN) FinBERT, para cuantificar la importancia que los subsectores y empresas turísticas otorgan a distintos aspectos relacionados con ESG (por ejemplo, cambio climático, capital humano, capital natural, entre otros).

Las contribuciones de esta investigación enriquecen la literatura turística actual, proporcionando un marco metodológico innovador y mejorando la comprensión de la RSC y la SC. Específicamente, se realizan las siguientes aportaciones:

- Se brinda una visión actualizada y detallada del campo teórico de RSC y SC en turismo, validando si existe un cambio de enfoque investigador desde la RSC hacia la SC. En el contexto del turismo sostenible esto es fundamental, debido a que una comprensión teórica sólida puede promover cambios coherentes que reducen la ambigüedad en la academia y en la implementación de prácticas empresariales (Miller & Twining-Ward, 2005).
- Se aporta una metodología innovadora que combina técnicas de ciencia de datos para el análisis de la sostenibilidad en turismo, siendo replicable en futuros estudios. En este sentido, se mejora el enfoque de la investigación turística, puesto que los estudios que aplican ciencia de datos y el análisis de *big data* son limitados (Egger, 2022; Mariani et al., 2018; Mariani & Baggio, 2022).
- Se aumenta el conocimiento científico sobre cómo las empresas turísticas españolas abordan la teoría de la RSC y SC. Investigaciones recientes, como la de Koh (2024) y Back (2024), evidencian que se requieren mayores estudios con un enfoque longitudinal y sectorial. Esto es útil para evaluar la evolución a lo largo del tiempo e identificar la importancia que distintos subsectores le otorgan a la sostenibilidad.

Por tanto, este análisis reduce ese *gap* de la literatura existente y comprueba si hay una transición de la estrategia empresarial turística centrada en la RSC hacia la SC.

## **2. Revisión de la literatura**

### **2.1. Responsabilidad Social Corporativa y Sostenibilidad Corporativa: ¿Cambio de paradigma?**

A lo largo de décadas, se han analizado y debatido los conceptos, interpretaciones y las delimitaciones entre sostenibilidad y responsabilidad social corporativa (RSC) (Damtoft et al.,

2024), revelando la complejidad y evolución de ambos términos en el ámbito empresarial y científico. En sus inicios, la sostenibilidad se enfocaba principalmente en cuestiones ambientales, poniendo énfasis en la gestión responsable de los recursos naturales (Caradonna, 2014). Por el contrario, las raíces de la RSC se vincularon más con la sociedad y los individuos (Strand et al., 2015). Ambos conceptos evolucionaron a lo largo del tiempo, dado que el crecimiento económico, impulsado por modelos de negocio que explotaban de manera intensiva los recursos naturales, evidenciaba las graves consecuencias para la sociedad y el planeta (Rajeev et al., 2017).

El Informe Brundtland de la Comisión Mundial sobre el Medio Ambiente y Desarrollo marcó un cambio en la historia y concibió el desarrollo sostenible como aquel que satisface las necesidades del presente sin comprometer la capacidad de satisfacción de las futuras generaciones (WCED, 1987). Diversos actores clave, como gobiernos y empresas, han asumido un papel activo en el desarrollo sostenible (Valbuena-Hernandez & Ortiz-de-Mandojana, 2022). Por tanto, las estrategias impulsadas desde el ámbito público y privado resultan fundamentales para alcanzar un equilibrio entre el crecimiento económico, la protección ambiental y el bienestar social. Las acciones públicas consolidan el marco institucional para promover políticas de cumplimiento obligatorio o voluntario, mientras que las acciones privadas reflejan el compromiso de las empresas u organizaciones con la sostenibilidad (Vidal, 2011). La aplicación del concepto de desarrollo sostenible en el ámbito empresarial dio lugar a la Sostenibilidad Corporativa (SC) (Steurer et al., 2005). Sin embargo, la RSC ha representado históricamente una estrategia empresarial para gestionar aspectos sociales, económicos y ambientales (Montiel, 2008; Van Marrewijk, 2003).

La discusión sobre la responsabilidad social de las organizaciones se ha dado a lo largo de la historia. Académicos como Clark (1939) y Weber (1992) decidieron demostrar a los empresarios que era posible implementar un modelo de gestión competitivo basado en acciones responsables desde el punto de vista económico, cumpliendo con los compromisos financieros adquiridos con empleados, accionistas y proveedores (Barrena Martínez et al., 2016). Bowen (1953) fue reconocido como el padre del término “responsabilidad social”, y su trabajo hizo que la temática ganara más relevancia e interés en el ámbito científico (Barrena Martínez et al., 2016). Este autor definió la responsabilidad social como las obligaciones morales y personales que una empresa debería seguir y que van en línea con los valores y objetivos deseados por la sociedad. Posteriormente, diversos académicos han contribuido a la evolución del concepto

(Costa et al., 2022). Por ejemplo, Carroll (1979) indicó que la RSC da respuesta a cuatro expectativas que la sociedad espera de una organización en un momento determinado (económicas, legales, éticas y filantrópicas). Freeman (1984) aportó la teoría de los *stakeholders*, en la que se afirma que las empresas tienen responsabilidades con sus grupos de interés, por lo que se deben gestionar adecuadamente los impactos generados, minimizando los negativos e incrementando los positivos.

Garriga & Melé (2004) clasifican en cuatro grupos las teorías más relevantes sobre RSC, incluyendo dentro de las teorías éticas la subcategoría de desarrollo sostenible:

- Teorías Instrumentales: Consideran a la RSC como una herramienta estratégica para lograr los objetivos económicos y la creación de riqueza. Existen tres sub-categorías: Maximización del valor de los accionistas; Estrategias para las ventajas competitivas; Marketing con causa.
- Teorías Políticas: Se centran en las interacciones y conexiones entre la empresa y la sociedad, en el poder y posición de la empresa y su responsabilidad inherente. Se dividen en: Constitucionalismo corporativo; Teoría de la integración del contrato social; Empresa ciudadana.
- Teorías Integradoras: Se centra en cómo las empresas integran las demandas sociales para obtener un mayor grado de legitimidad social y prestigio. Se considera que las empresas dependen de la sociedad para su existencia, permanencia y desarrollo. Las teorías de esta categoría son: Aspectos administrativos y de gestión; Responsabilidad Pública; Administración de los *stakeholders*; Desempeño Social Corporativo.
- Teorías Éticas: Se centra en los requisitos éticos que cimientan la relación entre empresa y sociedad. Se divide en cuatro sub-categorías: Teoría normativa de los grupos de interés; Derechos humanos; Desarrollo sostenible; Enfoque del bien común.

Durante las últimas décadas se ha evidenciado un aumento en la literatura sobre RSC y se estima que más del 40% de los artículos científicos sobre esta temática han sido publicados desde el año 2005 (Aguinis & Glavas, 2012). Sin embargo, la teoría de la RSC no ha tenido un desarrollo óptimo que le permita alcanzar todo su potencial (Andreu Pinillos et al., 2018). Esto se puede deber en gran parte a que su evolución ha sido lenta y se ha dado en diferentes lugares del mundo, originándose una pluralidad de significados sin un consenso sobre su definición

(Andreu Pinillos et al., 2018; Matten & Moon, 2008; Sheehy & Farneti, 2021). La Organización Internacional de Estandarización (*International Organization for Standardization, ISO*, en sus siglas en inglés) destaca por su iniciativa de crear un concepto universal de RSC (Villafán Vidales, 2020). Durante cinco años logró reunir tanto a organizaciones como a expertos y observadores de 99 países para consensuar una definición universal. El resultado de dicha iniciativa dio origen en el año 2010 a la Norma ISO26000 o Guía de Responsabilidad Social. Esta norma establece que la responsabilidad social de una empresa contempla la gestión de los impactos que genere en la sociedad y medio ambiente como consecuencia del desarrollo de sus actividades, mediante un comportamiento ético y transparente que favorezca al desarrollo sostenible; integre las expectativas de sus *stakeholders*; cumpla con la normatividad y legislación aplicable; interiorice la RSC en toda la entidad y se practique en sus relaciones (ISO, 2010).

Adicionalmente, se han atribuido muchos impactos negativos en el medio ambiente y sociedad a las empresas, por lo que éstas se han visto obligadas a implementar políticas de RSC y sostenibilidad en sus estrategias, siendo objeto de múltiples estudios para evaluar y determinar su compromiso con el desarrollo sostenible (Costa et al., 2022; Lozano, 2018). Esta situación ha generado un “efecto halo” en torno al concepto de RSC, limitando su alcance, asociándola más con la filantropía corporativa y siendo reconocida como una táctica para mejorar la reputación empresarial (Andreu Pinillos et al., 2018). En este contexto, y gracias al auge del Informe Brundtland (WCED, 1987), la Sostenibilidad Corporativa (SC) surge como un concepto capaz de superar las limitaciones de la RSC, reemplazando el modelo de gestión empresarial utilizado (Andreu Pinillos et al., 2018).

Van Marrewijk (2003) define la SC como un enfoque estratégico que integra las dimensiones económicas, sociales y ambientales dentro de la gestión y operaciones de la empresa, adaptándose al contexto particular de cada organización. Es decir, que ésta debe ajustarse según el nivel de desarrollo y las aspiraciones de cada empresa, abarcando desde el cumplimiento normativo básico hasta un enfoque holístico que integra la sostenibilidad en todas las áreas operativas y estratégicas. La definición propuesta por el *Dow Jones Sustainability Index (DJSI)*, indica que la SC es un enfoque de negocio que busca crear valor a largo plazo para los accionistas a través del aprovechamiento de oportunidades y la gestión eficaz de riesgos inherentes al desarrollo económico, social y medio ambiental (Andreu Pinillos & Fernández Fernández, 2011).

Las estrategias de SC se pueden clasificar en cuatro, según el grado de integración o aplicación que una organización tenga sobre diferentes aspectos sostenibles (Baumgartner,

2005, 2009; Baumgartner & Ebner, 2010; Dyllick, 2000; Hardtke & Prehn, 2001; Schaltegger & Dyllick, 2002):

- Estrategia de mitigación de riesgos o introvertida: Orientación al cumplimiento de la normatividad y legislación de ámbito social y medio ambiental, con el fin de reducir riesgos.
- Estrategia legitimadora o extrovertida: Enfoque en las relaciones con los grupos de interés para poder desarrollar la actividad empresarial.
- Estrategia de eficiencia o conservadora: Desarrollo de la actividad con un enfoque en la ecoeficiencia y la producción más limpia, con el objetivo de minimizar el impacto ambiental.
- Estrategia de sustentabilidad holística o visionaria: Integración de la sostenibilidad en todos los ámbitos de la empresa.

En ocasiones, las empresas desarrollan estrategias de sostenibilidad de mitigación de riesgos o legitimadoras para dar respuesta a las exigencias del mercado y el cumplimiento de normas. De acuerdo con la investigación de Othman et al. (2011), las nuevas tendencias en los informes de RSC o SC pueden surgir de: 1) Isomorfismo coercitivo: presión externa ejercida a las empresas por otras entidades de las que dependen, como el estado o las expectativas de la sociedad; 2) Isomorfismo normativo: intención normativa de la propia empresa por desempeñar un comportamiento responsable y transparente ante sus *stakeholders* o 3) Isomorfismo mimético: deseo de imitar a otras empresas exitosas.

Es habitual que los términos RSC y SC se utilicen indistintamente, sin una clara distinción (Bansal & Song, 2016; Montiel, 2008; Van Marrewijk, 2003). Según Van Marrewijk (2003), existe una relación jerárquica entre RSC y SC, donde la RSC es un paso intermedio hacia la SC. Este autor explica que la RSC se asocia con la “comuni3n”, es decir, la transparencia, la comunicaci3n con los grupos de inter3s y la responsabilidad social. Por el contrario, la SC hace referencia a la “agencia”, que abarca tanto la creaci3n de valor como la gesti3n medio ambiental y del capital humano. Sheehy & Farneti (2021) indican que la SC permite que las organizaciones contemplen un mayor n3mero de acciones para proteger el medio ambiente y minimiza los impactos sociales derivados de sus actividades con una visi3n a mediano y largo plazo.

Por su parte, Chang et al. (2017) identifican en orden cronol3gico las cuatro principales teor3as relacionadas con la sostenibilidad en el contexto empresarial: 1) Responsabilidad Social Corporativa, 2) Teor3a de las partes interesadas, 3) Sostenibilidad Corporativa, 4) Econom3a verde. Tambi3n se3alan que las teor3as m3s recientes son la de Coevoluci3n y el Enfoque

Multinivel. Los autores destacan que las teorías han pasado de centrarse en el “qué” hacia el “cómo”. Las teorías tradicionales, como la RSC y la SC, se enfocan en definir qué es la sostenibilidad y qué prácticas deberían adoptar las empresas para mejorar sus impactos sociales, ambientales y económicos. El enfoque más reciente se centra en “cómo” implementar la sostenibilidad de manera efectiva en las operaciones empresariales. Teorías como la Economía Verde y la Teoría de la Coevolución se enfocan en los mecanismos prácticos mediante los cuales las empresas pueden integrar la sostenibilidad en sus actividades diarias y adaptarse a un entorno en constante cambio.

Por su parte, Scherer & Palazzo (2011) señalan que la RSC asume un rol más amplio y político, en el que las empresas no solo siguen las leyes y regulaciones, sino que también juegan un papel activo en la gobernanza global. De esta forma, la RSC se convierte en una herramienta de gobernanza y de democracia, donde las empresas tienen una influencia directa en las decisiones políticas que afectan a la sociedad y al medio ambiente. No obstante, Burbano et al. (2023) precisan que la investigación sobre SC más reciente incluye temas ambientales, sociales y de gobernanza (*Environmental, Social and Governance, ESG*, en sus siglas en inglés).

## **2.2. RSC y SC en el sector turístico**

La aplicación de la RSC y la SC en turismo conlleva a que las empresas u organizaciones cumplan con sus objetivos económicos y respondan de manera efectiva a las demandas sociales y ambientales (Font et al., 2012; Font & Lynes, 2018; S. Lee et al., 2022). El turismo es una de las fuerzas más importantes que dan forma al mundo (Cohen & Kennedy, 2017). Es una industria que impacta la economía y sociedad, dado que proporciona una infraestructura, genera empleo y aumenta la calidad de vida de los individuos (Henderson, 2007). Sin embargo, impacta negativamente al planeta mediante el alto uso de recursos naturales, contaminación o concentraciones de residuos, entre otros. Asimismo, debido a las características propias del turismo se considera que la sociedad es parte del producto, por lo que las empresas turísticas adquieren responsabilidades con las sociedades y entornos de los lugares donde desarrollan su actividad (Henderson, 2007). Esto conlleva a que las organizaciones de este sector sean más sensibles a las políticas de RSC y SC, lo cual contribuye al desarrollo sostenible, la protección de los recursos amenazados y, a su vez, cumple con las nuevas expectativas de los consumidores, el movimiento ambiental y las normativas (Henderson, 2007). Garrigos-Simon et al. (2018) puntualizan que los aspectos morales, éticos y ambientales se convierten en componentes integrales de la política y estrategia turística.

Durante los últimos años, se evidencia un interés creciente en la literatura sobre RSC y SC en turismo (Font & Lynes, 2018; Lechuga Sancho et al., 2020; Madanaguli et al., 2022). Este interés creciente se observa a partir de 2007 pero, especialmente, desde 2018 (Lechuga Sancho et al., 2020; Madanaguli et al., 2022). Los estudios previos han abarcado diversidad de temas asociados con la RSC como la gestión ambiental, el rendimiento financiero, la perspectiva desde marketing, el comportamiento de turistas, la lealtad del cliente o la reputación corporativa (Aragon-Correa et al., 2015; Font et al., 2012; Martínez & Rodríguez del Bosque, 2013, 2014; Sánchez-Camacho et al., 2022; Wut et al., 2023). El estudio de Yoopetch et al. (2023) refleja cuatro escuelas de pensamiento de investigación sobre RSC: 1) desempeño o resultados de la RSC, 2) política de RSC, 3) ética empresarial, 4) turismo y medio ambiente sostenibles. Asimismo, la investigación de Sánchez-Camacho et al. (2022) revela que la investigación sobre RSC en turismo abarca ocho grandes áreas: 1) actividad empresarial y rendimiento financiero; 2) actitud y comportamiento; 3) perspectiva de marketing; 4) turismo sostenible; 5) sostenibilidad; 6) acción gubernamental; 7) conciencia y motivación ambiental; y 8) impacto de grandes eventos.

De acuerdo con Back (2024), la RSC en la industria turística se ha transformado en prácticas estándar de gestión sostenible bajo los criterios ESG. Este enfoque es una de las formas más comunes de implementar y materializar la SC. Sin embargo, autores como Legendre et al. (2024) sostienen que, aunque los criterios ESG han ido ganando popularidad, la transición de RSC a ESG no se ha dado por completa, y ambos términos a menudo se utilizan indistintamente. Por consiguiente, es necesario desarrollar más estudios que brinden una visión actualizada y detallada sobre la RSC y la SC en el sector turístico, evidenciando si existe un cambio de paradigma de RSC a SC.

### **2.3. Aplicación de la ciencia de datos en informes de RSC y SC**

Los informes emitidos por las empresas representan el canal más eficaz para dar a conocer su desempeño económico, social y ambiental. Durante las últimas décadas han surgido diversos tipos de reportes corporativos, como memorias anuales o de gestión; informes integrados; reportes de sostenibilidad, RSC o gobierno corporativo (Girella et al., 2019). Dentro de estos informes, sobresalen actualmente los de sostenibilidad, dado que se han convertido en una herramienta que mejora el diálogo con los *stakeholders* y la sociedad en general. En estos documentos las empresas declaran su compromiso con el desarrollo sostenible a la vez que comunican las acciones y estrategias sobre RSC (C. W. Hsu et al., 2013; Miralles-Quiros et al., 2017). Por tanto, las empresas han aumentado la divulgación de informes para transmitir

información clara, oportuna y comprensible a sus *stakeholders*. Según Font et al. (2012), la cantidad y calidad de los informes de RSC son un indicador esencial para conocer cómo de responsable es una industria. Esta situación ha permitido llevar a cabo investigaciones empíricas donde se analiza el auge y evolución de la sostenibilidad y la responsabilidad. Uno de los estudios que confirman el auge de la sostenibilidad en la comunicación de las empresas es el de Gatti & Seele (2014), puesto que examinaron la terminología de los títulos de 329 informes de RSC y SC de 50 empresas europeas entre 1998 y 2010. En ese estudio se considera que los títulos sintetizan los principales temas del informe y reflejan la forma en que las empresas interpretan el concepto de RSC y su enfoque.

También existen estudios con un enfoque multidisciplinar, en los que se ha aplicado la ciencia de datos a informes de RSC y SC. Liu et al. (2017) llevaron a cabo una investigación en la industria petroquímica con técnicas de minería de textos para extraer información valiosa de informes de RSC publicados entre 2012 y 2015 de las 50 principales empresas del sector. El estudio de Szekely & Vom Brocke (2017) también resalta la importancia y el potencial del procesamiento del lenguaje natural (PNL) y el análisis de *big data*. Se empleó la minería de textos para hacer un modelado de temas en 9.514 informes de sostenibilidad publicados entre 1999 y 2015, con el fin de identificar temas comunes y las prácticas de sostenibilidad de las empresas. Por otro lado, Landrum & Ohsowski (2018) buscaban comprender las cosmovisiones de la SC a través del análisis de contenidos de 230 informes de sostenibilidad y RSC de empresas con sede en América del Norte. Cada informe se posicionó en una etapa de SC mediante el conteo de la frecuencia relativa de las palabras claves asociadas a cada etapa y se utilizó un paquete de minería de textos del lenguaje de programación R. Asimismo, se han analizado los mensajes del CEO (*Chief Executive Officer*, en sus siglas en inglés) o director general de las memorias de sostenibilidad, combinando técnicas de minería de textos y de sentimientos (Na et al., 2020).

Con relación al sector turístico, Koseoglu et al. (2021) afirman que las empresas con mejor desempeño en RSC tienen una mayor propensión a emitir informes de RSC, adoptando el marco GRI (*Global Reporting Initiative*, en sus siglas en inglés). Esto apoya la teoría de la señalización, que sugiere que las empresas utilizan los informes de RSC para mostrar su compromiso social y ambiental. Algunos estudios previos evidencian que en la literatura turística existente sobre informes RSC y SC se tiende a usar en gran proporción el análisis de contenido tradicional y manual (Akmese et al., 2016; de Grosbois, 2016; Guix et al., 2025; Nyahunzvi, 2013). Sin embargo, no se emplean técnicas avanzadas o grandes volúmenes de información para analizar

los informes corporativos y extraer información relevante. Es notable que la investigación turística aún cuenta con un número limitado de estudios que utilicen la automatización en el análisis de contenido, lo cual podría ser una herramienta valiosa para superar las limitaciones inherentes al análisis y procesamiento manual. Dado el escaso número de estudios en esta línea, destaca el trabajo de Uyar et al. (2020). Los autores analizan la estructura temática de los informes de sostenibilidad en la industria turística desde 1999 hasta 2018, mediante técnicas de minería de textos y análisis de redes. Sus resultados muestran que durante el periodo de 1999 y 2009 los temas sociales dominaban sobre los temas ambientales. Posteriormente, en los últimos años 2015 a 2018 se refleja un equilibrio entre ambos. Los autores también subrayan que temas como la biodiversidad y la energía renovable siguen siendo tratados de manera superficial.

### **3. Preguntas e hipótesis de investigación**

Dado que esta investigación busca analizar la percepción y evolución de la Responsabilidad Social Corporativa (RSC) y la Sostenibilidad Corporativa (SC) en turismo desde el ámbito teórico y práctico, se ha dividido en dos partes y se han establecido distintas preguntas e hipótesis de investigación. La primera parte se enfoca en el ámbito teórico, proporcionando una base sólida y una visión detallada y actualizada de la literatura turística en torno a las temáticas de estudio. A través de la metodología de revisión computacional de la literatura, que emplea el modelado de temas, se identifican las principales corrientes de investigación y su evolución a lo largo del tiempo. Además, se realiza una validación estadística de la evolución de los temas identificados, específicamente en relación con la RSC y la SC, para examinar si ha habido cambios teóricos significativos entre ambos conceptos. Para alcanzar este propósito exploratorio de la literatura, se han formulado las siguientes dos preguntas de investigación:

**P1:** ¿Qué temas se han estudiado en la literatura existente sobre Responsabilidad Social Corporativa y Sostenibilidad Corporativa en turismo durante los últimos veinte años?

**P2:** ¿Existe un cambio de paradigma teórico desde la Responsabilidad Social Corporativa hacia la Sostenibilidad Corporativa en turismo?

La segunda parte del estudio se enfoca en un análisis empírico de 364 informes corporativos de empresas turísticas en España, empleando tanto métodos exploratorios como estadísticos. Primero, se aplica el modelado de temas con BERTopic para detectar las áreas clave relacionadas con la gestión de la RSC y SC en el sector. Luego, se utiliza la minería de texto para examinar la

frecuencia de términos asociados a cada temática, lo que permite comprender su presencia y relevancia en los informes a lo largo del tiempo y observar si existe una evolución en el enfoque empresarial desde la RSC hacia la SC. Por último, y con el objetivo de explorar en profundidad los informes, se analiza el contenido textual de las memorias corporativas con el modelo de procesamiento de lenguaje natural FinBERT, que cuantifica la importancia que los distintos subsectores y empresas turísticas otorgan a diversos criterios ESG (*Environmental, Social and Governance*, en sus siglas en inglés), por ejemplo, cambio climático, capital humano, capital natural, entre otros.

De este modo, se opta por formular preguntas e hipótesis de investigación, ya que ambas herramientas permiten abordar distintas dimensiones del análisis de manera complementaria. Las siguientes preguntas de investigación permiten guiar el análisis exploratorio de los informes corporativos, facilitando el descubrimiento de patrones y relaciones:

**P3:** ¿Cuáles son las áreas claves relacionadas con la gestión de la Responsabilidad Social Corporativa y la Sostenibilidad Corporativa de los informes corporativos de las empresas turísticas?

**P4:** ¿Cuál es la importancia que las empresas y subsectores turísticos otorgan a los criterios ESG en los informes corporativos del sector turístico español?

Por otro lado, las siguientes dos hipótesis de investigación permiten comprobar estadísticamente si ha habido un cambio significativo a lo largo del tiempo en las categorías de RSC y SC mediante el análisis de las palabras utilizadas en los informes.

**H1:** Existe un cambio significativo a lo largo del tiempo en la categoría “Responsabilidad Social Corporativa” a través de las palabras utilizadas en los informes corporativos en las empresas turísticas.

**H2:** Existe un cambio significativo a lo largo del tiempo en la categoría “Sostenibilidad Corporativa” a través de las palabras utilizadas en los informes corporativos en las empresas turísticas.

Este enfoque, que combina preguntas e hipótesis de investigación, permite un análisis tanto exploratorio como confirmatorio. Las preguntas facilitan la identificación de patrones y tendencias, mientras que las hipótesis aportan una base para analizar estadísticamente cambios significativos en el uso de términos asociados con cada temática (RSC y SC). Esto permite

comprender su presencia y relevancia en los informes corporativos de las empresas turísticas españolas a lo largo del tiempo.

#### 4. Metodología de la investigación

##### 4.1. Metodología aplicable al enfoque teórico de la revisión computacional de la literatura

Para el desarrollo del estudio teórico se emplea una revisión computacional de la literatura. Esta metodología emplea técnicas de minería de texto y algoritmos para analizar y sintetizar grandes volúmenes de literatura científica de un campo específico (Antons et al., 2021). Los pasos para llevarla a cabo son: 1) establecimiento de los objetivos conceptuales (explicar, relacionar, explorar, debatir; 2) planificación de la revisión especificando el ámbito de conocimiento, el corpus de textos, el alcance, el período de análisis y las bases de datos a utilizar; 3) elección de las técnicas computacionales (algoritmos de Inteligencia Artificial) según los objetivos y recursos; 4) realización de los análisis de contenido, limpieza de datos y verificación de resultados; 5) síntesis y validación de los resultados obtenidos y 6) presentación de los resultados de forma clara y replicable (Antons et al., 2021).

De esta forma, el análisis teórico tiene como fin explorar la estructura y evolución de la red de conocimiento científico sobre la Responsabilidad Social Corporativa (RSC) y la Sostenibilidad Corporativa (SC) en turismo. Por consiguiente, se descarga la información de la base de datos Scopus, reconocida por su alto nivel de calidad (Mortenson & Vidgen, 2016). La búsqueda se realiza en inglés empleando conceptos clave y comillas para limitar los resultados a aquellos que incluyan exactamente las expresiones seleccionadas. Se emplea el operador lógico *OR* para incluir variaciones de los conceptos, y el operador *AND* para combinar diferentes áreas de interés. La búsqueda queda configurada de la siguiente manera: (*“corporate social responsibility” OR “CSR”*) *OR* (*“corporate sustainability” OR “holistic sustainability” OR “ESG” OR “environmental, social and governance” OR “sustainable strategy” OR “Sustainable Business”*) *AND* (*“tourism” OR “travel” OR “hospitality”*). El marco temporal de análisis establecido es del 2004 al 2023, puesto que la información de años anteriores es muy limitada y podría sesgar el análisis. Como resultado, se obtienen 1.505 documentos compuestos por artículos de revistas, actas de congresos, libros y capítulos de libros.

Por último, se establece que la técnica computacional a emplear es el modelado de temas, ya que es una herramienta útil para la revisión de literatura en áreas con un elevado número de publicaciones, identificando de forma automatizada y precisa de los temas clave a partir de resúmenes (Mortenson & Vidgen, 2016). Concretamente, se aplica la técnica de *machine*

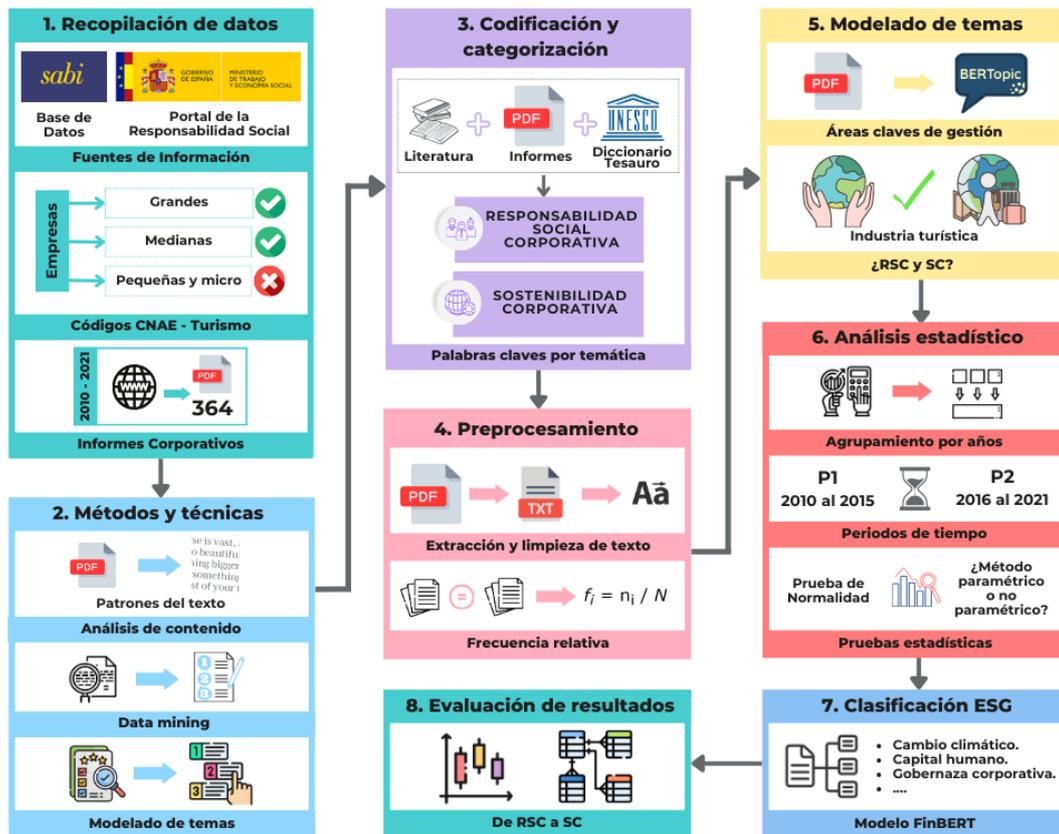
*learning* de modelado de temas BERTopic, dado que es capaz de capturar el contexto semántico (Grootendorst, 2021). El preprocesamiento y análisis posteriores de los datos se realiza con el lenguaje de programación Python. Después, de ejecutar el modelado de temas se elabora una categorización manual de los temas detectados y se agrupan por áreas de investigación más amplias. Una vez identificados los temas asociados con las temáticas de estudio (RSC y SC), se efectúan regresiones lineales para capturar los cambios longitudinales y establecer si hay una variación estadísticamente significativa.

#### **4.2. Metodología aplicable al enfoque empírico de la revisión de las memorias corporativas de las empresas turísticas españolas**

La investigación empírica se desarrolla a partir de las memorias anuales, reportes de RSC y sostenibilidad, e informes integrados de las empresas turísticas de España. Es decir, se contemplan diversos documentos estratégicos que contienen información relevante para este tipo de análisis (Ning et al., 2021). La metodología utilizada para examinar estos informes es el análisis de contenidos, ya que permite estudiar los mensajes o patrones de un texto para interpretar su significado (Landrum & Ohsowski, 2018). Este enfoque de análisis textual es una metodología indirecta que identifica la frecuencia de los conceptos, como palabras o frases de fuentes documentales ya existentes (Guix Oliver, 2008; Landrum & Ohsowski, 2018), como pueden ser los informes corporativos. Además, el análisis de contenidos se ha aplicado a informes empresariales (de RSC, anuales y sostenibilidad) en diferentes investigaciones sobre sostenibilidad y RSC (Asif et al., 2013; Campopiano & de Massis, 2015; Einwiller & Carroll, 2020; Ellili & Nobanee, 2022; Gatti & Seele, 2014; Torelli et al., 2020). No obstante, esta investigación se distingue por aplicar técnicas de ciencia de datos para llevar a cabo el análisis de contenido de manera automatizada y computacional.

Al igual que en el estudio teórico, el preprocesamiento, análisis y evaluación de la información se realiza mediante el lenguaje de programación Python y se desarrolla un script con diferentes órdenes y funciones para aplicar la técnica de minería de textos. Se utiliza una de las librerías más conocidas de Python denominada Pandas, dado que es un código abierto especializado en el manejo y análisis de estructura de datos (*Pandas - Python Data Analysis Library*, n.d.). Igualmente, se emplea la librería NLTK (Natural Language Toolkit) que contiene funciones para el análisis de datos de lenguaje humano y procesamiento de texto como clasificación, tokenización o derivación. En la Ilustración 24 se presenta el marco metodológico del estudio empírico.

Ilustración 24. Marco metodológico de la investigación empírica.



Fuente: Elaboración propia.

#### 4.2.1. Recopilación de datos

La selección de las empresas se lleva a cabo en abril de 2022 y se toma como referencia la base de datos del Sistema de Análisis de Balances Ibéricos (SABI, 2022). Se filtran los códigos CNAE asociados a las actividades de turismo (hoteles y alojamientos, transporte y restauración), así como las empresas activas y con página web. Posteriormente, se seleccionan las grandes y medianas empresas (435) para la búsqueda de informes corporativos, puesto que la probabilidad de que estas empresas publiquen informes es más alta que las pequeñas y micro, ya que son más visibles y los grupos de interés generan más presión para que divulguen sus resultados económicos, sociales y ambientales (Branco & Rodrigues, 2008). Además, otros estudios han demostrado una relación positiva entre el tamaño de la organización y su divulgación de la RSC, es decir, que cuanto más grande la empresa, mejor es la divulgación de la RSC y sostenibilidad (Branco & Rodrigues, 2008; H. Khan, 2010; Reverte, 2009). La búsqueda se complementa con la información del Portal de la Responsabilidad Social de España (Ministerio de Trabajo y Economía Social, 2022).

En total, se descargan 404 informes del año 2000 al 2021. Sin embargo, la información de 2000 al 2009 es escasa para aplicar los análisis, comparaciones y pruebas estadísticas pertinentes. Por tanto, se establece como período de tiempo de análisis del 2010 a 2021 con un total de 364 informes corporativos. Asimismo, se divide en dos periodos de tiempo: P1 del año 2010 al 2015 y P2 del año 2016 al 2021 por dos razones. La primera, para comparar periodos de tiempo con el mismo número de años. La segunda, es que en septiembre del año 2015 la Asamblea General de la Organización de las Naciones Unidas (ONU, 2015) marcó un cambio hacia la sostenibilidad a nivel mundial con la aprobación de la Agenda 2030 y se involucró al sector privado de manera directa con los Objetivos de Desarrollo Sostenible (ODS) (Pizzi et al., 2020). Por esta razón, durante ese año se produjo un cambio a nivel empresarial en políticas y estrategias sobre RSC y SC. Además, el sector turístico se considera como uno de los principales sectores más interesados en dicha iniciativa (Rosato et al., 2021).

#### 4.2.2. Métodos y técnicas de investigación

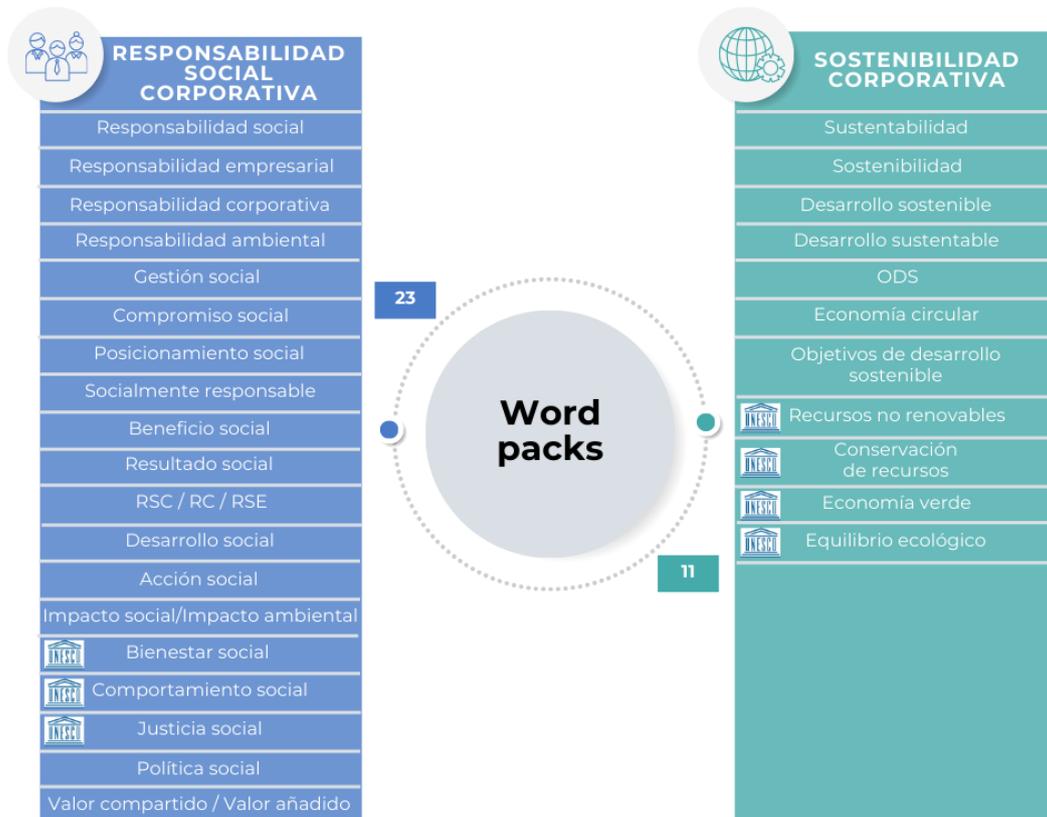
Se establece el modelado de temas como análisis complementario para identificar áreas claves asociadas a la RSC y SC, puesto que es una técnica que se encarga de extraer automáticamente los temas principales o patrones temáticos presentes en un conjunto de documentos o texto. Asimismo, se define que el método de minería de textos a emplear en el estudio será la frecuencia de palabras, debido a que se encarga de elaborar un conteo de todos los términos empleados en una determinada base de datos, obteniendo las palabras más frecuentes de dicha base de datos (Witten I. et al., 2011). Por esta razón, se considera que en un informe predomina una temática (RSC o SC) si una palabra o grupo de palabras asociadas con las mismas tiene una mayor aparición o frecuencia en el documento con respecto a las demás. Por último, se selecciona el modelo FinBERT para clasificar los textos de las memorias corporativas bajo los criterios ESG. FinBERT destaca al ser un modelo de lenguaje específico del dominio financiero, así como por su alto desempeño (Huang et al., 2022).

#### 4.2.3. Codificación y categorización

En la tercera fase se definen las palabras claves que representan la RSC y la SC mediante la codificación y categorización. Para la elaboración del sistema de categorías se utiliza el proceso deductivo-inductivo (Vives Varela & Hamui Sutton, 2021), partiendo de estudios análogos realizados previamente por autores como Karen (2008), Palazzo & Scherer (2006) o Van Marrewijk (2003b), de manera que se extraen palabras o patrones de los informes recopilados. Asimismo, se complementa con los conceptos relacionados definidos para los términos de

“Responsabilidad social” y “Desarrollo sostenible” del diccionario Tesauro (UNESCO, 2022). Estos conceptos relacionados proporcionan una lista de términos controlada, estructurada y actualizada para el análisis temático y búsqueda de documentos en diferentes campos. De tal manera, se aporta a la literatura una nueva categorización para RSC y SC (Ilustración 25). Se observa un mayor número de términos asociados a RSC que a SC, dado que los informes analizados son de empresas españolas y suelen utilizarse diversos sinónimos para referirse a la RSC, lo que a su vez puede confirmar la ausencia de un concepto unificado de RSC.

Ilustración 25. Categorización de RSC y SC.



Fuente: Elaboración propia.

#### 4.2.4. Preprocesamiento

En la cuarta fase se realiza un preanálisis de los informes corporativos para estandarizar y/o unificar los datos, ya que cada empresa tiene su propio formato y estilo. Este proceso de tratamiento de datos se considera prácticamente obligatorio para que los algoritmos de extracción de texto se puedan aplicar correctamente y no se obtengan resultados erróneos. La preparación de datos incluye técnicas de transformación, normalización, integración y limpieza de datos (García et al., 2016). Primero se extrae el texto de todos los archivos PDF a TXT y se eliminan las frases duplicadas (encabezados o pie de página). Luego se realiza una limpieza de

texto, convirtiéndolo a minúsculas y eliminando patrones no informativos (por ejemplo, símbolos, números, signos de puntuación o caracteres sueltos). Igualmente, para eliminar el sesgo de longitud de documentos, se emplea el método de la frecuencia relativa (Muller, 1973), puesto que permite reasignarle a un número de ocurrencias un valor en relación con el tamaño del texto o corpus. La frecuencia relativa es el resultado de dividir la frecuencia absoluta de un patrón o palabra en el tamaño total de la muestra (Salinas & Martínez, 2015). Este método también ha sido empleado en otras investigaciones (Landrum & Ohsowski, 2018) porque el tamaño de los informes (número de páginas) varía entre empresas o años.

#### 4.2.5. Modelado de temas basado en el enfoque BERTopic

Durante esta etapa se aplica el modelado de temas para identificar las áreas claves relacionadas con la gestión de la RSC y SC. Esta técnica de *machine learning* se utiliza para descubrir los temas centrales en documentos o textos y sobresale porque es capaz de capturar el contexto semántico. En esta investigación, se emplea el enfoque de BERTopic (Grootendorst, 2021), basado en *transformers* y el modelo de lenguaje BERT (*Bidirectional Encoder Representations from Transformers*) desarrollado por Google en 2018. BERTopic realiza cinco pasos para su proceso: 1) Incrustar documentos, 2) Reducción de dimensionalidad, 3) Agrupar documentos, 4) Bolsa de palabras, y 5) Representación del tema (Grootendorst, 2021). Dado que el estudio se centra en la RSC y SC, se utiliza el modelado de temas guiado de BERTopic. Esta técnica guía el proceso de modelado de temas al establecer temas semilla a los cuales el modelo convergerá (RSC y SC). Es decir, el modelado de temas se dirige hacia los temas semilla definidos, lo que garantiza que los temas de interés estén representados en los resultados. Esto mejora la interpretabilidad y utilidad de los temas generados, permitiendo obtener información más relevante y valiosa de los informes.

#### 4.2.6. Análisis estadístico de los datos

En la sexta fase se procede a calcular las frecuencias relativas para cada uno de los informes. Para esto se computa la frecuencia absoluta de las palabras asociadas con cada categoría y se divide en el total de palabras de cada documento. Luego, se consolidan o agrupan todas las frecuencias relativas por años y por categorías (RSC y SC). Con las frecuencias relativas y por medio del análisis estadístico se contrasta si hay diferencias significativas entre categorías. En primer lugar, es fundamental identificar si se deben aplicar métodos paramétricos o no paramétricos según la distribución que tengan los datos. Para ello, se aplican pruebas de normalidad dado que para poder emplear pruebas paramétricas se debe cumplir con este

supuesto. Por el contrario, si los datos no cumplen con el supuesto de normalidad, se deben aplicar métodos no paramétricos.

#### 4.2.7. Clasificación bajo los criterios ESG

Con el fin de detallar las prácticas sostenibles que las empresas turísticas han comunicado en sus informes corporativos a lo largo del tiempo, se aplica el modelo FinBERT para clasificar los textos en criterios ESG. Huang et al. (2022) desarrollan FinBERT, un modelo de lenguaje específico del dominio financiero basado en la arquitectura BERT y entrenado previamente con textos financieros, incluidos informes corporativos anuales y trimestrales, informes de analistas financieros y transcripciones de conferencias de resultados financieros. Este modelo fue ajustado (*fine-tuned*) para dos tareas específicas, clasificación de sentimientos y clasificación de temas ESG. El desempeño del modelo para esta última tarea alcanza una precisión del 89,5%, superando a otros algoritmos de aprendizaje automático como Naive Bayes (80,0%), SVM (Support Vector Machine, en sus siglas en inglés) (75,0%), CNN (Convolutional Neural Network, en sus siglas en inglés) (82,0%) y también ligeramente a BERT (87,0%). Por ende, FinBERT se considera un modelo idóneo para la clasificación de los criterios ESG en este estudio, ya que destaca por haber sido entrenado específicamente con textos financieros, por su robustez y alto rendimiento. Además, FinBERT se ha usado en estudios previos con éxito (Arami et al., 2023; Fedorova & Iasakova, 2024; Gao et al., 2023; Sahu et al., 2024; Sinha et al., 2022; Srijiranon et al., 2022; Unlu et al., 2023, 2024; Xiao & Ihnaini, 2023).

El modelo original de FinBERT tiene la capacidad de clasificar los textos en cuatro etiquetas: ambiental, social, gobernanza o ninguno<sup>3</sup>. Sin embargo, el modelo ajustado permite una clasificación más detallada, distribuyendo los textos en nueve categorías: cambio climático, capital natural, contaminación y residuos, capital humano, responsabilidad del producto, relaciones con la comunidad, gobernanza corporativa, ética y valores empresariales y no ESG<sup>4</sup>. Con el objetivo de cuantificar la importancia que las empresas turísticas españolas conceden a los criterios ESG y obtener un mayor nivel de detalle, se aplica el modelo ajustado a través de la librería *transformers* de *Hugging Face* en Python. Para ello, se divide el texto de las memorias corporativas por oraciones y se obtienen puntuaciones para cada una de las nueve categorías ESG.

---

<sup>3</sup> Modelo disponible en: <https://huggingface.co/yiyanghkust/finbert-esg>

<sup>4</sup> Modelo disponible en: <https://huggingface.co/yiyanghkust/finbert-esg-9-categories>

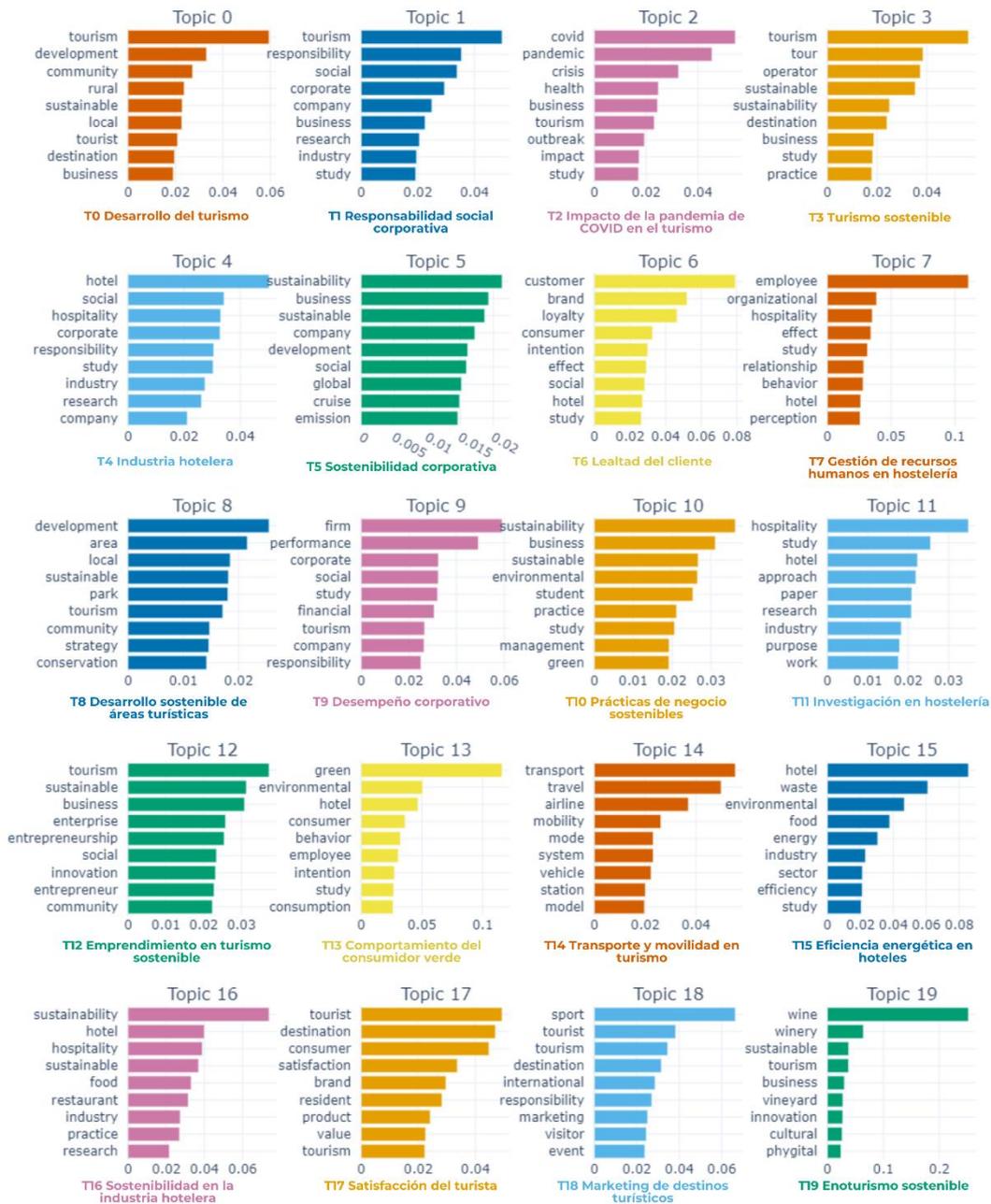
## **5. Resultados**

### **5.1. Resultados del enfoque teórico: Revisión computacional de la literatura**

Los resultados del modelado de temas en la revisión computacional de la literatura se muestran en la Ilustración 26. Las áreas de estudio de la RSC y la SC en turismo se identifican en 20 temas distintos. Por cada tema identificado se obtienen las palabras más representativas y el puntaje c-TF-IDF, que representa la importancia de una palabra dentro de un grupo (relevancia y frecuencia en ese grupo específico). Cuanto mayor sea la puntuación, más representativa es la palabra de ese tema. Por ende, los puntajes de importancia de las palabras varían entre los diferentes temas. El tema principal identificado (T0) está relacionado con el desarrollo del turismo, seguido por la responsabilidad social corporativa (T1). Esto sugiere una preocupación constante por integrar prácticas responsables y sostenibles en la actividad turística, manifestando la importancia del desarrollo rural y la participación de las comunidades locales hacia la sostenibilidad del sector.

El tema T2 refleja el fuerte impacto de la pandemia de covid-19 en la investigación del sector, posicionándolo como el tercer tema más central. El turismo sostenible (T3) y la industria hotelera (T4) ocupan posiciones destacadas, siendo objeto de las discusiones académicas de las últimas dos décadas. Por su parte, la sostenibilidad corporativa (T5) sugiere una evolución en las estrategias empresariales del sector, adoptando enfoques más integrales y a largo plazo. Temas como la lealtad del cliente (T6), la gestión de recursos humanos en hostelería (T7) y el desarrollo sostenible de áreas turísticas (T8) subrayan la diversidad de temas asociados con la investigación sobre sostenibilidad empresarial en turismo. La Tabla 6 sintetiza y consolida los temas identificados en ocho grandes áreas o categorías de investigación. Los temas centrales de esta investigación (RSC y SC) se establecen como categorías independientes y se destacan otras áreas sólidas como el “Turismo sostenible y desarrollo”, “Hostelería y su gestión”, “Clientes y marketing”, “Desempeño corporativo”, “Pandemia covid-19” y “Transporte y movilidad”.

Ilustración 26. Temas de investigación asociados a la RSC y la SC en turismo.



Fuente: Elaboración propia/procesamiento con BERTopic.

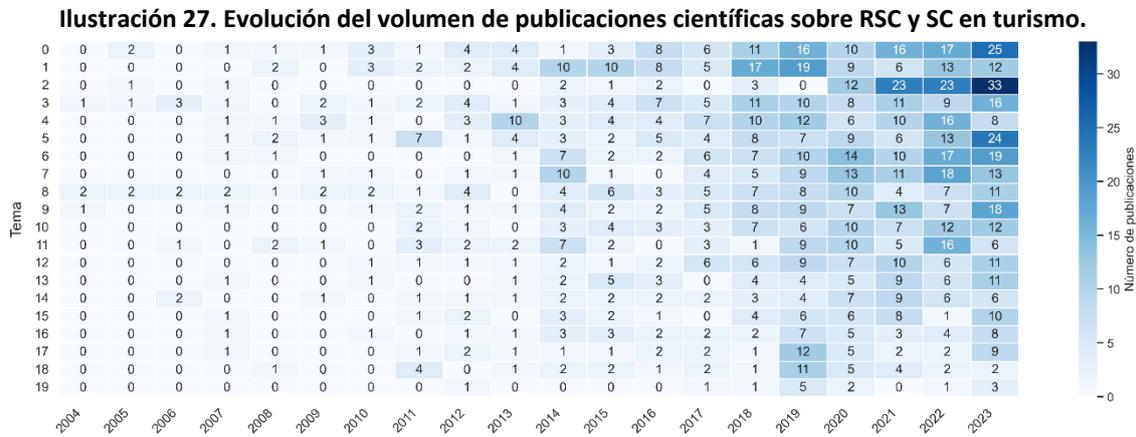
**Tabla 6. Categorías temáticas de la investigación sobre RSC y SC en turismo.**

Categoría		Tema	
C0	Responsabilidad social corporativa	T1	Responsabilidad social corporativa
C1	Sostenibilidad corporativa	T5	Sostenibilidad corporativa
		T10	Prácticas de negocio sostenibles
C2	Turismo Sostenible y Desarrollo	T0	Desarrollo del turismo
		T3	Turismo sostenible
		T8	Desarrollo sostenible de áreas turísticas
		T12	Emprendimiento en turismo sostenible
		T19	Enoturismo sostenible
C3	Hostelería y gestión hotelera	T4	Industria hotelera
		T7	Gestión de recursos humanos en hostelería
		T11	Investigación en hostelería
		T16	Sostenibilidad en la industria hotelera
		T15	Eficiencia energética en hoteles
C4	Clientes y marketing	T6	Lealtad del cliente
		T17	Satisfacción del turista
		T13	Comportamiento del consumidor verde
		T18	Marketing de destinos turísticos
C5	Desempeño corporativo	T9	Desempeño corporativo
C6	Pandemia de Covid-19	T2	Impacto de la pandemia de COVID en el turismo
C7	Transporte y movilidad	T14	Transporte y movilidad en turismo

**Fuente: Elaboración propia.**

Para entender mejor la estructura y evolución de la investigación turística sobre RSC y SC, en la Ilustración 27 se presenta la cantidad de publicaciones por tema a lo largo del tiempo. Esto permite visualizar la evolución del interés en determinados temas y su concentración anual en la producción científica. Se evidencia un aumento en el número de publicaciones en la mayoría de los temas a partir de 2018, alcanzando su punto máximo en los años más recientes (2021-2023). Este incremento predomina más en los temas 0 (Desarrollo del turismo), 2 (Impacto de la pandemia de COVID en el turismo) y 5 (Sostenibilidad corporativa), indicando un creciente interés y desarrollo de estos. Por el contrario, los temas 18 (Marketing de destinos turísticos) y 19 (Enoturismo sostenible) presentan el menor número de publicaciones durante todo el

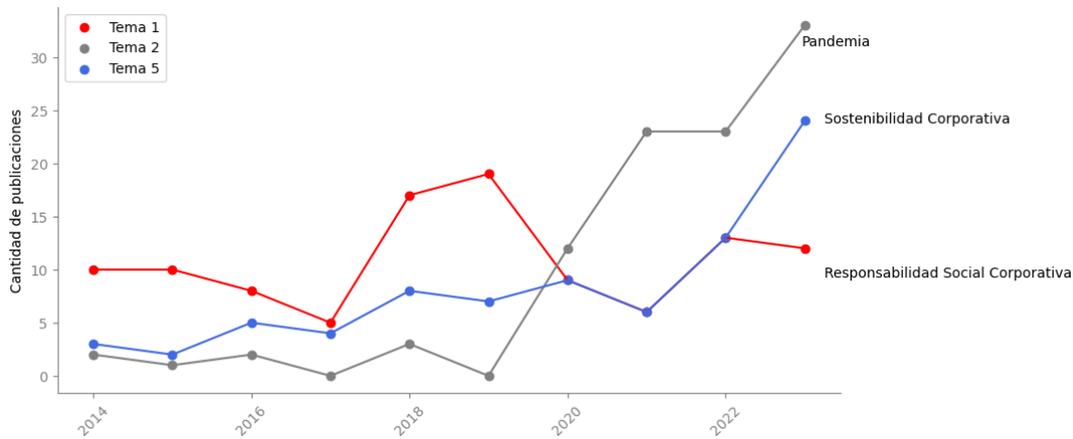
periodo de análisis. Esto podría indicar que son temas emergentes, menos explorados o con menor relevancia en la literatura existente.



**Fuente: Elaboración propia.**

La Ilustración 28 muestra la evolución de las publicaciones durante la última década (2014-2023) para los dos temas objeto de estudio, RSC (T1) y SC (T5), así como para el tema dos (T2) (Pandemia de covid-19), dada su elevada cifra de documentos de los recientes años. Los resultados indican una tendencia fluctuante pero generalmente estable a lo largo del tiempo para RSC, con determinados picos de interés, como en 2018 y 2019. No obstante, se da un leve descenso hacia 2022 y 2023. Esto podría sugerir que, aunque sigue siendo un tema sustancial, su presencia en la investigación ha alcanzado una cierta madurez, posiblemente debido a que se ha consolidado como un estándar en las prácticas empresariales en turismo. Por el contrario, la SC refleja una tendencia al alza con un crecimiento sostenido, experimentando un crecimiento significativo en 2022 y 2023, indicando un interés creciente en este tema en la investigación turística. Asimismo, la pandemia covid-19 se consolida como un tema dominante desde 2020, lo que podría sugerir que la comunidad académica ha centrado sus esfuerzos en comprender el impacto de esta crisis sanitaria global en el sector turístico.

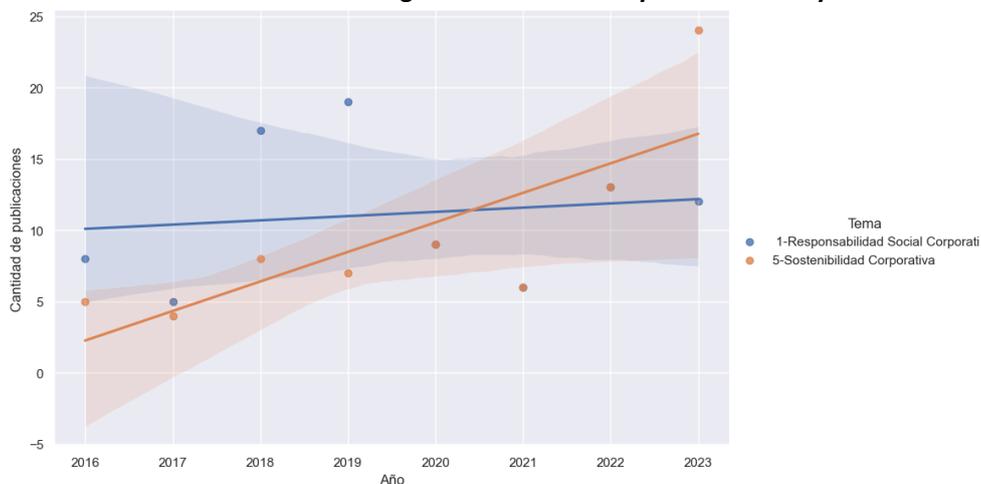
**Ilustración 28. Volumen de publicaciones científicas en la última década sobre RSC, SC y Pandemia de covid-19.**



Fuente: Elaboración propia.

Para identificar si realmente se está dando un cambio de paradigma de RSC hacia SC, se calcularon regresiones lineales para los últimos seis y ocho años. Este marco temporal es útil para captar los cambios en las prioridades de investigación, particularmente en lo que respecta a la adopción de prácticas sostenibles en el sector turístico, puesto que en 2015 se aprobó la Agenda 2030 de los Objetivos de Desarrollo Sostenible. La Ilustración 29 y la Tabla 7 detallan los resultados correspondientes a los últimos 8 años (2016-2023), donde se observa una pendiente positiva para la SC (T5) de 2,0714 y un *p-value* significativo (0,0216), lo que indica un crecimiento constante en la cantidad de publicaciones. Por otro lado, la RSC (T1) muestra una pendiente positiva moderada (0,2976), pero con un *p-value* no significativo (0,7333), lo que sugiere que no hay un crecimiento estadísticamente relevante en las publicaciones sobre este tema durante el mismo período.

**Ilustración 29. Análisis de regresión lineal de RSC y SC entre 2016 y 2023.**



Fuente: Elaboración propia.

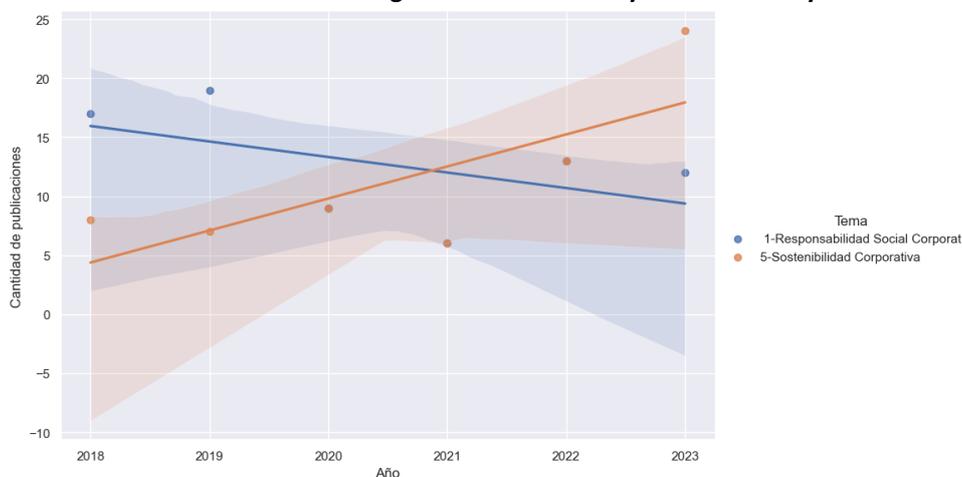
**Tabla 7. Resultados estadísticos de la regresión lineal de RSC y SC entre 2016 y 2023.**

Tema	Pendiente	<i>p-value</i>	R cuadrado	Error estándar
Responsabilidad Social Corporativa	0,297619	0,733323	0,020798	0,833702
Sostenibilidad Corporativa	2,071429	0,021589	0,612974	0,67196

Fuente: Elaboración propia.

Al reducir el período de análisis a los últimos seis años (2018-2023) (ver Ilustración 30 y Tabla 8), se observa una tendencia más marcada. La pendiente para la RSC (T1) ahora es negativa (-1,314), lo que indica una disminución en el número de publicaciones. Si bien el *p-value* de 0,3040 sugiere que esta tendencia no es estadísticamente significativa, esta pendiente negativa refleja una disminución moderada en el interés por este tema en los años más recientes. Por el contrario, la pendiente para la SC (T5) es aún más pronunciada en este período, con un valor de 2,714 y un *p-value* cercano al umbral de significancia (0,0833). A pesar de que el resultado no es estadísticamente significativo, el R cuadrado de 0,5684 sugiere que más de la mitad de la variabilidad en las publicaciones sobre SC puede explicarse por el tiempo, lo que refuerza la idea de que este tema ha ganado una relevancia considerable en los últimos años.

**Ilustración 30. Análisis de regresión lineal de RSC y SC entre 2018 y 2023.**



Fuente: Elaboración propia.

**Tabla 8. Resultados estadísticos de la regresión lineal de RSC y SC entre 2018 y 2023.**

Tema	Pendiente	<i>p-value</i>	R cuadrado	Error estándar
Responsabilidad Social Corporativa	-1,314286	0,304024	0,25763	1,115506
Sostenibilidad Corporativa	2,714286	0,083387	0,568385	1,182641

Fuente: Elaboración propia.

Por tanto, los resultados de ambos análisis revelan que la investigación turística sobre Sostenibilidad Corporativa ha mostrado un crecimiento constante y significativo, especialmente

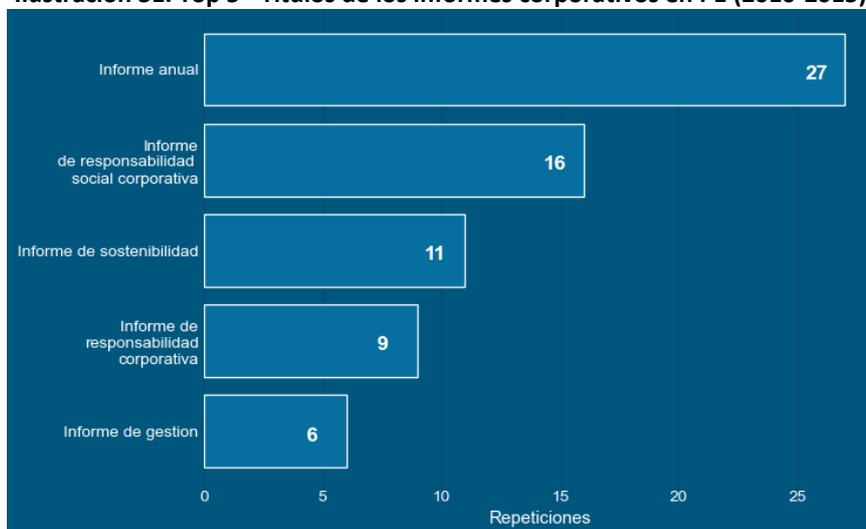
en los últimos años. Esto manifiesta un cambio de enfoque hacia la sostenibilidad como una prioridad estratégica tanto para la academia como para las empresas turísticas. Por otra parte, aunque la RSC continúa siendo un tema relevante, parece haber alcanzado un punto de estabilización o incluso una leve disminución en términos de publicaciones.

## 5.2. Resultados del enfoque empírico: Análisis de los informes corporativos de las empresas turísticas

### 5.2.1. Análisis de títulos y palabras frecuentes

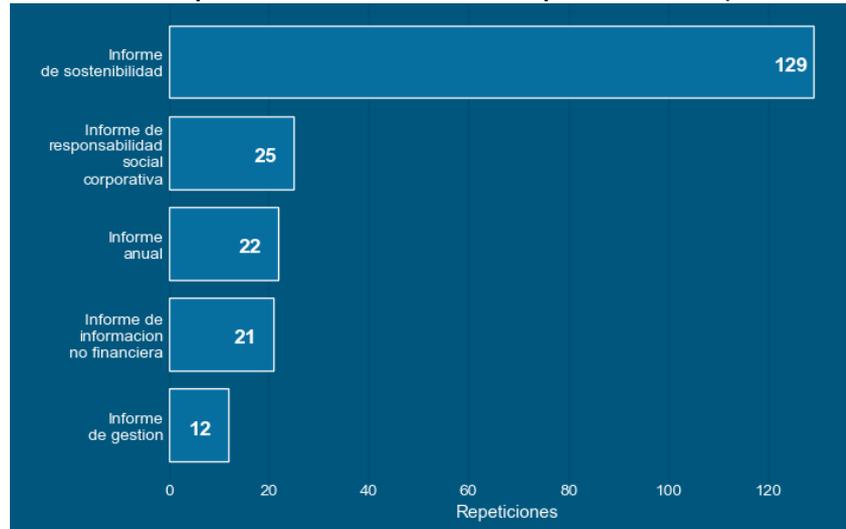
La información recopilada refleja que las empresas turísticas españolas utilizan múltiples nombres para sus informes corporativos. En la Ilustración 31 e Ilustración 32 se presentan los cinco títulos más utilizados por periodos, P1 (2010-2015) y P2 (2016-2021). Con respecto a P1, “Informe anual” se posiciona en primer lugar, seguido de “Informe de responsabilidad social corporativa” y en tercer lugar “Informe de sostenibilidad”. En P2 el título más usado es “Informe de sostenibilidad”, luego “Informe de responsabilidad social corporativa” y en tercer lugar “Informe anual”. Con el fin de sintetizar y facilitar el análisis de la información, se consideró como equivalentes los términos de memoria e informe.

**Ilustración 31. Top 5 - Títulos de los informes corporativos en P1 (2010-2015).**



Fuente: Elaboración propia.

**Ilustración 32. Top 5 - Títulos de los informes corporativos en P2 (2016-2021).**



Fuente: Elaboración propia.

De igual forma, resulta interesante analizar las palabras más frecuentes de la taxonomía de conceptos (Ilustración 25) establecida previamente por periodos de tiempo. En la Ilustración 33 y la Ilustración 34 se detallan las diez palabras más utilizadas para la categoría de RSC con el número de repeticiones, donde se observa que “Responsabilidad social”, “RSC” y “Responsabilidad Corporativa” están en los primeros lugares en ambos periodos. En P2 los términos de “Impacto ambiental” y “Valor añadido” tiene más uso que en P1.

**Ilustración 33. Palabras más frecuentes de RSC en P1 (2010-2015).**



Fuente: Elaboración propia.

**Ilustración 34. Palabras más frecuentes de RSC en P2 (2016-2021).**



Fuente: Elaboración propia.

Por otro lado, en la Ilustración 35 se representan las palabras más usadas para SC en P1. Nótese que de los once términos que componen el total de la categoría solo se visualizan ocho, dado que en este periodo los demás términos no tuvieron ninguna repetición en los informes de las empresas. En P1 la palabra “Sostenibilidad” ocupa el primer lugar, luego “Desarrollo sostenible” y “Sustentabilidad”.

**Ilustración 35. Palabras más frecuentes de SC en P1 (2010-2015).**



Fuente: Elaboración propia.

En P2 las empresas incluyeron más términos asociados con la SC en sus informes, dado que en este periodo todas las palabras de la taxonomía tuvieron frecuencia. Por lo que en la Ilustración 36 se muestran las diez palabras más usadas en P2. En primer lugar, se encuentra “Sostenibilidad”, seguido por “ODS” y luego “Desarrollo sostenible”. En este periodo el término de “Objetivos de desarrollo sostenible” tiene un aumento relevante con respecto al periodo anterior (P1), así como su abreviatura (ODS).

Ilustración 36. Palabras más frecuentes de SC en P2 (2016-2021).



Fuente: Elaboración propia.

Es importante señalar que el alto incremento de las repeticiones de los títulos y la frecuencia de palabras más usadas de P1 a P2, también se debe al mayor número de informes disponibles en el último periodo, puesto que, del total de 364 informes del estudio, hay 96 informes que corresponden a P1 (2010-2015) y 268 a P2 (2016-2021).

#### 5.2.2. Modelado de temas con el enfoque BERTopic

Para aplicar el modelado de temas guiado de BERTopic se establecen tres temas a los que convergerá el modelo. Estos temas corresponden a las temáticas de estudio (RSC, SC) y el ámbito económico, ya que la mayoría de los informes corporativos incluyen resultados económicos. Para cada uno de estos temas se definen catorce términos asociados, teniendo en cuenta la taxonomía propuesta anteriormente. Asimismo, es necesario dividir los informes en oraciones y se establecen parámetros específicos enfocados en dos aspectos. Primero, para la reducción de dimensionalidad, se crea un modelo UMAP (*Uniform Manifold Approximation and Projection*) con 15 vecinos para calcular las relaciones de proximidad, proyectando los datos en un espacio de 5 dimensiones y sin restricciones de compresión. Se emplea la métrica euclidiana para medir la similitud entre los puntos en el espacio de alta dimensión. Igualmente, se establece el parámetro “*random\_state=42*” para que el modelo se comporte de manera determinista y genere resultados reproducibles en cada ejecución con los mismos datos y parámetros. Esto facilita la comparación y replicabilidad de los resultados. Segundo, para la agrupación de los datos se utiliza el parámetro de k-Means con el objetivo de obtener 50 temas para cada período analizado (P1 y P2) y reducir el número de valores atípicos. Tercero, se asigna un título manualmente a cada uno de los temas identificados por BERTopic, basándose en las palabras más representativas de cada tema.

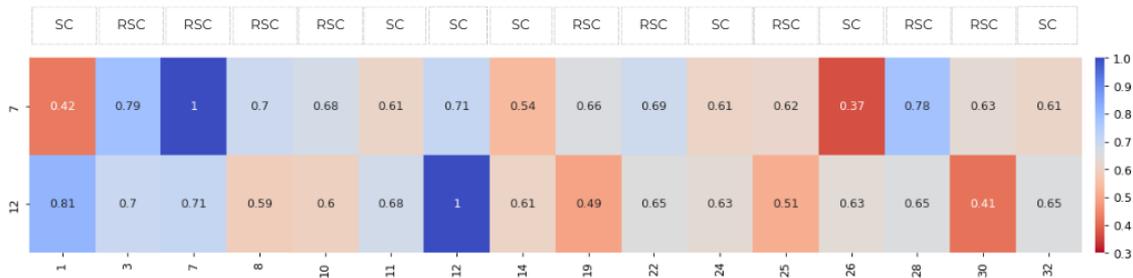
El Anexo 1 y Anexo 2 detallan los resultados del modelado de temas aplicado a los informes corporativos del primer periodo (P1, 2010-2015), mientras que el Anexo 3 y Anexo 4 muestran los resultados del segundo periodo (P2, 2016-2021). Se observa que en P1 la RSC se posiciona en el séptimo lugar y está representado principalmente por las palabras social, responsabilidad, discapacidad, personas, fundación. El tema 12 se asocia con SC y los términos que lo componen son ambiental, medioambiental, sostenibilidad, gestión y biodiversidad. Para P2 la SC se ubica en el tema 12 y la RSC con el tema 20. Los demás temas reflejan otros aspectos o áreas de gestión de las empresas, como los subsectores que componen la muestra (hoteles, transporte ferroviario, terrestre y marítimo). Igualmente, sobresalen temas asociados con el ámbito económico, clientes, gobernanza y empleados, ética, inclusión e igualdad.

### 5.2.3. Áreas claves de gestión en RSC y SC

Para profundizar en los resultados obtenidos y en las temáticas de estudio (RSC y SC) se calculó la matriz de similitud de temas para cada período. Los valores en la matriz indican qué tan similares son los temas entre sí. Valores cercanos a 1 indican alta similitud y valores cercanos a 0 indican baja similitud. Luego, se seleccionaron los temas más similares para RSC y SC, teniendo en cuenta los temas con valores superiores a 0,60. También, se considera que un tema está relacionado con RSC cuando su valor es mayor que el de SC, y viceversa (cuando el valor es mayor en SC se determina que ese tema es clave para esa temática específica).

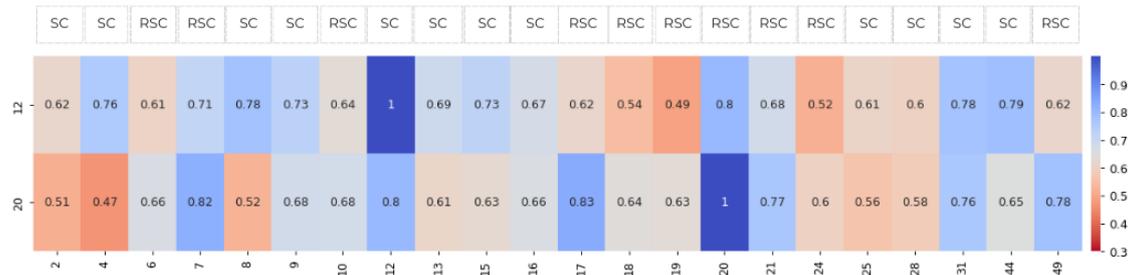
En la Ilustración 37 se detalla la similitud entre temas para P1 (2010-2015), donde la primera fila representa el tema 7 de RSC y la segunda fila el tema 12 de SC. Las columnas corresponden a cada uno de los temas con valores superiores a 0,60 con los que se relacionan la RSC y la SC. Se evidencia que hay un mayor número de temas vinculados con RSC (9 temas) que con la SC (7 temas). La Ilustración 38 muestra la similitud entre temas para P2 (2016-2021), donde la primera fila representa el tema 12 de SC y la segunda fila el tema 20 de RSC. Las columnas corresponden a cada uno de los temas con valores superiores a 0,60 que se asocian con las temáticas de estudio (RSC y SC). En este caso, sucede lo contrario que en P1, dado que hay más temas relacionados con SC (12 temas) que con RSC (10 temas).

**Ilustración 37. Matriz de similitud de temas para RSC y SC en P1 (2010-2015).**



Fuente: Elaboración propia.

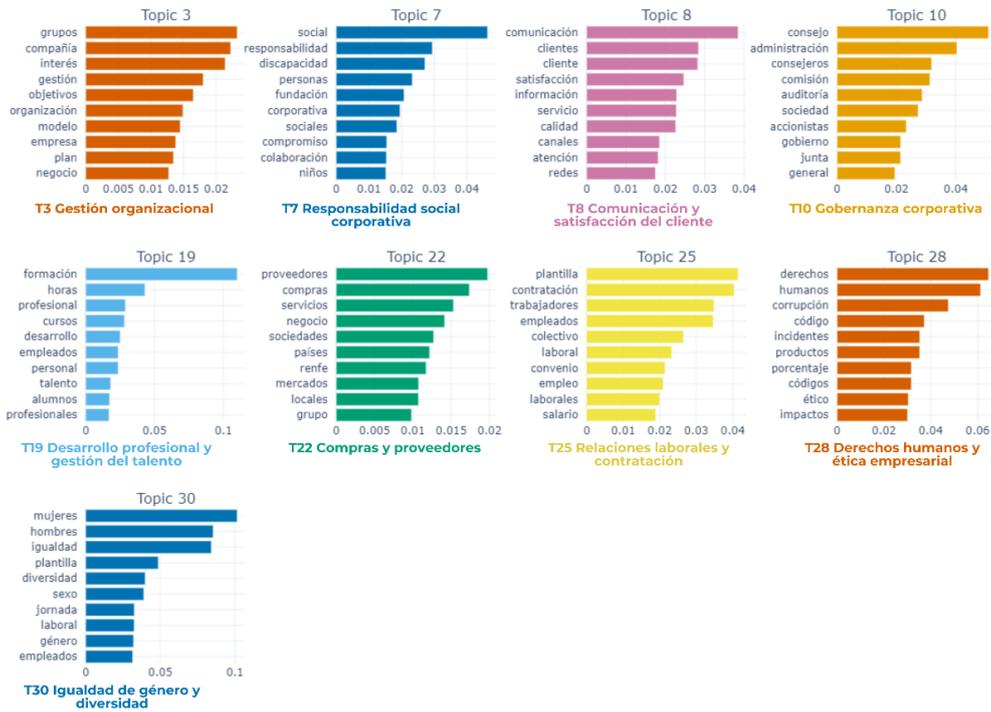
**Ilustración 38. Matriz de similitud de temas para RSC y SC en P2 (2016-2021).**



Fuente: Elaboración propia.

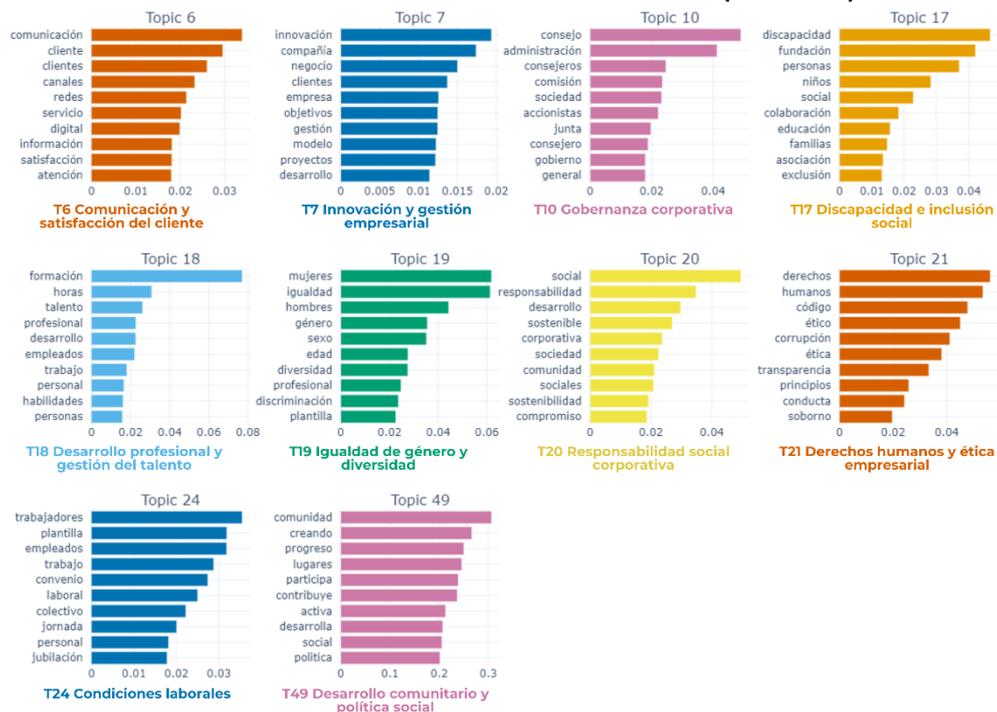
A continuación, en la Ilustración 39 y la Ilustración 40 se presentan los temas altamente vinculados con RSC para P1 y P2 (respectivamente) con sus palabras más representativas. Los resultados indican que las áreas o temas claves de gestión de RSC para ambos períodos son la gestión organizacional, filantropía, gobernanza, comunicación, clientes, ética, empleados (contratación, formación, condiciones laborales, igualdad, diversidad e inclusión). En P1 se menciona el grupo de interés de proveedores, mientras que en P2 no. Por el contrario, en P2 se posiciona el ámbito de contribución a la sociedad (palabras como comunidad, contribuye, desarrollo, etc.).

Ilustración 39. Temas relacionados con la RSC en P1 (2010-2015).



Fuente: Elaboración propia/procesamiento con BERTopic.

Ilustración 40. Temas relacionados con la RSC en P2 (2016-2021).

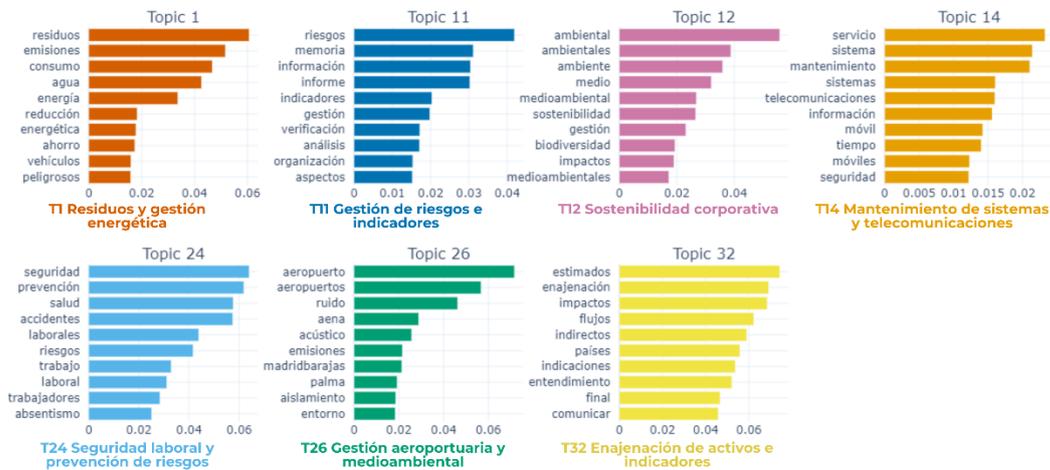


Fuente: Elaboración propia/procesamiento con BERTopic.

Con respecto a las áreas claves de gestión de SC (ver Ilustración 41 e Ilustración 42), se evidencia para P1 y P2 las empresas se han enfocado en diversos aspectos medioambientales

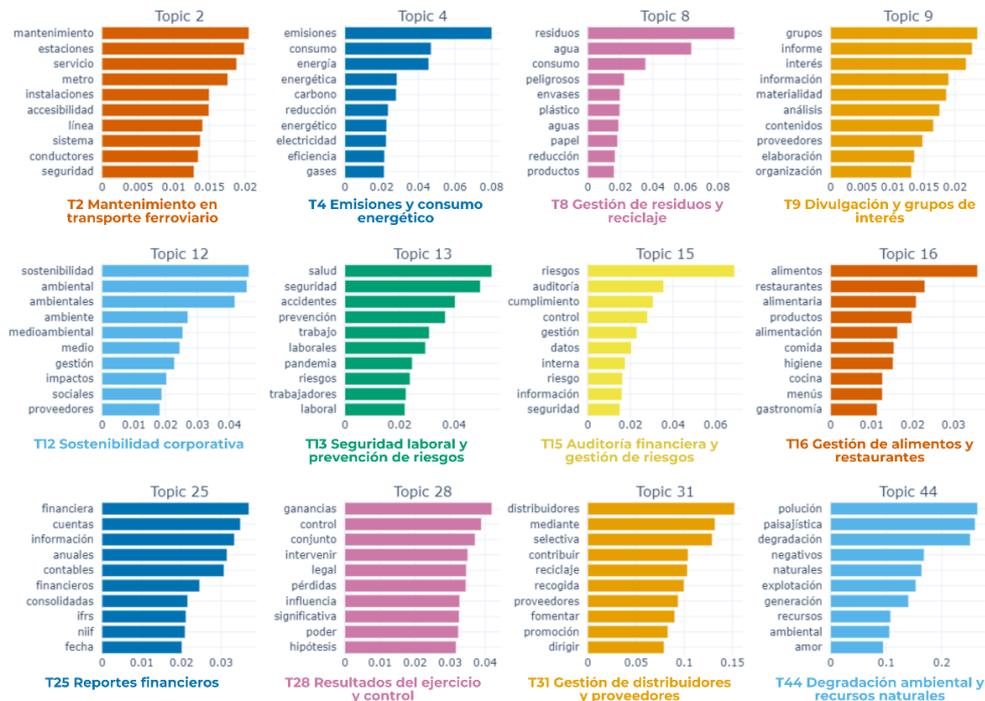
como residuos, emisiones, recursos naturales (agua, energía), biodiversidad. Asimismo, se destaca en ambos períodos la comunicación e indicadores (palabras como informes, memoria, indicadores, información) y la gestión de riesgos. En P1 se incluye los temas de servicios, telecomunicaciones, y seguridad. Por el contrario, en P2 las empresas incluyen dentro de sus áreas claves la reducción de huella de carbono (palabras como carbono, polución), reciclaje y la accesibilidad de sus instalaciones. En este segundo período se empiezan a asociar temas relevantes como las finanzas y cuentas anuales con la SC.

Ilustración 41. Temas relacionados con la SC en P1 (2010-2015).



Fuente: Elaboración propia/procesamiento con BERTopic.

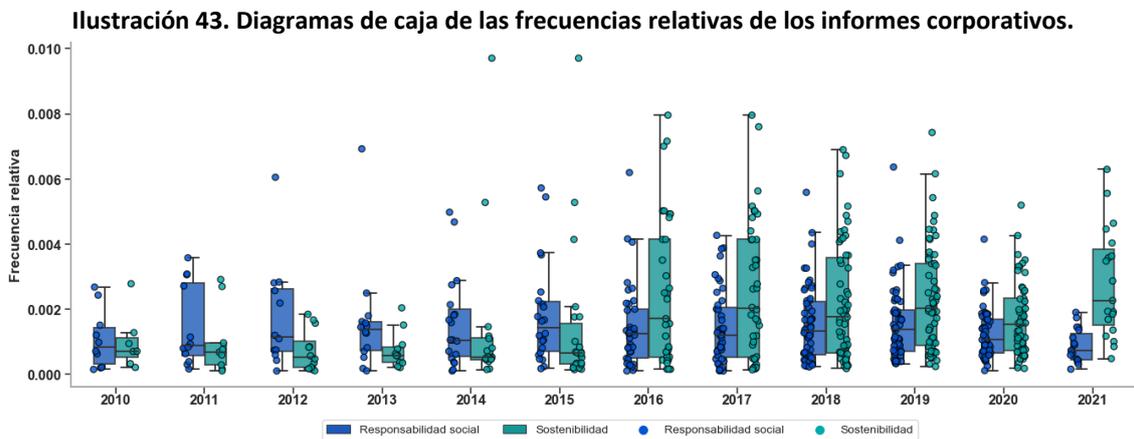
Ilustración 42. Temas relacionados con la SC en P2 (2016-2021).



Fuente: Elaboración propia/procesamiento con BERTopic.

#### 5.2.4. Análisis estadístico

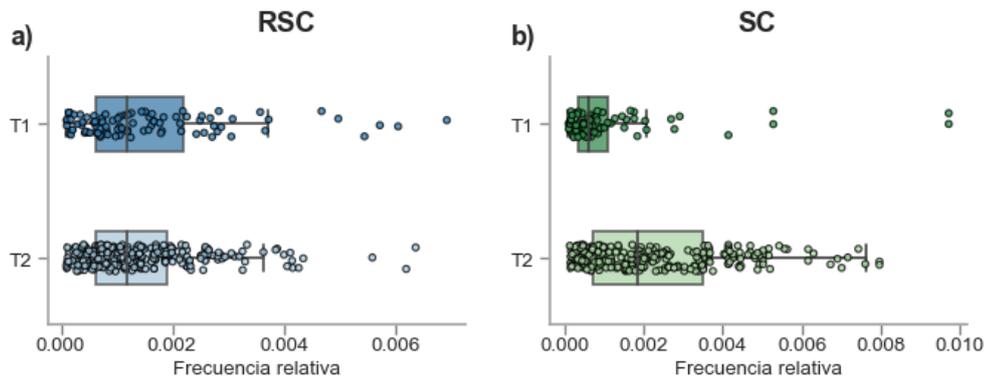
Siguiendo las fases de la metodología establecida, se calcularon las frecuencias relativas de los 364 informes corporativos de los años incluidos en el estudio (2010 al 2021). Como se observa en la Ilustración 43, los valores medianos de RSC de los años 2010 al 2015 son superiores a los de SC. Por el contrario, desde el año 2016 las medianas de SC son mayores que las de RSC. Del año 2015 al 2016 se puede ver un aumento en la cantidad de informes analizados (mayor densidad de puntos) en ambas categorías. Asimismo, a partir de 2016 se evidencia un aumento de la frecuencia relativa de los informes de SC, debido a que presentan un mayor número de palabras asociadas con la categoría. Igualmente, la mediana de SC aumenta desde el 2017 hasta 2019. Luego, en 2020 disminuye y finalmente en 2021 se marca el valor más alto de todo el periodo analizado.



**Fuente: Elaboración propia.**

En la Ilustración 44 se presenta por categoría y periodo (P1:2010-2015 y P2:2016-2021), los diagramas de caja. En a) se evidencia una mayor densidad de puntos en P2 comparado con P1. Cada punto representa la frecuencia relativa de un informe de RSC. No obstante, P2 presenta frecuencias relativas más altas con respecto a P1. En b) se observa una diferencia más amplia entre ambos periodos y la frecuencia relativa de los informes de P2 aumenta considerablemente con respecto a P1.

**Ilustración 44. Diagramas de caja de las frecuencias relativas de los informes corporativos.**

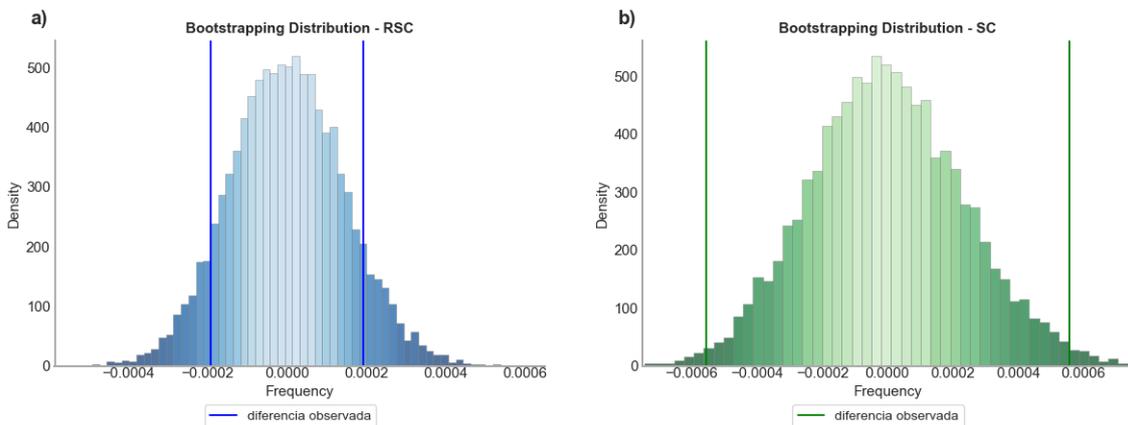


**Fuente: Elaboración propia.**

**5.2.5. Pruebas de normalidad y *Bootstrapping***

Después del preanálisis de los datos, se realizó por cada categoría la prueba de normalidad *D'Agostino* para contrastar la hipótesis nula de que los datos provenían de una población con distribución normal (*D'Agostino & Pearson, 1973*). Los resultados obtenidos indican que las hipótesis nulas de las pruebas de normalidad son rechazadas. Por ende, se aceptan las hipótesis alternativas para los grupos. Es decir, los datos de los periodos y categorías no cumplen la normalidad. Por tanto, se seleccionó el método no paramétrico de *bootstrapping* (*Efron, 1979*) para calcular significancia estadística (*p-value*) para la diferencia entre los periodos de tiempo P1 y P2 en cada categoría. En la Ilustración 45 se visualizan las distribuciones de *bootstrapping* para cada categoría, donde se representan las variaciones en las diferencias de medias como resultado del muestreo aleatorio.

**Ilustración 45. Distribuciones de *Bootstrapping*.**



**Fuente: Elaboración propia.**

La diferencia observada corresponde a la resta de las medias de P1 y P2 por cada categoría. El promedio de P1 en RSC fue de 0,001439 y en P2 de 0,001596, con una diferencia de 0,0001929

representadas por las líneas verticales azules en a) en la Ilustración 45. Para SC la media en P1 fue de 0,002081 y en P2 de 0,002251, dando como resultado una diferencia de -0,001111 graficadas por líneas verticales verdes en b) en la Ilustración 45. Nótese que la diferencia observada en cada categoría se representa en valor positivo y negativo, generándose un intervalo en el cual se observa un menor número de pseudomuestras en RSC y un mayor número de pseudomuestras en SC dentro del mismo. Asimismo, se calcula el *p-value* para contrastar las hipótesis nulas, que plantean la inexistencia de diferencias significativas entre ambos periodos (P1 y P2) para cada una de las temáticas de estudio. Tal como se indicó en la sección 3 de este capítulo, “Preguntas e hipótesis de investigación”, las hipótesis alternativas son las siguientes:

H1: Existe un cambio significativo a lo largo del tiempo en la categoría “Responsabilidad Social Corporativa” a través de las palabras utilizadas en los informes corporativos en las empresas turísticas.

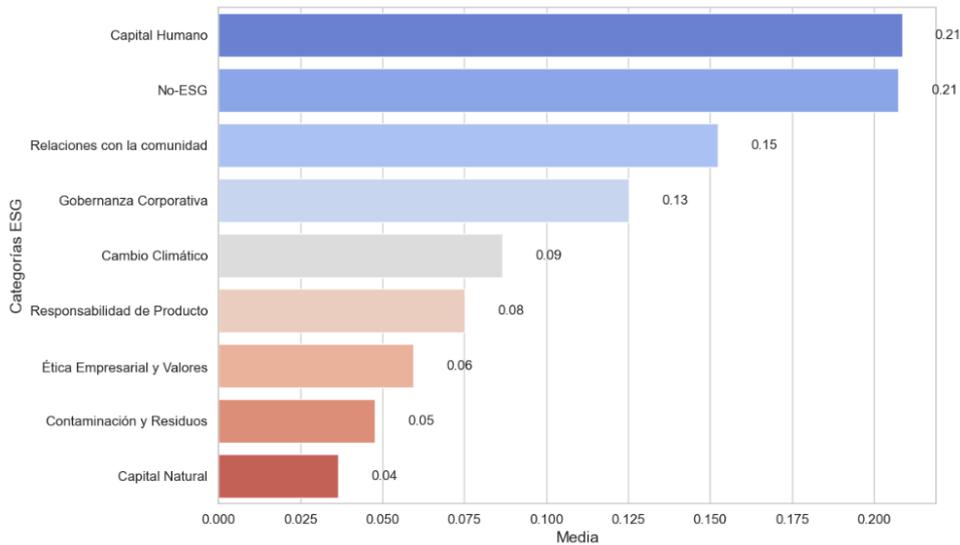
H2: Existe un cambio significativo a lo largo del tiempo en la categoría “Sostenibilidad Corporativa” a través de las palabras utilizadas en los informes corporativos en las empresas turísticas.

Los resultados de los valores de *p-value* obtenidos para la categoría de RSC son superiores a 0,05 por lo que no es posible rechazar la hipótesis nula, lo que indica que las medias de ambos periodos (P1 y P2) pueden ser iguales. Es decir, no hay diferencias significativas. Por el contrario, los valores de *p-value* para la categoría de SC son inferiores de 0,05. Por tanto, se rechaza la hipótesis nula y se acepta la hipótesis alternativa de que si existe diferencia significativa entre P1 y P2.

#### 5.2.6. Importancia de los criterios ESG en las empresas turísticas de España

Los resultados de la clasificación de textos utilizando el modelo FinBERT aplicado a los informes corporativos de las empresas turísticas se presenta en la Ilustración 46. Como se observa, los aspectos sociales y de gobernanza dominan el contenido de los informes. El *capital humano* y *No-ESG* son las categorías más representativas, cada una con un promedio de 0,21. Esto podría indicar que las empresas priorizan los intereses y necesidades de empleados, como su bienestar o desarrollo profesional. *Relaciones con la comunidad* y *Gobernanza corporativa* también sobresalen con promedios de 0,15 y 0,13, respectivamente. Por otro lado, las categorías asociadas con aspectos ambientales presentan promedios más bajos, *Cambio climático* (0,09), *Capital natural* (0,04) y *Contaminación y residuos* (0,05).

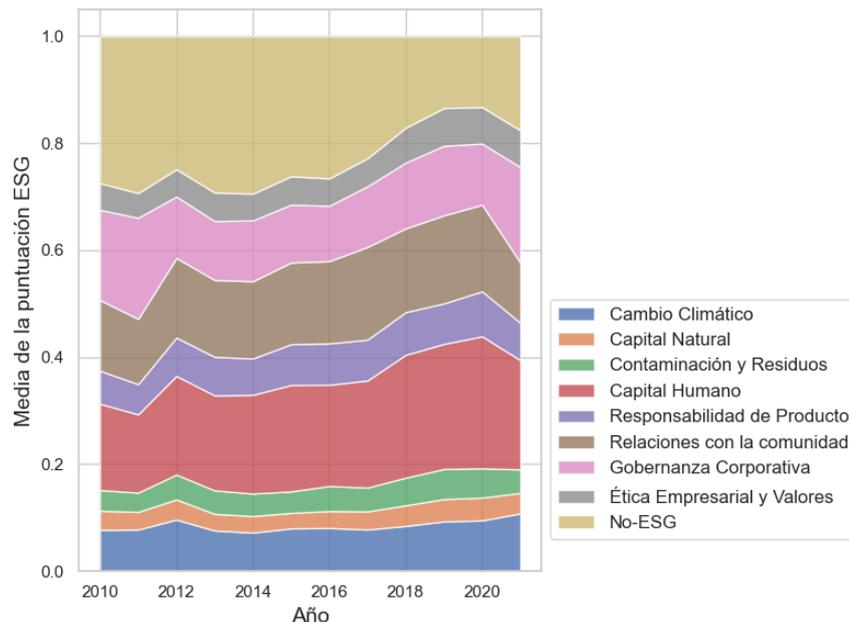
**Ilustración 46. Distribución de las categorías ESG (media) de los informes corporativos.**



Fuente: Elaboración propia.

La Ilustración 47 muestra la evolución de las puntuaciones ESG desde 2010 hasta 2021, observando un aumento gradual en las categorías. *Capital humano* ha mantenido una presencia constante y relevante a lo largo de los años. Las categorías ambientales como *Cambio climático* y *Capital natural* muestran un crecimiento más modesto en comparación con las categorías sociales y de gobernanza. No obstante, se observa un leve aumento en la importancia de *Cambio climático* en los últimos años. De igual forma, se destaca que a partir de 2016 se da un aumento más pronunciado en varias categorías, con una disminución relativa de los temas No-ESG.

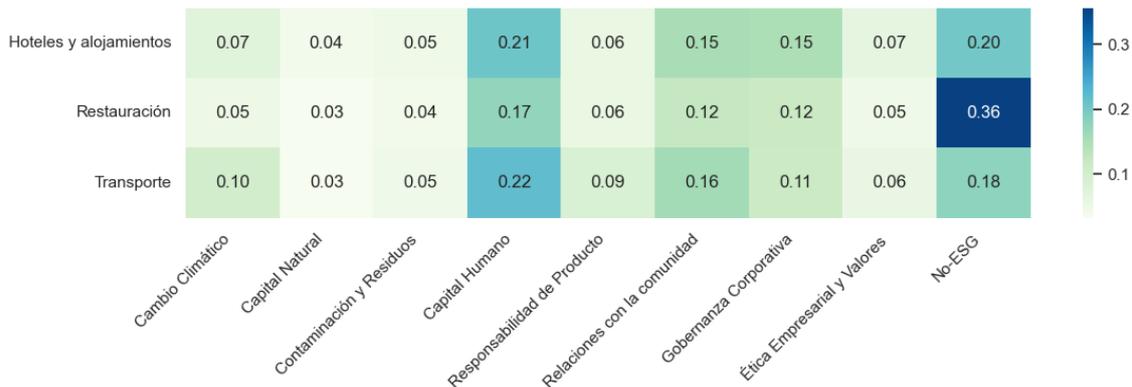
**Ilustración 47. Evolución de las categorías ESG a lo largo del tiempo.**



Fuente: Elaboración propia.

La Ilustración 48 muestra las variaciones de las categorías ESG entre los subsectores que conforman la muestra del estudio, permitiendo identificar las áreas de mayor enfoque en cada uno. Transporte es el subsector que otorga mayor importancia a *Cambio climático* (0,10), reflejando su preocupación por las emisiones de gases de efecto invernadero y el impacto ambiental de sus operaciones. Esto es coherente con el tipo de actividad prestada, puesto que las empresas de transporte aéreo, terrestre y marítimo tienden a estar más reguladas en cuanto a emisiones. En las categorías de *Capital natural*, y *Contaminación y residuos*, todos los subsectores presentan valores bajos. Por el contrario, en *Capital humano* los promedios son altos, Transporte (0,22), Hoteles y alojamientos (0,21) y Restauración (0,17). Con relación a la *Gobernanza Corporativa*, aunque es importante en todos los subsectores, Hoteles y alojamientos (0,15) sobresalen por su mayor puntuación.

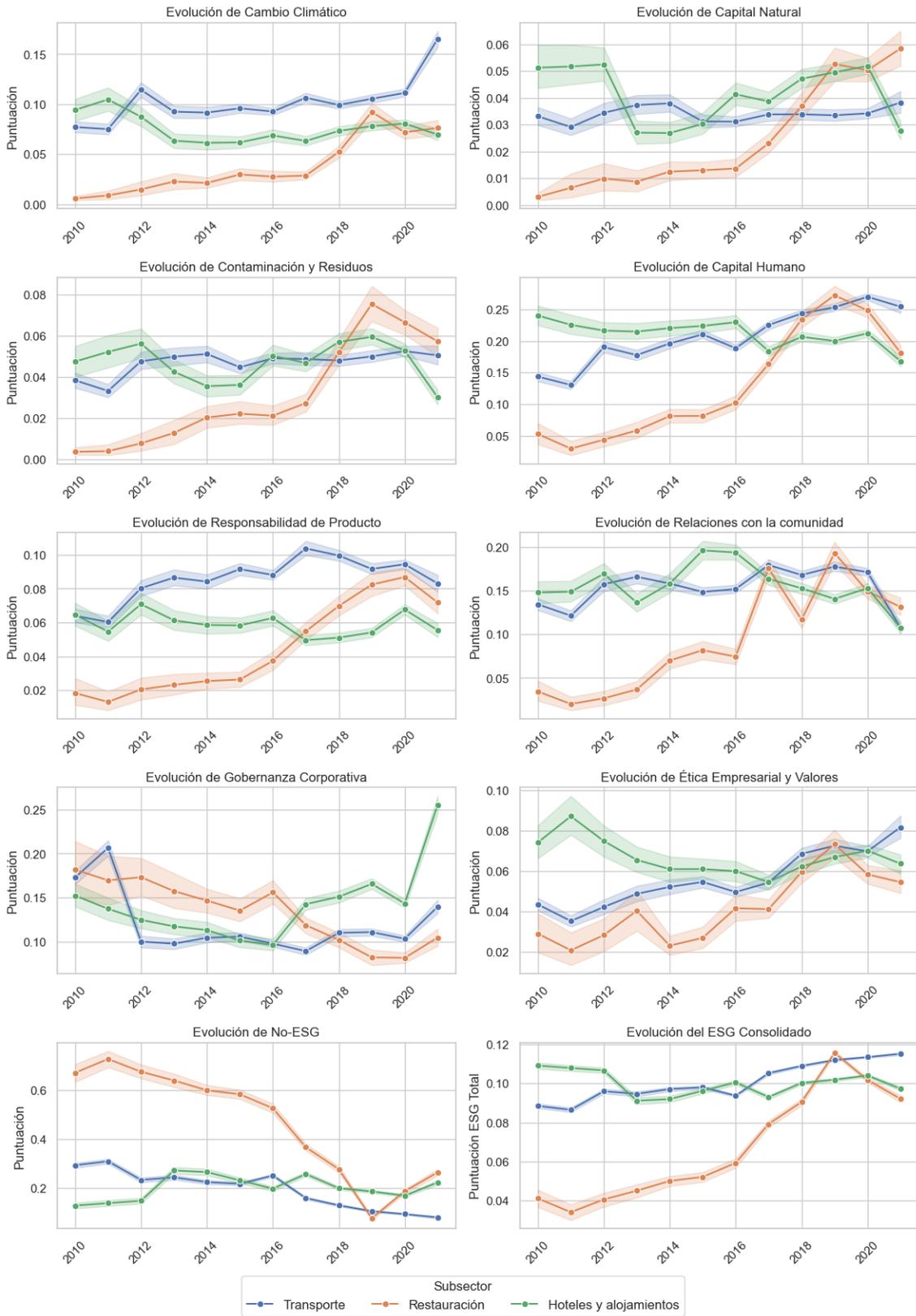
**Ilustración 48. Promedio de las categorías ESG por subsectores turísticos.**



**Fuente: Elaboración propia.**

Por último, en la Ilustración 49 se aprecia la evolución de las categorías ESG para los subsectores entre 2010 y 2021, detallando en gráficos individuales cómo ha cambiado la puntuación de una categoría ESG específica en cada subsector. El subsector transporte destaca por su fuerte incremento en áreas clave como el *Cambio climático* y el *Capital humano*, especialmente a partir de 2018. Por su parte, el subsector de hoteles y alojamientos muestra un notable aumento en la *Gobernanza corporativa*, mientras que el subsector de restauración refleja una mejora a partir de 2016 en categorías como *Capital humano*, *Relaciones con la comunidad* y *Responsabilidad del producto*. No obstante, es evidente que las puntuaciones en este subsector en comparación con los otros dos subsectores han sido relativamente más bajas en distintas categorías. Esto sugiere que la integración de prácticas sostenibles y responsables ha sido más pausada en este subsector.

**Ilustración 49. Evolución de cada categoría ESG a lo largo del tiempo.**



Fuente: Elaboración propia.

## 6. Conclusiones

Este estudio analiza la transición teórica y empírica de la Responsabilidad Social Corporativa a la Sostenibilidad Corporativa en empresas turísticas españolas, destacando la integración de criterios ESG. Para ello, se aplican distintas técnicas de ciencia de datos y pruebas estadísticas, adoptando un enfoque multidisciplinar con una metodología novedosa y replicable en posteriores estudios. Este enfoque permite reducir el *gap* de la literatura actual, dado que a pesar de las grandes oportunidades que ofrece la ciencia de datos a la investigación turística, son pocas las investigaciones en esta línea (Egger, 2022; Mariani et al., 2018; Mariani & Baggio, 2022). Igualmente, se pueden superar las limitaciones de lectura y procesamiento humanas, abordando grandes volúmenes de información. En consecuencia, los hallazgos permiten responder y validar las preguntas e hipótesis de investigación planteadas, generando implicaciones tanto teóricas como prácticas.

### 6.1. Conclusiones del enfoque teórico: Revisión computacional de la literatura

A nivel teórico, la presente investigación enriquece el conocimiento científico sobre RSC y SC en turismo, brindando un panorama actualizado y profundo a partir de una revisión computacional de la literatura, mediante la técnica de *machine learning* de modelado de temas. Se analizan 1.505 documentos de la base de datos Scopus (entre 2004 y 2023) y se complementan con análisis estadísticos. Este enfoque proporciona una respuesta clara a las dos primeras preguntas de investigación planteadas: P1: ¿Qué temas se han estudiado en la literatura existente sobre RSC y SC en turismo durante los últimos veinte años?; P2: ¿Existe un cambio de paradigma teórico desde la RSC hacia la SC en turismo?

A partir de los temas de investigación detectados en la literatura sobre RSC y SC, se puede concluir que ambos enfoques están conectados con aspectos fundamentales del turismo. El desarrollo turístico emerge como un tema central, que refleja la necesidad de equilibrar el crecimiento económico con la protección del medio ambiente y aspectos sociales. Además, se evidencia un fuerte cambio en las dinámicas de investigación de RSC y SC asociados con los efectos de la pandemia de covid-19 en el sector. Otros temas por destacar están relacionados con la industria hotelera, la lealtad del cliente y la gestión de recursos humanos, manifestando un enfoque más integral y centrado en distintos grupos de interés.

De igual modo, los resultados de las regresiones lineales permiten establecer una evolución teórica desde la RSC hacia la SC. Este cambio se da de forma gradual. La investigación sobre SC

ha mostrado un crecimiento constante y significativo, especialmente en los últimos años. Mientras que la RSC ha mostrado una estabilización o incluso una ligera disminución, no estadísticamente significativa. Esto muestra que las dinámicas en el campo científico del turismo están evolucionando hacia la SC, lo cual respalda la relación jerárquica entre RSC y SC establecida por Van Marrewijk (2003), en la que la RSC actúa como un paso intermedio para acercarse al objetivo final de la SC. Los resultados enriquecen la literatura, puesto que una comprensión teórica sólida puede promover cambios coherentes que reducen la ambigüedad en la academia y en las prácticas empresariales (Miller & Twining-Ward, 2005).

## **6.2. Conclusiones del enfoque empírico: Revisión de los informes corporativos de las empresas turísticas de España**

Las implicaciones prácticas de este estudio son múltiples, gracias al análisis detallado de 364 informes corporativos de empresas turísticas en España entre 2010 y 2021. La investigación también sobresale porque se lleva a cabo en empresas que utilizan el español como idioma corporativo. Esto contribuye significativamente al conocimiento científico, puesto que la literatura en este aspecto es escasa (Pérez Cañizares, 2021), dado que la mayoría de los estudios se enfocan en empresas que tiene como idioma corporativo el inglés. El estudio mejora la comprensión de las áreas clave de gestión asociadas a la RSC y la SC por parte de las empresas turísticas de España. De esta forma, se da respuesta a la tercera pregunta de investigación definida: ¿Cuáles son las áreas claves relacionadas con la gestión de la RSC y la SC de los informes corporativos de las empresas turísticas? Los resultados muestran que la industria turística a lo largo del tiempo ha implementado prácticas asociadas con la gestión organizacional, la ética, la igualdad, la diversidad, la inclusión y aspectos medioambientales claves (como residuos, emisiones, recursos naturales). No obstante, en años más recientes (2016-2021) se añaden áreas relacionadas con la huella de carbono y la gestión del riesgo y el reciclaje. Además, los temas de finanzas, contabilidad y cuentas anuales se empiezan a relacionar con la SC, lo cual puede llegar a ser objeto de investigación en futuros estudios.

Por otro lado, los hallazgos proporcionan una visión clara sobre cómo las empresas del sector están implementando la RSC y la SC, evidenciando un cambio de paradigma desde la RSC hacia la SC, alineándose con las nuevas exigencias de la humanidad. A través de pruebas estadísticas se validan las dos hipótesis propuestas: H1: Existe un cambio significativo a lo largo del tiempo en la categoría “Responsabilidad Social Corporativa” a través de las palabras utilizadas en los informes corporativos en las empresas turísticas; y H2: Existe un cambio

significativo a lo largo del tiempo en la categoría “Sostenibilidad Corporativa” a través de las palabras utilizadas en los informes corporativos en las empresas turísticas. Los hallazgos indican que al igual que en el ámbito científico, esta transición se produce de manera gradual, dado que es evidente que la RSC aún tiene un fuerte impacto en la gestión de las empresas turísticas. A pesar de que se observa una tendencia alcista en el uso de términos de los informes corporativos asociados con la SC y una tendencia bajista en la RSC, este cambio solo muestra significancia estadística en la categoría de SC y no en RSC. Por tanto, se puede inferir que las empresas continúan aplicando la RSC como estrategia de gestión central, evolucionando hacia un enfoque más centrado en la SC.

Bajo este nuevo enfoque de sostenibilidad, se cuantifica la importancia que los subsectores y empresas turísticas otorgan a los criterios ESG, lo que permite responder la cuarta pregunta de investigación establecida: ¿Cuál es la importancia que las empresas y subsectores turísticos otorgan a los criterios ESG en los informes corporativos del sector turístico español? Según Back (2024), la RSC en la industria turística se ha transformado en prácticas estándar de gestión sostenible bajo los criterios ESG. A través del uso del modelo FinBERT se logró analizar detalladamente nueve categorías de ESG (cambio climático, capital natural, contaminación y residuos, capital humano, responsabilidad del producto, relaciones con la comunidad, gobernanza corporativa, ética y valores empresariales y no-ESG). Los resultados revelan las diferencias entre subsectores en la forma en que abordan los ESG. El transporte sobresale como el subsector más proactivo en términos de sostenibilidad ambiental, con un fuerte aumento en temas relacionados con el cambio climático, responsabilidad del producto y capital humano, reflejando una adaptación a las exigencias regulatorias y sociales actuales. Por su parte, los hoteles y alojamientos han priorizado la gobernanza corporativa y las relaciones con la comunidad, lo que sugiere una respuesta estratégica hacia una mayor transparencia y responsabilidad social. Sin embargo, la restauración ha mostrado una integración más lenta de los criterios ESG, lo que podría indicar la necesidad de un mayor esfuerzo por parte de este subsector para alinearse con las prácticas sostenibles. Estos hallazgos resaltan las áreas donde las empresas han progresado e identifica aquellos ámbitos en los que aún se requiere una mayor integración de prácticas sostenibles, proporcionando información útil y clave para guiar tanto la toma de decisiones empresariales como la formulación de políticas públicas en el futuro. Del mismo modo, se enriquecen los estudios longitudinales y sectoriales sobre RSC y SC, disminuyendo el *gap* actual de la literatura (Back, 2024; Koh, 2024).

### **6.3. Limitaciones y futuras líneas de investigación**

Con respecto a las limitaciones de la investigación, cabe mencionar que el estudio teórico contempla únicamente una base de datos científica, por lo que futuros estudios podrían abarcar una gama más amplia de fuentes de información. Por otro lado, el estudio empírico se desarrolla en un solo sector (turismo), por lo que los resultados pueden no ser extrapolables para otros sectores o países. Además, se incluyeron únicamente dos grupos de empresas (grandes y medianas) por lo que no fue posible analizar la percepción de la RSC y SC en las demás organizaciones (pequeñas y micro). Sin embargo, este trabajo puede servir como referencia para que pequeñas empresas divulguen este tipo de información y fortalezcan las relaciones con sus *stakeholders*.

**CAPÍTULO IV. DISCURSO DE LOS TURISTAS SOBRE  
SOSTENIBILIDAD: INSIGHTS A PARTIR DE RESEÑAS EN LÍNEA DE  
ATRACCIONES TURÍSTICAS RURALES Y URBANAS DE ESPAÑA**



## **CAPÍTULO IV. DISCURSO DE LOS TURISTAS SOBRE SOSTENIBILIDAD: INSIGHTS A PARTIR DE RESEÑAS EN LÍNEA DE ATRACCIONES TURÍSTICAS RURALES Y URBANAS DE ESPAÑA**

A lo largo de este cuarto capítulo de la Tesis Doctoral se desarrolla la segunda parte del estudio empírico de España, referente a la perspectiva de los turistas. Este capítulo se centra en el tercer objetivo específico de analizar cómo el discurso de los turistas sobre sostenibilidad (económica, ambiental, social y cultural) influye en su satisfacción y valoración de los principales atractivos turísticos de España al escribir reseñas online. Específicamente, se explora la relación entre la importancia que los turistas otorgan a la sostenibilidad y su nivel de satisfacción reflejada en la calificación otorgada a los atractivos turísticos. Mediante un enfoque mixto que combina técnicas de *big data* con pruebas estadísticas, se procesan 207.706 comentarios de los principales atractivos turísticos de España. El contenido de este capítulo se divide en distintas secciones, incluyendo una introducción que contextualiza el estudio, seguida de una revisión de la literatura con los fundamentos teóricos. Posteriormente, se detalla el marco metodológico y los resultados de su aplicación. Finalmente, se presentan las conclusiones, que sintetizan las principales implicaciones teóricas y prácticas, así como las limitaciones y futuras líneas de investigación.

### **1. Introducción**

Las nuevas tecnologías, internet y la web 2.0 han permitido que los usuarios adopten un rol activo como generadores y difusores de contenido, dejando de ser únicamente receptores pasivos (Lo & Yao, 2019). El contenido generado por los usuarios (*User Generated Content*, UGC bajo sus siglas en inglés), se puede definir como aquella información disponible y de fácil acceso en internet difundida por usuarios no profesionales (Vickery & Wunsch-Vincent, 2007). Los turistas confían cada vez más en el UGC para reducir la incertidumbre, derivada de la propia naturaleza intangible del turismo y su enfoque experiencial (M. Kim & Kim, 2020). En efecto, las reseñas en línea constituyen una fuente de información relevante que refleja las percepciones de los turistas, influyendo en la toma de decisiones de otros turistas (Bigné et al., 2019; Jalilvand & Samiei, 2012; Z. Liu & Park, 2015; Xiang & Gretzel, 2010). Asimismo, las valoraciones *online* reflejan el grado de satisfacción con el servicio o producto turístico (Bulchand-Gidumal et al., 2013; Schuckert et al., 2015). La satisfacción del turista es un aspecto fundamental en la gestión del sector, ya que representa la diferencia entre las expectativas previas y la experiencia real (Jiang et al., 2018; Pizam et al., 1978).

El UGC se ha convertido en una alternativa para evaluar el grado de conciencia de los consumidores respecto a las prácticas de responsabilidad social y sostenibilidad de la industria turística (D'Acunto et al., 2020b). Recientemente, los estudios han comenzado a aplicar técnicas de *big data* al UGC para extraer información útil sobre temáticas específicas (Cai et al., 2024; D'Acunto et al., 2020b, 2023, 2024; Väisänen et al., 2023). Sin embargo, aún no se ha alcanzado el pleno potencial de estas técnicas en la investigación turística (Egger, 2022; Mariani et al., 2018; Mariani & Baggio, 2022). En consecuencia, se destaca la necesidad de realizar mayores esfuerzos para incorporar metodologías novedosas en la investigación turística que permitan aprovechar al máximo el UGC y la información disponible en la web.

De igual modo, comprender las percepciones sostenibles de los turistas contribuye a disminuir los impactos negativos del turismo y mejorar la competitividad de los destinos. Esto permite atraer más turistas, fomentar comportamientos más responsables y mejorar el grado de satisfacción (Palacios-Florencio et al., 2021). La sostenibilidad percibida en turismo es un elemento clave para el crecimiento de los destinos, puesto que refleja la evaluación afectivo-cognitiva de las políticas y prácticas de sostenibilidad que realizan los *stakeholders* (Cavalcante et al., 2021; Iniesta-Bonillo et al., 2016). Aunque se ha demostrado que la sostenibilidad percibida mejora la satisfacción del turista y ofrece una ventaja competitiva, aún se necesita más investigación en este ámbito (Conti et al., 2023; Galeone & Sebastiani, 2021; Sánchez-Fernández et al., 2019). Autores como Mior Shariffuddin et al. (2023) y Sreen et al. (2022) indican que es necesario desarrollar investigaciones que profundicen en la comprensión de la percepción de los turistas.

El objetivo de esta investigación es analizar cómo el discurso de los turistas sobre sostenibilidad influye en su satisfacción y valoración de los principales atractivos turísticos de España al escribir reseñas online. Se explora la relación entre la importancia que los turistas otorgan a la sostenibilidad y su nivel de satisfacción reflejada en la calificación otorgada a los atractivos turísticos, usando técnicas de *big data* y procesamiento del lenguaje natural (PLN). Se considera que el discurso sobre sostenibilidad incluye aspectos de las dimensiones económica, social, cultural y ambiental, dado que promueven un desarrollo turístico sostenible (Agyeiwaah et al., 2017). Esto incluye cualquier información o percepción que el turista escriba en su comentario de forma explícita o implícita. Asimismo, siguiendo la línea de estudios previos sobre satisfacción y *rating* de reseñas *online* (Bulchand-Gidumal et al., 2013; Schuckert et al., 2015), se considera como indicador de satisfacción la calificación asignada por el turista al atractivo

turístico. Los análisis se realizan para los atractivos turísticos tanto a nivel nacional (España), como diferenciándolos por zonas (rurales y urbanas), dado que pueden existir diferencias importantes en las dimensiones de sostenibilidad entre áreas turísticas rurales y urbanas (Rasoolimanesh et al., 2017).

La contribución de este estudio es significativa en varios aspectos. En primer lugar, se enriquece la literatura sobre la sostenibilidad percibida por los turistas, dado que ha sido poco investigada (Sánchez-Fernández et al., 2019). Una mayor comprensión de este grupo de interés resulta clave para fomentar comportamientos responsables y atraer turistas de calidad (Palacios-Florencio et al., 2021). En ocasiones, los comportamientos irrespetuosos de los turistas pueden afectar la sostenibilidad de los destinos, amenazando su viabilidad a largo plazo y reduciendo la calidad de la experiencia turística (Im & Kim, 2023). En segundo lugar, se amplían los estudios que aplican técnicas de *big data* al UGC para analizar la sostenibilidad, siguiendo la tendencia emergente de investigación en este ámbito (Cai et al., 2024). Dado que pocos estudios han adoptado este enfoque (D'Acunto et al., 2020b, 2023, 2024; Väisänen et al., 2023), se pretende ofrecer una metodología replicable en futuras investigaciones. En tercer lugar, y hasta donde se tiene conocimiento, este es el primer estudio que compara la sostenibilidad percibida por turistas entre zonas rurales y urbanas aplicando *big data* al UGC. En cuarto lugar, esta investigación es pionera en analizar la sostenibilidad percibida del destino turístico de España, dado que no ha sido posible encontrar una investigación similar en el mismo ámbito geográfico. Esto confiere un valor adicional a este estudio porque los resultados brindan a las partes interesadas oportunidades para comprender la percepción sostenible de los turistas en las atracciones turísticas para la toma de decisiones y mejora continua.

## **2. Revisión de la literatura**

### **2.1. Sostenibilidad percibida y satisfacción de los turistas: Enfoque de *big data* y UGC**

La sostenibilidad percibida es una construcción multidimensional compuesta por las dimensiones ambiental, sociocultural y económica (Iniasta-Bonillo et al., 2016). Sin embargo, los atributos sostenibles suelen asociarse de forma negativa con los atributos de los productos o servicios, ya que los consumidores tienden a relacionar los productos o servicios sostenible con un rendimiento inferior, un precio elevado y baja calidad (Font et al., 2017; Olson, 2013; Ottman et al., 2006; Peng & Chen, 2019). A pesar de que hay evidencia de que la sostenibilidad percibida de un destino turístico contribuye positivamente a la satisfacción del turista y puede derivarse en una ventaja competitiva, aún se necesita más investigación en este ámbito, puesto que es

escasa (Conti et al., 2023; Galeone & Sebastiani, 2021; Sánchez-Fernández et al., 2019). Igualmente, son pocos los estudios que han analizado la sostenibilidad percibida desde el punto de vista de los turistas (Sánchez-Fernández et al., 2019). Esto conlleva a la necesidad de desarrollar más investigaciones que brinden una mejor comprensión de la percepción de los turistas (Mior Shariffuddin et al., 2023; Sreen et al., 2022). Algunos estudios previos abordan únicamente una dimensión de la sostenibilidad y no suelen adoptar un enfoque de *big data* que aproveche la valiosa información disponible en el contenido generado por los usuarios (*User Generated Content*, UGC bajo sus siglas en inglés). Por ejemplo, Sanchez-Fernandez et al. (2016) utilizan una técnica de segmentación de clases latente para analizar la percepción de la sostenibilidad ambiental de cinco ciudades de la cuenca del Mar Mediterráneo. Penagos-Londoño et al. (2021) analizan la sostenibilidad y la confiabilidad percibidas a partir de una muestra de 438 turistas de Chile y Ecuador. Guizzardi et al. (2022) aplican un modelo de ecuaciones estructurales basados en covarianza para analizar el valor percibido del turismo rural y la sostenibilidad percibida de los turistas.

Durante la última década, se evidencia un crecimiento de las investigaciones en turismo que emplean técnicas de *big data* (Cai et al., 2024). Los estudios previos señalan que el análisis de *big data* mejora la investigación turística, puesto que permite superar las limitaciones de tamaño de muestra que suelen enfrentar los estudios que utilizan datos de encuestas (Yang et al., 2015), aportan nuevos conocimientos sobre variables previamente estudiadas en la literatura y ofrecen una mejor comprensión del comportamiento de los turistas al proporcionar datos abundantes y libres de sesgos de muestreo (X. Li et al., 2017; Xiang, Schwartz, Gerdes, et al., 2015). Los métodos de análisis del UGC han avanzado, pasando de extraer contenidos a análisis más sofisticados como distribución de temas, puntuaciones de sentimientos y relaciones causales (Cai et al., 2024). En turismo, se han aplicado distintas técnicas de Procesamiento del Lenguaje Natural (PLN) y modelos de aprendizaje profundo como BERT para analizar y clasificar reseñas online (Chai et al., 2021; Rey-Moreno et al., 2023; Viñán-Ludeña & de Campos, 2022; H. Zhang et al., 2023). De este modo, el UGC ofrece a los investigadores una oportunidad para detectar tendencias y patrones en los intereses de los turistas, algo que resultaría difícil de lograr mediante métodos de investigación tradicionales (Daugherty et al., 2008; Lu & Stepchenkova, 2015).

El análisis del UGC proporciona información valiosa a los investigadores para comprender las preferencias, los comportamientos y la satisfacción de los turistas (León et al., 2025). Las

reseñas *online* sobresalen porque su calificación representa un indicador de satisfacción general y su contenido ofrece información detallada que puede influir en la demanda futura de los consumidores (K. L. Xie et al., 2014; L. Zhou et al., 2014). Gerdt et al. (2019) señalan que, a diferencia de los estudios tradicionales basados en encuestas, las reseñas *online* son una fuente de datos original e imparcial. Las encuestas pueden generar una sobreestimación de la importancia de la sostenibilidad en turismo. A pesar de que diversos estudios han analizado la satisfacción del cliente a partir del UGC, son pocas las investigaciones que emplean el UGC como fuente de información para analizar la satisfacción y la sostenibilidad (Brazyté et al., 2016; Gerdt et al., 2019; Yu et al., 2017). Además, no suelen adoptar un enfoque de *big data* y la mayoría se centra en el subsector hotelero (Brazyté et al., 2016; D’Acunto et al., 2020, 2023; Gerdt et al., 2019; Väisänen et al., 2023; Yu et al., 2017). Los turistas demuestran actitudes positivas hacia las alternativas sostenibles (Brazyté et al., 2016; Väisänen et al., 2023). Sin embargo, su deseo de obtener experiencias memorables y placenteras puede contradecir los objetivos de sostenibilidad (Malone et al., 2014). En este contexto, comprender la percepción de sostenibilidad por parte de los turistas es crucial para fomentar comportamientos más responsables y minimizar los impactos negativos del turismo (Palacios-Florencio et al., 2021).

El desarrollo turístico sostenible debe centrarse en las cuatro áreas de la sostenibilidad: económica, ambiental, social y cultural (Agyeiwaah et al., 2017). Este enfoque se alinea con la teoría del “*Triple Bottom Line*” (TBL) (Elkington & Rowlands, 1999), donde los aspectos sociales se dividen en impactos en la sociedad y la cultura. A pesar de que en la literatura turística existen estudios previos que analizan la sostenibilidad desde sus dimensiones, algunos se enfocan en una sola dimensión (por ejemplo, Sanchez-Fernandez et al., 2016; Mishra & Gupta, 2019; D’Acunto et al., 2023). Por ende, es relevante desarrollar más investigaciones que aborden todas las dimensiones sostenibles, empleen técnicas de *big data* y extraigan información relevante del UGC. Esta investigación adopta un enfoque similar al utilizado por Väisänen et al. (2023), en el que se analiza cómo los turistas perciben e informan en sus reseñas online las cuatro dimensiones centrales de la sostenibilidad. De este modo, es necesario puntualizar cada una de estas dimensiones desde la perspectiva de los turistas.

La dimensión económica se refiere a la mejora del bienestar financiero y al crecimiento sostenido, tanto de las empresas como el propio destino turístico. Esto se refleja en la rentabilidad de las operaciones, que incluye un número suficiente de clientes, gestión de costos, estrategias de precios y la satisfacción de los clientes (Väisänen et al., 2023). Por tanto, los

clientes evalúan el precio frente a la calidad (Ye et al., 2014), lo que se deriva en una buena o mala relación calidad-precio que genera satisfacción o insatisfacción (Väisänen et al., 2023). Es decir, la percepción de los turistas sobre la relación entre calidad y precio influye en su satisfacción y afecta a largo plazo la rentabilidad de las empresas turísticas y el destino. En este contexto y siguiendo el enfoque de Väisänen et al. (2023), se considera que la dimensión económica se manifiesta en las reseñas online de los turistas a través del valor por el dinero, que se genera de dicha relación de calidad-precio. Además, el gasto turístico efectuado por los visitantes en un destino (compras, alojamiento y actividades) ayuda a los organismos institucionales locales y a la comunidad a movilizar esos fondos para el desarrollo del destino y mejorar la calidad de vida de los residentes (Jasrotia et al., 2019; Suhartanto, 2018). En lo que respecta a la dimensión ambiental, esta se encarga de proteger el medio ambiente para el uso de los humanos y otros seres vivos, garantizando el beneficio mutuo (Jones et al., 2016). Algunos ejemplos de medidas o acciones que abarcan esta dimensión en turismo pueden estar relacionadas con el consumo de energía, contaminación, entornos agradables, ruidos, mantenimientos de instalaciones, limpieza, transporte o materiales ecológicos (Väisänen et al., 2023).

La sostenibilidad social hace referencia a los derechos humanos y laborales, el bienestar, la inclusión, la accesibilidad y la seguridad (Swarbrooke, 1999; UNEP, 2005). En servicios la dimensión social se puede dividir entre factores físicos y sociales (Väisänen et al., 2023). Los atributos físicos ayudan a garantizar la salud y seguridad de los clientes, así como el apoyo a la economía local (Dewhurst & Thomas, 2003; Garay & Font, 2012; Jones et al., 2014). Estos pueden abarcar estructuras y señales seguras, medidas de accesibilidad, compras locales, opciones saludables de actividades o alimentación. Por otro lado, el factor social se refiere a las interacciones humanas que incluye atributos como emplear gente local o jóvenes; actividades de educación para promover la sostenibilidad en los clientes; trato igualitario, respetuoso y libre de discriminación hacia los clientes por parte de los empleados; masificación turística (Väisänen et al., 2023). Finalmente, la dimensión cultural asegura la preservación de la cultura local en el desarrollo de las actividades turísticas (Madanaguli et al., 2023). Esta dimensión también se puede clasificar en factores físicos y sociales (Väisänen et al., 2023). Los factores físicos se asocian con prácticas tradicionales relacionadas con el estilo de vida de los locales, los edificios patrimoniales o históricos y la gastronomía local. Los factores sociales se asocian con la comunidad local, puesto que desempeñan un rol importante en la experiencia del turista y promueven la cultura.

Las investigaciones previas sobre las dimensiones sostenibles en turismo han demostrado cómo las percepciones de los turistas se asocian con la calificación que otorgan y su satisfacción. El estudio de Wu & Yang (2023) emplea 5.832 comentarios *online* para identificar atributos de servicio clave de la sostenibilidad en destinos rurales en China y manifiesta las cuatro dimensiones sostenibles. Sus hallazgos evidencian que las dimensiones ambiental y cultural impactan de manera positiva en la satisfacción de los turistas, puesto que valoran el paisaje natural y las vistas, la autenticidad del lugar y el valor cultural e histórico de las construcciones. Por el contrario, algunos de los factores que generan insatisfacción se asocian con las dimensiones social y económica, resaltando precios elevados de las entradas y productos, bajo profesionalismo del personal y falta de mantenimiento de ciertas estructuras históricas.

Las dimensiones social y económica pueden vincularse con experiencias turísticas negativas que generan cierto grado de insatisfacción, emociones negativas y calificaciones bajas (D'Acunto et al., 2020). Estas experiencias pueden formarse por la descortesía del personal de servicio, generando al mismo tiempo una mala calidad del servicio (Kim et al., 2020) y afectando la percepción del turista frente a la relación de calidad-precio. De acuerdo con Rita et al. (2022) y Väisänen et al. (2023), las percepciones más negativas provienen de la dimensión económica, que generalmente está asociada con aspectos de dinero.

Además, Prakash et al. (2019) revelan que algunas de las causas de insatisfacción de los visitantes a los Parques Nacionales de Sri Lanka se relacionan con aspectos sociales, como la congestión y tráfico dentro de los parques, la masificación y la seguridad de los visitantes por las malas condiciones de senderos e instalaciones. En efecto, los atributos relacionados con la dimensión social de un destino como seguridad, accesibilidad, hospitalidad, información turística y prácticas comerciales poco éticas pueden desencadenar experiencias turísticas negativas e insatisfacción (Kim et al., 2020). En esta misma línea, Meneses et al. (2023) analizan el UGC para determinar los factores que generan satisfacción e insatisfacción en el turismo enológico. Sus resultados muestran que los factores de insatisfacción se asocian con la dimensión social, destacando los tiempos de espera extensos y un servicio al cliente deficiente. No obstante, los aspectos relacionados con la dimensión cultural, como la variedad y calidad del vino y las guías informativas con la historia de las bodegas, generan satisfacción en los visitantes.

La dimensión cultural tiene un impacto significativo en la satisfacción y percepciones positivas de los turistas (Chen & Rahman, 2018; Li & Liu, 2019; Väisänen et al., 2023). El contacto cultural es el deseo de los turistas por interactuar y sumergirse en la cultura local del destino

que visitan (Chen & Rahman, 2018). Éste impacta notablemente la satisfacción de los turistas y su intención de visitar nuevamente el destino (Nguyen Huu et al., 2024). Además, los atributos de un destino como la cultura e historia local, variedad de actividades, infraestructura y gestión ambiental, contribuyen a crear experiencias turísticas memorables y positivas (Kim, 2014). Por otro lado, los comentarios sobre temas ambientales se relacionan con emociones positivas y calificaciones más altas (D'Acunto et al., 2020; Mariani & Borghi, 2023). La mayoría de las actividades ecológicas son apreciadas por los turistas y tienen efectos favorables en su satisfacción (Merli et al., 2019; Oroian et al., 2015). Cong et al. (2014) han realizado un análisis de contenido sobre las reseñas online de TripAdvisor y las experiencias de turismo de vida silvestre, revelando que estas interacciones aumentan la satisfacción de los turistas. Asimismo, Prud'homme & Raymond (2013) puntualizan que las prácticas ecológicas como la reducción de residuos y el reciclaje, influye positivamente en la satisfacción de los clientes.

Para explorar cómo el discurso de los turistas sobre sostenibilidad (económica, ambiental, social y cultural) influye en su satisfacción y valoración de los atractivos turísticos al escribir reseñas online, y tomando como referencia la literatura previa, se plantean cuatro hipótesis de investigación.

**H1:** Las reseñas con un mayor discurso sobre la sostenibilidad ambiental influyen positivamente en la satisfacción de los turistas.

**H2:** Las reseñas con un mayor discurso sobre la dimensión cultural influyen positivamente en la satisfacción de los turistas.

**H3:** Las reseñas con un mayor discurso sobre la dimensión económica influyen negativamente en la satisfacción de los turistas.

**H4:** Las reseñas con un mayor discurso sobre la dimensión social influyen negativamente en la satisfacción de los turistas.

## **2.2. Diferencias de sostenibilidad en zonas turísticas rurales y urbanas**

Las actividades turísticas en zonas rurales y urbanas ofrecen experiencias diferentes a los visitantes, dependiendo sus recursos disponibles (naturales, culturales, etc.), generándose distintos impactos positivos y negativos. La Organización Mundial del Turismo (World Tourism Organization, 2024) define el turismo rural como la actividad turística que ofrece a las personas experiencias basadas en la agricultura, la naturaleza, la pesca, el estilo de vida, la cultura rural y

el turismo. Por otro lado, el turismo urbano se caracteriza por una economía no agrícola y por una amplia gama de experiencias y productos turísticos culturales, arquitectónicos, tecnológicos, sociales y naturales para el ocio y los negocios (World Tourism Organization, 2019b).

El turismo rural favorece la conservación de la cultura y patrimonio al promover estilos de vida locales entre los visitantes urbanos (S. Kumar & Shekhar, 2020). Contrario al turismo urbano, el turismo rural brinda a los visitantes la posibilidad de integrarse en un entorno idealizado, disminuyendo algunos de los factores negativos que generan estrés en zonas urbanas (contaminación, masificación, ruido, entre otros) (Kastenholz et al., 2012). El turismo en zonas rurales sirve para fomentar nuevas actividades socioeconómicas y contrarrestar los problemas de salud mental pública (Kabadayi et al., 2020; López-Sanz et al., 2021).

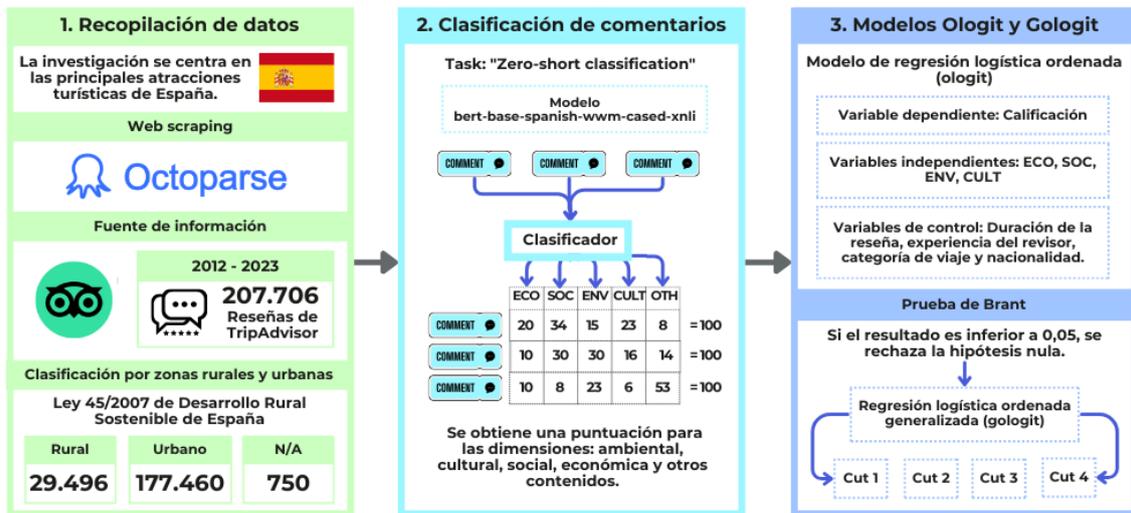
Rasoolimanesh et al. (2017) confirman que se pueden identificar notables diferencias en la sostenibilidad económica, social, ambiental y cultural entre áreas turísticas rurales y urbanas, dado que es posible que haya variaciones en la forma en que se materializan los impactos (Ashworth & Page, 2011; Gilbert & Clark, 1997; Wall & Mathieson, 2006). Por ejemplo, los hallazgos de Kazeminia et al. (2016) señalan que la disposición de los ecoturistas a pagar precios más altos por servicios sostenibles se relaciona positivamente con sus actitudes y creencias ambientales. Igualmente, Väisänen et al. (2023) resaltan la importancia de realizar en futuras investigaciones análisis para comparar destinos rurales y urbanos. Por consiguiente, se plantea la siguiente pregunta de investigación:

**P1:** ¿Existen diferencias en la contribución de las dimensiones sostenibles de las atracciones turísticas en zonas rurales y urbanas?

### **3. Metodología de investigación**

La Ilustración 50 muestra la metodología de investigación del estudio, compuesta por tres fases: 1) recopilación de datos, 2) clasificación y 3) planteamiento de modelos. El lenguaje de programación utilizado para la codificación es Python y el programa estadístico Stata.

Ilustración 50. Metodología de investigación.



Fuente: Elaboración propia.

### 3.1. Recopilación de datos

La fuente de información es el sitio web TripAdvisor, dado que continúa siendo una de las plataformas líderes de UGC por su facilidad de uso percibida, credibilidad y utilidad (Filiari et al., 2021). El índice de popularidad de TripAdvisor organiza las atracciones de un destino turístico según la calidad, número y novedad de las reseñas. Además, la popularidad y calificación de un atractivo turístico influye sobre la decisión de visita del turista (Hernández et al., 2021; Kemperman et al., 2019). La recopilación de datos se realizó entre diciembre de 2023 y febrero de 2024, descargando los primeros 2.000 comentarios (cuando estuvieron disponibles) de las 15 principales atracciones turísticas de cada Comunidad Autónoma de España con la herramienta de *web scraping*<sup>5</sup> Octoparse. Se obtuvieron 313.683 comentarios de 255 atracciones, abarcando el periodo desde 2007 hasta 2024. Para cada comentario se recopilaban datos sobre ubicación, fecha, categoría de viaje, nacionalidad del visitante, calificación (de uno a cinco) y número de contribuciones en TripAdvisor. Algunos campos, como nacionalidad y categoría de viaje, no siempre fueron completados por los usuarios. Se eliminó la categoría "Negocios" debido a que sus motivaciones difieren del turismo de ocio. Además, se filtraron las fechas de los comentarios, incluyendo solo los años de 2012 a 2023, lo que resultó en una muestra final de 207.706 comentarios.

<sup>5</sup> *Web scraping* (o raspado de páginas web) es una técnica utilizada para extraer información de sitios web de forma automatizada a través de programas o scripts (Mitchell, 2018). Esta técnica se usa comúnmente para obtener grandes volúmenes de información estructurada y no estructurada de manera rápida y eficiente.

Dado que también se tiene el propósito de comparar las percepciones sostenibles de los turistas en el ámbito urbano y rural, se realiza una clasificación entre zona rural y urbana según la Ley 45/2007, de Desarrollo Sostenible del Medio Rural en España. En consecuencia, se considera municipio rural aquel con menos de 30.000 habitantes y menos de 100 habitantes por km<sup>2</sup>. La muestra comprende 29.496 reseñas de zonas rurales, 177.460 reseñas de zonas urbanas y 750 reseñas sin clasificación, puesto que los turistas no han detallado el municipio, sino que han registrado el nombre de la Comunidad Autónoma. En la Tabla 9 se presenta la distribución de la muestra del estudio por Comunidad Autónoma, especificando el número de municipios, la cantidad total de comentarios y su desagregación por zonas rurales y urbanas.

Tabla 9. Detalle de la muestra de atracciones turísticas de España.

Comunidad Autónoma	Número de atracciones	Número de municipios	Comentarios				
			Total	%	Rural	Urbano	Sin ubicación
Andalucía	15	7	15.827	8%	-	15.827	-
Aragón	15	8	13.213	6%	5.986	7.227	-
Principado de Asturias	15	10	10.804	5%	5.407	5.310	87
Islas Baleares	15	12	11.046	5%	497	10.549	-
País Vasco	15	7	12.534	6%	652	11.882	-
Islas Canarias	15	14	11.053	5%	3.280	7.773	-
Cantabria	15	9	11.344	5%	1.711	9.633	-
Castilla y León	15	8	16.073	8%	192	15.881	-
Castilla-La Mancha	15	4	11.696	6%	727	10.969	-
Cataluña	15	2	16.322	8%	-	16.322	-
Comunidad de Madrid	15	1	17.453	8%	-	17.453	-
Extremadura	15	8	10.346	5%	3.238	7.108	-
Galicia	15	7	14.891	7%	1.352	13.431	108
La Rioja	15	9	5.668	3%	1.625	4.043	-
Comunidad Foral de Navarra	15	9	7.251	3%	3.416	3.280	555
Región de Murcia	15	6	5.856	3%	134	5.722	-
Comunidad Valenciana	15	4	16.329	8%	1.279	15.050	-
<b>Totales</b>	<b>255</b>	<b>125</b>	<b>207.706</b>	<b>100%</b>	<b>29.496</b>	<b>177.460</b>	<b>750</b>

Fuente: Elaboración propia.

### 3.2. Clasificación de los comentarios con Inteligencia Artificial

Para explorar las percepciones sostenibles de los turistas, se aplica el enfoque emergente de aprendizaje automático conocido como “*Zero-shot learning*”. Este método permite que un modelo de inteligencia artificial categorice clases o conceptos sin haber sido entrenado explícitamente en ellos (Alhoshan et al., 2023). Se emplea el modelo “es-BERT *cased*”, también conocido como “BETO *cased*” (Cañete et al., 2020), una versión adaptada al idioma español de BERT. Se selecciona este modelo por cuatro razones. En primer lugar, ha sido aplicado a otros estudios con éxito, tal como se detalló en la sección 2 de este capítulo, “Revisión de la literatura”. En segundo lugar, BETO *cased* es capaz de comprender el contexto de las palabras en una oración al predecir palabras en función de su contexto tanto a la izquierda como a la derecha (bidireccional). Esto lo hace especialmente útil para una variedad de tareas PLN, como la clasificación de texto, la generación de texto y la traducción. En tercer lugar, BETO *cased* es un modelo BERT entrenado con un gran corpus en español y la técnica *Whole Word Masking* (wwm) (Cañete et al., 2020). En cuarto lugar, BETO *cased* sobresale por su rendimiento superior en comparación con otros modelos. Cañete et al. (2020) evalúan el rendimiento de BETO *cased* en diversas tareas de PNL, frente a dos modelos: 1) “es-BERT *uncased*” (BETO *uncased*) y 2) mBERT (versión “multilingüe” de BERT) que destaca por su rendimiento en una amplia gama de contextos lingüísticos (S. Wu & Dredze, 2019).

En este estudio, se utiliza la versión de ajuste fino de BETO *cased* (*fine-tuned*), “*bert-base-spanish-wwm-cased-xnli*”, creada por la empresa Recognai (2021) y disponible en la plataforma de código abierto de *Hugging Face*<sup>6</sup>. Este modelo fue entrenado con la porción española del conjunto de datos XNLI (Cross-lingual Natural Language Inference). Esto ayuda a adaptar y mejorar el modelo para el idioma español y para la tarea específica de inferencia de lenguaje natural en ese idioma. El conjunto de datos XNLI se utiliza para entrenar y evaluar modelos en la tarea de inferencia de lenguaje natural cruzada. Este conjunto de datos contiene pares de oraciones traducidas a múltiples idiomas, incluida una versión en español, que permite a los modelos de PLN aprender a razonar sobre relaciones entre frases en distintos contextos lingüísticos.

La aplicación del modelo BETO *cased* permite categorizar el contenido de cada comentario en las dimensiones sostenibles de manera sistemática. Es decir, se identifican los aspectos

---

<sup>6</sup> Actualmente, *Hugging Face* es el repositorio de modelos y conjuntos de datos de Inteligencia Artificial (IA) de código abierto más extenso: <https://huggingface.co/>

ambientales, culturales, sociales y económicos dentro de cada comentario y se asigna una puntuación a cada uno ellos según su grado de importancia dentro del contexto del comentario. Asimismo, se añade la categoría “otros” para considerar el contenido de las reseñas que no estaba relacionado con aspectos sostenibles (satisfacción o temas generales). En la Tabla 10 se muestran algunos resultados obtenidos después de cargar y ejecutar el modelo. Cada comentario se clasifica y se puntúa según su discurso sostenible.

**Tabla 10. Ejemplos de comentarios con su clasificación y puntuación.**

<b>Reseña</b>	<b>ECO</b>	<b>SOC</b>	<b>ENV</b>	<b>CULT</b>	<b>OTROS</b>
Hay un tour con el que puedes visitar 5 sitios, es muy recomendado, es bastante económico, lo venden en la oficina de turismo.	95,97	0,96	0,70	0,85	1,52
Aunque 7 euros pueden parecer caros, la visita merece la pena; en la entrada te proporcionan una audioguía para entender mejor esta Catedral.	93,82	1,32	0,73	1,98	2,15
Mejorar las normas de seguridad, la distancia social en los espectáculos. Produce discusiones entre los espectadores. Deben señalarse los asientos que no se deben ocupar.	1,62	89,19	0,87	1,38	6,94
Todo en el Barrio Gótico tiene su encanto: iglesias escondidas entre callejuelas, músicos callejeros que te sacan una sonrisa, catedrales impresionantes, gente de todo tipo. En conjunto, uno de los mejores barrios de Barcelona para perderse.	4,54	68,95	3,39	14,07	9,05
Sin duda, si desea visitar un entorno natural representativo del bosque mediterráneo, debe explorar este maravilloso parque nacional.	0,64	1,10	94,55	1,60	2,11
Gran entorno, se oye el río Segura, hay espacio para barbacoa y puedes llevar a tu perro. Y además, el Santuario es precioso.	1,47	1,55	93,22	1,90	1,86
La catedral posee un importante valor cultural, que reconocen y destacan restaurándola y conservándola. En su interior, explican el significado de cada cosa, dando a conocer su importancia histórica y cultural.	0,51	1,45	0,81	95,53	1,69
Un ejemplo clásico de nuestra mezcla cultural. El Museo Sefardí, con excelentes textos de apoyo disponibles en la tienda de regalos, ayuda a los visitantes a comprender mejor la visita y a profundizar en su historia.	0,61	2,13	0,74	86,68	9,85

Reseña	ECO	SOC	ENV	CULT	OTROS
Otros lectores y visitantes lo han descrito mejor que yo. A mí, la verdad, me encantó. Destacaría dos detalles: el brazo de San Vicente y la Capilla del Santo Cáliz. Por lo demás, es un templo de visita obligada.	0,75	2,32	0,80	2,19	93,94
Te transporta a otra época en la que las cosas se hacían con más ingenio que tecnología. Los cañones son algo de otro mundo.	1,54	2,51	1,16	2,54	92,24
Cultura por doquier, testimonio de la Hispania romana. Es una visita que no puede perderse si su objetivo es comprender la historia de España y sus raíces.	0,39	5,30	0,36	89,79	4,16
¡¡¡Precioso museo, magníficamente cuidado, gran ambiente!!! ¡La entrada es muy razonable para todo lo que ofrece! Recomendable.	13,62	0,83	47,57	37,21	0,76

**Fuente: Elaboración propia.**

### 3.3. Modelos Ologit y Gologit

Para responder a las hipótesis de investigación en esta sección se emplea un modelo de regresión logística ordenada (*ologit*) para comprender cómo las dimensiones sostenibles de los comentarios turísticos influyen en la satisfacción, reflejada en las calificaciones de los atractivos. La variable dependiente es la calificación y las variables independientes son las dimensiones de sostenibilidad (económica, social, cultural y ambiental) calculadas previamente utilizando el modelo “BETO *cased*”. El modelo *ologit* fue considerado por primera vez por McCullagh (1980) y se utiliza cuando la variable dependiente es ordinal, es decir, tiene categorías ordenadas. Es una extensión de la regresión logística binaria y modela la probabilidad de pertenencia a una categoría ordinal específica. Al igual que D’Acunto et al. (2023), se contemplan como variables de control: la longitud de la reseña (número de palabras), la experiencia del revisor (contribución) y la categoría de viaje. También se añade a las variables de control la nacionalidad del turista. Estas variables se detallan en la Tabla 13.

Antes de aplicar el modelo *ologit*, es necesario validar y preparar los datos. Para ello, se evaluó la correlación entre variables independientes, se transformaron tanto las variables de las dimensiones sostenibles, como las contribuciones y número de palabras a logaritmo neperiano para equilibrar valores extremos y se convirtieron las variables de categoría de viaje y nacionalidad en categorías *dummy* (binarias). Posteriormente, se verificó el cumplimiento del supuesto de las rectas paralelas con la prueba de Brant. Esta prueba asume como hipótesis nula

que no hay diferencia significativa en la forma en que las variables independientes están asociadas con la variable dependiente en diferentes categorías de la variable dependiente. Si el resultado es menor que 0,05 se rechaza la hipótesis nula y se acepta la hipótesis alternativa de que si hay diferencias. En ese caso, es necesario aplicar la regresión logística ordenada generalizada (*gologit*) (Williams, 2006).

En la Tabla 11 se detallan las fórmulas de los modelos *ologit* y *gologit* (UCLA, 2024b; Williams, 2006).

**Tabla 11. Ecuaciones de los modelos *ologit* y *gologit*.**

Modelo	Ecuación	Descripción
Ologit	$\text{logit}(P(Y \leq j)) = \beta_{j0} + \beta_1 x_1 + \dots + \beta_p x_p$	Sea $Y$ un resultado ordinal con $j$ categorías. Entonces, $P(Y \leq j)$ es la probabilidad de estar en una categoría menor o igual a una categoría particular.
Gologit	$P(Y_i > j) = \frac{\exp(a_j + x_i \beta_j)}{1 + \{\exp(a_j + x_i \beta_j)\}}, j = 1, 2, \dots, M - 1$	<p><math>\beta_{j0} + \beta_1 + \dots + \beta_p</math> son los parámetros de los coeficientes del modelo (es decir, interceptos y pendientes) con <math>p</math> predictores.</p> <p><math>x_1 + \dots + x_p</math> son las variables independientes.</p> <p><math>M</math> es el número de categorías de la variable dependiente ordinal.</p> <p><math>a_j</math> es el término de intersección específico para la categoría <math>j</math>.</p>

**Fuente: Elaboración propia a partir de: (UCLA, 2024b; Williams, 2006).**

Las estimaciones de un modelo de regresión logística (*logit*) se obtienen mediante máxima verosimilitud y no minimizan la varianza como en la regresión de mínimos cuadrados ordinarios. Por lo tanto, el coeficiente R-cuadrado no es aplicable ni interpretable como medida de ajuste en modelos *logit*. En su lugar, se utiliza el Pseudo R-cuadrado, que, aunque no se puede interpretar independientemente ni comparar entre conjuntos de datos, es válido para evaluar modelos que predicen el mismo resultado en el mismo conjunto de datos. Un Pseudo R-cuadrado más alto indica un mejor modelo predictivo (Hemmert et al., 2016; UCLA, 2024a).

Otros dos parámetros importantes para la correcta interpretación de los modelos son los coeficientes y odds ratio. Un coeficiente positivo indica que los valores más altos de las variables independientes (o de control) hacen más probables valores más altos de la variable dependiente

(rating). Sin embargo, lo contrario sucede para los coeficientes negativos. Por otro lado, un odds ratio proporciona la relación entre las posibilidades de que ocurra un evento en dos grupos diferentes. Es decir, indica cómo un cambio en una variable independiente (o de control) afecta las probabilidades de pertenecer a una categoría particular de la variable dependiente, en comparación con una categoría de referencia. Un valor mayor a 1 indica una asociación positiva entre las variables, mientras que un valor menor a 1 indica una asociación negativa.

En la Tabla 12 se detalla la cantidad de comentarios por cada una de las categorías de la variable dependiente (calificación), que varía de 1 a 5. La mayor cantidad de comentarios se acumula en los rangos 4 y 5, evidenciando notables diferencias entre categorías. Por tanto, se elaboran cuatro modelos *ologit* como prueba de robustez. Para los modelos 1 y 2 (M1 y M2) se consideraron las 5 categorías de la calificación. Sin embargo, para los modelos 3 y 4 (M3 y M4) se agruparon tanto las categorías 1 y 2, como las categorías 3 y 4, dejando únicamente 3 categorías de calificación. Además, con el fin de buscar un mejor ajuste de los modelos, se varió el número de variables de control de los modelos. M1 y M3 incluyeron 2 variables de control (categoría de viaje y nacionalidad del turista). Sin embargo, M2 y M4 contemplaron 4 variables de control (categoría de viaje, nacionalidad del turista, longitud del comentario y experiencia del revisor). En la Tabla 13 se detallan las variables utilizadas para el estudio y cada uno de los modelos.

**Tabla 12. Número de reseñas por categoría de la variable dependiente.**

Categoría	Número de reseñas	M1 y M2	M3 y M4
1	2.503	2.503	5.209
2	2.706	2.706	67.587
3	12.693	12.693	134.910
4	54.894	54.894	
5	134.910	134.910	
<b>Totales</b>	<b>207.706</b>	<b>207.706</b>	<b>207.706</b>

Fuente: Elaboración propia.

**Tabla 13. Variables incluidas en los análisis de regresión.**

Constructo	Nombre de la variable	Tipo de la variable	Operacionalización	M1	M2	M3	M4
Rating Score	RATING1	Variable dependiente	Variables categóricas de 1 a 5	X	X		
Rating Score	RATING2	Variable dependiente	Variables categóricas de 1 a 3			X	X

Constructo	Nombre de la variable	Tipo de la variable	Operacionalización	M1	M2	M3	M4
Contenido económico de la reseña	ECO	Variable independiente	El logaritmo natural de la puntuación calculada del contenido económico	X	X	X	X
Contenido social de la reseña	SOC	Variable independiente	El logaritmo natural de la puntuación calculada del contenido social	X	X	X	X
Contenido ambiental de la reseña	ENV	Variable independiente	El logaritmo natural de la puntuación calculada del contenido ambiental	X	X	X	X
Contenido cultural de la reseña	CULT	Variable independiente	El logaritmo natural de la puntuación calculada del contenido cultural	X	X	X	X
Origen	ORI	Variable de control	Variable dummy: Turista nacional (0 extranjero)	X	X	X	X
Categoría de viaje	FRI, SOLO, FAM	Variable de control	Variable dummy: Amigos, Solo y Familia (0 parejas)	X	X	X	X
Longitud de la reseña	WC	Variable de control	El logaritmo natural del recuento de palabras de cada reseña		X		X
Contribuciones de los revisores	CONT	Variable de control	El logaritmo natural del número total de opiniones publicadas por el revisor		X		X

**Fuente: Elaboración propia.**

## 4. Resultados

### 4.1. Pruebas estadísticas

Los análisis estadísticos se llevaron a cabo con el programa Stata. En primer lugar, se aplican los modelos *ologit* a la totalidad de la muestra (207.706 comentarios), con el objetivo de validar las hipótesis planteadas. En segundo lugar, se realiza un análisis separado aplicando modelos *ologit* tanto a las reseñas rurales (29.496 comentarios) como a las urbanas (177.460 comentarios), para abordar la pregunta de investigación. En tercer lugar, se verifica si los modelos son adecuados mediante la prueba de Brant. Los resultados de esta prueba indican la violación de los supuestos de proporcionalidad en todos los modelos, por lo que fue necesario aplicar modelos *gologit* (regresión logística ordenada generalizada) a los datos. Esto significa

que hay diferencias significativas en la forma en que las variables independientes están asociadas con la variable dependiente (rating) en sus diferentes categorías. En consecuencia, tanto la validación de las hipótesis como la respuesta a la pregunta de investigación se basan en los resultados obtenidos de las regresiones *gologit*.

Para los modelos de regresión logística ordenada generalizada, se utiliza el módulo *gologit2* escrito por Williams (2006) con la opción de autoajuste, puesto que permite relajar la restricción de líneas paralelas únicamente en las variables donde se viole el supuesto. En los siguientes apartados se puntualizan los resultados de los modelos *gologit* con las intersecciones o puntos de corte entre los diferentes rangos de las variables dependientes. Es importante aclarar que, como se mencionó previamente en la Tabla 13, el constructo “categoría de viaje” es una variable *dummy* que consta de cuatro opciones distintas (amigos, en solitario, familia y parejas). Sin embargo, la categoría “parejas” se excluye de los modelos *ologit* y *gologit* estimados, siendo asignada el valor de cero. Por tanto, se convierte en el punto de referencia para la comparación de todos los resultados de las demás opciones (amigos, en solitario y familia).

#### 4.1.1. Influencia del discurso sostenible de los turistas sobre la calificación

La Tabla 14 muestra los resultados de los modelos *gologit* para toda la muestra (207.706 reseñas). Al observar los modelos, se encuentra que los resultados son consistentes en los cuatro, indicando la confiabilidad de los hallazgos. Sin embargo, M2 y M4 superan a M1 y M3 debido a sus Pseudo R-cuadrados más altos, indicando que los modelos con todas las variables de control proyectan mejores resultados. Cabe destacar que los Pseudo R-cuadrados no deben interpretarse como medidas de bondad de ajuste en el mismo sentido utilizado en el método de mínimos cuadrados ordinarios, ya que las estimaciones de un modelo *logit* son de máxima verosimilitud y no se calculan para minimizar la varianza.

Tabla 14. Resultados de los modelos gologit.

Modelo		M1		M2		M3		M4	
Pseudo R2		0.0219		0.0316		0.0223		0.0335	
Variable	Corte	Coeff.	Odds Ratio						
<b>_cons</b>	Cut 1	2.3823***	10.8299	4.8488***	127.5879	1.4490***	4.2591	4.1258***	61.9174
	Cut 2	1.3149***	3.7245	3.9547***	52.1787	-.5270***	.59037	.3433***	1.4095
	Cut 3	.2103***	1.2341	1.8208***	6.1769				
	Cut 4	-.5290***	1.0312	.3364***	1.3999				
<b>ECO</b>	Cut 1	-.1303***	.8778	-.1830***	.8327	-.1325***	.8759	-.1854***	.8308
	Cut 2	-.1303***	.8778	-.1900***	.8270	-.1325***	.8759	-.1349***	.8738
	Cut 3	-.1303***	.8778	-.1401***	.8693				
	Cut 4	-.1303***	.8778	-.1344***	.8743				
<b>SOC</b>	Cut 1	-.6039***	.5467	-.5192***	.5950	-.4198***	.6572	-.3894***	.6775
	Cut 2	-.4105***	.6633	-.3517***	.7035	-.0087	.9913	-.0114	.9887
	Cut 3	-.1465***	.8637	-.1422***	.8675				
	Cut 4	-.0098	.9902	-.0117	.9883				
<b>ENV</b>	Cut 1	.5642***	1.7581	.5014***	1.6511	.4772***	1.6115	.4274***	1.5333
	Cut 2	.4995***	1.6479	.4314***	1.5394	.1600***	1.1735	.1442***	1.1551
	Cut 3	.3562***	1.4279	.3183***	1.3748				
	Cut 4	.1599***	1.1734	.1438***	1.1547				
<b>CULT</b>	Cut 1	1.0300***	2.8010	.8383***	2.3125	.9533***	2.5944	.8024***	2.2310
	Cut 2	.9587***	2.6082	.7973***	2.2196	.3290***	1.3896	.3254***	1.3846
	Cut 3	.6831***	1.9800	.6295***	1.8767				
	Cut 4	.3295***	1.3902	.3247***	1.3836				
<b>FRI</b>	Cut 1	.0952***	1.0998	.0831***	1.0867	.0158	1.0159	.0826***	1.0861
	Cut 2	.0952***	1.0998	.0831***	1.0867	.0983***	1.1032	.0826***	1.0861
	Cut 3	.0952***	1.0998	.0831***	1.0867				
	Cut 4	.0952***	1.0998	.0831***	1.0867				
<b>SOLO</b>	Cut 1	-.0599	.9419	-.1125	.8936	-.0133	.9868	.0051	1.0052

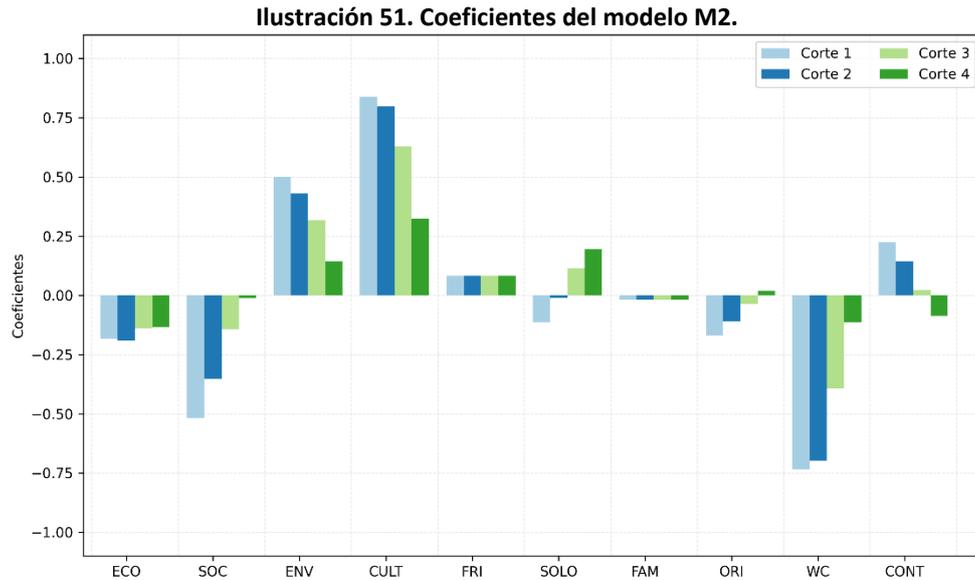
Modelo		M1		M2		M3		M4	
Pseudo R2		0.0219		0.0316		0.0223		0.0335	
Variable	Corte	Coeff.	Odds Ratio						
	Cut 2	.0208	1.0210	-.0107	.9894	.1979***	1.2188	.1965***	1.2171
	Cut 3	.1034**	1.1090	.1146***	1.1214				
	Cut 4	.1965***	1.2171	.1947***	1.2150				
FAM	Cut 1	-.1943***	.8234	-.0199	.9803	-.1609***	.8514	-.0181	.9820
	Cut 2	-.1302***	.8779	-.0199	.9803	.0073	1.0073	-.0181	.9820
	Cut 3	-.0630***	.9390	-.0199	.9803				
	Cut 4	.0061	1.0061	-.0199	.9803				
ORI	Cut 1	-.2650***	.7672	-.1705**	.8432	-.1994***	.8192	-.1098**	.8960
	Cut 2	-.1885***	.8282	-.1091**	.8966	.0309*	1.0314	.0201	1.0203
	Cut 3	-.0657***	.9364	-.0356	.9650				
	Cut 4	.0307**	1.0312	.0198	1.0200				
WC	Cut 1			-.7353***	.4794			-.7320***	.4809
	Cut 2			-.6984***	.4974			-.1165***	.8901
	Cut 3			-.3941***	.6742				
	Cut 4			-.1133***	.8929				
CONT	Cut 1			.2238***	1.2508			.1563***	1.1692
	Cut 2			.1427***	1.1534			-.0858***	.9178
	Cut 3			.0228***	1.0231				
	Cut 4			-.0861***	.9175				

Los niveles de significación se indican como: \*\*\* para  $p$ -value < 0,001, \*\* para  $p$ -value < 0,01, y \* para  $p$ -value < 0,05.

Fuente: Elaboración propia.

La probabilidad de recibir calificaciones más altas para las revisiones disminuyó cuando el discurso de los turistas contiene aspectos económicos y sociales (ECO y SOC), mientras que las revisiones enriquecidas con aspectos ambientales (ENV) y culturales (CULT) tienden a aumentar las puntuaciones de calificación, indicando que estos aspectos influyen positivamente en las percepciones de los usuarios. Además, la naturaleza del viaje juega un papel fundamental. Las personas que viajan con amigos (FRI) o viajeros en solitario (SOLO), a diferencia de los viajes en pareja, tienden a dar calificaciones más favorables a las atracciones turísticas. Por el contrario, los usuarios domésticos (ORI) muestran una propensión a ser más críticos, a menudo clasifican las atracciones turísticas de manera menos positiva en comparación con los viajeros internacionales (Tabla 14, modelos 1 y 3, respectivamente). Además, las revisiones que se caracterizan por un mayor número de palabras (WC) tienden a producir puntuaciones más bajas, lo que refleja la inclinación a elaborar sobre experiencias negativas que requieren largas explicaciones. Esto confirma los resultados de Filieri et al. (2019), debido a que demuestran que las calificaciones negativas se asocian con reseñas largas y los revisores que realizan muchas contribuciones son más exigentes al momento de valorar sus experiencias. Del mismo modo, los usuarios expertos (CONT) tienden a asignar calificaciones más altas a los atractivos turísticos. Este comportamiento podría reflejar que, con mayor experiencia, los usuarios desarrollan una apreciación más crítica pero positiva hacia ciertos aspectos de las atracciones. Sin embargo, estos resultados difieren de investigaciones previas, que suelen asociar una mayor experiencia con reseñas más negativas, dado que el efecto de la calidad percibida del destino es más débil en viajeros expertos (T. L. Wang et al., 2017; Z. Zhang et al., 2016).

De los cuatro modelos analizados, M2 se convierte en la opción más sólida para reflejar los resultados de esta investigación, puesto que es el modelo que contempla todas las variables y puntos de corte (Cortes 1 - 4), a su vez, su Pseudo R-cuadrado esta entre los dos más altos. Si se analizan las discrepancias más significativas de los *odds* ratio de las variables independientes (ver Tabla 14), se encuentra que un aumento de una unidad de la variable SOC está asociado con una disminución del 41% de pertenecer a una categoría más alta ( $1 - 0,5950 \cdot 100$ ) en el corte 1 (transición de 1 a 2 en el *rating*). Por el contrario, el *odds* ratio de CULT en el corte 1 se refleja que un aumento de una unidad en la variable CULT, incrementa un 131% la probabilidad de pertenecer a una categoría más alta. En la Ilustración 51 se muestran los coeficientes de M2 detallados en la Tabla 14 y explicados previamente.



Fuente: Elaboración propia.

Para la validación de las hipótesis y teniendo en cuenta que no se cumple con el supuesto de probabilidades proporcionales, se ha optado por aceptar una hipótesis cuando los coeficientes de todos los puntos de corte muestren una significancia estadística en concordancia con la relación planteada, ya sea positiva o negativa. Por tanto, se acepta H1, H2 y H3, puesto que los coeficientes de ENV y CULT son positivos y estadísticamente significativos, mientras que los de ECO son negativos y muestran una significación estadística. Esto prueba que las reseñas con un mayor discurso sobre las dimensiones ambiental y cultural influyen positivamente en la calificación de los atractivos turísticos. Sin embargo, las reseñas que contienen un mayor discurso sobre la dimensión económica tienden a obtener calificaciones más bajas. Estos hallazgos se alinean con investigaciones previas (Mariani & Borghi, 2023; Rita et al., 2022; Väisänen et al., 2023).

Por otro lado, H4 correspondiente a la variable independiente SOC se rechaza, ya que, a pesar de que todos sus coeficientes son negativos, el coeficiente de la última intersección (corte 4) no es estadísticamente significativo. Por consiguiente, no se puede afirmar que las reseñas con un mayor discurso sobre la dimensión social influyen negativamente en la calificación asignada por el turista. Estos resultados difieren del estudio de D'Acunto et al. (2020). Sin embargo, se cumplió la relación planteada con respecto a la calificación, siendo negativa en todos los coeficientes. Es decir, se evidencia que los comentarios que contienen un mayor discurso social tienden a influir negativamente en la calificación de los turistas, pero no se pudo validar estadísticamente. Esto se debe a que un modelo *gologit* muestra cómo las variables

independientes afectan la probabilidad de transición entre las distintas categorías de la variable dependiente en cada punto de corte. Las diferencias entre las intersecciones no reflejan necesariamente cambios globales en la relación entre las variables, puesto que se pudo comprobar que las relaciones entre las variables (dependientes e independientes) no son paralelas. Por consiguiente, se podría decir que la dimensión social no influye significativamente en la probabilidad de que un atractivo turístico sea calificado con una puntuación de cuatro en lugar de una puntuación de cinco.

#### 4.1.2. Diferencias de percepción sostenible del turista en zonas rurales y urbanas

Considerar los modelos bajo el carácter urbano/rural de las atracciones turísticas produce resultados interesantes (ver Tabla 15 y Tabla 16). Al igual que antes, los modelos 2 y 4 son los que obtienen los mejores resultados, puesto que tienen los Pseudo R-cuadrados más altos. Si bien la probabilidad de dar una calificación más alta es también disminuida por el discurso asociado con las dimensiones económica (ECO) y social (SOC) en atracciones urbanas, la probabilidad es mayor cuando las revisiones contienen referencias sociales, pero se suaviza cuando la revisión contiene referencias económicas. Las dimensiones ambientales (ENV) y culturales (CULT) en las reseñas tienen un efecto similar en las atracciones turísticas rurales y urbanas, aunque la dimensión ambiental muestra un impacto mayor en atractivos rurales y la dimensión cultural en atractivos urbanos.

Los viajes realizados con amigos (FRI) dan puntuaciones más altas a las atracciones turísticas urbanas, a diferencia de los viajes de parejas. En cambio, los resultados en atractivos rurales son menos consistentes y no significativos, lo que sugiere que la compañía de amigos no tiene un impacto en las calificaciones de las atracciones rurales. Las personas que viajan solas (SOLO) suelen ser más críticas y exigentes que las parejas en zonas rurales, otorgando calificaciones más altas en zonas urbanas. Los viajes familiares (FAM) no valoran positivamente las atracciones rurales y tienden a dar calificaciones bajas, mientras que en las urbanas su impacto no es significativo, lo que sugiere que este tipo de viaje no influye de manera clara en la valoración de atractivos urbanos. Por otro lado, los viajeros domésticos (ORI) dan puntuaciones más bajas a los entornos urbanos en comparación con los extranjeros, lo que podría sugerir que los nacionales tienen expectativas más altas sobre estos atractivos turísticos. Por el contrario, esta variable no tiene influencia en las calificaciones de los atractivos rurales, ya que su efecto no es significativo.

La necesidad de explicar las experiencias negativas no parece estar relacionada con el tipo de entorno, aunque la extensión de la reseña (WC) influye en las puntuaciones bajas en los entornos rurales más que en las atracciones turísticas urbanas. Además, los viajeros experimentados (CONT), en general, califican positivamente las atracciones tanto en ámbitos rurales como urbanos, aunque tienden a otorgar calificaciones aún más altas en entornos urbanos. Esto sugiere que, a medida que aumenta la experiencia de los usuarios, su percepción de valor en las atracciones urbanas se ve favorecida, posiblemente debido a una mayor diversidad de servicios o a una oferta turística que responde mejor a sus expectativas. Para una interpretación más sencilla de los resultados, se han graficado en la Ilustración 52 los coeficientes por cada punto de corte de M2 para toda la muestra, así como para las áreas rurales y urbanas.

Tabla 15. Modelos gologit por zonas rurales y urbanas de M1 y M2.

Modelo		M1				M2			
Pseudo R2		0.0169	0.0233			0.0255	0.0341		
Variable	Corte	Coeff.		Odds Ratio		Coeff.		Odds Ratio	
		Rural	Urbano	Rural	Urbano	Rural	Urbano	Rural	Urbano
<b>_cons</b>	Cut 1	2.3961***	2.3415***	10.9798	10.3973	6.2360***	4.7690	510.8377	117.8068
	Cut 2	1.4242***	1.2807***	4.1547	3.5991	4.6923***	3.9417	109.108	51.5074
	Cut 3	.3068	.1612***	1.3590	1.1750	2.1347***	1.8596	8.4548	6.4209
	Cut 4	-.5189***	-.5729**	.5952	.5639	.5332***	.3118	1.7044	1.3658
<b>ECO</b>	Cut 1	-.0847***	-.1333***	.9188	.8752	-.1578*	-.1946***	.8540	.8231
	Cut 2	-.0847***	-.1333***	.9188	.8752	-.1569***	-.1969***	.8547	.8213
	Cut 3	-.0847***	-.1333***	.9188	.8752	-.0643*	-.1469***	.9377	.8633
	Cut 4	-.0847***	-.1333***	.9188	.8752	-.0993***	-.1359***	.9055	.8730
<b>SOC</b>	Cut 1	-.7878***	-.5771***	.4549	.5615	-.6436***	-.4999***	.5254	.6066
	Cut 2	-.5715***	-.3855***	.5647	.6801	-.4338***	-.3432***	.6480	.7095
	Cut 3	-.1328***	-.1455***	.8756	.8646	-.0833*	-.1482***	.9201	.8622
	Cut 4	.0132	-.0088	1.0133	.99123	.0218	-.0092	1.0220	.9908
<b>ENV</b>	Cut 1	.7008***	.5359***	2.0154	1.7090	.5770***	.4703***	1.7806	1.6004
	Cut 2	.6130***	.4738***	1.8459	1.6061	.4932***	.4036***	1.6376	1.4972
	Cut 3	.4289***	.3381***	1.5355	1.4022	.3640***	.2966***	1.4390	1.3453
	Cut 4	.2346***	.1446***	1.2644	1.1555	.2055***	.1272***	1.2282	1.1357
<b>CULT</b>	Cut 1	1.0222***	1.040***	2.7793	2.8302	.7920***	.8313***	2.2079	2.2963
	Cut 2	.9397***	.9683***	2.5591	2.6336	.7478***	.8056***	2.1123	2.2380
	Cut 3	.5735***	.7078***	1.7744	2.0295	.4828***	.6557***	1.6206	1.9265
	Cut 4	.2537***	.3506***	1.2888	1.4200	.2101***	.3526***	1.2338	1.4228
<b>FRI</b>	Cut 1	.0081	.1138***	1.0081	1.1205	.0022	.1012***	1.0022	1.1065
	Cut 2	.0081	.1138***	1.0081	1.1205	.0022	.1012***	1.0022	1.1065
	Cut 3	.0081	.1138***	1.0081	1.1205	.0022	.1012***	1.0022	1.1065
	Cut 4	.0081	.1138***	1.0081	1.1205	.0022	.1012***	1.0022	1.1065

Modelo		M1				M2			
Pseudo R2		0.0169	0.0233			0.0255	0.0341		
Variable	Corte	Coeff.		Odds Ratio		Coeff.		Odds Ratio	
		Rural	Urbano	Rural	Urbano	Rural	Urbano	Rural	Urbano
<b>SOLO</b>	Cut 1	.0359	-.0091	1.0365	.9909	-.7116*	-.0128	.4909	.9872
	Cut 2	.0359	.0449	1.0365	1.0460	-.1874	.0583	.8291	1.0600
	Cut 3	.0359	.1322***	1.0365	1.1413	-.0226	.1643***	.9776	1.1786
	Cut 4	.0359	.2236***	1.0365	1.2506	.0491	.2167***	1.0503	1.2420
<b>FAM</b>	Cut 1	-.0795**	-.1831***	.9236	.8327	-.0708**	-.0089	.9317	.9911
	Cut 2	-.0795**	-.1218***	.9236	.8853	-.0708**	-.0089	.9317	.9911
	Cut 3	-.0795**	-.0553**	.9236	.9461	-.0708**	-.0089	.9317	.9911
	Cut 4	-.0795**	.0215	.9236	1.0218	-.0708**	-.0089	.9317	.9911
<b>ORI</b>	Cut 1	.0149	-.3046***	1.0150	.7374	.0173	-.1818**	1.0175	.8337
	Cut 2	.0149	-.2305***	1.0150	.7941	.0173	-.1328**	1.0175	.8756
	Cut 3	.0149	-.0814***	1.0150	.9218	.0173	-.0531*	1.0175	.9483
	Cut 4	.0149	.0208	1.0150	1.0211	.0173	.0059	1.0175	1.0059
<b>WC</b>	Cut 1					-.8686***	-.7493***	.4195	.4727
	Cut 2					-.7668***	-.7152***	.4645	.4891
	Cut 3					-.4567***	-.4136***	.6334	.6613
	Cut 4					-.1921***	-.1105***	.8252	.8953
<b>CONT</b>	Cut 1					.1310***	.2577***	1.1400	1.2939
	Cut 2					.1354***	.1575***	1.1450	1.1706
	Cut 3					.0542***	.0241***	1.0557	1.0244
	Cut 4					-.0284***	-.0971***	.9720	.9074

Los niveles de significación se indican como: \*\*\* para  $p\text{-value} < 0,001$ , \*\* para  $p\text{-value} < 0,01$ , y \* para  $p\text{-value} < 0,05$ .

Fuente: Elaboración propia.

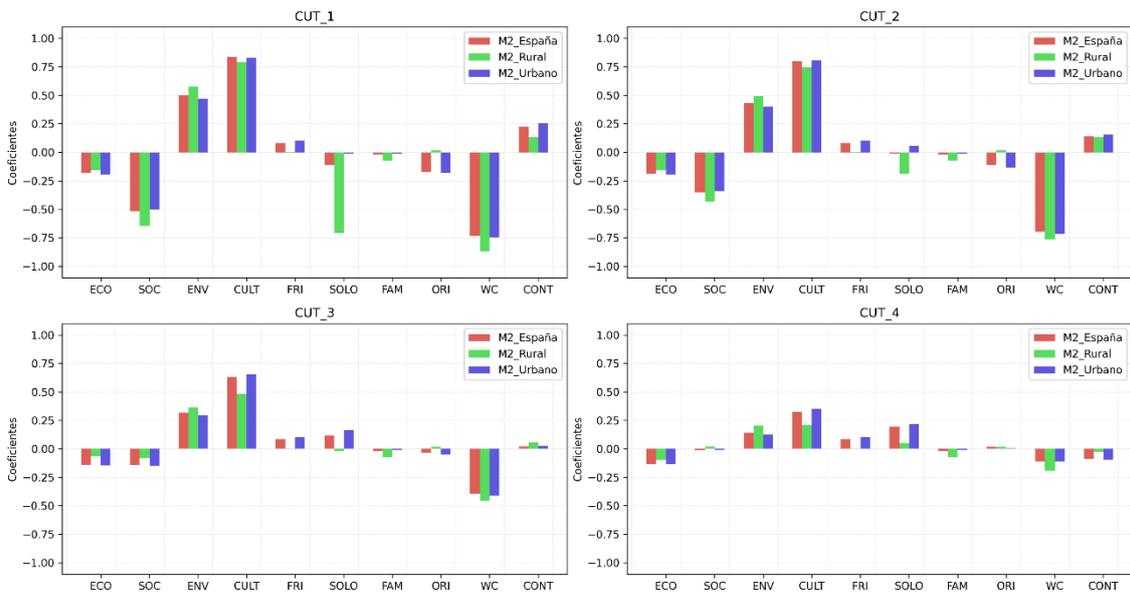
Tabla 16. Modelos gologit por zonas rurales y urbanas de M3 y M4.

Modelo		M3				M4			
Pseudo R2		0.0187	0.0236			0.0288	0.0359		
Variable	Corte	Coeff.		Odds Ratio		Coeff.		Odds Ratio	
		Rural	Urbano	Rural	Urbano	Rural	Urbano	Rural	Urbano
<b>_cons</b>	Cut 1	1.4289***	1.4008***	4.1742	4.0585	4.7178***	4.1624***	111.926	64.2286
	Cut 2	-.5135***	-.5690***	.5983	.5661	.5284***	.3201***	1.6963	1.3772
<b>ECO</b>	Cut 1	-.0911***	-.1352***	.9129	.8736	-.1014***	-.1929***	.9036	.8246
	Cut 2	-.0911***	-.1352***	.9129	.8736	-.1014***	-.1365***	.9036	.8723
<b>SOC</b>	Cut 1	-.6221***	-.4008***	.5368	.6698	-.5027***	-.3754***	.6049	.6870
	Cut 2	.0164	-.0074	1.0165	.9926	.0243	-.0092	1.0246	.9908
<b>ENV</b>	Cut 1	.6107***	.4531***	1.8417	1.5732	.5088***	.3957***	1.6632	1.4855
	Cut 2	.2339***	.1443***	1.2636	1.1552	.2065***	.1276***	1.2294	1.1360
<b>CULT</b>	Cut 1	.9966***	.9644***	2.7090	2.6233	.8058***	.8001***	2.2385	2.2259
	Cut 2	.2514***	.3496***	1.2858	1.4186	.2098***	.3535***	1.2334	1.4241
<b>FRI</b>	Cut 1	.0031	.1146***	1.0031	1.1214	-.0041	.1017***	.9959	1.1071
	Cut 2	.0031	.1146***	1.0031	1.1214	-.0041	.1017***	.9959	1.1071
<b>SOLO</b>	Cut 1	.0351	.0435	1.0357	1.0445	.0392	.0831	1.0399	1.0867
	Cut 2	.0351	.2240***	1.0357	1.2511	.0392	.2197***	1.0399	1.2457
<b>FAM</b>	Cut 1	-.0795**	-.1244***	.9236	.8831	.1512	-.0061	1.1632	.9939
	Cut 2	-.0795**	.0219	.9236	1.0221	-.0812**	-.0061	.9220	.9939
<b>ORI</b>	Cut 1	.0265	-.2474***	1.0268	.7809	.0295	-.1472***	1.0299	.863
	Cut 2	.0265	.0213	1.0268	1.0215	.0295	.0061	1.0299	1.0061
<b>WC</b>	Cut 1					-.8517***	-.7556***	.4267	.4697
	Cut 2					-.1934***	-.1141***	.8242	.8922
<b>CONT</b>	Cut 1					.1631***	.1736***	1.1772	1.1895
	Cut 2					-.0285***	-.0968***	.9719	.9077

Los niveles de significación se indican como: \*\*\* para  $p$ -value < 0,001, \*\* para  $p$ -value < 0,01, y \* para  $p$ -value < 0,05.

Fuente: Elaboración propia.

Ilustración 52. Coeficientes del modelo M2 para España, zonas rurales y urbanas.



Fuente: Elaboración propia.

## 5. Conclusiones

Esta investigación se centra en la importancia de la percepción de los turistas sobre las dimensiones sostenibles y cómo influye en la satisfacción de los principales atractivos turísticos de España. Los hallazgos tienen relevancia tanto a nivel teórico como práctico. Primero, se amplían los estudios sobre sostenibilidad en turismo que aplican de forma práctica los conceptos teóricos e involucran a las partes interesadas (Bramwell et al., 2017; Ruhanen et al., 2015; Sharpley, 2020). Este estudio se enfoca en las percepciones sostenibles de los turistas debido a la limitada investigación y la necesidad de comprender el punto de vista de este grupo de interés (Mior Shariffuddin et al., 2023; Sánchez-Fernández et al., 2019; Sreen et al., 2022). Comprendiendo las percepciones sostenibles de los turistas es posible promover comportamientos responsables, minimizar los impactos negativos del turismo, atraer turistas de calidad y mejorar la competitividad de los destinos (Palacios-Florencio et al., 2021). Además, son pocos los estudios que analizan varias dimensiones de la sostenibilidad, puesto que la mayoría centran en una única dimensión (Conti et al., 2023; Galeone & Sebastiani, 2021; Sánchez-Fernández et al., 2019).

Segundo, se enriquece la literatura siguiendo la línea de investigación turística emergente, en la que se aplican técnicas de *big data* al UGC para analizar y extraer información útil sobre temas específicos (sostenibilidad, privacidad o uso de robots de servicios) (Cai et al., 2024). Por ende, se aporta una metodología novedosa que emplea técnicas de *big data* y pruebas

estadísticas para analizar la sostenibilidad percibida por los turistas a partir del UGC. Esto contribuye a aumentar las investigaciones turísticas en este campo, ya que los estudios actuales son escasos (D'Acunto et al., 2020b, 2023, 2024; Väisänen et al., 2023). Tercero, este estudio representa un avance significativo al ser pionero en examinar las discrepancias en la sostenibilidad percibida por el turista entre áreas urbanas y rurales, desde un enfoque de *big data* y UGC. Esto también amplía los estudios de turismo rural y *big data*. Wu & Yang (2023) señalan que las investigaciones anteriores se basan en métodos tradicionales y que la mayoría de los profesionales en turismo rural son agricultores, por lo que carecen de habilidades para el manejo de grandes volúmenes de información y técnicas de *big data*. Esta situación impide maximizar el potencial del sector, desaprovechando la información útil proporcionada por los viajeros en sus reseñas.

Cuarto, esta investigación sobresale por ser la primera en abordar la percepción de la sostenibilidad en el destino turístico de España. Hasta donde se tiene conocimiento, no se ha realizado una investigación comparable en este ámbito. Los resultados señalan que la cultura es un factor clave para el turismo en España, por lo que debe seguir fortaleciéndose y promoviéndose, dado que influye de manera significativa y positiva en la satisfacción de los turistas. Los académicos resaltan que el turismo puede ser impulsado por los recursos culturales y viceversa (Cellini & Cuccia, 2013; Cuccia et al., 2016; Guccio et al., 2017). Es decir, los turistas pueden seleccionar determinado destino ya sea por la diversidad y riqueza cultural, o, por su compromiso e interés cultural propio (aprendizaje, participar en eventos culturales, entre otros).

Con respecto a las implicaciones prácticas, se destaca que los hallazgos son útiles para los gestores de los atractivos turísticos, ya que permiten comprender la importancia de las dimensiones sostenibles para los usuarios. Saber cómo influyen los aspectos sostenibles (económicos, sociales, ambientales y culturales) en la satisfacción puede ayudar a los gestores de la siguiente manera: 1) orientar sus estrategias y potencializar sus ventajas competitivas, dependiendo la naturaleza del atractivo; 2) evaluar si la comunicación realizada sobre los aspectos sostenibles es acorde con el valor percibido por parte de los usuarios finales; 3) mejorar la reputación de la atracción turística, para generar más *eWOM* (*Electronic Word of Mouth*, en sus siglas en inglés, o boca a boca electrónico) positivo, lo que puede derivarse en mayores ingresos; y 4) atraer turistas con comportamientos respetuosos hacia el medioambiente, fomentando la conservación y protección de los recursos.

Finalmente, se puntualiza que esta investigación no está exenta de limitaciones. En primer lugar, los datos recopilados provienen de una sola fuente de información (TripAdvisor), por lo que para próximas investigaciones se podría incluir un mayor número de plataformas de opinión. En segundo lugar, el origen del turista se clasificó en nacional y extranjero, por tanto, la información detallada del país exacto o comunidad autónoma proporcionada por el turista no se utilizó. En futuros estudios, se podría evaluar el impacto de la nacionalidad (o comunidad autónoma) sobre las dimensiones sostenibles y la satisfacción de los atractivos turísticos. En tercer lugar, este estudio no abordó los motivos detrás de la asociación de los turistas entre las dimensiones sostenibles y la satisfacción. Por tanto, resulta interesante llevar a cabo investigaciones con métodos cualitativos y otras técnicas de *big data* que profundicen en los aspectos subyacentes.



## **CAPÍTULO V. CONCLUSIONES DE LA TESIS DOCTORAL**



## CAPÍTULO V. CONCLUSIONES DE LA TESIS DOCTORAL

En este capítulo final se presentan las conclusiones, que sintetizan los principales hallazgos y sus implicaciones tanto teóricas como empíricas. La presente Tesis Doctoral integra dos campos en pleno auge y constante evolución: la sostenibilidad y la ciencia de datos. Ambos ámbitos son esenciales para abordar los desafíos actuales y guían la sociedad futura hacia un desarrollo más inteligente, eficiente y consciente con el medio ambiente. Por un lado, la sostenibilidad representa el estado ideal de equilibrio entre naturaleza y humanidad, garantizando el bienestar de las generaciones presentes y futuras. Por otro lado, la ciencia de datos emerge como aquel campo novedoso e interdisciplinario que analiza grandes cantidades de datos y apoya la toma de decisiones más precisas y estratégicas.

La Tesis Doctoral se centra en el sector turístico, abordando la sostenibilidad desde un enfoque teórico y empírico, a través de la aplicación de técnicas de ciencia de datos. En primer lugar, se aporta una revisión de la literatura detallada y actualizada, dado que un marco teórico conciso y sólido puede reducir la ambigüedad de los temas de estudio (Miller & Twining-Ward, 2005). En segundo lugar, existen pocos estudios que emplean metodologías basadas en la ciencia de datos y el análisis de *big data* en turismo (Egger, 2022; Köseoglu et al., 2020; van der Zee et al., 2020). Esta Tesis Doctoral se presenta como una aportación más en esta línea, ya que el marco metodológico integra diferentes técnicas de ciencia de datos y analiza grandes volúmenes de información. En tercer lugar, se lleva a cabo una investigación empírica que examina las perspectivas sostenibles de las empresas y turistas en España. Esto brinda una comprensión más profunda de los desafíos y oportunidades que surgen en el camino hacia la adopción de un modelo turístico más sostenible. Además, la literatura sobre sostenibilidad y turismo ha evolucionado de un enfoque teórico a uno práctico, incluyendo a distintos grupos de interés (Bramwell et al., 2017; Ruhanen et al., 2015; Sharpley, 2020).

También, es pertinente destacar que la realización y culminación de la presente Tesis Doctoral ha sido posible gracias al desarrollo del proyecto de investigación “TED2021-131314B-I00 SOSTENIBILIDAD CORPORATIVA Y TURISMO INTELIGENTE EN COMUNIDADES RURALES: INFLUENCIA EN EL DESARROLLO ECONÓMICO Y SOCIAL DEL TERRITORIO”, financiado por MICIU/AEI /10.13039/501100011033 y por la Unión Europea NextGenerationEU/ PRTR.

A continuación, se detallan las principales aportaciones teóricas y empíricas de la Tesis Doctoral, que enriquecen la literatura existente y minimizan los *gaps* identificados.

## **1. Conclusiones teóricas**

A lo largo de esta Tesis Doctoral se presentan diversas aportaciones teóricas, derivadas tanto del marco conceptual general que sustenta los dos pilares fundamentales: la sostenibilidad y la ciencia de datos (capítulo I, página 39), como del análisis de la red de conocimiento sobre sostenibilidad y turismo (capítulo II, página 71). Además, cada capítulo incluye apartados específicos que profundizan en aspectos clave, permitiendo una mayor comprensión y reflexión teórica sobre los temas abordados, como la Responsabilidad Social Corporativa (RSC) y la Sostenibilidad Corporativa (SC) (capítulo III, página 99).

### **1.1. Conclusiones relativas al concepto e interpretación de la sostenibilidad turística**

El turismo y la sostenibilidad mantienen una relación compleja y contradictoria, puesto que el progreso económico impulsado por este sector genera considerables impactos negativos en el medio ambiente y la sociedad (R. Buckley, 2012). Al igual que en cualquier otro sector, la sostenibilidad en el turismo es esencial y presenta grandes desafíos para su implementación (Casagrandi & Rinaldi, 2002). Esta investigación permite, en primer lugar, establecer una distinción clara de los términos de turismo sostenible y sostenibilidad turística, dada su ambigüedad conceptual y variedad de interpretaciones (Butler, 1999; Espiner et al., 2017; C. Hunter, 1997; Miller & Torres-Delgado, 2023; Weaver, 2007; Wheeller, 2007; S. Zhang & Chan, 2020). Por tanto, se puede concluir que el turismo sostenible hace referencia al conjunto de procesos, prácticas y estrategias implementadas para alcanzar el objetivo final de la sostenibilidad turística. Estas diferencias se alinean con los conceptos originales que los fundamentan, debido a que la sostenibilidad se logra mediante el proceso de desarrollo sostenible. Una comprensión clara de estos conceptos puede impulsar cambios coherentes y reducir la confusión en su aplicación, tanto en la academia como en la formulación de políticas y prácticas empresariales (Miller & Twining-Ward, 2005).

En consecuencia, la sostenibilidad turística representa aquel enfoque holístico que integra las tres dimensiones (económica, sociocultural y ambiental) en la planificación y desarrollo de cualquier actividad turística a largo plazo. Por el contrario, el turismo sostenible es un tipo de turismo que se enfoca en acciones y actividades que no afectan negativamente el entorno ni la sociedad en general. Este último pretende satisfacer las necesidades de los visitantes, la industria, el entorno y las comunidades locales, contemplando las consecuencias actuales y futuras (World Tourism Organization, 2019a). En este contexto, también sobresale el rol activo de los *stakeholders* clave (turistas, gobierno, empresas y residentes), dado que la sostenibilidad

solo se alcanza cuando estos grupos de interés contribuyen positivamente en las dimensiones social, económica y ambiental (Ng et al., 2017). Estas dimensiones están interrelacionadas, ya que cualquier cambio en una de ellas puede generar efectos significativos en las demás.

De este modo, es posible concluir que ambos términos “turismo sostenible” y “sostenibilidad turística” tienen una aplicabilidad particular según el grupo de interés y su rol dentro del sector. Para los turistas, el concepto de turismo sostenible es más relevante, ya que puede relacionarse con prácticas, actitudes y comportamientos específicos que pueden adoptar para minimizar su impacto ambiental, social y cultural a lo largo de sus viajes. Estas prácticas incluyen, por ejemplo, selección de alojamientos ecológicos, respetar la cultura local, y contribuir a la economía local (Bernini et al., 2020; Chilufya et al., 2019; Dolnicar et al., 2008; Im & Kim, 2023; Kazeminia et al., 2016; Väisänen et al., 2023). Por otro lado, para el gobierno, el concepto de sostenibilidad turística tiene una aplicabilidad más amplia y estratégica. Este grupo de interés puede ser representado por las entidades públicas u organismos locales, encargados de gestionar tanto las políticas y estrategias turísticas, como de administrar y promocionar el turismo. Por tanto, estos actores deben brindar las condiciones adecuadas para garantizar la sostenibilidad de un destino (Ferrer-Roca et al., 2020; Kapera, 2018). Esto puede incluir la regulación de la capacidad de carga de los destinos, la protección de los recursos naturales, el fomento de la economía local o la mejora de la calidad de vida de los residentes. El éxito de estas prácticas de turismo sostenible depende de una adecuada planificación y gestión, equilibrando la protección del medio ambiente y los recursos naturales con la demanda turística (Bramwell & Lane, 2000; Hall & Page, 2006).

Para las empresas turísticas, ambos conceptos son aplicables, ya que para garantizar su existencia, permanencia y desarrollo dentro de la sociedad, deben alcanzar un alto grado de legitimidad y prestigio social (Garriga & Melé, 2004). Por lo tanto, la implementación de estrategias de Responsabilidad Social Corporativa (RSC) y Sostenibilidad Corporativa (SC), permite a las empresas cumplir sus objetivos económicos y responder de manera efectiva a las demandas ambientales y sociales (Font et al., 2012; Font & Lynes, 2018; S. Lee et al., 2022). Las empresas pueden integrar la sostenibilidad de acuerdo con sus objetivos y aspiraciones (Van Marrewijk, 2003). Estas pueden llevar a cabo prácticas de turismo sostenible específicas en sus operaciones y actividades diarias (como gestión de residuos, consumo energético eficiente, etc.), o desarrollar un enfoque holístico en el que integran la sostenibilidad en todas las áreas operativas y estratégicas. Con respecto a los residentes, la sostenibilidad turística también tiene

gran relevancia, puesto que perciben el turismo de forma dual, como una fuente de beneficios (oportunidades de empleo, mejora de la infraestructura) y, al mismo tiempo, como una amenaza o conjunto de desventajas (impactos negativos en la sociedad y medio ambiente). Las perspectivas, actitudes y comportamientos de los residentes de un destino turístico pueden fortalecer la sostenibilidad ambiental y sociocultural (C. Y. Hsu et al., 2020; Rasoolimanesh et al., 2019), siendo reconocidas como parte integral de la planificación y desarrollo del turismo (H. S. C. Choi & Sirakaya, 2005; G. Kim et al., 2020).

## **1.2. Conclusiones relativas a la ciencia de datos en la investigación turística**

Esta investigación pone de manifiesto la relevancia actual de la ciencia de datos en la investigación científica. La gestión de los datos se presenta como un desafío para los académicos y profesionales en la era digital, puesto que los datos crecen de forma exponencial. Según George et al. (2016) el uso de la ciencia de datos en el ámbito académico puede brindar respuestas más precisas a preguntas teóricas existentes y facilita la formulación de nuevas preguntas con un alcance más amplio a partir del análisis de grandes cantidades de datos. Este cambio disruptivo en la investigación ha empezado a transformar múltiples áreas y campos de conocimiento. En consecuencia, se han desarrollado nuevas metodologías y enfoques multidisciplinares que permiten aprovechar la capacidad computacional para extraer información útil de los datos y generar resultados más robustos y exactos.

Dentro de la investigación turística se reconoce el gran potencial de la ciencia de datos como marco metodológico (Egger, 2022), puesto que permite superar las limitaciones de tamaño de muestra (Yang et al., 2015) y reducir los sesgos de muestreo (X. Li et al., 2017; Xiang, Schwartz, Gerdes, et al., 2015). Sin embargo, autores como Mariani & Baggio (2022) y Egger (2022) sugieren que el uso de técnicas de ciencia de datos y el análisis de *big data* en los estudios turísticos aún no ha alcanzado su máximo desarrollo y potencial. Esto en parte puede derivarse de una resistencia de los académicos de esta área a emplear enfoques avanzados y algoritmos computacionales, ya sea por falta de conocimiento o de experiencia (Mariani et al., 2018). Igualmente, la naturaleza caótica de los datos turísticos puede generar gran confusión e incertidumbre (Egger, 2022). La presente Tesis Doctoral contribuye en esta línea, aportando diferentes marcos metodológicos novedosos y basados en la ciencia de datos para analizar la perspectiva teórica, empresarial y de turistas. El enfoque multidisciplinar del estudio permite enriquecer la literatura existente y pone en relieve la importancia de la sostenibilidad y la ciencia de datos en turismo.

De este modo, se aprovecha la capacidad computacional para analizar grandes cantidades de datos, aprovechando la información accesible y disponible en la red. Los estudios empíricos centrados en las empresas y turistas incrementan los estudios relacionados con la gestión de grandes cantidades de datos. Dado que se analizan los informes corporativos de las empresas y se enriquece una línea de investigación turística emergente, en la que se aplican técnicas de *big data* al contenido generado por los usuarios (*User Generated Content*, UGC por sus siglas en inglés) para analizar y extraer información útil sobre temas específicos, como la sostenibilidad, privacidad o uso de robots de servicios (Cai et al., 2024). La aportación de las metodologías de estos estudios puede replicarse en futuros estudios en el campo turístico, disminuyendo el *gap* entre la implementación de novedosos enfoques y métodos avanzados en este campo de investigación.

### **1.3. Conclusiones relativas al análisis de la red de conocimiento sobre sostenibilidad y turismo**

Como resultado del estudio teórico, se proporciona una visión general y actualizada de la literatura sobre sostenibilidad y turismo. El enfoque metodológico utilizado permite analizar 9.449 documentos científicos donde se combina el análisis bibliométrico junto con la revisión computacional de la literatura mediante el modelado de temas. En primer lugar, se confirma que la teoría de los *stakeholders* de Freeman (1984) ha tenido una gran influencia en la literatura turística sobre sostenibilidad. La aplicación de esta teoría al ámbito del turismo pone de manifiesto que todas las partes interesadas o afectadas por el desarrollo de la actividad turística deberían tener la posibilidad de influir en su gestión (Sautter & Leisen, 1999; UNEP & UNWTO, 2005). Los resultados de esta investigación demuestran que los principales grupos de interés en turismo (gobiernos, residentes, turistas y empresas) identificados por McIntosh et al. (1995) y Salee et al. (2022), han sido ampliamente estudiados en la literatura sobre sostenibilidad en turismo.

En efecto, se identifican cuatro áreas de investigación relacionadas con estos stakeholders. El gobierno se puede asociar con el área de *Desarrollo y gestión de destinos turísticos*, la más predominante en la literatura, reflejando la importancia de la planificación y gestión sostenible del sector. De acuerdo con Vieira et al. (2016), el liderazgo político del gobierno es primordial para implementar políticas y estrategias efectivas de turismo sostenible. Por otro lado, el área de *Turismo y comunidad local* se puede relacionar con los residentes, representado por dos temas: “Percepción de residentes y apoyo al desarrollo turístico” y “Desarrollo del turismo comunitario y rural”. Este grupo de interés ejerce una influencia significativa en la planificación

y el desarrollo del turismo, dependiendo de cómo perciba los impactos derivados de la actividad turística (Lundberg, 2015). Según Lee (2013), y con base en la teoría del intercambio social (Ap, 1992), las percepciones tanto negativas como positivas de los residentes influyen en la aceptación o rechazo del turismo. De igual modo, se destaca la importancia del turismo para la comunidad local y rural, actuando como una fuente de progreso económico y mejora en la calidad de vida.

Respecto a los turistas, sobresale el área de *Comportamiento del turista*, la cual ha sido considerada como un aspecto esencial para fomentar un desarrollo turístico sostenible (Dolnicar et al., 2019; S. Kim & Filimonau, 2017; Lee, 2011). Esta área se interrelaciona con dos temas clave: Actitudes del turista y comportamiento ambiental. En efecto, el estudio de las actitudes de los turistas en la literatura es primordial, dado que estas pueden fomentar comportamientos específicos (por ejemplo, ecológicos o proambientales) e influir en su satisfacción (Leonidou et al., 2015). Chiu et al. (2014) señalan que proporcionar experiencias satisfactorias turísticas puede mejorar el conocimiento sobre el medio ambiente e impulsar comportamientos más responsables y sostenibles en los turistas. Por último, se ha identificado el área de *Prácticas sostenibles en empresas*. La literatura ha demostrado que la implementación de prácticas responsables y sostenibles beneficia a las empresas (Berezan et al., 2013; González-Rodríguez et al., 2019; Martínez & Rodríguez del Bosque, 2014; Wut et al., 2023). Algunos de los beneficios incluyen tanto la mejora del desempeño económico, la satisfacción de los clientes, la reputación corporativa y la retención de talento, así como la reducción de costes y riesgos medioambientales o laborales. También se debe puntualizar que la temática asociada a esta área corresponde a las acciones verdes de la industria hotelera. Por tanto, esto podría sugerir que la literatura actual se ha enfocado en su mayoría en dicho subsector, prestando menos atención a otros subsectores del turismo, como el transporte (Hopkins, 2020), o la restauración (Rhou et al., 2016).

Otra conclusión relevante del análisis de la red de conocimiento sobre sostenibilidad y turismo es la identificación de tendencias emergentes y en declive. Esto permite aportar una comprensión detallada de las dinámicas de la literatura existente, que puede servir de guía para investigadores, profesionales y formuladores de políticas. Los resultados demuestran que los temas que han ganado mayor atención en la investigación reciente se enfocan en el comportamiento de los turistas; los desafíos sostenibles relacionados con la gestión del agua, la masificación y la movilidad turística; el impacto de la pandemia de covid-19; el empoderamiento

de género y emprendimiento; y distintos tipos de turismo especializado, como el gastronómico, enológico y religioso. Esto confirma la alta sensibilidad del sector hacia la sostenibilidad y crisis, reflejando de manera implícita la necesidad de gestionar adecuadamente las dimensiones ambiental, social y cultural. De igual modo, se destaca un cambio en las preferencias de los viajeros, quienes buscan nuevas modalidades de turismo que les brinden experiencias más enriquecedoras y significativas.

Por otro lado, las tendencias en declive pueden indicar que ciertas temáticas han alcanzado un grado de madurez o un menor interés de investigación. Entre estas se incluyen el ecoturismo, políticas del destino, indicadores de sostenibilidad y gobernanza del destino. Esto puede sugerir que han sido ampliamente investigadas en la literatura. Además, es importante mencionar que, aunque no se pudo comprobar estadísticamente la tendencia positiva de las temáticas de cambio climático y de turismo patrimonial, cultural y arqueológico, estas podrían considerarse como temáticas emergentes. Al igual que las mencionadas anteriormente, estas temáticas están vinculadas tanto a la creciente problemática del cambio climático como al interés de los turistas en diferentes tipos de turismo especializado.

Finalmente, este estudio teórico confirma que la integración de métodos avanzados de ciencia de datos junto con el análisis bibliométrico permite una exploración más exhaustiva y detallada de la literatura. Esta combinación metodológica facilita la identificación de temáticas y áreas de investigación, así como sus tendencias, proporcionando una visión más completa y estructurada de las dinámicas del campo de estudio. De esta forma, los investigadores pueden aprovechar las oportunidades que ofrecen la ciencia de datos y los métodos avanzados de análisis para sintetizar grandes corpus de literatura y extraer conocimientos clave.

#### **1.4. Conclusiones relativas a la evolución teórica de la RSC a la SC en el turismo**

Esta Tesis Doctoral ha permitido observar un cambio de paradigma teórico de Responsabilidad Social Corporativa (RCS) a Sostenibilidad Corporativa (SC) en la literatura turística. No obstante, esta transición se da de forma gradual. Aunque la teoría de RSC ha alcanzado una sólida madurez y cuenta con una extensa base de investigación en turismo, su popularidad muestra signos de declive, indicando un interés decreciente en la investigación reciente. Por el contrario, la SC muestra una tendencia alcista, captando mayor interés académico. Estos resultados indican que las dinámicas en el campo científico del turismo están evolucionando hacia la SC, lo cual es coherente con la relación jerárquica entre RSC y SC

establecida por Van Marrewijk (2003), en la que la RSC actúa como un paso intermedio para alcanzar el objetivo final de la SC. La confirmación de esta transición a nivel teórico resulta fundamental para orientar futuras investigaciones, ya que permite establecer un marco más preciso. La distinción entre RSC y SC es esencial para el análisis e implementación adecuada de estas teorías (Bansal & Song, 2016; Montiel, 2008; Van Marrewijk, 2003). Ambas, contribuyen a la gestión de los impactos económicos, sociales y ambientales de las empresas. Sin embargo, la SC abarca un enfoque más integral con una perspectiva a mediano y largo plazo, promoviendo un crecimiento equilibrado y comprometido con los distintos *stakeholders*.

Por otra parte, los resultados de este estudio identifican las principales áreas de investigación sobre RSC y SC en turismo, reflejando aspectos clave del turismo. Se destaca el área de *Clientes y marketing*, englobando temas como el comportamiento del consumidor, satisfacción y lealtad. Esto sugiere y confirma la importancia de la satisfacción y lealtad del cliente en la industria turística, así como su relación positiva con las prácticas sostenibles llevadas a cabo por las empresas (Berezan et al., 2013; Martínez & Rodríguez del Bosque, 2013). Además, se identifican dos áreas de investigación relacionadas con los subsectores de hotelería y transporte. Estos subsectores, al generar altos niveles de emisiones y residuos, y consumir grandes cantidades de energía y agua (Bohdanowicz, 2005; Cowper-Smith & de Grosbois, 2011; Verma & Chandra, 2018), han atraído el interés de los investigadores en el estudio de prácticas responsables y sostenibles. Su análisis es fundamental para abordar los desafíos ambientales en la industria turística y fomentar estrategias que mitiguen los efectos negativos en el planeta y los recursos naturales.

## **2. Conclusiones empíricas**

El análisis de las percepciones sostenibles de las empresas y los turistas realizado en esta Tesis Doctoral permite resaltar la importancia de estos dos grupos de interés en el desarrollo del turismo. Ambos grupos de interés ejercen una influencia directa en la oferta y la demanda turística, lo que convierte sus percepciones en una fuente valiosa para identificar oportunidades y desafíos en el contexto de la sostenibilidad. Al comprender sus percepciones sostenibles, es posible obtener una visión más detallada de las áreas o dimensiones en las que es necesario mejorar las prácticas sostenibles, favoreciendo así un desarrollo turístico más equilibrado. Por un lado, los comportamientos irrespetuosos de los turistas afectan negativamente la sostenibilidad de los destinos y amenazan su viabilidad a largo plazo (Im & Kim, 2023). Sin embargo, por otro lado, las empresas turísticas pueden promover en sus clientes

comportamientos más sostenibles y respetuosos con el medioambiente (Font et al., 2016). A continuación, se detallan las implicaciones resultantes tanto para la gestión de las empresas turísticas, a partir del análisis de las memorias corporativas, como para los gestores de atractivos turísticos, derivadas del estudio de los comentarios online de los turistas.

### **2.1. Implicaciones para la gestión empresarial turística**

Este trabajo aporta una comprensión profunda de cómo las empresas turísticas han implementado la Responsabilidad Social Corporativa (RSC) y han avanzado hacia un enfoque más integrado de Sostenibilidad Corporativa (SC). En primer lugar, se confirma que la industria turística española ha adoptado distintas prácticas sostenibles y responsables a lo largo del tiempo. Los informes corporativos de las empresas turísticas, analizados entre 2010 y 2015, revelan su compromiso por el bienestar social, la protección del medio ambiente y el progreso económico, incluso antes de la aprobación de la Agenda 2030 de los Objetivos de Desarrollo Sostenible (ODS). Algunas de las prácticas que sobresalen son la gestión energética y de residuos, las acciones de filantropía corporativa y las políticas de igualdad de género, diversidad y discriminación. Este comportamiento empresarial puede haber estado motivado por el reconocimiento de su significativo impacto ambiental, el crecimiento del consumo verde y la necesidad de optimizar costes y adaptar los servicios o productos a las nuevas demandas sociales y sostenibles (Rodríguez-Antón et al., 2012). Los resultados de los informes entre 2016 y 2021 incorporan nuevas prácticas relacionadas con la huella de carbono, el reciclaje y la gestión del riesgo. Igualmente, los temas de finanzas, contabilidad y cuentas anuales se empiezan a asociar con la SC. Esto sugiere la necesidad de incorporar procesos o herramientas que permitan a las empresas turísticas medir el impacto real que generan en sus *stakeholders*, como el valor social integrado o la contabilidad social. Este enfoque permite a las empresas evaluar e informar su desempeño social y ambiental en relación con sus objetivos económicos (Gray, 2002; Retolaza & San-Jose, 2021).

Es importante destacar que la implementación de prácticas sostenibles y responsables en las empresas del sector es fundamental para cumplir con las expectativas del mercado y de los *stakeholders*. Sin embargo, este compromiso debe ir acompañado de una adecuada comunicación y divulgación, que permita fortalecer las relaciones con sus grupos de interés y refuerce la legitimidad social (Crane & Glozer, 2016; Ning et al., 2021; Schuler & Cording, 2006). Hackston & Milne (1996) puntualiza que los informes de sostenibilidad reflejan la forma cómo la empresa interactúa con su entorno. Asimismo, Coebergh (2011), basado en la teoría de la

legitimidad, señala que los informes de RSC son una herramienta de legitimación y gestión de la reputación corporativa, puesto que evidencian el cumplimiento de requisitos básicos de responsabilidad corporativa para operar en la sociedad y favorecen la construcción de la marca y la diferenciación del servicio o producto. En consecuencia, la elaboración de este tipo de informes permite a las empresas monitorear y reportar prácticas responsables y sostenibles, identificando áreas de mejora para ajustar sus estrategias según los resultados y las expectativas del mercado. De este modo, las empresas pueden destacarse frente a competidores que no tienen el mismo nivel de compromiso con la sostenibilidad, reforzando el posicionamiento de la empresa en el mercado, entre consumidores e inversionistas.

En segundo lugar, se demuestra que existe una evolución gradual desde la RSC hacia la SC en las estrategias de las empresas turísticas de España. La teoría de la RSC aún tiene un fuerte impacto en la gestión del sector de turismo, evidenciándose una transición hacia un enfoque más centrado en la SC. Esto implica una adopción de criterios ambientales, sociales y de gobernanza (*Environmental, Social and Governance*, en sus siglas en inglés), que abordan la sostenibilidad desde una perspectiva más holística y estratégica. La confirmación del cambio de paradigma de RSC a SC desde el ámbito práctico facilita una mejor comprensión de ambas temáticas y promueve un enfoque estratégico del sector. Este cambio de enfoque exige que las empresas turísticas adapten sus estrategias y operaciones diarias, integrando la sostenibilidad como un elemento central en su toma de decisiones. Asimismo, desafía a las empresas tanto a establecer objetivos claros de sostenibilidad y alineados con los ODS, como a promover una cultura organizacional coherente con estos valores. A pesar de que la muestra de estudio se centra exclusivamente en grandes empresas, los resultados pueden servir como un incentivo para que medianas, pequeñas y microempresas adopten prácticas sostenibles, las comuniquen adecuadamente y mejoren sus estrategias a largo plazo.

En tercer lugar, se cuantifica la importancia que los subsectores de transporte, hotelería y restauración otorgan a los criterios ESG, detallándolos en nueve categorías: cambio climático, capital natural, contaminación y residuos, capital humano, responsabilidad del producto, relaciones con la comunidad, gobernanza corporativa, ética y valores empresariales y no-ESG. Los resultados de la investigación evidencian diferencias entre subsectores en la forma en que abordan los ESG. Por tanto, se resalta que las empresas de transporte españolas obtienen las puntuaciones más altas en criterios asociados con el cambio climático. Esto indica un fuerte enfoque en la gestión de su impacto ambiental y un compromiso hacia la reducción de emisiones

y la adopción de prácticas sostenibles. Por otro lado, el subsector de hotelería ha puesto énfasis en la gobernanza corporativa y en fortalecer sus relaciones con la comunidad, reflejando un enfoque orientado a la transparencia y la responsabilidad social. No obstante, según los resultados obtenidos de los informes analizados, el subsector de la restauración ha mostrado una integración más gradual de los criterios ESG. De esta forma, se destaca la necesidad de realizar un esfuerzo por parte de este subsector para alinearse con las prácticas sostenibles. La comparación entre los subsectores del turismo en España permite evidenciar los desafíos y oportunidades a los que se enfrentan, así como enriquecer los estudios longitudinales y sectoriales sobre RSC y SC, disminuyendo el *gap* actual de la literatura (Back, 2024; Koh, 2024).

## **2.2. Implicaciones para la gestión de atractivos turísticos**

La segunda parte del estudio empírico de esta Tesis Doctoral se centra en comprender cómo los turistas perciben la sostenibilidad en atractivos turísticos en España. A través del análisis de las reseñas online, se pueden captar las percepciones y valoraciones de los turistas sobre las diferentes dimensiones de sostenibilidad (económica, social, ambiental y cultural). Los hallazgos de esta investigación brindan información útil para que los gestores de atractivos turísticos alineen sus estrategias con las expectativas y valores de los visitantes. De este modo, se resaltan varias implicaciones a nivel práctico. En primer lugar, se pueden mejorar y potencializar las ventajas competitivas de los atractivos turísticos de naturaleza y cultura de España, ya que estas dimensiones suelen incrementar las calificaciones y niveles de satisfacción de los turistas. Según Correia et al. (2013), la cultura es uno de los factores más influyentes para los turistas al momento de elegir un destino. Las motivaciones para realizar turismo cultural se relacionan con el aprendizaje, la identidad y la experiencia transformadora (Richards, 2018). Además, se ha demostrado que el turismo de naturaleza puede contribuir tanto a la salud mental como al bienestar general de las personas (Akhoundogli & Buckley, 2023; R. Buckley, 2020, 2023; R. Buckley et al., 2021; R. C. Buckley, 2019; Cooper & Buckley, 2022). Esto no excluye a los atractivos turísticos que no están estrictamente vinculados a aspectos culturales o naturales, ya que estos también pueden incorporar y potenciar elementos culturales y ambientales en su oferta, promoviendo así una sostenibilidad más equilibrada.

En segundo lugar, se evidencia una necesidad por realizar esfuerzos para mejorar la percepción de los turistas sobre los aspectos sociales y económicos de los atractivos turísticos de España. Los resultados del estudio demuestran que es más probable que los turistas asignen calificaciones más bajas a los atractivos, cuando sus reseñas incluyen o reflejan aspectos

económicos y sociales. Esto puede implicar fortalecer o desarrollar iniciativas que destaquen el impacto positivo de los atractivos turísticos en las comunidades locales y la economía, a través de estrategias de comunicación o marketing apropiadas. Es importante señalar que, la dimensión económica de la sostenibilidad se manifiesta en las reseñas online como el valor por el dinero, dado que los turistas evalúan el precio frente a la calidad, influyendo en su satisfacción y en la rentabilidad de las empresas y destinos a largo plazo (Väisänen et al., 2023; Ye et al., 2014). A su vez, la dimensión social incluye elementos como la seguridad de clientes e infraestructuras, apoyo a la empleabilidad local, opciones saludables de alimentación y actividades, educación de los clientes sobre la sostenibilidad, entre otras (Dewhurst & Thomas, 2003; Garay & Font, 2012; Jones et al., 2014; Väisänen et al., 2023). Al identificar las dimensiones sostenibles que los turistas valoran más, los gestores de atractivos pueden aumentar las fortalezas del atractivo turístico en función de su naturaleza y diseñar una propuesta de valor diferenciada.

En tercer lugar, y en línea con lo anterior, los gestores de atractivos turísticos pueden realizar una autoevaluación de la comunicación e imagen proyectada sobre la sostenibilidad. Esto puede facilitar una mayor comprensión de las diferencias entre la sostenibilidad proyectada y la percibida por los turistas, permitiéndoles ajustar sus estrategias de comunicación o marketing para reflejar de forma auténtica y precisa sus compromisos en este ámbito. La comunicación de sostenibilidad busca informar a los consumidores sobre productos o servicios sostenibles, mostrando cómo satisfacen sus necesidades y cumplen con criterios de sostenibilidad, para así fomentar compras responsables (Tölkes, 2018). En este sentido, la comunicación de sostenibilidad hace parte de la estrategia de marketing de sostenibilidad de la empresa u organización (Belz & Ken Peattie, 2012), añadiendo nuevos retos a la gestión.

En cuarto lugar, el análisis de los comentarios online de atractivos turísticos en entornos rurales y urbanos resalta diferencias clave que pueden ser útiles para los gestores, evidenciando tanto fortalezas como debilidades. En ambos contextos, las dimensiones ambiental y cultural se asocian con calificaciones altas, aunque la dimensión ambiental muestra un impacto mayor en atractivos rurales y la dimensión cultural en atractivos urbanos. Por el contrario, las dimensiones económica y social se relacionan con calificaciones más bajas. En este caso, la dimensión económica ejerce una influencia más fuerte en los atractivos urbanos, mientras que la dimensión social tiene un mayor impacto en los atractivos rurales. En consecuencia, los gestores pueden continuar optimizando sus estrategias de comunicación y marketing al destacar los

aspectos ambientales y culturales que se alinean con las motivaciones específicas de sus visitantes, según la ubicación del atractivo. En áreas rurales, se podría enfatizar en el contacto con la naturaleza, la desconexión digital o la tranquilidad. De acuerdo con Yin et al. (2022), los turistas valoran positivamente la inmersión en la naturaleza de áreas rurales, dado que les permite escapar de la vida urbana y experimentar una conexión profunda con la naturaleza. Con respecto a las áreas urbanas, se podrían destacar las motivaciones para realizar turismo cultural, como el aprendizaje, la identidad y la experiencia transformadora (Richards, 2018). Por otro lado, es recomendable que los gestores de los atractivos dediquen esfuerzos en mejorar la percepción que los turistas tienen sobre los aspectos sociales y económicos, dado que ambos suelen ser valorados negativamente y generar insatisfacción, especialmente el económico en áreas urbanas y el social en áreas rurales. Para ello, es necesario mejorar la percepción de la relación calidad-precio del turista y poner en valor los aspectos relacionados con la comunidad local, la infraestructura, la accesibilidad, la seguridad, entre otros. De este modo, los gestores podrían desarrollar campañas de comunicación novedosas que abarquen estos aspectos, personalizar productos o servicios turísticos, realizar colaboraciones con los residentes, mejorar la infraestructura y la seguridad, etc.

En quinto lugar, al mejorar las percepciones de los turistas sobre aspectos de sostenibilidad, la reputación *online* de las atracciones turísticas puede verse beneficiada. Las reseñas de otros viajeros se han convertido en una fuente distinguida para la toma de decisiones turísticas, ya que ofrecen opiniones confiables (Ayeh, 2015; Ayeh et al., 2013; Sparks & Browning, 2011; Z. Zhang et al., 2016). En este contexto, el contenido generado por el usuario (*User Generated Content*, UGC, en sus siglas en inglés) brinda una visión directa del grado de conciencia y percepción de los consumidores sobre las prácticas sostenibles de las empresas turísticas (D'Acunto et al., 2020a). Por lo tanto, una reputación fortalecida en sostenibilidad impulsa tanto la satisfacción del cliente como el eWOM (*Electronic Word of Mouth*, en sus siglas en inglés, o boca a boca electrónico) positivo, lo cual es un factor clave para atraer a nuevos visitantes.

### 3. Limitaciones de la investigación

A continuación, se presentan las principales limitaciones que se derivan del desarrollo de esta Tesis Doctoral tanto a nivel teórico como empírico.

#### 3.1. Limitaciones asociadas a los estudios teóricos

Una primera limitación respecto a los análisis de revisión de literatura turística sobre sostenibilidad, responsabilidad social corporativa y sostenibilidad corporativa es el uso de una única base de datos científica. Aunque *Scopus* es una base de datos reconocida por la alta calidad de sus fuentes académicas, limitar el estudio exclusivamente a esta fuente podría restringir el alcance y la diversidad de la información descargada. Es posible que, algunas investigaciones científicas relevantes, disponibles en otras bases de datos como *Web of Science*, no hayan sido consideradas. Una segunda limitación se relaciona con la naturaleza y el objetivo de los análisis teóricos, dado que estos están orientados principalmente a la exploración y mapeo de la literatura existente. Por consiguiente, no se realiza un análisis exhaustivo ni detallado de los temas, áreas y tendencias de investigación identificados. Finalmente, desde una perspectiva metodológica, se emplea únicamente la técnica de *machine learning* de modelado de temas. A pesar de que este enfoque permite detectar y agrupar temáticas clave de manera exhaustiva, el uso exclusivo de esta técnica puede limitar el análisis al no integrar métodos adicionales que podrían enriquecer la interpretación y validación de los resultados obtenidos.

#### 3.2. Limitaciones asociadas a los estudios empíricos

Como resultado de los estudios empíricos se observan varias limitaciones. En primer lugar, se lleva a cabo en el sector turístico de España, tanto a empresas (hoteles, restaurantes, transporte) como a reseñas online de atractivos turísticos. Este hecho podría limitar la generalización de algunos hallazgos, ya que estos dependen de las particularidades y naturaleza de las empresas y los atractivos turísticos españoles. En segundo lugar, la muestra de empresas se compuso exclusivamente por dos grupos (grandes y medianas), lo cual impidió analizar la percepción de la responsabilidad social corporativa y la sostenibilidad corporativa en otros tipos de organizaciones, como las pequeñas y microempresas. Esta limitación también puede influir en la generalización de los resultados, ya que las prácticas y percepciones en sostenibilidad podrían variar considerablemente en empresas de menor tamaño, debido a diferencias en recursos, estructura y enfoque estratégico.

En tercer lugar, y en relación con el análisis de las percepciones de turistas a partir de comentarios en línea, los datos fueron obtenidos de una sola plataforma, TripAdvisor. Aunque esta plataforma de viajes es una fuente de información distinguida y ampliamente utilizada en el sector turístico, limitar el estudio a una única plataforma puede introducir sesgos o limitar la diversidad de opiniones analizadas. En cuarto lugar, la investigación no aborda los motivos detrás de la asociación de los turistas entre las dimensiones sostenibles (económica, social, ambiental y cultural) y la satisfacción. Esto puede implicar la ausencia de métodos cualitativos y otras técnicas avanzadas de ciencia de datos que permitan profundizar en los aspectos subyacentes de los turistas. Finalmente, otra limitación de esta investigación radica en la falta de integración y vinculación entre los datos propios de la investigación, es decir, los informes corporativos y el UGC, con los indicadores empresariales reflejados en las cuentas anuales y los indicadores ESG externos. La falta de integración de estas fuentes puede limitar la capacidad de obtener una visión completa y detallada sobre la relación entre las prácticas de sostenibilidad, la percepción de los turistas y el desempeño empresarial.

#### **4. Futuras líneas de investigación**

A partir de las conclusiones y limitaciones de esta Tesis Doctoral, es posible establecer una serie de cuestiones o líneas de investigación futuras relacionadas con el análisis de la sostenibilidad en el sector turístico mediante la aplicación de técnicas de ciencia de datos. En primer lugar, futuros estudios podrían ampliar el alcance y profundidad de los análisis de revisión de la literatura. Se podrían emplear las mismas metodologías e incluir múltiples bases de datos científicas para mejorar la representatividad de los hallazgos y validar las tendencias observadas. Asimismo, resultaría de gran interés desarrollar estudios que profundicen y detallen las tendencias emergentes identificadas, analizando sus características específicas y posibles impactos a largo plazo. Esto permitiría comprender mejor los factores que impulsan estas tendencias y su relevancia en el contexto de la sostenibilidad en el sector turístico.

En segundo lugar, convendría explorar las percepciones sobre sostenibilidad en pequeñas y microempresas españolas, adaptando las metodologías y técnicas de análisis a sus formas específicas de divulgación de información sostenible, como redes sociales, plataformas de viajes, sitios web y otros canales digitales. Esto permitiría capturar de manera más precisa el enfoque de sostenibilidad en este tipo de empresas que a diferencia de las grandes y medianas empresas, suelen comunicar sus prácticas sostenibles de manera más directa y accesible. En tercer lugar, podría ser enriquecedor analizar los comentarios online de turistas en otras plataformas

(distintas de TripAdvisor) utilizando el mismo enfoque metodológico, con el fin de verificar si existen discrepancias en las percepciones y valoraciones de los atractivos turísticos. Un análisis comparativo entre plataformas podría revelar si existen sesgos particulares en cada una, enriqueciendo el conocimiento sobre cómo los turistas perciben la sostenibilidad en atractivos a partir de reseñas online. En cuarto lugar, resultaría de gran interés desarrollar un marco metodológico que permita la integración y vinculación de distintos tipos de datos internos y externos de las empresas turísticas, como pueden ser las memorias de sostenibilidad y el UGC, con los indicadores empresariales y financieros, así como con los indicadores ESG (*Environmental, Social and Governance*, en sus siglas en inglés) externos. Este enfoque permitiría un análisis más unificado y exhaustivo de la relación entre las prácticas de sostenibilidad de las empresas, la percepción de los turistas y el desempeño empresarial.

En quinto lugar, resultaría beneficioso ampliar esta investigación al resto de agentes que integran el enfoque holístico. Tal como se ha mencionado a lo largo de la Tesis Doctoral y de acuerdo con los resultados obtenidos, existen cuatro grupos de interés clave en la sostenibilidad del sector turístico: empresas, turistas, gobiernos y residentes. Si bien esta investigación se centra en empresas y turistas, la inclusión de los gobiernos y los residentes permitiría obtener una perspectiva más completa del ecosistema turístico sostenible. El estudio de estos dos grupos de interés podría llevarse a cabo bajo el enfoque multidisciplinario e innovador propuesto en esta investigación, incorporando la ciencia de datos como marco metodológico. Esto permitiría analizar grandes volúmenes de datos de manera sistemática, identificando patrones y tendencias relevantes en las percepciones sostenibles de gobiernos y residentes. Para los gobiernos, por ejemplo, se podrían emplear los planes de sostenibilidad turística por Comunidades Autónomas y la información tanto de redes sociales como de las páginas web oficiales. En el caso de los residentes, resultaría útil la información de redes sociales y noticias, puesto que podrían capturar la perspectiva de la comunidad local sobre el impacto del turismo en su entorno, incluyendo aspectos como el bienestar social y ambiental.

Por último, otra línea de investigación interesante sería desarrollar un análisis comparativo entre distintos tipos de destinos turísticos, considerando las percepciones y el impacto en los cuatro stakeholders clave: empresas, turistas, residentes y gobiernos. Esto podría incluir comparaciones entre Comunidades Autónomas, entre zonas costeras y de interior, o entre destinos rurales y urbanos. Un enfoque de este tipo permitiría identificar variaciones y similitudes en la implementación y percepción de la sostenibilidad turística en función de las

características y desafíos específicos de cada destino. Asimismo, este análisis podría ampliarse a nivel internacional, permitiendo realizar comparaciones entre diferentes destinos turísticos. Siguiendo el enfoque empleado en esta Tesis Doctoral, sería posible aprovechar la gran cantidad de información disponible en internet, incluyendo memorias de sostenibilidad y el UGC, para llevar a cabo análisis robustos. Este tipo de comparación internacional podría ofrecer perspectivas valiosas sobre cómo las prácticas de sostenibilidad y la percepción de los turistas varían en distintos contextos culturales y geográficos, contribuyendo a identificar mejores prácticas y adaptar estrategias de sostenibilidad específicas a cada destino turístico.



## **REFERENCIAS**



## REFERENCIAS

- Agarwal, R., Mehrotra, A., Mishra, A., Rana, N. P., Nunkoo, R., & Cho, M. (2024). Four decades of sustainable tourism research: Trends and future research directions. *International Journal of Tourism Research*, 26(2), e2643. <https://doi.org/10.1002/JTR.2643>
- Aguinis, H., & Glavas, A. (2012). What We Know and Don't Know About Corporate Social Responsibility: A Review and Research Agenda. *Journal of Management*, 38(4), 932–968. <https://doi.org/10.1177/0149206311436079>
- Agyeiwaah, E., McKercher, B., & Suntikul, W. (2017). Identifying core indicators of sustainable tourism: A path forward? *Tourism Management Perspectives*, 24, 26–33. <https://doi.org/10.1016/J.TMP.2017.07.005>
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Akhoundogli, M., & Buckley, R. (2023). Outdoor tourism to escape social surveillance: health gains but sustainability costs. *Journal of Ecotourism*, 22(1), 4–24. <https://doi.org/10.1080/14724049.2021.1934688>
- Akmese, H., Cetin, H., & Akmese, K. (2016). Corporate Social Responsibility Reporting: A Comparative Analysis of Tourism and Finance Sectors of G8 Countries. *Procedia Economics and Finance*, 39, 737–745. [https://doi.org/10.1016/S2212-5671\(16\)30273-8](https://doi.org/10.1016/S2212-5671(16)30273-8)
- Alhoshan, W., Ferrari, A., & Zhao, L. (2023). Zero-shot learning for requirements classification: An exploratory study. *Information and Software Technology*, 159, 107202. <https://doi.org/10.1016/J.INFSOF.2023.107202>
- Alonso-Arévalo, J. (2019). La gestión de datos de investigación en el horizonte de las bibliotecas universitarias y de investigación. *Cuadernos de Documentación Multimedia*, 30, 75–88. <https://doi.org/10.5209/CDMU.62806>
- Alonso-Muñoz, S., Torrejón-Ramos, M., Medina-Salgado, M. S., & González-Sánchez, R. (2023). Sustainability as a building block for tourism – future research: Tourism Agenda 2030. *Tourism Review*, 78(2), 461–474. <https://doi.org/10.1108/TR-12-2021-0568>

- Andreu Pinillos, A., & Fernández Fernández, J. L. (2011). De la RSC a la sostenibilidad corporativa: una evolución necesaria para la creación de valor. *Harvard Deusto Business Review*, 207, 5–21. <https://www.researchgate.net/publication/297757785>
- Andreu Pinillos, A., Fernández-Fernández, J. L., & Fernández Mateo, J. (2018). A critical review of the Corporate Governance dimension in the Sustainability Indexes questionnaires. *Revista de Comunicacion*, 17(2), 9–40. <https://doi.org/10.26441/RC17.2-2018-A1>
- António, N., & Rita, P. (2023). Twenty-two years of International Journal of Hospitality Management: A bibliometric analysis 2000–2021. *International Journal of Hospitality Management*, 114. <https://doi.org/10.1016/j.ijhm.2023.103578>
- Antons, D., Breidbach, C. F., Joshi, A. M., & Salge, T. O. (2021). Computational Literature Reviews: Method, Algorithms, and Roadmap. *Organizational Research Methods*, 26(1), 107–138. <https://doi.org/10.1177/1094428121991230>
- Ap, J. (1992). Residents' perceptions on tourism impacts. *Annals of Tourism Research*, 19(4), 665–690. [https://doi.org/10.1016/0160-7383\(92\)90060-3](https://doi.org/10.1016/0160-7383(92)90060-3)
- Aragon-Correa, J. A., Martin-Tapia, I., & de la Torre-Ruiz, J. (2015). Sustainability issues and hospitality and tourism firms' strategies: Analytical review and future directions. *International Journal of Contemporary Hospitality Management*, 27(3), 498–522. <https://doi.org/https://doi.org/10.1108/IJCHM-11-2014-0564>
- Arami, M., Balina, S., Chang, V., Kim, J., Kim, H.-S., & Choi, S.-Y. (2023). Forecasting the S&P 500 Index Using Mathematical-Based Sentiment Analysis and Deep Learning Models: A FinBERT Transformer Model and LSTM. *Axioms*, 12(9), 835. <https://doi.org/10.3390/AXIOMS12090835>
- Ashworth, G., & Page, S. J. (2011). Urban tourism research: Recent progress and current paradoxes. *Tourism Management*, 32(1), 1–15. <https://doi.org/10.1016/J.TOURMAN.2010.02.002>
- Asif, M., Searcy, C., Santos, P. dos, & Kensah, D. (2013). A Review of Dutch Corporate Sustainable Development Reports. *Corporate Social Responsibility and Environmental Management*, 20(6), 321–339. <https://doi.org/10.1002/csr.1284>

- Ateljevic, I. (2020). Transforming the (tourism) world for good and (re)generating the potential 'new normal.' *Tourism Geographies*, 22(3), 467–475. <https://doi.org/10.1080/14616688.2020.1759134>
- Ayeh, J. K. (2015). Travellers' acceptance of consumer-generated media: An integrated model of technology acceptance and source credibility theories. *Computers in Human Behavior*, 48, 173–180. <https://doi.org/10.1016/J.CHB.2014.12.049>
- Ayeh, J. K., Au, N., & Law, R. (2013). "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. <https://doi.org/https://doi.org/10.1177/0047287512475217>
- Back, K. J. (2024). ESG for the hospitality and tourism research: Essential demanded research area for all. *Tourism Management*, 105, 104954. <https://doi.org/10.1016/J.TOURMAN.2024.104954>
- Baggio, R. (2020). Tourism destinations: A universality conjecture based on network science. *Annals of Tourism Research*, 82, 102929. <https://doi.org/10.1016/J.ANNALS.2020.102929>
- Baggio, R., & Sainaghi, R. (2011). Complex and chaotic tourism systems: Towards a quantitative approach. *International Journal of Contemporary Hospitality Management*, 23(6), 840–861. <https://doi.org/https://doi.org/10.1108/09596111111153501>
- Bansal, P., & Song, H. C. (2016). Similar But Not the Same: Differentiating Corporate Sustainability from Corporate Responsibility. *Academy of Management Annals*, 11(1), 105–149. <https://doi.org/10.5465/ANNALS.2015.0095>
- Barrena Martínez, J., López Fernández, M., & Romero Fernández, P. M. (2016). Corporate social responsibility: Evolution through institutional and stakeholder perspectives. *European Journal of Management and Business Economics*, 25(1), 8–14. <https://doi.org/10.1016/j.redde.2015.11.002>
- Baumgartner, R. J. (2005). Sustainable Business Management : Grundlagen, Strategien und Instrumente einer nachhaltigen Unternehmensführung. *Rainer Hampp: München*, 51–72.
- Baumgartner, R. J. (2009). Organizational culture and leadership: Preconditions for the development of sustainable corporation. *Sustainable Development*, 17(2), 102–113. <https://doi.org/10.1002/sd.405>

- Baumgartner, R. J., & Ebner, D. (2010). Corporate sustainability strategies: Sustainability profiles and maturity levels. *Sustainable Development*, 18(2), 76–89. <https://doi.org/10.1002/sd.447>
- Bellato, L., Frantzeskaki, N., Lee, E., Cheer, J. M., & Peters, A. (2024). Transformative epistemologies for regenerative tourism: towards a decolonial paradigm in science and practice? *Journal of Sustainable Tourism*, 32(6), 1161–1181. <https://doi.org/10.1080/09669582.2023.2208310>
- Bellato, L., Frantzeskaki, N., & Nygaard, C. A. (2022). Regenerative tourism: a conceptual framework leveraging theory and practice. *Tourism Geographies*, 25(4), 1026–1046. <https://doi.org/10.1080/14616688.2022.2044376>
- Bellato, L., & Pollock, A. (2023). Regenerative tourism: a state-of-the-art review. *Tourism Geographies*. <https://doi.org/10.1080/14616688.2023.2294366>
- Belz, F.-M., & Ken Peattie. (2012). *Sustainability marketing: A global perspective*. John Wiley & Sons.
- Berezan, O., Raab, C., Yoo, M., & Love, C. (2013). Sustainable hotel practices and nationality: The impact on guest satisfaction and guest intention to return. *International Journal of Hospitality Management*, 34(1), 227–233. <https://doi.org/10.1016/J.IJHM.2013.03.010>
- Bernini, C., Emili, S., & Vici, L. (2020). Are mass tourists sensitive to sustainability? *Tourism Economics*, 27(7), 1375–1397. <https://doi.org/https://doi.org/10.1177/1354816620923212>
- Bigné, E., Oltra, E., & Andreu, L. (2019). Harnessing stakeholder input on Twitter: A case study of short breaks in Spanish tourist cities. *Tourism Management*, 71, 490–503. <https://doi.org/10.1016/J.TOURMAN.2018.10.013>
- Bird, S., Loper, E., & Klein, E. (2009). *Natural Language Processing with Python*. O'Reilly Media Inc.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer.
- Bohdanowicz, P. (2005). European Hoteliers' Environmental Attitudes. *Cornell Hospitality Quarterly*, 46(2), 188–204. <https://doi.org/10.1177/0010880404273891>

- Bowen, H. (1953). *Social responsibilities of the businessman*. New York: Harper.
- Boyd, D., & Crawford, K. (2012). CRITICAL QUESTIONS FOR BIG DATA. *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Bramwell, B., Henry, I., Jackson, G., Prat, A. G., Richards, G. W., & Straaten, J. van der. (1996). *Sustainable Tourism Management: Principles and Practice*. Tilburg University Press. <https://research.tilburguniversity.edu/en/publications/sustainable-tourism-management-principles-and-practice>
- Bramwell, B., Higham, J., Lane, B., & Miller, G. (2017). Twenty-five years of sustainable tourism and the Journal of Sustainable Tourism: looking back and moving forward. *Journal of Sustainable Tourism*, 25(1), 1–9. <https://doi.org/10.1080/09669582.2017.1251689>
- Bramwell, B., & Lane, B. (1993). Sustainable tourism: An evolving global approach. *Journal of Sustainable Tourism*, 1(1), 1–5. <https://doi.org/https://doi.org/10.1080/09669589309450696>
- Bramwell, B., & Lane, B. (2000). Collaboration and Partnerships in Tourism Planning. In *Tourism Collaboration and Partnerships* (pp. 1–19). Channel View Publications.
- Bramwell, B., & Lane, B. (2011). Critical research on the governance of tourism and sustainability. *Journal of Sustainable Tourism*, 19(4–5), 411–421. <https://doi.org/10.1080/09669582.2011.580586>
- Branco, M. C., & Rodrigues, L. L. (2008). Factors Influencing Social Responsibility Disclosure by Portuguese Companies. *Journal of Business Ethics*, 83(4), 685–701. <https://doi.org/10.1007/s10551-007-9658-z>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706QP0630A>
- Brazyté, K., Weber, F., & Schaffner, D. (2016). Sustainability Management of Hotels: How Do Customers Respond in Online Reviews? *Journal of Quality Assurance in Hospitality & Tourism*, 18(3), 282–307. <https://doi.org/10.1080/1528008X.2016.1230033>
- Buckley, R. (2012). Sustainable tourism: Research and reality. *Annals of Tourism Research*, 39(2), 528–546. <https://doi.org/10.1016/J.ANNALS.2012.02.003>

- Buckley, R. (2020). Nature tourism and mental health: parks, happiness, and causation. *Journal of Sustainable Tourism*, 28(9), 1409–1424. <https://doi.org/10.1080/09669582.2020.1742725>
- Buckley, R. (2023). Tourism and Mental Health: Foundations, Frameworks, and Futures. *Journal of Travel Research*, 62(1), 3–20. <https://doi.org/https://doi.org/10.1177/00472875221087669>
- Buckley, R. C. (2019). Therapeutic mental health effects perceived by outdoor tourists: A large-scale, multi-decade, qualitative analysis. *Annals of Tourism Research*, 77, 164–167. <https://doi.org/10.1016/J.ANNALS.2018.12.017>
- Buckley, R., Zhong, L., & Martin, S. (2021). Mental health key to tourism infrastructure in China's new megapark. *Tourism Management*, 82, 104169. <https://doi.org/10.1016/J.TOURMAN.2020.104169>
- Bulchand-Gidumal, J., Melián-González, S., & González Lopez-Valcarcel, B. (2013). A social media analysis of the contribution of destinations to client satisfaction with hotels. *International Journal of Hospitality Management*, 35, 44–47. <https://doi.org/10.1016/J.IJHM.2013.05.003>
- Burbano, V. C., Delmas, M. A., & Cobo, M. J. (2023). The Past and Future of Corporate Sustainability Research. *Organization & Environment*, 37(2), 133–158. <https://doi.org/10.1177/10860266231213105>
- Butler, R. (1980). The concept of a tourist area cycle of evolution: Implications for management of resources. *Canadian Geographer / Le Géographe Canadien*, 24(1), 5–12. <https://doi.org/10.1111/J.1541-0064.1980.TB00970.X>
- Butler, R. (1999). Sustainable tourism: A state-of-the-art review. *Tourism Geographies*, 1(1), 7–25. <https://doi.org/10.1080/14616689908721291>
- Butler, R. (2005). Developing the destination: Difficulties in achieving sustainability. In W. Aleziak & R. Winiarski (Eds.), *Tourism in scientific research* (pp. 33–45). Wydawnictwo akademii wychowania fizycznego w krakowie i wyższej szkoły informatyki i zarządzania w rzeszowie.
- Butler, R. (2018). Sustainable Tourism in Sensitive Environments: A Wolf in Sheep's Clothing? *Sustainability*, 10(6), 1789. <https://doi.org/10.3390/SU10061789>

- Butler, R., Nelson, J. G., & Wall, G. (1993). Tourism-an evolutionary perspective. In Waterloo (Ed.), *Tourism-and sustainable development: monitoring, planning, managing* (p. 284). U. of Waterloo.
- Butowski, L. (2019). Tourist sustainability of destination as a measure of its development. *Current Issues in Tourism*, 22(9), 1043–1061. <https://doi.org/10.1080/13683500.2017.1351926>
- Byrd, E. T. (2007). Stakeholders in sustainable tourism development and their roles: Applying stakeholder theory to sustainable tourism development. *Tourism Review*, 62(2), 6–13. <https://doi.org/10.1108/16605370780000309>
- Cai, Y., Li, G., Wen, L., & Liu, C. (2024). Intellectual landscape and emerging trends of big data research in hospitality and tourism: A scientometric analysis. *International Journal of Hospitality Management*, 117, 103633. <https://doi.org/10.1016/J.IJHM.2023.103633>
- Caloffi, A., Colovic, A., Rizzoli, V., & Rossi, F. (2023). Innovation intermediaries' types and functions: A computational analysis of the literature. *Technological Forecasting and Social Change*, 189, 122351. <https://doi.org/10.1016/J.TECHFORE.2023.122351>
- Campón-Cerro, A. M., Hernández-Mogollón, J. M., & Alves, H. (2017). Sustainable improvement of competitiveness in rural tourism destinations: The quest for tourist loyalty in Spain. *Journal of Destination Marketing & Management*, 6(3), 252–266. <https://doi.org/10.1016/J.JDMM.2016.04.005>
- Campopiano, G., & de Massis, A. (2015). Corporate Social Responsibility Reporting: A Content Analysis in Family and Non-family Firms. *Journal of Business Ethics*, 129(3), 511–534. <https://doi.org/10.1007/s10551-014-2174-z>
- Cañete, J., Chaperon, G., Fuentes, R., Ho, J.-H., Kang, H., & Pérez, J. (2020). Spanish Pre-trained BERT Model and Evaluation Data. *ArXiv*. <https://doi.org/https://doi.org/10.48550/arXiv.2308.02976>
- Cao, L. (2017). Data Science: A Comprehensive Overview. *ACM Computing Surveys (CSUR)*, 50(3). <https://doi.org/10.1145/3076253>
- Caradonna, J. L. (2014). *Sustainability: A History*. Oxford University Press.

- Carroll, A. B. (1979). The Pyramid of Corporate Social Responsibility: Toward the Moral Management of Organizational Stakeholders. *Business Horizons*, 34(4), 39–48.
- Carvalho, P., & Alves, H. (2023). Customer value co-creation in the hospitality and tourism industry: a systematic literature review. *International Journal of Contemporary Hospitality Management*, 35(1), 250–273. <https://doi.org/10.1108/IJCHM-12-2021-1528>
- Casagrandi, R., & Rinaldi, S. (2002). A Theoretical Approach to Tourism Sustainability. *Conservation Ecology*, 6(1), 1. <https://doi.org/https://www.jstor.org/stable/26271854>
- Cavalcante, W. Q. de F., Coelho, A., & Bairrada, C. M. (2021). Sustainability and tourism marketing: A bibliometric analysis of publications between 1997 and 2020 using vosviewer software. *Sustainability*, 13(9). <https://doi.org/10.3390/su13094987>
- Ceballos-Lascurain, H. (1996). *Tourism, ecotourism, and protected areas: The state of nature-based tourism around the world and guidelines for its development*.
- Cellini, R., & Cuccia, T. (2013). Museum and monument attendance and tourism flow: a time series analysis approach. *Applied Economics*, 45(24), 3473–3482. <https://doi.org/10.1080/00036846.2012.716150>
- Cervera, D. de J., de Esteban Curiel, J., & Pérez-Bustamante Yábar, D. C. (2024). Machine Learning for short-term property rental pricing based on seasonality and proximity to food establishments. *British Food Journal*, 126(13), 332–352. <https://doi.org/10.1108/BFJ-07-2023-0634>
- Chai, C., Song, Y., & Qin, Z. (2021). A Thousand Words Express a Common Idea? Understanding International Tourists' Reviews of Mt. Huangshan, China, through a Deep Learning Approach. *Land*, 10(6), 549. <https://doi.org/10.3390/LAND10060549>
- Chaminade, C. (2020). Innovation for What? Unpacking the Role of Innovation for Weak and Strong Sustainability. *Journal of Sustainability Research*, 2(1). <https://doi.org/10.20900/jsr20200007>
- Chang, R. D., Zuo, J., Zhao, Z. Y., Zillante, G., Gan, X. L., & Soebarto, V. (2017). Evolving theories of sustainability and firms: History, future directions and implications for renewable energy research. In *Renewable and Sustainable Energy Reviews* (Vol. 72, pp. 48–56). Elsevier Ltd. <https://doi.org/10.1016/j.rser.2017.01.029>

- Chen, H., & Rahman, I. (2018). Cultural tourism: An analysis of engagement, cultural contact, memorable tourism experience and destination loyalty. *Tourism Management Perspectives*, 26, 153–163. <https://doi.org/10.1016/J.TMP.2017.10.006>
- Chen, Y., Fay, S., & Wang, Q. (2011). The Role of Marketing in Social Media: How Online Consumer Reviews Evolve. *Journal of Interactive Marketing*, 25(2), 85–94. <https://doi.org/10.1016/J.INTMAR.2011.01.003>
- Chhabra, D., Healy, R., & Sills, E. (2003). Staged authenticity and heritage tourism. *Annals of Tourism Research*, 30(3), 702–719. [https://doi.org/10.1016/S0160-7383\(03\)00044-6](https://doi.org/10.1016/S0160-7383(03)00044-6)
- Chilufya, A., Hughes, E., & Scheyvens, R. (2019). Tourists and community development: corporate social responsibility or tourist social responsibility? *Journal of Sustainable Tourism*, 27(10), 1513–1529. <https://doi.org/https://doi-org.unican.idm.oclc.org/10.1080/09669582.2019.1643871>
- Chiu, Y. T. H., Lee, W. I., & Chen, T. H. (2014). Environmentally responsible behavior in ecotourism: Antecedents and implications. *Tourism Management*, 40, 321–329. <https://doi.org/10.1016/J.TOURMAN.2013.06.013>
- Choi, H. C., & Murray, I. (2010). Resident attitudes toward sustainable community tourism. *Journal of Sustainable Tourism*, 18(4), 575–594. <https://doi.org/10.1080/09669580903524852>
- Choi, H. S. C., & Sirakaya, E. (2005). Measuring Residents' Attitude toward Sustainable Tourism: Development of Sustainable Tourism Attitude Scale. *Journal of Travel Research*, 43(4), 380–394. <https://doi.org/10.1177/0047287505274651>
- Choi, H. S. C., & Sirakaya, E. (2006). Sustainability indicators for managing community tourism. *Tourism Management*, 27(6), 1274–1289. <https://doi.org/10.1016/J.TOURMAN.2005.05.018>
- Clark, J. M. (1939). *Social control of business*. Chicago: University of Chicago Press.
- Clarke, J. (1997). A Framework of Approaches to Sustainable Tourism. *Journal of Sustainable Tourism*, 5(3), 224–233. <https://doi.org/10.1080/09669589708667287>

- Coccosis, H. (1996). Tourism and sustainability: Perspectives and implications. In G. K. Priestley, J. A. Edwards, & H. Coccosis (Eds.), *In Sustainable Tourism? European Experience* (pp. 1–21). CAB International.
- Coebergh, dr mr P. H. (2011). Voluntary Disclosure of Corporate Strategy: Determinants and Outcomes - An Empirical Study into the Risks and Payoffs of Communicating Corporate Strategy. *SSRN Electronic Journal*. <https://doi.org/10.2139/SSRN.1965029>
- Cohen, R., & Kennedy, P. (2017). *Global sociology*. Bloomsbury Publishing.
- Cong, L., Wu, B., Morrison, A. M., Shu, H., & Wang, M. (2014). Analysis of wildlife tourism experiences with endangered species: An exploratory study of encounters with giant pandas in Chengdu, China. *Tourism Management*, 40, 300–310. <https://doi.org/10.1016/J.TOURMAN.2013.07.005>
- Conti, S., Dias, Á., & Pereira, L. (2023). Perceived City Sustainability and Tourist Behavioural Intentions. *Smart Cities 2023, Vol. 6, Pages 692-708*, 6(2), 692–708. <https://doi.org/10.3390/SMARTCITIES6020033>
- Cooper, M. A., & Buckley, R. (2022). Tourist Mental Health Drives Destination Choice, Marketing, and Matching. *Journal of Travel Research*, 61(4), 786–799. <https://doi.org/10.1177/00472875211011548>
- Corral-Gonzalez, L., Cavazos-Arroyo, J., & García-Mestanza, J. (2023). Regenerative tourism: A bibliometric analysis. *Journal of Tourism, Heritage and Services Marketing*, 9(2), 41–54. <https://doi.org/10.5281/zenodo.10539722>
- Correia, A., Kozak, M., & Ferradeira, J. (2013). From tourist motivations to tourist satisfaction. *International Journal of Culture, Tourism, and Hospitality Research*, 7(4), 411–424. <https://doi.org/10.1108/IJCTHR-05-2012-0022>
- Costa, A. J., Curi, D., Bandeira, A. M., Ferreira, A., Tomé, B., Joaquim, C., Santos, C., Góis, C., Meira, D., Azevedo, G., Inácio, H., Jesus, M., Teixeira, M. G., Monteiro, P., Duarte, R., & Marques, R. P. (2022). Literature Review and Theoretical Framework of the Evolution and Interconnectedness of Corporate Sustainability Constructs. In *Sustainability (Switzerland)* (Vol. 14, Issue 8). MDPI. <https://doi.org/10.3390/su14084413>
- Cottrell, S. P., & Vaske, J. J. (2006). A framework for monitoring and modeling sustainable tourism. *E-Review of Tourism Research*, 4(9), 74–84.

- Cottrell, S. P., Vaske, J. J., & Roemer, J. M. (2013). Resident satisfaction with sustainable tourism: The case of Frankenwald Nature Park, Germany. *Tourism Management Perspectives*, 8, 42–48. <https://doi.org/10.1016/J.TMP.2013.05.005>
- Countryside Commission. (1995). *Sustaining Rural Tourism*. Countryside Commission (CCP 483).
- Cowper-Smith, A., & de Grosbois, D. (2011). The adoption of corporate social responsibility practices in the airline industry. *Journal of Sustainable Tourism*, 19(1), 59–77. <https://doi.org/10.1080/09669582.2010.498918>
- Crane, A., & Glozer, S. (2016). Researching Corporate Social Responsibility Communication: Themes, Opportunities and Challenges. In *Journal of Management Studies* (Vol. 53, Issue 7, pp. 1223–1252). Blackwell Publishing Ltd. <https://doi.org/10.1111/joms.12196>
- Cuccia, T., Guccio, C., & Rizzo, I. (2016). The effects of UNESCO World Heritage List inscription on tourism destinations performance in Italian regions. *Economic Modelling*, 53, 494–508. <https://doi.org/10.1016/J.ECONMOD.2015.10.049>
- D’Acunto, D., Filieri, R., & Amato, S. (2023). Hotels’ environmentally-framed eWOM. The moderating role of environmental culture. *Tourism Management*, 98. <https://doi.org/10.1016/j.tourman.2023.104776>
- D’Acunto, D., Filieri, R., & Amato, S. (2024). Who is sharing green eWOM? Big data evidence from the travel and tourism industry. *Journal of Sustainable Tourism*, 1–23. <https://doi.org/10.1080/09669582.2024.2328103>
- D’Acunto, D., Tuan, A., Dalli, D., Viglia, G., & Okumus, F. (2020a). Do consumers care about CSR in their online reviews? An empirical analysis. *International Journal of Hospitality Management*, 85, 102342. <https://doi.org/10.1016/J.IJHM.2019.102342>
- D’Acunto, D., Tuan, A., Dalli, D., Viglia, G., & Okumus, F. (2020b). Do consumers care about CSR in their online reviews? An empirical analysis. *International Journal of Hospitality Management*, 85. <https://doi.org/10.1016/j.ijhm.2019.102342>
- D’Agostino, R., & Pearson, E. S. (1973). Tests for Departure from Normality. Empirical Results for the Distributions of  $b_2$  and  $v$   $b_1$ . *Biometrika*, 60(3), 613–622. <https://doi.org/https://doi.org/10.2307/2335012>

- Damtoft, N. F., van Liempd, D., & Lueg, R. (2024). Sustainability performance measurement – a framework for context-specific applications. *Journal of Global Responsibility, ahead-of-print*(ahead-of-print). <https://doi.org/10.1108/JGR-05-2023-0082>
- Daugherty, T., Eastin, M. S., & Bright, L. (2008). Exploring Consumer Motivations for Creating User-Generated Content. *Journal of Interactive Advertising, 8*(2), 16–25. <https://doi.org/10.1080/15252019.2008.10722139>
- David, L. K., Wang, J., Angel, V., & Luo, M. (2024). China’s ESG scorecard: A predictive machine learning model. *Corporate Social Responsibility and Environmental Management, 31*(4), 3468–3486. <https://doi.org/10.1002/CSR.2746>
- de Grosbois, D. (2016). Corporate social responsibility reporting in the cruise tourism industry: a performance evaluation using a new institutional theory based model. *Journal of Sustainable Tourism, 24*(2), 245–269. <https://doi.org/10.1080/09669582.2015.1076827>
- Della Corte, V., Del Gaudio, G., Sepe, F., & Sciarelli, F. (2019). Sustainable Tourism in the Open Innovation Realm: A Bibliometric Analysis. *Sustainability, 11*(21), 6114. <https://doi.org/10.3390/SU11216114>
- Demeter, C., Fechner, D., & Dolnicar, S. (2023). Progress in field experimentation for environmentally sustainable tourism – A knowledge map and research agenda. *Tourism Management, 94*, 104633. <https://doi.org/10.1016/J.TOURMAN.2022.104633>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *ArXiv*. <https://doi.org/https://doi.org/10.48550/arXiv.1810.04805>
- Dewhurst, H., & Thomas, R. (2003). Encouraging Sustainable Business Practices in a Non-regulatory Environment: A Case Study of Small Tourism Firms in a UK National Park. *Journal of Sustainable Tourism, 11*(5), 383–403. <https://doi.org/10.1080/09669580308667212>
- Dias, B. D. (2019). Regenerative development—building evolutive capacity for healthy living systems. In *Management and Applications of Complex Systems* (8th ed., Vol. 13, pp. 315–323).
- Diesendorf, Mark. (2000). Sustainability and sustainable development. In *Sustainability: the Corporate Challenge of the 21st Century* (pp. 19–37).

- Diwan, H., & Amarayil Sreeraman, B. (2023). From financial reporting to ESG reporting: a bibliometric analysis of the evolution in corporate sustainability disclosures. *Environment, Development and Sustainability* 2023 26:6, 26(6), 13769–13805. <https://doi.org/10.1007/S10668-023-03249-2>
- Dolnicar, S., Crouch, G. I., & Long, P. (2008). Environment-friendly Tourists: What Do We Really Know About Them? *Journal of Sustainable Tourism*, 16(2), 197–210. <https://doi.org/10.2167/JOST738.0>
- Dolnicar, S., Knezevic Cvelbar, L., & Grün, B. (2019). A Sharing-Based Approach to Enticing Tourists to Behave More Environmentally Friendly. *Journal of Travel Research*, 58(2), 241–252. <https://doi.org/https://doi.org/10.1177/0047287517746013>
- Donoho, D. (2017). 50 Years of Data Science. *Journal of Computational and Graphical Statistics*, 26(4), 745–766. <https://doi.org/10.1080/10618600.2017.1384734>
- Donohoe, H. M. (2011). Defining culturally sensitive ecotourism: a Delphi consensus. *Current Issues in Tourism*, 14(1), 27–45. <https://doi.org/10.1080/13683500903440689>
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285–296. <https://doi.org/10.1016/J.JBUSRES.2021.04.070>
- Dredge, D. (2022). Regenerative tourism: transforming mindsets, systems and practices. *Journal of Tourism Futures*, 8(3), 269–281. <https://doi.org/10.1108/JTF-01-2022-0015>
- Dyllick, T. (2000). Strategischer Einsatz von Umweltmanagementsystemen. *Umweltwirtschaftsforum*, 8(3), 64–68.
- Eber, S. (1992). *Beyond the Green Horizon: A Discussion Paper on Principles for Sustainable Tourism*. World Wide Fund for Nature.
- Efron, B. (1979). Bootstrap methods: Another look at the Jackknife. *The Annals of Statistics*, 7(1), 1–26. <https://doi.org/DOI: 10.1214/aos/1176344552>
- Egger, R. (2022). *Tourism on the verge. Applied data science in tourism: Interdisciplinary approaches, methodologies, and applications*. Springer.

- Einwiller, S. A., & Carroll, C. E. (2020). Negative disclosures in corporate social responsibility reporting. *Corporate Communications*, 25(2), 319–337. <https://doi.org/10.1108/CCIJ-05-2019-0054>
- Elkington, J., & Rowlands, I. H. (1999). Cannibals with Forks: The Triple Bottom Line of Sustainability. *Alternatives Journal*, 25(4), 42.
- Ellili, N. O. D., & Nobanee, H. (2022). Impact of economic, environmental, and corporate social responsibility reporting on financial performance of UAE banks. *Environment, Development and Sustainability*. <https://doi.org/10.1007/s10668-022-02225-6>
- Espiner, S., Orchiston, C., & Higham, J. (2017). Resilience and sustainability: a complementary relationship? Towards a practical conceptual model for the sustainability–resilience nexus in tourism. *Journal of Sustainable Tourism*, 25(10), 1385–1400. <https://doi.org/10.1080/09669582.2017.1281929>
- Esteban Curiel, J. De, Delgado Jalón, M. L., Flores Ureba, S., & Rivero Menéndez, J. Á. (2018). Neural network analysis for hotel service design in Madrid: the 3Ps methodology and the frontline staff. *Tourism & Management Studies*, 14(S11), 83–94. <https://doi.org/10.18089/tms.2018.14S1109>
- Exceltur. (2024). *Perspectivas Turísticas: Valoración turística empresarial de 2023 y perspectivas para 2024*. <https://www.exceltur.org/wp-content/uploads/2024/02/Informe-Perspectivas-N87-Balance-del-ano-2023-y-previsiones-para-2024.pdf>
- Farmaki, A., Christou, P., Saveriades, A., & Spanou-Tripinioti, E. (2019). Perceptions of Pafos as European Capital of Culture: Tourism stakeholder and resident perspectives. *International Journal of Tourism Research*, 21(2), 234–244. <https://doi.org/10.1002/JTR.2257>
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. *AI Magazine*, 17(3), 37–37. <https://doi.org/10.1609/AIMAG.V17I3.1230>
- Fedorova, E., & Iasakova, P. (2024). The impact of climate change news on the US stock market. *Journal of Risk Finance*, 25(2), 293–320. <https://doi.org/10.1108/JRF-06-2023-0133>
- Fenniak, M. (2023). *PyPDF2: PDF toolkit*. <https://pypi.org/project/PyPDF2/>

- Fernandes, A. J. C., Rodrigues, R. G., & Ferreira, J. J. (2022). National innovation systems and sustainability: What is the role of the environmental dimension? *Journal of Cleaner Production*, 347. <https://doi.org/10.1016/j.jclepro.2022.131164>
- Ferrer-Roca, N., Weston, R., Guia, J., Mihalic, T., Blasco, D., Prats, L., Lawler, M., & Jarratt, D. (2020). Back to the future: challenges of European tourism of tomorrow. *Journal of Tourism Futures*, 7(2), 184–191. <https://doi.org/10.1108/JTF-10-2019-0114>
- Filieri, R., Acikgoz, F., Ndou, V., & Dwivedi, Y. (2021). Is TripAdvisor still relevant? The influence of review credibility, review usefulness, and ease of use on consumers' continuance intention. *International Journal of Contemporary Hospitality Management*, 33(1), 199–223. <https://doi.org/10.1108/IJCHM-05-2020-0402>
- Filieri, R., Raguseo, E., & Vitari, C. (2019). What moderates the influence of extremely negative ratings? The role of review and reviewer characteristics. *International Journal of Hospitality Management*, 77, 333–341. <https://doi.org/10.1016/J.IJHM.2018.07.013>
- Font, X., Elgammal, I., & Lamond, I. (2017). Greenhushing: the deliberate under communicating of sustainability practices by tourism businesses. *Journal of Sustainable Tourism*, 25(7), 1007–1023. <https://doi.org/10.1080/09669582.2016.1158829>
- Font, X., English, R., Gkritzali, A., & Tian, W. (Stella). (2021). Value co-creation in sustainable tourism: A service-dominant logic approach. *Tourism Management*, 82, 104200. <https://doi.org/10.1016/J.TOURMAN.2020.104200>
- Font, X., Garay, L., & Jones, S. (2016). Sustainability motivations and practices in small tourism enterprises in European protected areas. *Journal of Cleaner Production*, 137, 1439–1448. <https://doi.org/10.1016/J.JCLEPRO.2014.01.071>
- Font, X., & Lynes, J. (2018). Corporate social responsibility in tourism and hospitality. *Journal of Sustainable Tourism*, 26(7), 1027–1042. <https://doi.org/10.1080/09669582.2018.1488856>
- Font, X., & McCabe, S. (2017). Sustainability and marketing in tourism: its contexts, paradoxes, approaches, challenges and potential. *Journal of Sustainable Tourism*, 25(7), 869–883. <https://doi.org/10.1080/09669582.2017.1301721>
- Font, X., Walmsley, A., Cogotti, S., McCombes, L., & Häusler, N. (2012). Corporate social responsibility: The disclosure-performance gap. *Tourism Management*, 33(6), 1544–1553. <https://doi.org/10.1016/j.tourman.2012.02.012>

- Freeman, E. (1984). *Strategic Management: A Stakeholder Approach*. Pitman: Boston.
- Galeone, A., & Sebastiani, R. (2021). Transformative Service Research in Hospitality. *Tourism Management*, 87, 104366. <https://doi.org/10.1016/J.TOURMAN.2021.104366>
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137–144. <https://doi.org/10.1016/J.IJINFOMGT.2014.10.007>
- Gao, Y., Zheng, W., & Wang, Y. (2023). Sectoral risk contagion and quantile network connectedness on Chinese stock sectors after the COVID-19 outbreak. *China Finance Review International*, ahead-of-print(ahead-of-print). <https://doi.org/10.1108/CFRI-02-2023-0039>
- Garay, L., & Font, X. (2012). Doing good to do well? Corporate social responsibility reasons, practices and impacts in small and medium accommodation enterprises. *International Journal of Hospitality Management*, 31(2), 329–337. <https://doi.org/10.1016/J.IJHM.2011.04.013>
- García, S., Ramírez-Gallego, S., Luengo, J., & Herrera, F. (2016). *Big Data: Preprocesamiento y calidad de datos*. [https://sci2s.ugr.es/sites/default/files/ficherosPublicaciones/2133\\_Nv237-Digital-sramirez.pdf](https://sci2s.ugr.es/sites/default/files/ficherosPublicaciones/2133_Nv237-Digital-sramirez.pdf)
- Garijo, D., Alper, P., Belhajjame, K., Corcho, O., Gil, Y., & Goble, C. (2014). Common motifs in scientific workflows: An empirical analysis. *Future Generation Computer Systems*, 36, 338–351. <https://doi.org/10.1016/J.FUTURE.2013.09.018>
- Garriga, E., & Melé, D. (2004). Corporate social responsibility theories: Mapping the territory. *Journal of Business Ethics*, 53(1), 51–71.
- Garrigos-Simon, F. J., Narangajavana-Kaosiri, Y., & Lengua-Lengua, I. (2018). Tourism and sustainability: A bibliometric and visualization analysis. *Sustainability*, 10(6). <https://doi.org/10.3390/su10061976>
- Gatti, L., & Seele, P. (2014). Evidence for the prevalence of the sustainability concept in European corporate responsibility reporting. *Sustainability Science*, 9(1), 89–102. <https://doi.org/10.1007/s11625-013-0233-5>

- George, G., Haas, M. R., & Pentland, A. (2014). Big Data and Management. *Academy of Management Journal*, 57(2), 321–326. <https://doi.org/10.5465/AMJ.2014.4002>
- George, G., Osinga, E., Lavie, D., & Scott, B. (2016). Big Data and Data Science Methods for Management Research. *Academy of Management Journal*, 59(5), 1493–1507. <https://doi.org/10.5465/AMJ.2016.4005>
- Gerdt, S. O., Wagner, E., & Schewe, G. (2019). The relationship between sustainability and customer satisfaction in hospitality: An explorative investigation using eWOM as a data source. *Tourism Management*, 74, 155–172. <https://doi.org/10.1016/j.tourman.2019.02.010>
- Gilbert, D., & Clark, M. (1997). An exploratory examination of urban tourism impact, with reference to residents attitudes, in the cities of Canterbury and Guildford. *Cities*, 14(6), 343–352. [https://doi.org/10.1016/S0264-2751\(97\)00025-5](https://doi.org/10.1016/S0264-2751(97)00025-5)
- Girella, L., Zambon, S., & Rossi, P. (2019). Reporting on sustainable development: A comparison of three Italian small and medium-sized enterprises. *Corporate Social Responsibility and Environmental Management*, 26(4), 981–996. <https://doi.org/10.1002/csr.1738>
- González-Rodríguez, M. R., Martín-Samper, R. C., Köseoglu, M. A., & Okumus, F. (2019). Hotels' corporate social responsibility practices, organizational culture, firm reputation, and performance. *Journal of Sustainable Tourism*, 27(3), 398–419. <https://doi.org/10.1080/09669582.2019.1585441>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. The MIT Press.
- Google, & Deloitte. (2024). *Viajeros y destinos NextGen: nuestra visión sobre la transformación del sector turístico*. [https://perspectivas.deloitte.com/l/915781/2024-06-11/t9wxf/915781/17180970624JfWZbST/20240305\\_Viajeros\\_Destinos\\_NextGen\\_Google.pdf](https://perspectivas.deloitte.com/l/915781/2024-06-11/t9wxf/915781/17180970624JfWZbST/20240305_Viajeros_Destinos_NextGen_Google.pdf)
- Gössling, S. (2002). Global environmental consequences of tourism. *Global Environmental Change*, 12(4), 283–302. [https://doi.org/10.1016/S0959-3780\(02\)00044-4](https://doi.org/10.1016/S0959-3780(02)00044-4)
- Gössling, S., Peeters, P., Hall, C. M., Ceron, J. P., Dubois, G., Lehmann, L. V., & Scott, D. (2012). Tourism and water use: Supply, demand, and security. An international review. *Tourism Management*, 33(1), 1–15. <https://doi.org/10.1016/J.TOURMAN.2011.03.015>

- Gössling, S., Scott, D., & Hall, C. M. (2020). Pandemics, tourism and global change: a rapid assessment of COVID-19. *Journal of Sustainable Tourism*, 1–20. <https://doi.org/10.1080/09669582.2020.1758708>
- Gray, R. (2002). The social accounting project and Accounting Organizations and Society Privileging engagement, imaginings, new accountings and pragmatism over critique? *Accounting, Organizations and Society*, 27(7), 687–708. [https://doi.org/10.1016/S0361-3682\(00\)00003-9](https://doi.org/10.1016/S0361-3682(00)00003-9)
- Grootendorst, M. (2021). *MaartenGr/BERTopic: Fix embedding parameter*. <https://doi.org/10.5281/ZENODO.4430182>
- Grosbois, D. (2012). Corporate social responsibility reporting by the global hotel industry: Commitment, initiatives and performance. *International Journal of Hospitality Management*, 31(3), 896–905. <https://doi.org/10.1016/j.ijhm.2011.10.008>
- Guccio, C., Lisi, D., Martorana, M., & Mignosa, A. (2017). On the role of cultural participation in tourism destination performance: an assessment using robust conditional efficiency approach. *Journal of Cultural Economics*, 41(2), 129–154. <https://doi.org/10.1007/S10824-017-9295-Z>
- Guerrero-Rodriguez, R., Álvarez-Carmona, M., Aranda, R., & López-Monroy, A. P. (2023). Studying Online Travel Reviews related to tourist attractions using NLP methods: the case of Guanajuato, Mexico. *Current Issues in Tourism*, 26(2), 289–304. <https://doi.org/10.1080/13683500.2021.2007227>
- Guix, M., Nájera Sánchez, J. J., Bonilla Priego, M. J., & Font, X. (2025). The changing institutional logics behind sustainability reports from the largest hotel groups in the world in 2014, 2018 and 2021. *Tourism Management*, 106, 105031. <https://doi.org/10.1016/J.TOURMAN.2024.105031>
- Guix Oliver, J. (2008). El análisis de contenidos: ¿Qué nos están diciendo? *Revista de Calidad Asistencial*, 23(1), 26–30. [https://doi.org/10.1016/S1134-282X\(08\)70464-0](https://doi.org/10.1016/S1134-282X(08)70464-0)
- Guizzard, A., Stacchini, A., & Costa, M. (2022). Can sustainability drive tourism development in small rural areas? Evidences from the Adriatic. *Journal of Sustainable Tourism*, 30(6), 1280–1300. <https://doi.org/10.1080/09669582.2021.1931256>

- Guler, N., Kirshner, S. N., & Vidgen, R. (2024). A literature review of artificial intelligence research in business and management using machine learning and ChatGPT. *Data and Information Management*, 100076. <https://doi.org/10.1016/J.DIM.2024.100076>
- Hackston, D., & Milne, M. J. (1996). Some determinants of social and environmental disclosures in New Zealand companies. *Accounting, Auditing & Accountability Journal*, 9(1), 77–108. <https://doi.org/10.1108/09513579610109987>
- Hall, C. M., & Page, S. J. (2006). *The Geography of Tourism and Recreation Environment, Place and Space*. Routledge.
- Han, H. (2021). Consumer behavior and environmental sustainability in tourism and hospitality: a review of theories, concepts, and latest research. *Journal of Sustainable Tourism*, 29(7), 1021–1042. <https://doi.org/10.1080/09669582.2021.1903019>
- Hardtke, A., & Prehn, M. (2001). *Perspektiven der Nachhaltigkeit* (Vol. 1). Gabler: Wiesbaden.
- Hardy, A., Beeton, R. J. S., & Pearson, L. (2002). Sustainable Tourism: An Overview of the Concept and its Position in Relation to Conceptualisations of Tourism. *Journal of Sustainable Tourism*, 10(6), 475–496. <https://doi.org/10.1080/09669580208667183>
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., ... Oliphant, T. E. (2020). Array programming with NumPy. *Nature* 2020 585:7825, 585(7825), 357–362. <https://doi.org/10.1038/s41586-020-2649-2>
- Hector, D. C., Christensen, C. B., & Petrie, J. (2014). Sustainability and Sustainable Development: Philosophical Distinctions and Practical Implications. *Environmental Values*, 23(1), 7–28. <https://doi.org/10.3197/096327114X13851122268963>
- Hemmert, G. A. J., Schons, L. M., Wieseke, J., & Schimmelpfennig, H. (2016). Log-likelihood-based Pseudo-R2 in Logistic Regression. <https://doi.org/10.1177/0049124116638107>, 47(3), 507–531. <https://doi.org/10.1177/0049124116638107>
- Henderson, J. C. (2007). Corporate social responsibility and tourism: Hotel companies in Phuket, Thailand, after the Indian Ocean tsunami. *International Journal of Hospitality Management*, 26(1), 228–239. <https://doi.org/10.1016/j.ijhm.2006.02.001>

- Henry, I. P., & Jackson, G. A. M. (1996). Sustainability of Management Processes and Tourism Products and Contexts. *Journal of Sustainable Tourism*, 4(1), 17–28. <https://doi.org/10.1080/09669589608667256>
- Hernández, J. M., Santana-Jiménez, Y., & González-Martel, C. (2021). Factors influencing the co-occurrence of visits to attractions: The case of Madrid, Spain. *Tourism Management*, 83, 104236. <https://doi.org/10.1016/J.TOURMAN.2020.104236>
- Hey, T., Tansley, S., Tolle, K., & Gray, J. (2009). *The fourth paradigm: data-intensive scientific discovery, volume 1*. Microsoft Research.
- Honnibal, M., Montani, I., Van Landeghem, S., & Boyd, A. (2020). *spaCy: Industrial-strength Natural Language Processing in Python*. <https://doi.org/10.5281/zenodo.121230>
- Hooghiemstra, R. (2000). Corporate Communication and Impression Management – New Perspectives Why Companies Engage in Corporate Social Reporting. *Journal of Business Ethics* 2000 27:1, 27(1), 55–68. <https://doi.org/10.1023/A:1006400707757>
- Höpken, W., Eberle, T., Fuchs, M., & Lexhagen, M. (2021). Improving Tourist Arrival Prediction: A Big Data and Artificial Neural Network Approach. *Journal of Travel Research*, 60(5), 998–1017. <https://doi.org/10.1177/0047287520921244>
- Hopkins, D. (2020). Sustainable mobility at the interface of transport and tourism. *Journal of Sustainable Tourism*, 28(2), 225–239. <https://doi.org/10.1080/09669582.2019.1691800>
- Howie, F. (2003). *Managing the tourist destination*. Cengage Learning EMEA.
- Hsu, C. W., Lee, W. H., & Chao, W. C. (2013). Materiality analysis model in sustainability reporting: A case study at Lite-On Technology Corporation. *Journal of Cleaner Production*, 57, 142–151. <https://doi.org/10.1016/j.jclepro.2013.05.040>
- Hsu, C. Y., Chen, M. Y., Nyaupane, G. P., & Lin, S. H. (2020). Measuring sustainable tourism attitude scale (SUS-TAS) in an Eastern island context. *Tourism Management Perspectives*, 33, 100617. <https://doi.org/10.1016/J.TMP.2019.100617>
- Huang, A., Wang, H., & Yang, Y. (2022). FinBERT: A Large Language Model for Extracting Information from Financial Text. *Contemporary Accounting Research*, 40(2), 806–841. <https://doi.org/10.1111/1911-3846.12832>

- Huayhuaca, C., Cottrell, S., Raadik, J., & Gradl, S. (2010). Resident perceptions of sustainable tourism development: Frankenwald nature park, Germany. *International Journal of Tourism Policy*, 3(2), 125–141. <https://doi.org/10.1504/IJTP.2010.034207>
- Hugging Face. (2024). <https://huggingface.co/> Accessed 16 Feb 2024.
- Hunter, C. (1997). Sustainable tourism as an adaptive paradigm. *Annals of Tourism Research*, 24(4), 850–867. [https://doi.org/10.1016/S0160-7383\(97\)00036-4](https://doi.org/10.1016/S0160-7383(97)00036-4)
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science and Engineering*, 9(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>
- Im, A. Y., & Kim, D. Y. (2023). Understanding Disrespectful Tourist Behavior Through Tourists' Stereotyped Perceptions. *Journal of Travel Research*. <https://doi.org/10.1177/00472875231209978>
- Iniesta-Bonillo, M. A., Sánchez-Fernández, R., & Jiménez-Castillo, D. (2016). Sustainability, value, and satisfaction: Model testing and cross-validation in tourist destinations. *Journal of Business Research*, 69(11), 5002–5007. <https://doi.org/10.1016/J.JBUSRES.2016.04.071>
- ISO. (2010). *ISO 26000:2010 Guía de responsabilidad social*.
- Jalilvand, M. R., & Samiei, N. (2012). The impact of electronic word of mouth on a tourism destination choice: Testing the theory of planned behavior (TPB). *Internet Research*, 22(5), 591–612. <https://doi.org/10.1108/10662241211271563>
- Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3), 685–695. <https://doi.org/10.1007/S12525-021-00475-2>
- Janssen, M., van der Voort, H., & Wahyudi, A. (2017). Factors influencing big data decision-making quality. *Journal of Business Research*, 70, 338–345. <https://doi.org/10.1016/J.JBUSRES.2016.08.007>
- Jasrotia, S. S., Kamila, M. K., & Patel, V. K. (2023). Impact of Sustainable Tourism on Tourist's Satisfaction: Evidence from India. *Business Perspectives and Research*, 11(2), 173–189. <https://doi.org/10.1177/22785337211043960>
- Jasrotia, S. S., Mishra, H. G., & Koul, S. (2019). Brick or Click? Channel Choice Disruptions in Travel Industry. *Asia-Pacific Journal of Management Research and Innovation*, 15(1–2), 16–26. <https://doi.org/10.1177/2319510x19829308>

- Jiang, J., Zhang, J., Zhang, H., & Yan, B. (2018). Natural soundscapes and tourist loyalty to nature-based tourism destinations: the mediating effect of tourist satisfaction. *Journal of Travel & Tourism Marketing*, 35(2), 218–230. <https://doi.org/10.1080/10548408.2017.1351415>
- Jones, P., Hillier, D., & Comfort, D. (2014). Sustainability in the global hotel industry. *International Journal of Contemporary Hospitality Management*, 26(1), 5–17. <https://doi.org/10.1108/IJCHM-10-2012-0180>
- Jones, P., Hillier, D., & Comfort, D. (2016). Sustainability in the hospitality industry: Some personal reflections on corporate challenges and research agendas. *International Journal of Contemporary Hospitality Management*, 28(1), 36–67. <https://doi.org/10.1108/IJCHM-11-2014-0572>
- Jovicic, D. Z. (2017). From the traditional understanding of tourism destination to the smart tourism destination. *Current Issues in Tourism*, 22(3), 276–282. <https://doi.org/10.1080/13683500.2017.1313203>
- Kabadayi, S., O'Connor, G. E., & Tuzovic, S. (2020). Viewpoint: The impact of coronavirus on service ecosystems as service mega-disruptions. *Journal of Services Marketing*, 34(6), 809–817. <https://doi.org/10.1108/JSM-03-2020-0090>
- Kapera, I. (2018). Sustainable tourism development efforts by local governments in Poland. *Sustainable Cities and Society*, 40, 581–588. <https://doi.org/10.1016/J.SCS.2018.05.001>
- Karen, P. (2008). Corporate sustainability, citizenship and social responsibility reporting: A website study of 100 model corporations. *The Journal of Corporate Citizenship*, 32, 63–78. <https://doi.org/https://www.jstor.org/stable/jcorpciti.32.63>
- Kastenholz, E., Carneiro, M. J., Peixeira Marques, C., & Lima, J. (2012). Understanding and managing the rural tourism experience — The case of a historical village in Portugal. *Tourism Management Perspectives*, 4, 207–214. <https://doi.org/10.1016/J.TMP.2012.08.009>
- Kazeminia, A., Hultman, M., & Mostaghel, R. (2016). Why pay more for sustainable services? The case of ecotourism. *Journal of Business Research*, 69(11), 4992–4997. <https://doi.org/10.1016/J.JBUSRES.2016.04.069>

- Kemperman, A., Arentze, T., & Aksenov, P. (2019). Tourists' City Trip Activity Program Planning: A Personalized Stated Choice Experiment. *Tourism, Hospitality and Event Management*, 53–70. [https://doi.org/10.1007/978-3-030-11160-1\\_4](https://doi.org/10.1007/978-3-030-11160-1_4)
- Khan, H. (2010). The effect of corporate governance elements on corporate social responsibility (CSR) reporting: Empirical evidence from private commercial banks of Bangladesh. *International Journal of Law and Management*, 52(2), 82–109.
- Khan, M., Uddin, M. F., & Gupta, N. (2014). Seven V's of Big Data understanding Big Data to extract value. *Proceedings of the 2014 Zone 1 Conference of the American Society for Engineering Education - "Engineering Education: Industry Involvement and Interdisciplinary Trends", ASEE Zone 1 2014*. <https://doi.org/10.1109/ASEEZONE1.2014.6820689>
- Kim, G., Duffy, L. N., & Moore, D. W. (2020). Tourist attractiveness: measuring residents' perception of tourists. *Journal of Sustainable Tourism*, 28(6), 898–916. <https://doi.org/10.1080/09669582.2019.1708919>
- Kim, J. H. (2014). The antecedents of memorable tourism experiences: The development of a scale to measure the destination attributes associated with memorable experiences. *Tourism Management*, 44, 34–45. <https://doi.org/10.1016/J.TOURMAN.2014.02.007>
- Kim, J. H., Wang, Y., & Song, H. (2020). Understanding the causes of negative tourism experiences. *Current Issues in Tourism*, 24(3), 304–320. <https://doi.org/10.1080/13683500.2020.1711711>
- Kim, M., & Kim, J. (2020). The Influence of Authenticity of Online Reviews on Trust Formation among Travelers. *Journal of Travel Research*, 59(5), 763–776. <https://doi.org/10.1177/0047287519868307>
- Kim, S., & Filimonau, V. (2017). On linguistic relativity and pro-environmental attitudes in tourism. *Tourism Management*, 63, 158–169. <https://doi.org/10.1016/J.TOURMAN.2017.06.012>
- Kim, S. H., Lee, K., & Fairhurst, A. (2017). The review of "green" research in hospitality, 2000–2014: Current trends and future research directions. *International Journal of Contemporary Hospitality Management*, 29(1), 226–247. <https://doi.org/10.1108/IJCHM-11-2014-0562>

- Koh, Y. (2024). The industry-academia gap on the corporate governance issues in tourism and hospitality: A critical review and future research suggestions. *Journal of Travel & Tourism Marketing*, 41(1), 3–19. <https://doi.org/10.1080/10548408.2023.2229364>
- Kok, J. N., Boers, E. J. W., Kosters, W. A., Putten, P. van der, & Poel, M. (2009). Artificial intelligence: definition, trends, techniques, and cases. *Artificial Intelligence*, 270–299.
- Konietzko, J., Das, A., & Bocken, N. (2023). Towards regenerative business models: A necessary shift? *Sustainable Production and Consumption*, 38, 372–388. <https://doi.org/10.1016/J.SPC.2023.04.014>
- Köseoglu, M. A., Mehraliyev, F., Altin, M., & Okumus, F. (2020). Competitor intelligence and analysis (CIA) model and online reviews: integrating big data text mining with network analysis for strategic analysis. *Tourism Review*, 76(3), 529–552. <https://doi.org/10.1108/TR-10-2019-0406>
- Koseoglu, M. A., Uyar, A., Kilic, M., Kuzey, C., & Karaman, A. S. (2021). Exploring the connections among CSR performance, reporting, and external assurance: Evidence from the hospitality and tourism industry. *International Journal of Hospitality Management*, 94, 102819. <https://doi.org/10.1016/J.IJHM.2020.102819>
- Kumar, P., Aggarwal, B., Kumar, V., & Saini, H. (2024). Sustainable tourism progress: a 10-year bibliometric analysis. *Cogent Social Sciences*, 10(1), 2299614. <https://doi.org/10.1080/23311886.2023.2299614>
- Kumar, S., & Shekhar. (2020). Technology and innovation: Changing concept of rural tourism-A systematic review. *Open Geosciences*, 12(1), 737–752. <https://doi.org/10.1515/geo-2020-0183>
- Kwon, W. (2023). Reading customers' minds through textual big data: Challenges, practical guidelines, and proposals. *International Journal of Hospitality Management*, 111. <https://doi.org/10.1016/j.ijhm.2023.103473>
- Labrinidis, A., & Jagadish, H. V. (2012). Challenges and opportunities with big data. *Proceedings of the VLDB Endowment*, 5(12), 2032–2033. <https://doi.org/10.14778/2367502.2367572>
- Lalicic, L., Marine-Roig, E., Ferrer-Rosell, B., & Martin-Fuentes, E. (2021). Destination image analytics for tourism design: An approach through Airbnb reviews. *Annals of Tourism Research*, 86, 103100. <https://doi.org/10.1016/J.ANNALS.2020.103100>

- Landrum, N. E., & Ohsowski, B. (2018). Identifying Worldviews on Corporate Sustainability: A Content Analysis of Corporate Sustainability Reports. *Business Strategy and the Environment*, 27(1), 128–151. <https://doi.org/10.1002/bse.1989>
- Laney, D. (2001). *3-D data management: Controlling data volume, velocity and variety*. <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>
- Lechuga Sancho, M. P., Martín-Navarro, A., & Ramos-Rodríguez, A. R. (2020). Information Systems Management Tools: An Application of Bibliometrics to CSR in the Tourism Sector. *Sustainability*, 12(20), 8697. <https://doi.org/10.3390/su12208697>
- Lee, S. W., & Xue, K. (2020). A model of destination loyalty: integrating destination image and sustainable tourism. *Asia Pacific Journal of Tourism Research*, 25(4), 393–408. <https://doi.org/10.1080/10941665.2020.1713185>
- Lee, S., Yeon, J., & Song, H. J. (2022). Current status and future perspective of the link of corporate social responsibility–corporate financial performance in the tourism and hospitality industry. *Tourism Economics*, 29(7), 1703–1735. <https://doi.org/10.1177/13548166221140505>
- Lee, T. H. (2011). How recreation involvement, place attachment and conservation commitment affect environmentally responsible behavior. *Journal of Sustainable Tourism*, 19(7), 895–915. <https://doi.org/10.1080/09669582.2011.570345>
- Lee, T. H. (2013). Influence analysis of community resident support for sustainable tourism development. *Tourism Management*, 34, 37–46. <https://doi.org/10.1016/J.TOURMAN.2012.03.007>
- Legendre, T. S., Ding, A., & Back, K. J. (2024). A bibliometric analysis of the hospitality and tourism environmental, social, and governance (ESG) literature. *Journal of Hospitality and Tourism Management*, 58, 309–321. <https://doi.org/10.1016/J.JHTM.2024.01.003>
- Leiper, N. (2000). Are Destinations “The Heart of Tourism”? The Advantages of an Alternative Description. *Current Issues in Tourism*, 3(4), 364–368. <https://doi.org/10.1080/13683500008667878>
- León, C. J., Suárez-Rojas, C., Cazorla-Artiles, J. M., & González Hernández, M. M. (2025). Satisfaction and sustainability concerns in whale-watching tourism: A user-generated

- content model. *Tourism Management*, 106, 105019. <https://doi.org/10.1016/J.TOURMAN.2024.105019>
- Leonidou, L. C., Coudounaris, D. N., Kvasova, O., & Christodoulides, P. (2015). Drivers and Outcomes of Green Tourist Attitudes and Behavior: Sociodemographic Moderating Effects. *Psychology & Marketing*, 32(6), 635–650. <https://doi.org/10.1002/MAR.20806>
- Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management*, 68, 301–323. <https://doi.org/10.1016/J.TOURMAN.2018.03.009>
- Li, N., Yang, X., Wong, I. K. A., Law, R., Xu, J. Y., & Zhang, B. (2023). Automating tourism online reviews: a neural network based aspect-oriented sentiment classification. *Journal of Hospitality and Tourism Technology*, 14(1), 1–20. <https://doi.org/10.1108/JHTT-03-2021-0099>
- Li, X., & Law, R. (2020). Network analysis of big data research in tourism. *Tourism Management Perspectives*, 33, 100608. <https://doi.org/10.1016/J.TMP.2019.100608>
- Li, X., Li, H., Pan, B., & Law, R. (2021). Machine Learning in Internet Search Query Selection for Tourism Forecasting. *Journal of Travel Research*, 60(6), 1213–1231. <https://doi.org/10.1177/0047287520934871>
- Li, X., Pan, B., Law, R., & Huang, X. (2017). Forecasting tourism demand with composite search index. *Tourism Management*, 59, 57–66. <https://doi.org/10.1016/J.TOURMAN.2016.07.005>
- Li, Y. Q., & Liu, C. H. (2019). Impact of cultural contact on satisfaction and attachment: mediating roles of creative experiences and cultural memories. *Journal of Hospitality Marketing & Management*, 29(2), 221–245. <https://doi.org/10.1080/19368623.2019.1611516>
- Liang, X., Li, X., Shu, L., Wang, X., & Luo, P. (2024). Tourism demand forecasting using graph neural network. *Current Issues in Tourism*. <https://doi.org/10.1080/13683500.2024.2320851>
- Liu, S. H., Chen, S. Y., & Li, S. T. (2017). Text-Mining Application on CSR Report Analytics: A Study of Petrochemical Industry. *Proceedings - 2017 6th IIAI International Congress on Advanced Applied Informatics, IIAI-AAI 2017*, 76–81. <https://doi.org/10.1109/IIAI-AAI.2017.164>

- Liu, Z. (2003). Sustainable Tourism Development: A Critique. *Journal of Sustainable Tourism*, 11(6), 459–475. <https://doi.org/10.1080/09669580308667216>
- Liu, Z., & Park, S. (2015). What makes a useful online review? Implication for travel product websites. *Tourism Management*, 47, 140–151. <https://doi.org/10.1016/J.TOURMAN.2014.09.020>
- Lo, A. S., & Yao, S. S. (2019). What makes hotel online reviews credible?: An investigation of the roles of reviewer expertise, review rating consistency and review valence. *International Journal of Contemporary Hospitality Management*, 31(1), 41–60. <https://doi.org/10.1108/IJCHM-10-2017-0671>
- López-Sanz, J. M., Penelas-Leguía, A., Gutiérrez-Rodríguez, P., & Cuesta-Valiño, P. (2021). Sustainable Development and Rural Tourism in Depopulated Areas. *Land 2021, Vol. 10, Page 985, 10(9)*, 985. <https://doi.org/10.3390/LAND10090985>
- Lozano, R. (2018). Sustainable business models: Providing a more holistic perspective. *Business Strategy and the Environment*, 27(8), 1159–1166. <https://doi.org/10.1002/bse.2059>
- Lu, W., & Stepchenkova, S. (2015). User-Generated Content as a Research Mode in Tourism and Hospitality Applications: Topics, Methods, and Software. *Journal of Hospitality Marketing & Management*, 24(2), 119–154. <https://doi.org/10.1080/19368623.2014.907758>
- Lundberg, E. (2015). The Level of Tourism Development and Resident Attitudes: A Comparative Case Study of Coastal Destinations. *Scandinavian Journal of Hospitality and Tourism*, 15(3), 266–294. <https://doi.org/10.1080/15022250.2015.1005335>
- Madanaguli, A., Dhir, A., Joseph, R. P., Albishri, N. A., & Srivastava, S. (2023). Environmental sustainability practices and strategies in the rural tourism and hospitality sector: a systematic literature review and suggestions for future research. *Scandinavian Journal of Hospitality and Tourism*, 23(1), 1–28. <https://doi.org/10.1080/15022250.2023.2174179>
- Madanaguli, A., Srivastava, S., Ferraris, A., & Dhir, A. (2022). Corporate social responsibility and sustainability in the tourism sector: A systematic literature review and future outlook. *Sustainable Development*, 30(3), 447–461. <https://doi.org/10.1002/sd.2258>
- Major, J., & Clarke, D. (2022). Regenerative tourism in Aotearoa New Zealand – A new paradigm for the VUCA world. *Journal of Tourism Futures*, 8(2), 194–199. <https://doi.org/10.1108/JTF-09-2021-0233>

- Malone, S., McCabe, S., & Smith, A. P. (2014). The role of hedonism in ethical tourism. *Annals of Tourism Research*, 44(1), 241–254. <https://doi.org/10.1016/J.ANNALS.2013.10.005>
- Manning, C. D., Raghavan, P., & Schütze, H. (2009). Introduction to Information Retrieval. In *Introduction to Information Retrieval*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511809071>
- Mariani, M., & Baggio, R. (2022). Big data and analytics in hospitality and tourism: a systematic literature review. *International Journal of Contemporary Hospitality Management*, 34(1), 231–278. <https://doi.org/10.1108/IJCHM-03-2021-0301>
- Mariani, M., Baggio, R., Fuchs, M., & Höepken, W. (2018). Business intelligence and big data in hospitality and tourism: a systematic literature review. *International Journal of Contemporary Hospitality Management*, 30(12), 3514–3554. <https://doi.org/10.1108/IJCHM-07-2017-0461>
- Mariani, M., & Borghi, M. (2023). Exploring environmental concerns on digital platforms through big data: the effect of online consumers' environmental discourse on online review ratings. *Journal of Sustainable Tourism*, 31(11), 2592–2611. <https://doi.org/10.1080/09669582.2022.2033982>
- Marine-Roig, E., & Anton Clavé, S. (2015). Tourism analytics with massive user-generated content: A case study of Barcelona. *Journal of Destination Marketing & Management*, 4(3), 162–172. <https://doi.org/10.1016/J.JDMM.2015.06.004>
- Marine-Roig, E., & Clave, S. A. (2015). A Method for Analysing Large-Scale UGC Data for Tourism: Application to the Case of Catalonia. *Information and Communication Technologies in Tourism 2015*, 3–17. [https://doi.org/10.1007/978-3-319-14343-9\\_1](https://doi.org/10.1007/978-3-319-14343-9_1)
- Martínez García de Leaniz, P., Herrero Crespo, Á., & Gómez López, R. (2018). Customer responses to environmentally certified hotels: the moderating effect of environmental consciousness on the formation of behavioral intentions. *Journal of Sustainable Tourism*, 26(7), 1160–1177. <https://doi.org/10.1080/09669582.2017.1349775>
- Martínez, P., & Rodríguez del Bosque, I. (2013). CSR and customer loyalty: The roles of trust, customer identification with the company and satisfaction. *International Journal of Hospitality Management*, 35, 89–99. <https://doi.org/10.1016/J.IJHM.2013.05.009>

- Martínez, P., & Rodríguez del Bosque, I. (2014). Sustainability Dimensions: A Source to Enhance Corporate Reputation. *Corporate Reputation Review* 2014 17:4, 17(4), 239–253. <https://doi.org/10.1057/CRR.2014.13>
- Mastercard, & Turespaña. (2023). *Expectativas y comportamiento de los viajeros europeos en materia de sostenibilidad*. <https://www.tourspain.es/es/notas-prensa-turismo/espana-mejores-destinos-turisticos-viajeros-sostenibles/>
- Matten, D., & Moon, J. (2008). “Implicit” and “explicit” CSR: A conceptual framework for a comparative understanding of corporate social responsibility. *Academy of Management Review*, 33(2), 404–424.
- Mazanec, J. A., Wöber, K., & Zins, A. H. (2007). Tourism Destination Competitiveness: From Definition to Explanation? *Journal of Travel Research*, 46(1), 86–95. <https://doi.org/10.1177/0047287507302389>
- McCool, S. F. (1995). *Linking tourism, the environmental, and concepts of sustainability: Setting the stage* (S. F. McCool & A. E. Watson, Eds.). Gen. Tech. Rep. INNNT -GTR-323, USDA, Forest Service, Intermountain Research Station.
- McCullagh, P. (1980). Regression Models for Ordinal Data. *Journal of the Royal Statistical Society: Series B (Methodological)*, 42(2), 109–127. <https://doi.org/10.1111/J.2517-6161.1980.TB01109.X>
- McIntosh, R. W., Goeldner, C. R., & Ritchie, J. R. B. (1995). *Tourism: Principles, practices, philosophies* (7th Edition). John Wiley and Sons.
- McKinney, W. (2010). Data Structures for Statistical Computing in Python. *Proceedings of the 9th Python in Science Conference*, 51–56. <https://doi.org/10.25080/Majora-92bf1922-00a>
- Medrado, L., & Jackson, L. A. (2015). Corporate nonfinancial disclosures: An illuminating look at the corporate social responsibility and sustainability reporting practices of hospitality and tourism firms. *Tourism and Hospitality Research*, 16(2), 116–132. <https://doi.org/10.1177/1467358415600210>
- Meneses, R., Brito, C., Lopes, B., & Correia, R. (2023). Satisfaction and Dissatisfaction in Wine Tourism: A User-Generated Content Analysis. *Tourism and Hospitality Research*. <https://doi.org/10.1177/14673584231191989>

- Merli, R., Preziosi, M., Acampora, A., Lucchetti, M. C., & Ali, F. (2019). The impact of green practices in coastal tourism: An empirical investigation on an eco-labelled beach club. *International Journal of Hospitality Management*, 77, 471–482. <https://doi.org/10.1016/J.IJHM.2018.08.011>
- Mihalic, T. (2020). Conceptualising overtourism: A sustainability approach. *Annals of Tourism Research*, 84, 103025. <https://doi.org/10.1016/J.ANNALS.2020.103025>
- Mihalic, T. (2024). Metaversal sustainability: conceptualisation within the sustainable tourism paradigm. *Tourism Review*. <https://doi.org/10.1108/TR-09-2023-0609>
- Mihalič, T., Šegota, T., Knežević Cvelbar, L., & Kuščer, K. (2016). The influence of the political environment and destination governance on sustainable tourism development: a study of Bled, Slovenia. *Journal of Sustainable Tourism*, 24(11), 1489–1505. <https://doi.org/10.1080/09669582.2015.1134557>
- Miller, G., & Torres-Delgado, A. (2023). Measuring sustainable tourism: a state of the art review of sustainable tourism indicators. *Journal of Sustainable Tourism*, 31(7), 1483–1496. <https://doi.org/10.1080/09669582.2023.2213859>
- Miller, G., & Twining-Ward, L. (2005). *Monitoring for a sustainable tourism transition: the challenge of developing and using indicators*. CABI Publishing. <https://doi.org/10.1079/9780851990514.0000>
- Ministerio de Trabajo y Economía Social. (2022). *El portal de la Responsabilidad Social*. <https://www.mites.gob.es/es/rse/>
- Mior Shariffuddin, N. S., Azinuddin, M., Hanafiah, M. H., & Wan Mohd Zain, W. M. A. (2023). A comprehensive review on tourism destination competitiveness (TDC) literature. *Competitiveness Review*, 33(4), 787–819. <https://doi.org/10.1108/CR-04-2021-0054>
- Miralles-Quiros, M. del M., Miralles-Quiros, J. L., & Arraiano, I. G. (2017). Are Firms that Contribute to Sustainable Development Valued by Investors? *Corporate Social Responsibility and Environmental Management*, 24(1), 71–84. <https://doi.org/10.1002/csr.1392>
- Mishra, A., & Gupta, A. (2019). Green hotel servicescape: attributes and unique experiences. *Current Issues in Tourism*, 22(20), 2566–2578. <https://doi.org/10.1080/13683500.2018.1502259>

- Mitchell, R. (2018). *Web scraping with Python: Collecting more data from the modern web*. O'Reilly Media, Inc.
- Molina-Collado, A., Santos-Vijande, M. L., Gómez-Rico, M., & Madera, J. M. (2022). Sustainability in hospitality and tourism: a review of key research topics from 1994 to 2020. *International Journal of Contemporary Hospitality Management*, 34(8), 3029–3064. <https://doi.org/10.1108/IJCHM-10-2021-1305>
- Montiel, I. (2008). Corporate social responsibility and corporate sustainability: Separate pasts, common futures. *Organization & Environment*, 21(3), 245–269. <https://doi.org/https://doi.org/10.1177/1086026608321329>
- Mor, M., Dalyot, S., & Ram, Y. (2023). Who is a tourist? Classifying international urban tourists using machine learning. *Tourism Management*, 95, 104689. <https://doi.org/10.1016/J.TOURMAN.2022.104689>
- Mortenson, M. J., & Vidgen, R. (2016). A computational literature review of the technology acceptance model. *International Journal of Information Management*, 36(6), 1248–1259. <https://doi.org/10.1016/J.IJINFOMGT.2016.07.007>
- Moyle, B., Moyle, C. Lee, Ruhanen, L., Weaver, D., & Hadinejad, A. (2020). Are we really progressing sustainable tourism research? A bibliometric analysis. *Journal of Sustainable Tourism*, 29(1), 106–122. <https://doi.org/10.1080/09669582.2020.1817048>
- Muller, C. (1973). *Estadística lingüística*. Gredos S.A.
- Muñoz, P., & Branzei, O. (2021). Regenerative Organizations: Introduction to the Special Issue. *Organization & Environment*, 34(4), 507–516. <https://doi.org/10.1177/10860266211055740>
- Na, H. J., Lee, K. C., Choi, S. U., & Kim, S. T. (2020). Exploring CEO messages in sustainability management reports: Applying sentiment mining and sustainability balanced scorecard methods. *Sustainability (Switzerland)*, 12(2). <https://doi.org/10.3390/su12020590>
- Natural Language Toolkit*. (n.d.).
- Ng, S. I., Chia, K. W., Ho, J. A., & Ramachandran, S. (2017). Seeking tourism sustainability – A case study of Tioman Island, Malaysia. *Tourism Management*, 58, 101–107. <https://doi.org/10.1016/J.TOURMAN.2016.10.007>

- Nguyen Huu, T., Nguyen Ngoc, H., Nguyen Dai, L., Nguyen Thi Thu, D., Ngo Truc, L., Nguyen Trong, L., nguyen huu, tri, nguyen ngoc, han, nguyen Dai, loi, nguyen thi thu, D., ngo truc, ly, & nguyen trong, luan. (2024). Effect of tourist satisfaction on revisit intention in Can Tho City, Vietnam. *Cogent Business & Management*, 11(1), 2322779. <https://doi.org/10.1080/23311975.2024.2322779>
- Nikolay, A., Anindya, G., & Panagiotis, G. I. (2011). Deriving the Pricing Power of Product Features by Mining Consumer Reviews. *Management Science*, 57(8), 1485–1509. <https://doi.org/10.1287/MNSC.1110.1370>
- Niñerola, A., Sánchez-Rebull, M. V., & Hernández-Lara, A. B. (2019). Tourism Research on Sustainability: A Bibliometric Analysis. *Sustainability*, 11(5), 1377. <https://doi.org/10.3390/SU11051377>
- Ning, X., Yim, D., & Khuntia, J. (2021). Online sustainability reporting and firm performance: lessons learned from text mining. *Sustainability*, 13(3), 1–15. <https://doi.org/10.3390/su13031069>
- Nunkoo, R., Sharma, A., Rana, N. P., Dwivedi, Y. K., & Sunnassee, V. A. (2023). Advancing sustainable development goals through interdisciplinarity in sustainable tourism research. *Journal of Sustainable Tourism*, 31(3), 735–759. <https://doi.org/10.1080/09669582.2021.2004416>
- Nyahunzvi, D. K. (2013). CSR reporting among Zimbabwe’s hotel groups: a content analysis. *International Journal of Contemporary Hospitality Management*. <https://doi.org/https://doi-org.unican.idm.oclc.org/10.1108/09596111311322943>
- Octoparse. (2023). *Octoparse: A web scraping tool for data extraction*. <https://www.octoparse.com>
- Olson, E. L. (2013). It’s not easy being green: The effects of attribute tradeoffs on green product preference and choice. *Journal of the Academy of Marketing Science*, 41(2), 171–184. <https://doi.org/10.1007/S11747-012-0305-6>
- ONU. (2015, September 25). *La Asamblea General adopta la Agenda 2030 para el Desarrollo Sostenible*. Organización de Las Nacionales Unidas . <https://www.un.org/sustainabledevelopment/es/2015/09/la-asamblea-general-adopta-la-agenda-2030-para-el-desarrollo-sostenible/>

- Oroian, M., Raiu, F.-R., & Ghereș, M. (2015). Tourists' perception of the likely impact of implementing CSR practices in Romanian hotels. *Academica Science Journal Geographica Series, 1*, 2015.
- Othman, S., Darus, F., & Arshad, R. (2011). The influence of coercive isomorphism on corporate social responsibility reporting and reputation. *Social Responsibility Journal, 7*(1), 118–135. <https://doi.org/https://doi.org/10.1108/174711111111114585>
- Ottman, J. A., Stafford, E. R., & Hartman, C. L. (2006). Avoiding green marketing myopia: Ways to improve consumer appeal for environmentally preferable products. *Environment, 48*(5), 22–36. <https://doi.org/https://doi.org/10.3200/ENVT.48.5.22-36>
- Owais, S. S., & Hussein, N. S. (2016). Extract Five Categories CPIVW from the 9V's Characteristics of the Big Data. (*IJACSA*) *International Journal of Advanced Computer Science and Applications, 7*(3). [www.ijacsa.thesai.org](http://www.ijacsa.thesai.org)
- Palacios-Florencio, B., Santos-Roldán, L., Berbel-Pineda, J. M., & Castillo-Canalejo, A. M. (2021). Sustainable Tourism as a Driving force of the Tourism Industry in a Post-Covid-19 Scenario. *Social Indicators Research, 158*(3), 991–1011. <https://doi.org/10.1007/S11205-021-02735-2>
- Palazzo, G., & Scherer, A. G. (2006). Corporate legitimacy as deliberation: A communicative framework. *Journal of Business Ethics, 66*(1), 71–88. <https://doi.org/10.1007/s10551-006-9044-2>
- Panapanaan, V. M., Linnanen, L., Karvonen, M. M., & Phan, V. T. (2003). Roadmapping Corporate Social Responsibility in Finnish Companies. *Journal of Business Ethics, 44*(2–3), 133–148. <https://doi.org/10.1023/A:1023391530903>
- pandas* - *Python Data Analysis Library*. (n.d.). Último Acceso: Enero de 2023. <https://pandas.pydata.org/>
- Pande, K., & Shi, F. (2023). Managing visitor experience at religious heritage sites. *Journal of Destination Marketing & Management, 29*, 100800. <https://doi.org/10.1016/J.JDMM.2023.100800>
- Panimalar, A., Shree, V., & Veneshia, K. (2017). The 17 V's Of Big Data. *International Research Journal of Engineering and Technology*. [www.irjet.net](http://www.irjet.net)

- Payne, R. (1993). Sustainable tourism: Suggested indicators and monitoring techniques. In J. G. Nelson, R. W. Butler, & G. Wall (Eds.), *In Tourism and Sustainable Development: Monitoring, Planning, Managing* (pp. 249–254). University of Waterloo (Department of Geography Publication 37).
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830. <http://scikit-learn.sourceforge.net>.
- Penagos-Londoño, G. I., Rodríguez-Sánchez, C., Ruiz-Moreno, F., & Torres, E. (2021). A machine learning approach to segmentation of tourists based on perceived destination sustainability and trustworthiness. *Journal of Destination Marketing & Management*, *19*, 100532. <https://doi.org/10.1016/J.JDMM.2020.100532>
- Peng, N., & Chen, A. (2019). Luxury hotels going green – the antecedents and consequences of consumer hesitation. *Journal of Sustainable Tourism*, *27*(9), 1374–1392. <https://doi.org/10.1080/09669582.2019.1622710>
- Pennington, D., Ebert-Uphoff, I., Freed, N., Martin, J., & Pierce, S. A. (2020). Bridging sustainability science, earth science, and data science through interdisciplinary education. *Sustainability Science*, *15*(2), 647–661. <https://doi.org/10.1007/S11625-019-00735-3>
- Pérez Cañizares, P. (2021). “Corporate Sustainability” or “Corporate Social Responsibility”? A Comparative Study of Spanish and Latin American Companies’ Websites. *Business and Professional Communication Quarterly*, *84*(4), 361–385. <https://doi.org/10.1177/23294906211023799>
- Pizam, A., Neumann, Y., & Reichel, A. (1978). Dimensions of tourist satisfaction with a destination area. *Annals of Tourism Research*, *5*(3), 314–322. [https://doi.org/10.1016/0160-7383\(78\)90115-9](https://doi.org/10.1016/0160-7383(78)90115-9)
- Pizzi, S., Caputo, A., Corvino, A., & Venturelli, A. (2020). Management research and the UN sustainable development goals (SDGs): A bibliometric investigation and systematic review. *Journal of Cleaner Production*, *276*. <https://doi.org/10.1016/j.jclepro.2020.124033>

- Polewsky, M., Hankammer, S., Kleer, R., & Antons, D. (2024). Degrowth vs. Green Growth. A computational review and interdisciplinary research agenda. *Ecological Economics*, 217, 108067. <https://doi.org/10.1016/J.ECOLECON.2023.108067>
- Prakash, S. L., Perera, P., Newsome, D., Kusuminda, T., & Walker, O. (2019). Reasons for visitor dissatisfaction with wildlife tourism experiences at highly visited national parks in Sri Lanka. *Journal of Outdoor Recreation and Tourism*, 25, 102–112. <https://doi.org/10.1016/J.JORT.2018.07.004>
- Prerana, P., Kapoor, D., & Jain, A. (2024). Sustainable tourism and its future research directions: a bibliometric analysis of twenty-five years of research. *Tourism Review*, 79(3), 541–567. <https://doi.org/10.1108/TR-11-2022-0540>
- Prud'homme, B., & Raymond, L. (2013). Sustainable development practices in the hospitality industry: An empirical study of their impact on customer satisfaction and intentions. *International Journal of Hospitality Management*, 34(1), 116–126. <https://doi.org/10.1016/J.IJHM.2013.03.003>
- Python Software Foundation. (2023a). *os* — *Miscellaneous operating system interfaces*. <https://docs.python.org/3/library/os.html>
- Python Software Foundation. (2023b). *Python programming language*. <https://www.python.org/>
- Python Software Foundation. (2023c). *re* — *Regular expression operations*. <https://docs.python.org/3/library/re.html>
- Rahmadian, E., Feitosa, D., & Zwitter, A. (2022). A systematic literature review on the use of big data for sustainable tourism. *Current Issues in Tourism*, 25(11), 1711–1730. <https://doi.org/10.1080/13683500.2021.1974358>
- Rajeev, A., Pati, R. K., Padhi, S. S., & Govindan, K. (2017). Evolution of sustainability in supply chain management: A literature review. *Journal of Cleaner Production*, 162, 299–314. <https://doi.org/10.1016/J.JCLEPRO.2017.05.026>
- Ramasamy, D., Sarasua, C., Bacchelli, A., & Bernstein, A. (2023). Workflow analysis of data science code in public GitHub repositories. *Empirical Software Engineering*, 28(1), 1–47. <https://doi.org/10.1007/S10664-022-10229-Z>

- Rasoolimanesh, S. M., Noor, S. M., & Jaafar, M. (2019). Positive and Negative Perceptions of Residents Toward Tourism Development: Formative or Reflective. *Perspectives on Asian Tourism, Part F184*, 247–271. [https://doi.org/10.1007/978-981-13-2463-5\\_12](https://doi.org/10.1007/978-981-13-2463-5_12)
- Rasoolimanesh, S. M., Ramakrishna, S., Hall, C. M., Esfandiar, K., & Seyfi, S. (2023). A systematic scoping review of sustainable tourism indicators in relation to the sustainable development goals. *Journal of Sustainable Tourism*, 31(7), 1497–1517. <https://doi.org/10.1080/09669582.2020.1775621>
- Rasoolimanesh, S. M., Ringle, C. M., Jaafar, M., & Ramayah, T. (2017). Urban vs. rural destinations: Residents' perceptions, community participation and support for tourism development. *Tourism Management*, 60, 147–158. <https://doi.org/10.1016/J.TOURMAN.2016.11.019>
- Recognai. (2021). *Model: bert-base-spanish-wwm-cased-xnli*. <https://huggingface.co/Recognai/bert-base-spanish-wwm-cased-xnli>
- Retolaza, J. L., & San-Jose, L. (2021). Understanding Social Accounting Based on Evidence. *SAGE Open*, 11(2). <https://doi.org/10.1177/21582440211003865>
- Reverte, C. (2009). Determinants of corporate social responsibility disclosure ratings by Spanish listed firms. *Journal of Business Ethics*, 88(2), 351–366. <https://doi.org/10.1007/s10551-008-9968-9>
- Rey-Moreno, M., Sánchez-Franco, M. J., De La, M., & Rey-Tienda, S. (2023). Examining transaction-specific satisfaction and trust in Airbnb and hotels. An application of BERTopic and Zero-shot text classification. *Tourism & Management Studies*, ISSN 2182-8458, ISSN-e 2182-8466, Vol. 19, No. 2, 2023, Págs. 21-37, 19(2), 21–37. <https://doi.org/10.18089/tms.2023.190202>
- Rhou, Y., Singal, M., & Koh, Y. (2016). CSR and financial performance: The role of CSR awareness in the restaurant industry. *International Journal of Hospitality Management*, 57, 30–39. <https://doi.org/10.1016/J.IJHM.2016.05.007>
- Richards, G. (2018). Cultural tourism: A review of recent research and trends. *Journal of Hospitality and Tourism Management*, 36, 12–21. <https://doi.org/10.1016/J.JHTM.2018.03.005>

- Ristić, V., Trišić, I., Štetić, S., Maksin, M., Nechita, F., Candrea, A. N., Pavlović, M., & Hertanu, A. (2024). Institutional, Ecological, Economic, and Socio-Cultural Sustainability—Evidence from Ponjavica Nature Park. *Land*, *13*(5), 669. <https://doi.org/10.3390/LAND13050669>
- Rita, P., Ramos, R., Borges-Tiago, M. T., & Rodrigues, D. (2022). Impact of the rating system on sentiment and tone of voice: A Booking.com and TripAdvisor comparison study. *International Journal of Hospitality Management*, *104*, 103245. <https://doi.org/10.1016/J.IJHM.2022.103245>
- Ritchie, B. J. R., & Crouch, G. I. (2003). *The competitive destination: A sustainable tourism perspective*. CABI Publishing.
- Rockey, D. Jr., & Lankford, S. (2012). Attitudes of local residents to recreation and tourism development of Fredericksburg/Spotsylvania: A case study. *E-Review of Tourism Research*, *10*(1).
- Rodríguez-Antón, J. M., Del Mar Alonso-Almeida, M., Celemín, M. S., & Rubio, L. (2012). Use of different sustainability management systems in the hospitality industry. The case of Spanish hotels. *Journal of Cleaner Production*, *22*(1), 76–84. <https://doi.org/10.1016/J.JCLEPRO.2011.09.024>
- Rosato, P. F., Caputo, A., Valente, D., & Pizzi, S. (2021). 2030 Agenda and sustainable business models in tourism: A bibliometric analysis. In *Ecological Indicators* (Vol. 121). Elsevier B.V. <https://doi.org/10.1016/j.ecolind.2020.106978>
- Rosenow, J. E., & Pulsipher, G. L. (1979). *Tourism: The Good, the Bad and the Ugly*. Media Productions and Marketing.
- Ruggerio, C. A. (2021). Sustainability and sustainable development: A review of principles and definitions. *Science of The Total Environment*, *786*, 147481. <https://doi.org/10.1016/J.SCITOTENV.2021.147481>
- Ruhanen, L. (2013). Local government: facilitator or inhibitor of sustainable tourism development? *Journal of Sustainable Tourism*, *21*(1), 80–98. <https://doi.org/10.1080/09669582.2012.680463>
- Ruhanen, L., Weiler, B., Moyle, B. D., & McLennan, C. lee J. (2015). Trends and patterns in sustainable tourism research: a 25-year bibliometric analysis. *Journal of Sustainable Tourism*, *23*(4), 517–535. <https://doi.org/10.1080/09669582.2014.978790>

- Russell, S. J., & Norvig, P. (2021). *Artificial intelligence: A modern approach*. Pearson.
- SABI. (2022). *Sistema de Análisis de Balances Ibéricos*. Último Acceso a La Web. <http://sabi.bvdep.com>
- Saeed, N., & Husamaldin, L. (2021). Big Data Characteristics (V's) in Industry. *Iraqi Journal of Industrial Research*, 8(1), 1–9. <https://doi.org/10.53523/ijoirvol8i1id52>
- Sahu, A. K., Debata, B., & Dash, S. R. (2024). Manager sentiment, policy uncertainty, ESG disclosure and firm performance: a large language model in corporate landscape. *International Journal of Accounting and Information Management*, 32(5), 858–882. <https://doi.org/10.1108/IJAIM-08-2023-0206>
- Salee, A., Mostafa Rasoolimanesh, S., & May Li Kusumo, C. (2022). Tourism Impacts on Destinations: Insights from a Systematic Review of Literature. *Asia-Pacific Journal of Innovation in Hospitality and Tourism*, 11(1), 121–148.
- Salinas, C. M., & Martínez, G. E. S. (2015). Hacia una normalización de la frecuencia de los corpus CREA y CORDE. *Revista Signos*, 48(89), 307–331. <https://doi.org/10.4067/S0718-09342015000300002>
- Sánchez-Camacho, C., Carranza, R., Martín-Consuegra, D., & Díaz, E. (2022). Evolution, trends and future research lines in corporate social responsibility and tourism: A bibliometric analysis and science mapping. *Sustainable Development*, 30(3), 462–476. <https://doi.org/10.1002/sd.2260>
- Sanchez-Fernandez, R., Iniesta-Bonillo, M. A., & Cervera-Taulet, A. (2016). Environmental sustainability in the Mediterranean destinations: a latent class segmentation analysis. *Environmental Engineering & Management Journal (EEMJ)*, 7.
- Sánchez-Fernández, R., Iniesta-Bonillo, M. Á., & Cervera-Taulet, A. (2019). Exploring the concept of perceived sustainability at tourist destinations: a market segmentation approach. *Journal of Travel & Tourism Marketing*, 36(2), 176–190. <https://doi.org/10.1080/10548408.2018.1505579>
- Sann, R., & Lai, P. C. (2020). Understanding homophily of service failure within the hotel guest cycle: Applying NLP-aspect-based sentiment analysis to the hospitality industry. *International Journal of Hospitality Management*, 91, 102678. <https://doi.org/10.1016/J.IJHM.2020.102678>

- Santos, M. C., Veiga, C., Santos, J. A. C., & Águas, P. (2022). Sustainability as a success factor for tourism destinations: a systematic literature review. *Worldwide Hospitality and Tourism Themes*, 14(1), 20–37. <https://doi.org/10.1108/WHATT-10-2021-0139>
- Sautter, E. T., & Leisen, B. (1999). Managing stakeholders a Tourism Planning Model. *Annals of Tourism Research*, 26(2), 312–328. [https://doi.org/10.1016/S0160-7383\(98\)00097-8](https://doi.org/10.1016/S0160-7383(98)00097-8)
- Schaltegger, S., & Dyllick, T. (2002). *Nachhaltig managen mit der Balanced Scorecard*. Wiesbaden: Gabler.
- Scherer, A. G., & Palazzo, G. (2011). The New Political Role of Business in a Globalized World: A Review of a New Perspective on CSR and its Implications for the Firm, Governance, and Democracy. *Journal of Management Studies*, 48(4), 899–931. <https://doi.org/10.1111/J.1467-6486.2010.00950.X>
- Scheyvens, R. (1999). Ecotourism and the empowerment of local communities. *Tourism Management*, 20(2), 245–249. [https://doi.org/10.1016/S0261-5177\(98\)00069-7](https://doi.org/10.1016/S0261-5177(98)00069-7)
- Scheyvens, R. (2002). *Tourism for Development: Empowering Communitie*. Prentice Hall.
- Schilling, M., & Chiang, L. (2011). The effect of natural resources on a sustainable development policy: The approach of non-sustainable externalities. *Energy Policy*, 39(2), 990–998. <https://doi.org/10.1016/j.enpol.2010.11.030>
- Schuckert, M., Liu, X., & Law, R. (2015). A segmentation of online reviews by language groups: How English and non-English speakers rate hotels differently. *International Journal of Hospitality Management*, 48, 143–149. <https://doi.org/10.1016/J.IJHM.2014.12.007>
- Schuler, D. A., & Cording, M. (2006). A Corporate Social Performance–Corporate Financial Performance Behavioral Model for Consumers. *Academy of Management Review*, 31(3), 540–558. <https://doi.org/https://doi.org/10.5465/amr.2006.21318916>
- Scott, D., Gössling, S., & Hall, C. M. (2012). International tourism and climate change. *WIREs Climate Change*, 3(3), 213–232. <https://doi.org/10.1002/WCC.165>
- Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and Statistical Modeling with Python. *Proceedings of the 9th Python in Science Conference*. <https://doi.org/10.25080/Majora-92bf1922-011>

- Serra, L. M., Lozano, M. A., Ramos, J., Ensinas, A. V., & Nebra, S. A. (2009). Polygeneration and efficient use of natural resources. *Energy*, 34(5), 575–586. <https://doi.org/10.1016/j.energy.2008.08.013>
- Serrano, L., Sianes, A., & Ariza-Montes, A. (2019). Using Bibliometric Methods to Shed Light on the Concept of Sustainable Tourism. *Sustainability*, 11(24), 6964. <https://doi.org/10.3390/SU11246964>
- Sharma, A., Nunkoo, R., Rana, N. P., & Dwivedi, Y. K. (2021). On the intellectual structure and influence of tourism social science research. *Annals of Tourism Research*, 91, 103142. <https://doi.org/10.1016/J.ANNALS.2021.103142>
- Sharpley, R. (2000). Tourism and Sustainable Development: Exploring the Theoretical Divide. *Journal of Sustainable Tourism*, 8(1), 1–19. <https://doi.org/10.1080/09669580008667346>
- Sharpley, R. (2020). Tourism, sustainable development and the theoretical divide: 20 years on. *Journal of Sustainable Tourism*, 28(11), 1932–1946. <https://doi.org/10.1080/09669582.2020.1779732>
- Sheehy, B. (2014). Defining CSR: Problems and Solutions. *Journal of Business Ethics*, 131(3), 625–648. <https://doi.org/10.1007/S10551-014-2281-X>
- Sheehy, B., & Farneti, F. (2021). Corporate social responsibility, sustainability, sustainable development and corporate sustainability: What is the difference, and does it matter? *Sustainability*, 13(11). <https://doi.org/10.3390/su13115965>
- Shen, F., & Cottrell, S. P. (2008). A sustainable tourism framework for monitoring residents' satisfaction with agritourism in Chongdugou Village, China. *International Journal of Tourism Policy*, 1(4), 368–375. <https://doi.org/10.1504/IJTP.2008.019277>
- Shoval, N., & Ahas, R. (2016). The use of tracking technologies in tourism research: the first decade. *Tourism Geographies*, 18(5), 587–606. <https://doi.org/10.1080/14616688.2016.1214977>
- Sims, R. (2009). Food, place and authenticity: local food and the sustainable tourism experience. *Journal of Sustainable Tourism*, 17(3), 321–336. <https://doi.org/10.1080/09669580802359293>

- Sinha, A., Kedas, S., Kumar, R., & Malo, P. (2022). SEntFiN 1.0: Entity-aware sentiment analysis for financial news. *Journal of the Association for Information Science and Technology*, 73(9), 1314–1335. <https://doi.org/10.1002/ASI.24634>
- Sparks, B. A., & Browning, V. (2011). The impact of online reviews on hotel booking intentions and perception of trust. *Tourism Management*, 32(6), 1310–1323. <https://doi.org/10.1016/J.TOURMAN.2010.12.011>
- Sreen, N., Mukherjee, S., Jebarajakirthy, C., Kumar, S., & Sharma, H. (2022). Role of personal values on revisiting green hotels: a hierarchical value map approach. *Tourism Recreation Research*. <https://doi.org/10.1080/02508281.2022.2076190>
- Srijiranon, K., Lertratanakham, Y., & Tanantong, T. (2022). A Hybrid Framework Using PCA, EMD and LSTM Methods for Stock Market Price Prediction with Sentiment Analysis. *Applied Sciences*, 12(21), 10823. <https://doi.org/10.3390/APP122110823>
- Stankey, G. H. (1973). Visitor perception of wilderness recreation carrying capacity. In *U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station* (Research Paper (Res. Pap. INT-RP-142); Vol. 142). <https://research.fs.usda.gov/treesearch/56243>
- StataCorp. (2023). *Stata Statistical Software*. <https://www.stata.com/>
- Steurer, R., Langer, M. E., Konrad, A., & Martinuzzi, A. (2005). Corporations, stakeholders and sustainable development I: A theoretical exploration of business-society relations. *Journal of Business Ethics*, 61(3), 263–281. <https://doi.org/10.1007/S10551-005-7054-0>
- Strand, R., Freeman, R. E., & Hockerts, K. (2015). Corporate Social Responsibility and Sustainability in Scandinavia: An Overview. *Journal of Business Ethics*, 127(1), 1–15. <https://doi.org/10.1007/S10551-014-2224-6>
- Streimikiene, D., Svagzdiene, B., Jasinskas, E., & Simanavicius, A. (2021). Sustainable tourism development and competitiveness: The systematic literature review. In *Sustainable Development* (Vol. 29, Issue 1, pp. 259–271). John Wiley and Sons Ltd. <https://doi.org/10.1002/sd.2133>
- Stylos, N., & Vassiliadis, C. (2015). Differences in Sustainable Management Between Four- and Five-Star Hotels Regarding the Perceptions of Three-Pillar Sustainability. *Journal of*

- Hospitality Marketing & Management*, 24(8), 791–825.  
<https://doi.org/10.1080/19368623.2015.955622>
- Sudhir, A., & Amartya, S. (2000). Human Development and Economic Sustainability. *World Development*, 28(12), 2029–2049. [https://doi.org/10.1016/S0305-750X\(00\)00071-1](https://doi.org/10.1016/S0305-750X(00)00071-1)
- Suhartanto, D. (2018). Tourist satisfaction with souvenir shopping: evidence from Indonesian domestic tourists. *Current Issues in Tourism*, 21(6), 663–679.  
<https://doi.org/10.1080/13683500.2016.1265487>
- Swarbrooke, J. (1999). Sustainable tourism management. In *Cabi*. CABI Publishing.
- Szekely, N., & Vom Brocke, J. (2017). What can we learn from corporate sustainability reporting? Deriving propositions for research and practice from over 9,500 corporate sustainability reports published between 1999 and 2015 using topic modelling technique. *PLoS ONE*, 12(4), 1–27. <https://doi.org/10.1371/journal.pone.0174807>
- Tian, Y. (2013). Engagement in online hotel reviews: A comparative study. *Discourse, Context & Media*, 2(4), 184–191. <https://doi.org/10.1016/J.DCM.2013.10.002>
- Tivy, J. (1973). *The Concept and Determination of Carrying Capacity of Recreational Land in the USA* (Occasional Paper Number 3).
- Tölkes, C. (2018). Sustainability communication in tourism – A literature review. *Tourism Management Perspectives*, 27, 10–21. <https://doi.org/10.1016/J.TMP.2018.04.002>
- Torelli, R., Balluchi, F., & Furlotti, K. (2020). The materiality assessment and stakeholder engagement: A content analysis of sustainability reports. *Corporate Social Responsibility and Environmental Management*, 27(2), 470–484. <https://doi.org/10.1002/csr.1813>
- Torres-Delgado, A., López Palomeque, F., Elorrieta Sanz, B., & Font Urgell, X. (2023). Monitoring sustainable management in local tourist destinations: performance, drivers and barriers. *Journal of Sustainable Tourism*, 31(7), 1672–1693.  
<https://doi.org/10.1080/09669582.2021.1937190>
- Trang, H. L. T., Lee, J. S., & Han, H. (2019). How do green attributes elicit pro-environmental behaviors in guests? The case of green hotels in Vietnam. *Journal of Travel & Tourism Marketing*, 36(1), 14–28. <https://doi.org/10.1080/10548408.2018.1486782>

- Trišić, I., Štetić, S., Candrea, A. N., Nechita, F., Apetrei, M., Pavlović, M., Stojanović, T., & Perić, M. (2024). The Impact of Sustainable Tourism on Resident and Visitor Satisfaction—The Case of the Special Nature Reserve “Titelski Breg”, Vojvodina. *Sustainability*, *16*(7), 2720. <https://doi.org/10.3390/SU16072720>
- Tsaur, S. H., Lin, Y. C., & Lin, J. H. (2006). Evaluating ecotourism sustainability from the integrated perspective of resource, community and tourism. *Tourism Management*, *27*(4), 640–653. <https://doi.org/10.1016/J.TOURMAN.2005.02.006>
- Tukey, J. W. (1962). The Future of Data Analysis. *The Annals of Mathematical Statistics*, *33*(1), 1–67. <https://doi.org/10.1214/AOMS/1177704711>
- UCLA. (2024a). *FAQ: What are Pseudo R-squareds?* <https://stats.oarc.ucla.edu/other/mult-pkg/faq/general/faq-what-are-pseudo-r-squareds/> Accessed 4 March 2024
- UCLA. (2024b). *How do I interpret the coefficients in an ordinal logistic regression in stata? | Stata FAQ.* <https://stats.oarc.ucla.edu/stata/faq/ologit-coefficients/> Accessed 1 March 2024
- UNEP. (2005). *Making Tourism More Sustainable - A Guide for Policy Makers*. Books; World Tourism Organization (UNWTO). <https://doi.org/10.18111/9789284408214>
- UNEP, & UNWTO. (2005). *Making Tourism More Sustainable-A Guide for Policy Makers (English version)*. <https://www.e-unwto.org/doi/abs/10.18111/9789284408214>
- UNESCO. (2022). *Tesaurus de la UNESCO*. <https://vocabularies.unesco.org/>
- Unlu, A., Truong, S., Sawhney, N., Sivelä, J., & Tammi, T. (2024). Tracing the dynamics of misinformation and vaccine stance in Finland amid COVID-19. *Information, Communication & Society*. <https://doi.org/10.1080/1369118X.2024.2331756>
- Unlu, A., Truong, S., Tammi, T., & Lohiniva, A. L. (2023). Exploring Political Mistrust in Pandemic Risk Communication: Mixed-Method Study Using Social Media Data Analysis. *Journal of Medical Internet Research*, *25*(1), e50199. <https://doi.org/10.2196/50199>
- Uyar, A., Köseoglu, M. A., Kılıç Karamahmutoğlu, M., & Mehraliyev, F. (2020). Thematic structure of sustainability reports of the hospitality and tourism sector: A periodical, regional, and format-based analysis. *Current Issues in Tourism*, *24*(18), 2602–2627. <https://doi.org/https://doi.org/10.1080/13683500.2020.1847050>

- Väisänen, H. M., Uusitalo, O., & Rynnänen, T. (2023). Towards sustainable servicescape – tourists' perspectives of accommodation service attributes. *International Journal of Hospitality Management*, *110*, 103449. <https://doi.org/10.1016/J.IJHM.2023.103449>
- Valbuena-Hernandez, J. P., & Ortiz-de-Mandojana, N. (2022). Encouraging corporate sustainability through effective strategic partnerships. *Corporate Social Responsibility and Environmental Management*, *29*(1), 124–134. <https://doi.org/10.1002/csr.2188>
- van der Zee, E., Bertocchi, D., & Vanneste, D. (2020). Distribution of tourists within urban heritage destinations: a hot spot/cold spot analysis of TripAdvisor data as support for destination management. *Current Issues in Tourism*, *23*(2), 175–196. <https://doi.org/10.1080/13683500.2018.1491955>
- van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538. <https://doi.org/10.1007/S11192-009-0146-3>
- Van Marrewijk, M. (2003). Concepts and definitions of CSR and corporate sustainability: Between agency and communion. *Journal of Business Ethics*, *44*(2), 95–105. <https://doi.org/https://doi.org/10.1023/A:1023331212247>
- Varian, H. R. (2014). Big Data: New Tricks for Econometrics. *Journal of Economic Perspectives*, *28*(2), 3–28. <https://doi.org/10.1257/JEP.28.2.3>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention Is All You Need. *ArXiv*, *2017-December*, 5999–6009. <https://doi.org/https://doi.org/10.48550/arXiv.1706.03762>
- Verma, V. K., & Chandra, B. (2018). An application of theory of planned behavior to predict young Indian consumers' green hotel visit intention. *Journal of Cleaner Production*, *172*, 1152–1162. <https://doi.org/10.1016/J.JCLEPRO.2017.10.047>
- Vickery, G., & Wunsch-Vincent, S. (2007). Participative web and user-created content: Web 2.0, wikis and social networking. In *Participative web and user-created content: Web 2.0*. Organisation for Economic Cooperation and Development (OECD). <https://doi.org/10.1787/9789264037472-EN>

- Vidal, I. (2011). *Sostenibilidad y Responsabilidad Social en la empresa*. [https://ideas.llorenteycuencia.com/download/Informes/110922\\_d+iLL&C\\_CIES\\_Informe\\_Especial\\_RSE.pdf](https://ideas.llorenteycuencia.com/download/Informes/110922_d+iLL&C_CIES_Informe_Especial_RSE.pdf)
- Vieira, I., Rodrigues, A., Fernandes, D., & Pires, C. (2016). The role of local government management of tourism in fostering residents' support to sustainable tourism development: Evidence from a Portuguese historic town. *International Journal of Tourism Policy*, 6(2), 109–135. <https://doi.org/10.1504/IJTP.2016.077967>
- Villafán Vidales, K. B. (2020). Evolución conceptual de la Responsabilidad Social. *Revista de La Facultad de Contaduría y Ciencias Administrativas*, 5(10), 69–80. <https://www.researchgate.net/publication/349324089>
- Villena, E. G., Villar, S. G., Dzul López, L. A., Álvarez, R. M., Noya, I. D., & Vidal Mazón, J. L. (2021). Approach to a project framework in the environment of sustainability and corporate social responsibility (CSR): Case study of a training proposal to a group of students in a higher education institution. *Sustainability*, 13(19). <https://doi.org/10.3390/su131910880>
- Viñán-Ludeña, M. S., & de Campos, L. M. (2022). Discovering a tourism destination with social media data: BERT-based sentiment analysis. *Journal of Hospitality and Tourism Technology*, 13(5), 907–921. <https://doi.org/10.1108/JHTT-09-2021-0259>
- Vives Varela, T., & Hamui Sutton, L. (2021). La codificación y categorización en la teoría fundamentada, un método para el análisis de los datos cualitativos. *Investigación En Educación Médica*, 10(40), 97–104. <https://doi.org/10.22201/fm.20075057e.2021.40.21367>
- Wall, G., & Mathieson, A. (2006). *Tourism: change, impacts, and opportunities*. Pearson Education.
- Wang, S., Wang, J., Li, J., & Yang, F. (2020). Do motivations contribute to local residents' engagement in pro-environmental behaviors? Resident-destination relationship and pro-environmental climate perspective. *Journal of Sustainable Tourism*, 28(6), 834–852. <https://doi.org/10.1080/09669582.2019.1707215>
- Wang, T. L., Tran, P. T. K., & Tran, V. T. (2017). Destination perceived quality, tourist satisfaction and word-of-mouth. *Tourism Review*, 72(4), 392–410. <https://doi.org/10.1108/TR-06-2017-0103>

- Waskom, M. L. (2021). seaborn: statistical data visualization. *Journal of Open Source Software*, 6(60), 3021. <https://doi.org/10.21105/JOSS.03021>
- WCED. (1987). *Our Common Future: a Report from the United Nations World Commission on Environment and Development*. <http://www.un-documents.net/our-common-future.pdf>,
- Weaver, D. (2007). Sustainable tourism. In *Sustainable Tourism*. Routledge. <https://doi.org/10.4324/9780080474526>
- Weber, M. (1992). *The theory of social and economic organization*. Simon and Schuster.
- Wheeller, B. (2007). Sustainable Mass Tourism: More Smudge than Nudge The Canard Continues. *Tourism Recreation Research*, 32(3), 73–75. <https://doi.org/10.1080/02508281.2007.11081543>
- Williams, R. (2006). Generalized Ordered Logit/Partial Proportional Odds Models for Ordinal Dependent Variables. <https://doi.org/10.1177/1536867X0600600104>, 6(1), 58–82. <https://doi.org/10.1177/1536867X0600600104>
- Witten I., Frank, E., & Hall, M. A. (2011). *Data Mining: Practical Machine Learning Tools and Techniques. 3rd Edition* (Morgan Kaufmann). Burlington.
- Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., Von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Le Scao, T., Gugger, S., ... Rush, A. M. (2020). Transformers: State-of-the-Art Natural Language Processing. *EMNLP 2020 - Conference on Empirical Methods in Natural Language Processing, Proceedings of Systems Demonstrations*, 38–45. <https://doi.org/10.18653/v1/2020.EMNLP-DEMOS.6>
- Woodley, S. (1993). Tourism and sustainable development in parks and protected areas. In J. G. Nelson, R. W. Butler, & G. Wall (Eds.), *In Tourism and Sustainable Development: Monitoring, Planning, Managing* (pp. 83–96). University of Waterloo (Department of Geography Publication 37).
- World Economic Forum. (2024). *Travel & Tourism Development Index 2024*. [https://www3.weforum.org/docs/WEF\\_Travel\\_and\\_Tourism\\_Development\\_Index\\_2024.pdf](https://www3.weforum.org/docs/WEF_Travel_and_Tourism_Development_Index_2024.pdf)

- World Tourism Organization. (1993). *Desarrollo turístico sostenible: Guía para planificadores locales*. <https://www.e-unwto.org/doi/book/10.18111/9789284401321>
- World Tourism Organization. (2019a). UNWTO Tourism Definitions. In *Books*. World Tourism Organization (UNWTO). <https://doi.org/10.18111/9789284420858>
- World Tourism Organization. (2019b). *UNWTO Tourism Definitions*. UNWTO, Madrid. <https://doi.org/https://doi.org/10.18111/9789284420858>
- World Tourism Organization. (2024). *Rural tourism*. <https://www.unwto.org/rural-tourism>  
Accessed 1 Feb 2024
- Wu, J., & Yang, T. (2023). Service attributes for sustainable rural tourism from online comments: Tourist satisfaction perspective. *Journal of Destination Marketing & Management*, 30, 100822. <https://doi.org/10.1016/J.JDMM.2023.100822>
- Wu, S., & Dredze, M. (2019). Beto, Bentz, Becas: The Surprising Cross-Lingual Effectiveness of BERT. *ArXiv*, 833–844. <https://doi.org/10.18653/v1/d19-1077>
- Wut, T. M., Xu, J. (Bill), & Wong, S. M. H. (2023). Timing and congruence effects of corporate social responsibility practices on social media crises in the tourism industry. *Journal of Hospitality and Tourism Technology*, 14(2), 154–171. <https://doi.org/10.1108/JHTT-02-2022-0043>
- Xia, H., Muskat, B., Karl, M., Li, G., & Law, R. (2024). Destination competitiveness research over the past three decades: a computational literature review using topic modelling. *Journal of Travel & Tourism Marketing*, 41(5), 726–742. <https://doi.org/10.1080/10548408.2024.2332278>
- Xiang, Z., Du, Q., Ma, Y., & Fan, W. (2017). A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. *Tourism Management*, 58, 51–65. <https://doi.org/10.1016/J.TOURMAN.2016.10.001>
- Xiang, Z., & Gretzel, U. (2010). Role of social media in online travel information search. *Tourism Management*, 31(2), 179–188. <https://doi.org/10.1016/J.TOURMAN.2009.02.016>
- Xiang, Z., Schwartz, Z., Gerdes, J. H., & Uysal, M. (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction? *International Journal of Hospitality Management*, 44, 120–130. <https://doi.org/10.1016/J.IJHM.2014.10.013>

- Xiang, Z., Schwartz, Z., & Uysal, M. (2015). What Types of Hotels Make Their Guests (Un)Happy? Text Analytics of Customer Experiences in Online Reviews. *Information and Communication Technologies in Tourism 2015*, 33–45. [https://doi.org/10.1007/978-3-319-14343-9\\_3](https://doi.org/10.1007/978-3-319-14343-9_3)
- Xiao, Q., & Ilnaini, B. (2023). Stock trend prediction using sentiment analysis. *PeerJ Computer Science*, 9, e1293. <https://doi.org/10.7717/PEERJ-CS.1293/SUPP-2>
- Xie, J., & Jain, T. (2024). Unpacking micro-CSR through a computational literature review: An identity heterogeneity view of internal stakeholders. *Journal of Business Research*, 172, 114451. <https://doi.org/10.1016/J.JBUSRES.2023.114451>
- Xie, K. L., Zhang, Z., & Zhang, Z. (2014). The business value of online consumer reviews and management response to hotel performance. *International Journal of Hospitality Management*, 43, 1–12. <https://doi.org/10.1016/J.IJHM.2014.07.007>
- Yang, X., Pan, B., Evans, J. A., & Lv, B. (2015). Forecasting Chinese tourist volume with search engine data. *Tourism Management*, 46, 386–397. <https://doi.org/10.1016/J.TOURMAN.2014.07.019>
- Ye, Q., Li, H., Wang, Z., & Law, R. (2014). The Influence of Hotel Price on Perceived Service Quality and Value in E-Tourism: An Empirical Investigation Based on Online Traveler Reviews. *Journal of Hospitality and Tourism Research*, 38(1), 23–39. <https://doi.org/10.1177/1096348012442540>
- Yin, Y., Zou, Y., Zhang, J., Wang, R., & Zhang, H. (2022). Temporarily returning to an ideal life: Research on Chinese rural tourists' home experience. *Tourism Management Perspectives*, 41, 100915. <https://doi.org/10.1016/J.TMP.2021.100915>
- Yoopetch, C., & Nimsai, S. (2019). Science Mapping the Knowledge Base on Sustainable Tourism Development, 1990–2018. *Sustainability*, 11(13), 3631. <https://doi.org/10.3390/SU11133631>
- Yoopetch, C., Nimsai, S., & Kongarchapatara, B. (2023). Bibliometric Analysis of Corporate Social Responsibility in Tourism. *Sustainability*, 15(1). <https://doi.org/10.3390/su15010668>
- Yu, Y., Li, X., & Jai, T. M. (Catherine). (2017). The impact of green experience on customer satisfaction: evidence from TripAdvisor. *International Journal of Contemporary Hospitality Management*, 29(5), 1340–1361. <https://doi.org/10.1108/IJCHM-07-2015-0371>

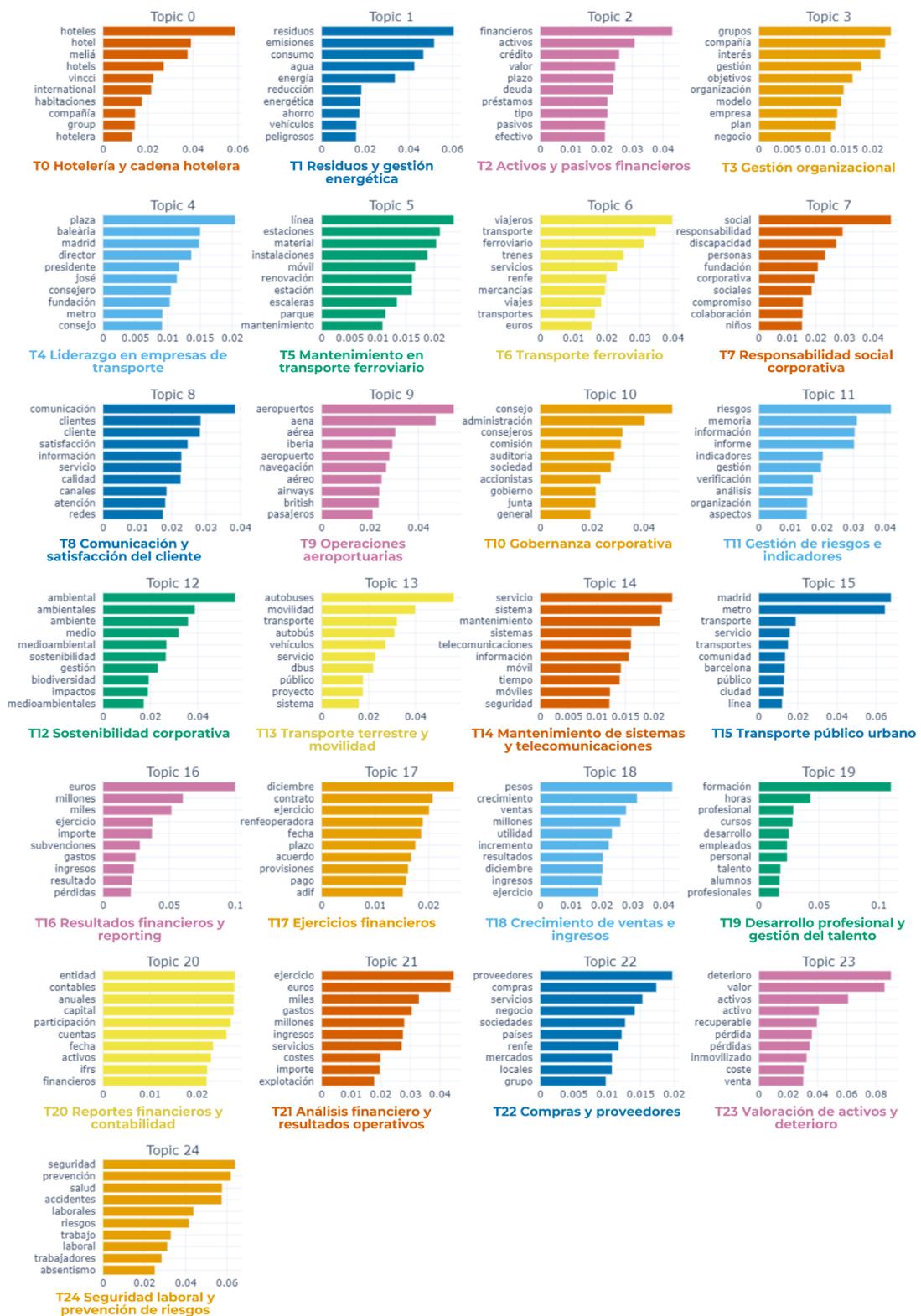
- Zhang, H., Liu, R., & Egger, R. (2023). Unlocking Uniqueness: Analyzing Online Reviews of Airbnb Experiences Using BERT-based Models. *Journal of Travel Research*. <https://doi.org/10.1177/00472875231197381>
- Zhang, S., & Chan, E. S. W. (2020). A modernism-based interpretation of sustainable tourism. *International Journal of Tourism Research*, 22(2), 223–237. <https://doi.org/10.1002/JTR.2330>
- Zhang, Z., Zhang, Z., & Yang, Y. (2016). The power of expert identity: How website-recognized expert reviews influence travelers' online rating behavior. *Tourism Management*, 55, 15–24. <https://doi.org/10.1016/J.TOURMAN.2016.01.004>
- Zhou, L., Ye, S., Pearce, P. L., & Wu, M. Y. (2014). Refreshing hotel satisfaction studies by reconfiguring customer review data. *International Journal of Hospitality Management*, 38, 1–10. <https://doi.org/10.1016/J.IJHM.2013.12.004>
- Zhou, W., Faturay, F., Driml, S., & Sun, Y. Y. (2024). Meta-analysis of the climate change-tourism demand relationship. *Journal of Sustainable Tourism*, 32(9), 1762–1783. <https://doi.org/10.1080/09669582.2024.2354882>



## **ANEXOS**

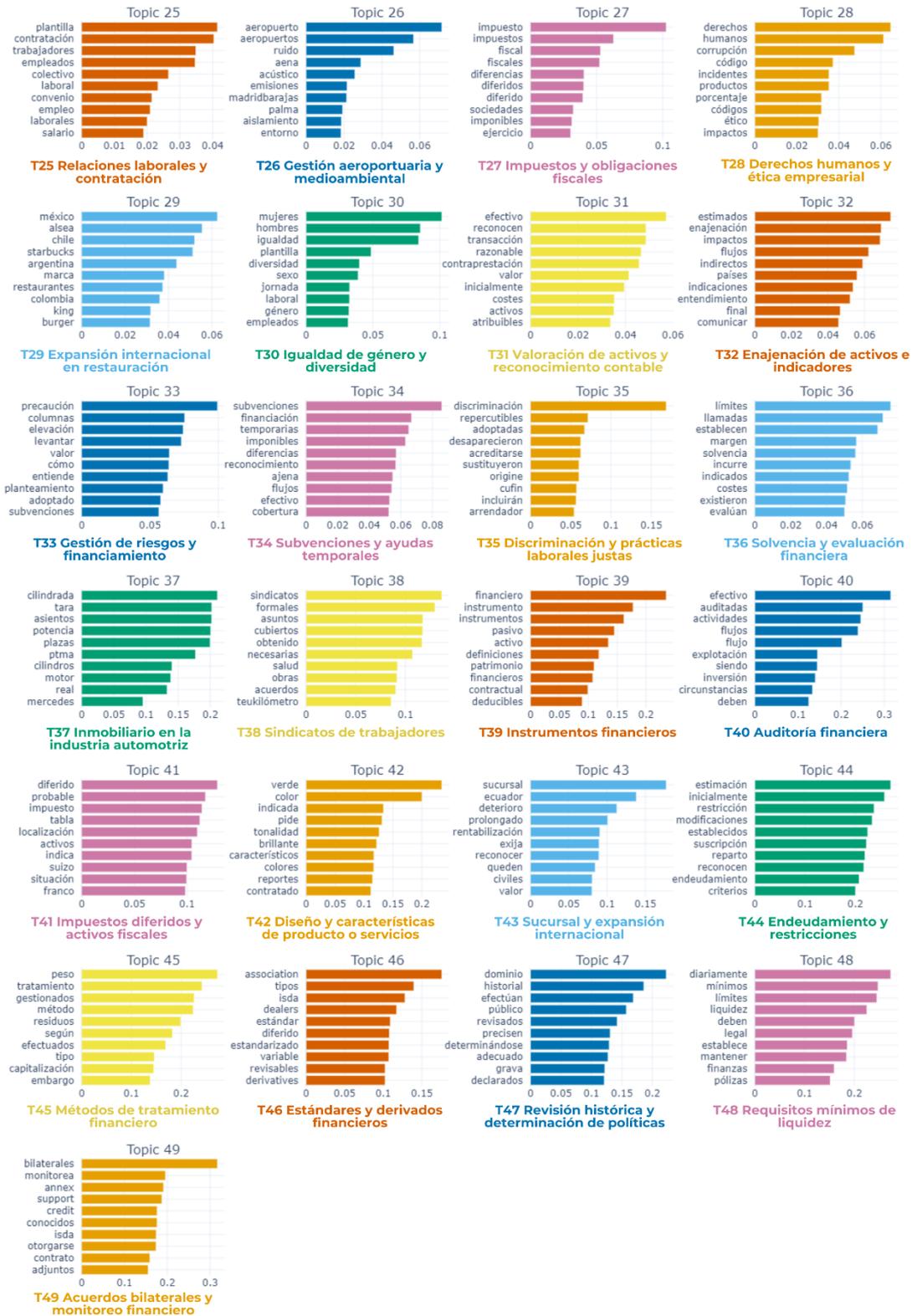


Anexo 1. Temas identificados en los informes de las empresas turísticas de 2010 a 2015 (del 0 al 24).



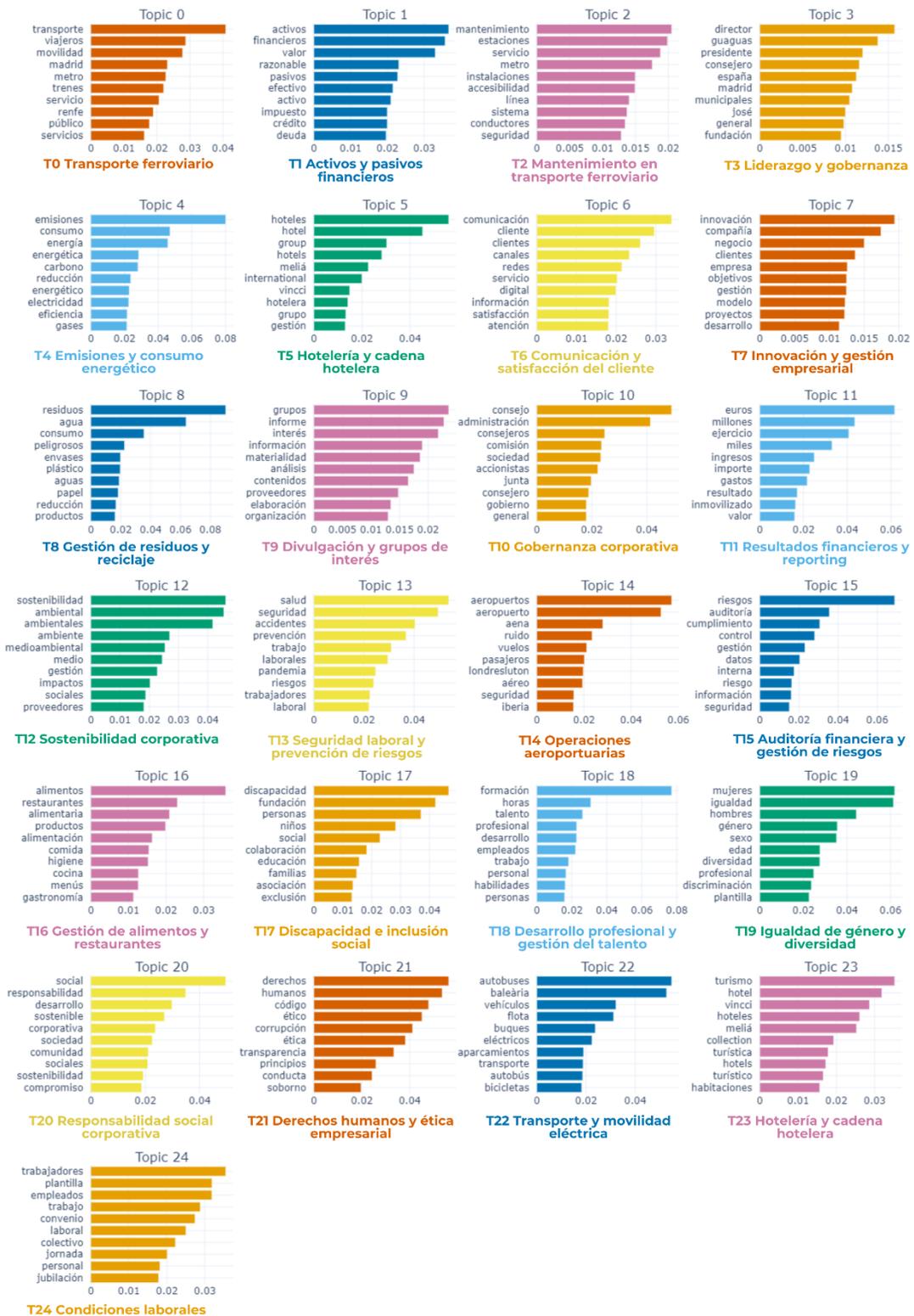
Fuente: Elaboración propia/procesamiento con BERTopic.

Anexo 2. Temas identificados en los informes de las empresas turísticas de 2010 a 2015 (del 25 al 49).



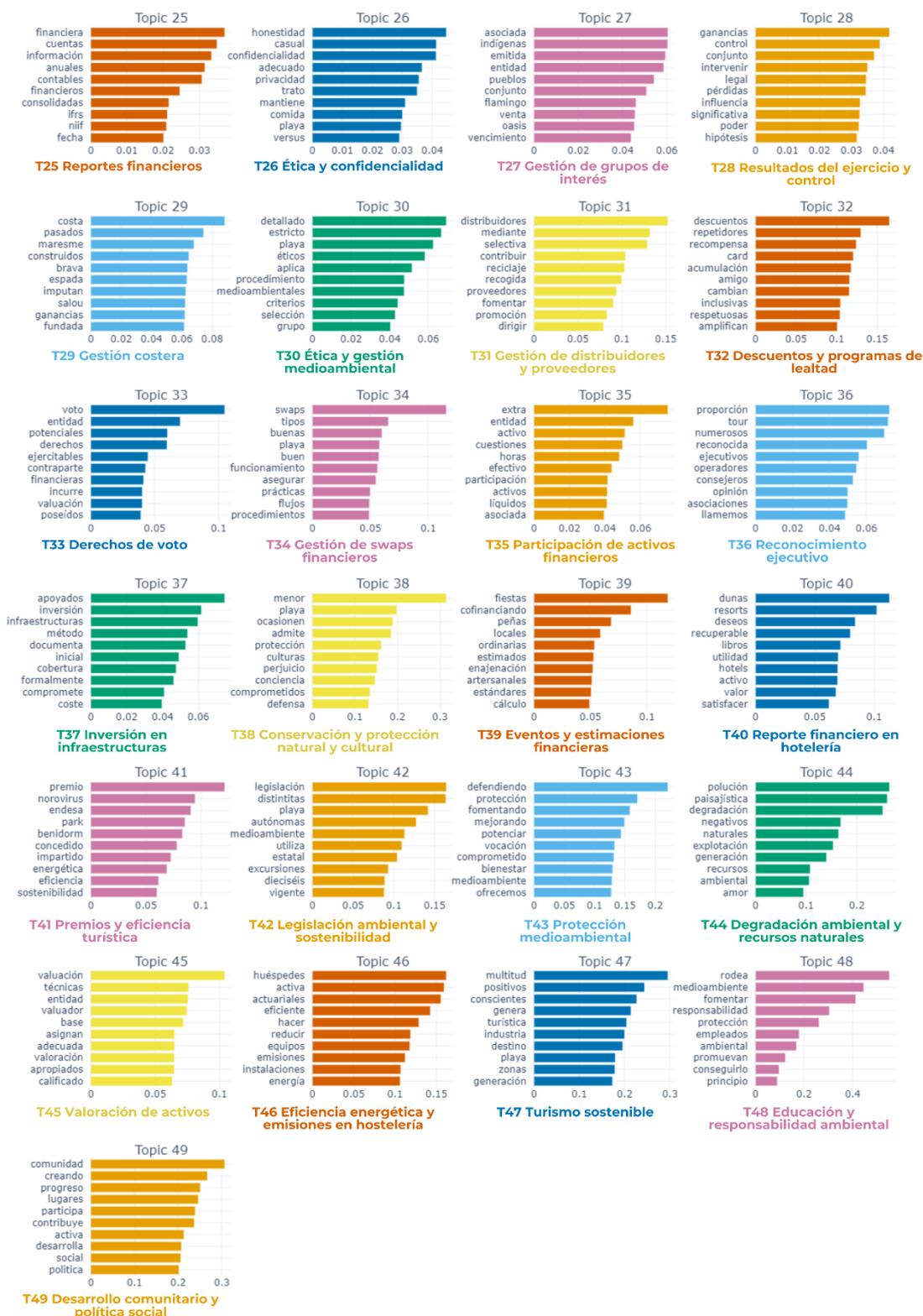
Fuente: Elaboración propia/procesamiento con BERTopic.

Anexo 3. Temas identificados en los informes de las empresas turísticas de 2016 a 2021 (del 0 al 24).



Fuente: Elaboración propia/procesamiento con BERTopic.

## Anexo 4. Temas identificados en los informes de las empresas turísticas de 2016 a 2021 (del 25 al 49).



Fuente: Elaboración propia/procesamiento con BERTopic.

**Anexo 5. Software, herramientas, metodologías y lenguajes de programación utilizados en la Tesis Doctoral.**

<b>Nombre</b>	<b>Uso en la Tesis Doctoral</b>	<b>Enlace o referencia</b>	<b>Descripción</b>
<b>Octoparse</b>	Recopilación de datos	<a href="https://www.octoparse.com/">https://www.octoparse.com/</a>	Herramienta de extracción de datos web (web scraping) que permite recopilar información de forma automatizada y estructurada (Octoparse, 2023).
<b>Python</b>	Manejo, preprocesamiento, análisis y visualización de datos	<a href="https://python.org/">https://python.org/</a>	Lenguaje de programación de alto nivel, versátil y ampliamente utilizado en áreas como desarrollo de software, ciencia de datos, automatización y análisis de datos. Destaca por su simplicidad, flexibilidad y amplia comunidad de soporte (Python Software Foundation, 2023b).
<b>Stata</b>	Análisis estadísticos	<a href="https://www.stata.com/">https://www.stata.com/</a>	Software estadístico especializado en el análisis de datos, manipulación de datos, visualización, estadísticas e informes (StataCorp, 2023).
<b>VOSviewer</b>	Mapeo bibliométrico	<a href="https://www.vosviewer.com/">https://www.vosviewer.com/</a>	Herramienta de software gratuita empleada para construir y visualizar redes bibliométricas. VOSviewer permite analizar redes de colaboración entre autores, países y revistas. Además, ofrece funcionalidades de minería de textos para construir y visualizar redes de coocurrencia de términos extraídos del corpus científico (van Eck & Waltman, 2010).
<b>Computational Literature Review (CLR)</b>	Exploración de temas y tendencias	Antons et al. (2021)	Metodología que utiliza técnicas de minería de texto y algoritmos de inteligencia artificial para analizar y sintetizar grandes volúmenes de literatura científica de un campo específico (Antons et al., 2021).
<b>Hugging Face</b>	Búsqueda de modelos de Inteligencia Artificial	<a href="https://huggingface.co/">https://huggingface.co/</a>	Plataforma de código abierto que alberga modelos y conjuntos de datos de Inteligencia Artificial (IA). Es reconocida como uno de los repositorios de IA más grandes a nivel global (Hugging Face, 2024).

**Fuente: Elaboración propia.**

**Anexo 6. Principales técnicas de ciencia de datos aplicadas en la Tesis Doctoral.**

Nombre		Descripción
<b>Minería de Texto y Procesamiento de Lenguaje Natural (PLN)</b>	Preprocesamiento lingüístico	Transformación y normalización del texto mediante técnicas como tokenización, lematización y eliminación de stopwords (palabras vacías). Estas técnicas estandarizan y preparan el texto para análisis posteriores, facilitando su estructura y manejo.
	Extracción de información	Técnicas para identificar patrones y conocimiento en grandes volúmenes de texto, permitiendo obtener información útil a partir de datos no estructurados.
	Análisis de frecuencia	Cálculo de frecuencias de palabras o frases para identificar términos clave y patrones recurrentes en el texto.
<b>Machine Learning</b>	Aprendizaje no supervisado	<u>Modelado de temas</u> : Identificación automática de temas y patrones en grandes cantidades de texto, mediante algoritmos probabilísticos.
	Aprendizaje supervisado	<p><u>Modelos de predicción</u>: Métodos predictivos que incluye la regresión lineal, así como modelos de clasificación ordinal (ologit y gologit) que permiten predecir tanto valores continuos como categorías ordinales en función de relaciones entre variables.</p> <p><u>Clasificación de texto</u>: Métodos para categorizar automáticamente documentos en clases predefinidas o emergentes, utilizando modelos preentrenados. Se aplica el enfoque de “Zero-shot learning”, puesto que permite que un modelo de inteligencia artificial categorice clases o conceptos sin haber sido entrenado explícitamente en ellos (Alhoshan et al., 2023).</p>

**Fuente: Elaboración propia.**

**Anexo 7. Modelos de Procesamiento de Lenguaje Natural (PLN) aplicados en la Tesis Doctoral.**

Nombre	Descripción
<b>BERTopic</b>	Técnica avanzada de modelado de temas capaz de capturar el contexto semántico. Se basa en los modelos <i>Transformers</i> <sup>7</sup> y el modelo de lenguaje BERT ( <i>Bidirectional Encoder Representations from Transformers</i> ) <sup>8</sup> . Este enfoque combina embeddings (representaciones numéricas de datos) generados por BERT con técnicas de clustering y el método c-TF-IDF para identificar automáticamente temas en grandes colecciones de documentos (Grootendorst, 2021).
<b>BETO <i>cased</i></b>	El modelo “es-BERT <i>cased</i> ”, también conocido como “BETO <i>cased</i> ” (Cañete et al., 2020), es una versión de BERT adaptada al español, entrenada con un extenso corpus de textos en este idioma. Específicamente, se emplea la versión de ajuste fino “ <i>bert-base-spanish-wwm-cased-xnli</i> ”, creada por la empresa Recognai (2021) y disponible en la plataforma de <i>Hugging Face</i> . Este modelo fue entrenado con la porción en español del conjunto de datos XNLI ( <i>Cross-lingual Natural Language Inference</i> ), optimizándolo para la inferencia de lenguaje natural en español. XNLI es un conjunto de datos multilingüe que contiene pares de oraciones traducidas, lo que permite a los modelos de PLN aprender a razonar sobre relaciones entre frases en distintos contextos lingüísticos.
<b>FinBERT</b>	Modelo especializado en el dominio financiero basado en la arquitectura BERT y entrenado previamente con textos financieros, incluidos informes corporativos, informes de analistas financieros y transcripciones de conferencias de resultados financieros (Huang et al., 2022). FinBERT fue ajustado ( <i>fine-tuned</i> ) para dos tareas específicas: clasificación de sentimientos y clasificación de temas ESG ( <i>Environmental, Social and Governance</i> , en sus siglas en inglés). El modelo original de FinBERT tiene la capacidad de clasificar los textos en cuatro etiquetas: ambiental, social, gobernanza o ninguno <sup>9</sup> . Sin embargo, el modelo ajustado permite una clasificación en nueve categorías: cambio climático, capital natural, contaminación y residuos, capital humano, responsabilidad del producto, relaciones con la comunidad, gobernanza corporativa, ética y valores empresariales y no ESG <sup>10</sup> .

**Fuente: Elaboración propia.**

<sup>7</sup> Los modelos *Transformers* son redes neuronales que comprenden el significado contextual al examinar cómo se relacionan entre sí las palabras en una secuencia. Utilizan técnicas matemáticas como la atención o atención propia para identificar estas conexiones y capturar el contexto de la información (Vaswani et al., 2017).

<sup>8</sup> BERT es un modelo de lenguaje desarrollado por Google en el año 2018 que utiliza la arquitectura de *Transformers* (Devlin et al., 2018).

<sup>9</sup> Modelo disponible en: <https://huggingface.co/yiyanghkust/finbert-esg>

<sup>10</sup> Modelo disponible en: <https://huggingface.co/yiyanghkust/finbert-esg-9-categories>

Anexo 8. Principales bibliotecas de Python empleadas en el desarrollo de la Tesis Doctoral.

Uso en la Tesis Doctoral	Nombre	Enlace	Descripción
Procesamiento de Datos	pandas	<a href="https://pandas.pydata.org/">https://pandas.pydata.org/</a>	Biblioteca especializada en la manipulación y análisis de datos en Python. Ofrece estructuras de datos y operaciones para manipular tablas numéricas y series temporales (McKinney, 2010).
	NumPy	<a href="https://numpy.org/">https://numpy.org/</a>	Biblioteca para realizar operaciones matemáticas con arrays multidimensionales (colecciones ordenadas de números) (Harris et al., 2020).
Procesamiento de Texto	re	<a href="https://docs.python.org/3/library/re.html">https://docs.python.org/3/library/re.html</a>	Módulo para el procesamiento de texto mediante expresiones regulares. Una expresión regular (o regex) es una secuencia de caracteres que define un patrón de búsqueda en texto. Este módulo de Python es útil tanto para la limpieza de datos, como para buscar y reemplazar patrones específicos de texto (Python Software Foundation, 2023c).
	PyPDF2	<a href="https://pypdf2.readthedocs.io/en/3.x/">https://pypdf2.readthedocs.io/en/3.x/</a>	Biblioteca especializada en manejo de archivos PDF, empleada para la extracción de texto y la fusión o división de documentos (Fenniak, 2023).
Manipulación de archivos y sistema operativo	os	<a href="https://docs.python.org/3/library/os.html">https://docs.python.org/3/library/os.html</a>	Módulo de Python que provee funciones para interactuar con el sistema operativo, como puede ser la gestión de archivos, directorios o rutas (Python Software Foundation, 2023a).

Uso en la Tesis Doctoral	Nombre	Enlace	Descripción
<b>Procesamiento de Lenguaje Natural (PLN)</b>	NLTK ( <i>Natural Language Toolkit</i> )	<a href="https://www.nltk.org/">https://www.nltk.org/</a>	Biblioteca útil para llevar a cabo tareas de análisis de texto complejos, como la tokenización, la lematización, el análisis gramatical y el procesamiento lingüístico básico (Bird et al., 2009).
	spaCy	<a href="https://spacy.io/">https://spacy.io/</a>	Biblioteca optimizada para tareas de PLN a gran escala, incluye análisis sintáctico, reconocimiento de entidades y procesamiento lingüístico avanzado (Honnibal et al., 2020).
	transformers	<a href="https://huggingface.co/">https://huggingface.co/</a>	Biblioteca desarrollada por Hugging Face que facilita el acceso a modelos avanzados de lenguaje preentrenados basados en la arquitectura <i>Transformer</i> y permite realizar fine-tuning (ajuste fino) de estos modelos. Es decir, adaptarlos a aplicaciones específicas (Wolf et al., 2020).
<b>Machine Learning</b>	scikit-learn	<a href="https://scikit-learn.org/">https://scikit-learn.org/</a>	Biblioteca que contiene herramientas y algoritmos eficientes para tareas como de clasificación, regresión, clustering (agrupación), reducción de dimensionalidad y selección de modelos (Pedregosa et al., 2011).
	statsmodels	<a href="https://www.statsmodels.org/">https://www.statsmodels.org/</a>	Biblioteca desarrollada para análisis estadísticos, econométricos y modelado de series temporales en Python. También es útil para explorar, describir y visualizar datos (Seabold & P(Seabold & Perktold, 2010).

Uso en la Tesis Doctoral	Nombre	Enlace	Descripción
Visualización de resultados	Matplotlib	<a href="https://matplotlib.org/">https://matplotlib.org/</a>	Biblioteca fundamental para la visualización de datos que permite generar gráficos de alta calidad con gran flexibilidad y personalización (J. D. Hunter, 2007).
	Seaborn	<a href="https://seaborn.pydata.org/">https://seaborn.pydata.org/</a>	Biblioteca de visualización de estadística avanzada basada en Matplotlib. Brinda una interfaz de alto nivel para generar gráficos estadísticos atractivos e informativos (Waskom, 2021).

Fuente: Elaboración propia.

