



Facultad de Ciencias Económicas y Empresariales
ICADE

TRABAJO FIN DE GRADO

humanidades digitales

Autor: Bastien Yefess
Director: María Eugenia Ramos Fernández

MADRID | 06/2026

1. Presentación

- 1.1. Propósitos
- 1.2. Razonamiento del análisis
- 1.3. Método
- 1.4. Organización del trabajo

2. Marco teórico: Conducta del consumidor y humanidades digitales

- 2.1. Idea y desarrollo de las humanidades digitales
- 2.2. Huellas digitales y datos textuales/sociales
- 2.3. Comportamiento humano y cultural en la era digital
- 2.4. Métodos y herramientas para el análisis de datos
 - 2.4.1. Análisis de sentimientos
 - 2.4.2. Minería de textos
 - 2.4.3. Big Data y visualización
- 2.5. Ética, privacidad y enfoque interdisciplinario en la evaluación de datos

3. Contexto: las tácticas de marketing y la estrategia corporativa

- 3.1. Vista general de la estrategia de marketing digital y los negocios
- 3.2. Usos de la explotación de las huellas digitales para entender la conducta y cultura digital
- 3.3. Retos y corrientes actuales en la toma de decisiones estratégicas

4. Investigación empírica: aplicación de las huellas digitales

- 4.1. Selección de empresas o sectores
- 4.2. Recolección de datos: desde las plataformas de redes sociales, la opinión de los clientes y las fuentes públicas
- 4.3. Procedimiento de análisis
 - 4.3.1. Preprocesamiento de datos (eliminación de datos incorrectos, tokenización, lematización)
 - 4.3.2. Análisis de textos y de la sociedad (análisis de las emociones, minería textual)
 - 4.3.3. Representación gráfica de datos (gráficos, tablas, tableros)
- 4.4. Resultados y discusión: patrones de comportamiento y tendencias culturales
- 4.5. Obstáculos, limitaciones y oportunidades en la recolección

5. Conclusión

- 5.1. Resultados principales
- 5.2. Consecuencias prácticas para la comprensión cultural y estrategias digitales

6. Limitaciones de la investigación y caminos futuros para las indagaciones

7. Fuentes bibliográficas

8. Apéndices

1. Presentación

1.1. Propósitos

El propósito principal de este trabajo es analizar de qué manera ciertos métodos propios del campo de las humanidades digitales pueden contribuir al estudio del comportamiento del consumidor y de las representaciones culturales en el entorno digital. No se trata de considerar las humanidades digitales como una herramienta sino como un área interdisciplinaria que permite combinar las tecnologías digitales con análisis crítico de los fenómenos culturales y sociales cuya metodología puede aplicarse al campo del marketing.

Así, la investigación busca una manera de estudiar cómo los consumidores producen huellas digitales (discursos, opiniones) en diferentes plataformas digitales como TikTok, Trustpilot y Google Reviews. Además, estas huellas digitales permiten observar no solo los comportamientos de compra, sino también percepciones, emociones sobre un sujeto o alrededor de una marca.

El trabajo toma como caso de estudio a Louis Vuitton, una marca que juega un papel central tanto en el sector del lujo como en el imaginario cultural contemporáneo. El objetivo de este caso de estudio es comprender cómo los consumidores expresan su relación con el lujo en el espacio digital, prestando atención a temas como la aspiración social, la autenticidad, el país de fabricación, la experiencia en tienda y la satisfacción o insatisfacción del cliente.

1.2. Razonamiento del análisis

La elección de este tema se justifica por la creciente importancia de los datos digitales en la comprensión de la sociedad contemporánea. Actualmente, una gran parte de los consumidores expresa sus experiencias, opiniones y reseñas en plataformas digitales. Así, los comentarios, reseñas, publicaciones y reacciones constituyen huellas digitales que pueden ser analizadas para comprender mejor las dinámicas sociales y culturales del consumo.

En este contexto, las humanidades digitales, no solo se limitan al uso técnico de herramientas computacionales, sino que permiten interpretar fenómenos humanos, culturales y sociales gracias a métodos digitales y mirada crítica. Además, estas huellas digitales no deben ser utilizado como datos cuantitativos sino también como fragmentos de discursos que refleja una imagen real de la sociedad con expectativas, valores y tensiones sociales.

Así, el sector del lujo se presenta especialmente pertinente para este tipo de análisis porque el hecho de comprar productos de lujo no responde únicamente a una lógica funcional o económica. De hecho, comprar un producto de lujo responde también a una lógica de identidad, prestigio y pertenencia social. Por ello, estudiar las huellas digitales producidas alrededor de esta marca permite observar como se construye y se discute el significado del lujo en la era digital, sin olvidar que se trata de una marca cuya legitimidad se ha construido también fuera del entorno digital.

Este trabajo se sitúa, por tanto, en un punto de encuentro entre el análisis cultural, el estudio del comportamiento del consumidor y los métodos digitales de tratamiento textual. Su interés reside en mostrar como una aproximación inspirada en las humanidades digitales, y mas concretamente en el análisis computacional de textos y en la minería de datos textuales, puede enriquecer la comprensión de fenómenos que, aunque vinculados al marketing, son también profundamente culturales y sociales.

1.3. Método

La metodología adoptada se basa en un corpus de comentarios y opiniones procedentes de varias plataformas digitales. Los datos fueron extraídos directamente de TikTok, Trustpilot y Google Reviews con el objetivo de combinar fuentes complementarias entre sí: Por un lado, los comentarios publicados en una red social centrada en la visibilidad, la aspiración y la interacción comunitaria; por otro lado, reseñas de clientes vinculadas a experiencias concretas de compra o de servicio

En segundo lugar, los datos pasaron por diferentes fases para que se puede ser analizado, el preprocesamiento textual. Esta etapa consiste en limpiar los comentarios, eliminar los elementos irrelevantes (eliminar los elementos que no aportan información semántica tales como menciones de usuarios, URLs) y por fin la tokenización y lematización. La tokenización consiste en dividir cada texto en unidades lingüísticas mínimas llamadas token suele ser palabras o expresiones. Por su parte, la lematización, consiste en reducir cada palabra a su forma canónica. El objetivo de este proceso fue transformar textos heterogéneos en un corpus más estructurado que pudiera ser analizado gracias a herramientas digitales.

En tercer lugar, la aplicación de técnicas de análisis de sentimientos y minería textual. El análisis de sentimiento permite distinguir los contenidos que son positivos, negativo y neutro. Por su parte, la minería textual permite evaluar la frecuencia de determinados términos, las asociaciones entre palabras y la aparición de temas recurrentes para determinar su importancia.

Finalmente, los resultados fueron interpretados de manera crítica teniendo en cuenta el contexto de producción de los datos, las características propias de cada plataforma y las limitaciones del análisis. Esta etapa interpretativa es fundamental, ya que permite evitar una lectura puramente automática de los resultados y situar los datos dentro de un marco social, cultural y comunicacional más amplio.

1.4. Organización del trabajo

Este trabajo se presenta en varias partes. En primer lugar, el marco teórico, centrado en la conducta del consumidor y en las humanidades digitales. Esta parte permite definir los conceptos clave del estudio tal como huellas digitales, el comportamiento del consumidor, minería textual, análisis de sentimiento, el Big Data, la visualización de datos y las cuestiones éticas relacionadas con el uso de la información digital.

En segundo lugar, el trabajo se centra en el contexto empresarial y estratégico. Así, esta parte se centra en las tácticas de marketing digital, la explotación de las huellas digitales por parte de las empresas y los retos actuales relacionados con la toma de decisiones estratégicas basada en datos.

En tercer lugar, la investigación empírica. En esta parte, permite poner en práctica los conceptos presentados anteriormente a través de un caso de estudio real, Louis Vuitton. Este ejemplo permite observar en concreto, la recolección de datos procedentes de TikTok, Trustpilot y Google Reviews. También se presenta el procedimiento completo del análisis incluyendo el preprocesamiento de los datos, el análisis textual, el análisis de sentimientos y la representación gráfica de los resultados.

Por último, el trabajo se termina con una conclusión general, con un resumen de los principales resultados obtenidos, sus consecuencias en prácticas para la comprensión cultural y las estrategias digitales. También se incluyen las limitaciones de la investigación y las posibles líneas futuras de estudio.

2. Marco teórico: Conducta del consumidor y humanidades digitales

El mundo se vuelve cada vez un poco más digital. De hecho, según las proyecciones de eMarketer recogidas por Statista (2024), basadas en el análisis de las tasas de crecimiento del comercio electrónico mundial, las ventas en línea deberían representar el 23% del total de las ventas minoristas mundiales en 2027. Sin embargo, no todos los sectores son afectados de manera uniforme. De hecho, por ejemplo el sector del lujo tiene una digitalización más progresiva, debido a que la experiencia física en tienda sigue siendo un elemento central del posicionamiento de las marcas (Kapferer & Bastien, 2012). Desde un punto de vista de las empresas, comprender el comportamiento del consumidor en este nuevo entorno ocupa un lugar central en su desarrollo. De hecho, considerar que las decisiones de compra se basan únicamente en factores económicos, es decir, en la búsqueda de maximizar la satisfacción bajo una restricción presupuestaria tal como la describe la teoría neoclásica del consumidor racional (Marshall, 1890), sería una visión reduccionista que no refleja la complejidad real del comportamiento del consumidor contemporáneo. Actualmente, las decisiones de compra están condicionadas también por dimensiones sociales tales como la búsqueda de reconocimiento y pertenencia a un grupo (Bourdieu, 1984), culturales (valores, las normas y las identidades colectivas) y tecnológicas con la influencia de los algoritmos y las redes sociales en los comportamientos de consumo (Sunstein, 2017).

Es esta idea que Kotler y Keller (2016) desarrollan. Entonces, el comportamiento del consumidor evoluciona en función de su entorno y de los intercambios a los que se enfrenta. Hoy en día, con el desarrollo constante de lo digital, los individuos producen un flujo constante de datos digitales. Esta información constituye así una nueva fuente de análisis que permite comprender mejor estos comportamientos. Es en este contexto que las humanidades digitales, como campo académico, ofrecen un marco metodológico pertinente para analizar estos comportamientos.

Así, esta parte tiene como objetivo presentar los fundamentos teóricos del trabajo, apoyándose particularmente en tres dimensiones. Las humanidades digitales, las huellas digitales y el comportamiento de los consumidores, así como los métodos de análisis y los desafíos éticos asociados.

2.1 Idea y desarrollo de las humanidades digitales

Las humanidades digitales se pueden definir como un campo de estudio interdisciplinario que combina las ciencias humanas y sociales con las herramientas digitales. Schreibman, Siemens y Unsworth describen las humanidades digitales como la aplicación de métodos digitales a la investigación en ciencias humanas. Por su parte Berry (2012) va un poco más allá de esta definición para introducir el cambio de método especialmente en la forma de recopilar, analizar e interpretar los datos a gran escala.

Por lo tanto, los avances tecnológicos han permitido, a través de las humanidades digitales el estudio de contenidos, comportamientos y fenómenos culturales. De hecho, el análisis de sentimientos y la minería textual son algunos ejemplos de los métodos analíticos que ha desarrollado esta disciplina y que tienen usos en distintos campos, entre ellos el marketing y la comunicación.

Concretamente, los métodos desarrollados en el marco de las humanidades digitales permiten analizar un volumen importante de datos provenientes de diversas fuentes. Hoy en día es posible estudiar miles de tuits o publicaciones casi en tiempo real con el fin de identificar la evolución de la opinión pública. El caso de Cambridge Analítica ilustra perfectamente fenómeno, aunque de manera muy controvertida por la recopilación de millones de perfiles de Facebook sin el consentimiento explícito de los usuarios. La empresa ha utilizado la publicidad segmentada y contenidos personalizados con el fin de influir en el comportamiento de los electores y orientar campañas políticas a través del análisis masivo de perfiles digitales.

Sin embargo, esta automatización del análisis tiene una limitación importante. De hecho, para que la IA o un algoritmo puedan realizar esta automatización antes de todo tiene que simplificar el texto. Schreibman, Siemens y Unsworth (2004) recuerdan que las herramientas digitales deben ayudar pero no reemplazar la interpretación humana. Este punto es fundamental: confiar completamente el análisis a un algoritmo es aceptar perder parte de la riqueza y la complejidad del lenguaje humano.

Drucker (2011) va más allá al señalar que los datos nunca son neutrales: siempre son el resultado de elecciones interpretativas humanas, ya sea en la manera en que se recogen, estructuran o analizan. Un algoritmo puede detectar la frecuencia de una palabra, pero no puede comprender el contexto cultural, histórico o emocional que le da su verdadero sentido. La palabra "crisis", por ejemplo, aparecerá con alta frecuencia tanto en un artículo de economía como en un mensaje de pánico personal, el algoritmo no hace la diferencia entre ambos. Del mismo modo, una frase irónica como "qué excelente servicio..." después de una mala experiencia con el cliente será muy probablemente interpretada como un comentario positivo por una herramienta de análisis automático.

Cambria y White (2014) confirman esta limitación al mostrar que las sutilezas propias del lenguaje humano. La ironía, el sarcasmo, las insinuaciones culturales, constituyen un desafío importante para los algoritmos de procesamiento del lenguaje natural. Estas sutilezas, aunque esenciales para comprender realmente lo que piensa un consumidor, a menudo escapan a las herramientas automatizadas.

Para las empresas, esto implica no conformarse nunca con una lectura bruta de los datos. El análisis automatizado debe ser sistemáticamente complementado por el juicio humano de los equipos de marketing, que son los únicos que disponen del contexto necesario para interpretar correctamente los resultados. Ignorar esta complementariedad supone exponerse a decisiones estratégicas basadas en interpretaciones erróneas, con consecuencias potencialmente costosas.

Las humanidades digitales ofrecen un marco conceptual y metodológico útil para analizar el comportamiento de los consumidores en línea. Por eso, hay que identificar con mayor precisión cuáles son los elementos que conforman este análisis: las huellas digitales que los consumidores dejan (consciente o inconscientemente) durante sus interacciones en línea.

2.2 Huellas digitales y datos textuales/sociales

El desarrollo de internet ha intensificado la producción de datos digitales, pero también ha favorecido innovaciones como las "cookies". Las cookies son pequeños archivos almacenados en el dispositivo, que permiten almacenar y recopilar datos sobre el usuario que son clics, tiempo en una página, los pedidos... (Mayer & Mitchell, 2012). Existen dos categorías: las cookies de sesión, que desaparecen al cerrar el navegador, y las cookies persistentes, que recopilan datos durante un período de tiempo determinado de antemano. Estas huellas digitales representan una fuente de información esencial para comprender los patrones de consumo.

Según Boyd y Crawford (2012), estos datos permiten observar las prácticas sociales a gran escala. Por ejemplo, Amazon analiza de manera sistemática las búsquedas de sus clientes en el sitio web y sus compras con el fin de recomendar productos personalizados en las próximas visitas.

Cuando se trata de estadística o de conexiones, es muy importante diferenciar entre correlación y relación de causa y efecto. Mayer-Schönberger y Cukier (2013) lo destacan particularmente tomando el ejemplo de un supermercado. Puede haber una correlación entre las personas que compran a menudo mantequilla y las que compran a menudo un cartón de jugo. Sin embargo, esto no significa que la compra de mantequilla sea la causa de la compra de zumo: se trata simplemente de productos de consumo corriente en un hogar (comprado de manera independiente).

Así, las huellas digitales permiten un acceso privilegiado a una base de datos. Sin embargo, requieren una interpretación prudente y mostrar un espíritu crítico. Una mala lectura puede conducir a decisiones estratégicas erróneas. Por ejemplo, en 2013, Google Flu Trends alertó sobre el número de búsquedas en internet relacionadas con la gripe, desencadenando un movimiento de pánico por una nueva epidemia. Sin embargo, resultó ser una mala

interpretación de los datos, ya que sobrestimó en un 50% el número de personas con síntomas de gripe.

Sin embargo, no basta con entender la naturaleza y el origen de estos rastros. Para aprovecharlos al máximo, hay que entender cómo generan los consumidores esos datos. En otras palabras, hay que comprender los procesos culturales y sociales que determinan su comportamiento en el entorno digital.

2.3 Comportamiento humano y cultural en la era digital

El comportamiento del consumidor ha cambiado provocado por la parte creciente del comercio electrónico y el desarrollo de las plataformas digitales. Actualmente, las decisiones de compra no se toman a partir de criterios económicos o si el consumidor ha visto un anuncio. De hecho, los consumidores intercambian constantemente en blogs, foros de discusión o redes sociales. Esto influye fuertemente en la manera en que descubren los productos, evalúan las marcas y forman sus opiniones. Las redes sociales desempeñan papel central en el éxito de una venta por los contenidos generados por los usuarios.

Kaplan y Haenlein (2010) muestran que, en el entorno digital, el consumidor ya no se limita a recibir información de manera pasiva, sino que puede participar más activamente. Es a través de comentarios, opiniones o incluso publicaciones que los consumidores se vuelven activos en la comunicación comercial de las marcas, pudiendo influir directamente en la percepción de un producto o de toda una marca. El poder de influencia pasa de ser centralizado a descentralizado, una mala experiencia puede ser compartida y desalentar así a potenciales nuevos clientes. Este fenómeno se intensifica con el desarrollo de plataformas como TikTok donde la información se comparte casi en tiempo real.

Las plataformas digitales también tienen un papel que desempeñar en lo que respecta al acceso a la información. De hecho, una gran parte de los contenidos visualizados por los usuarios proviene de las recomendaciones realizadas por un algoritmo que no solo tiene en cuenta los intereses de cada usuario, sino que también los modifica. Las plataformas como Netflix o Spotify integran sistemas de recomendación avanzados que proponen contenidos en función de los artistas, actores o estilos previamente visionados. Aunque estos algoritmos buscan mejorar la experiencia del cliente, también tienen limitaciones, especialmente en lo que respecta a la diversidad de los programas ofrecidos. Es en esta idea que Sunstein (2017) desarrolla el concepto de «burbujas de filtrado», definido como entornos digitales donde los individuos están principalmente expuestos a contenidos que corresponden a sus preferencias anteriores, reduciendo así la confrontación con opiniones o perspectivas diferentes.

Pierre Bourdieu (1984), por su parte, permite aportar el aspecto sociológico. De hecho, subraya que los consumos de los clientes no dependen únicamente de las elecciones individuales propias de cada uno, sino también de factores sociales y culturales. El consumo se transforma en una forma de expresión social e identitaria. En plataformas como Instagram, por ejemplo, ciertos productos, estilos de vida o marcas funcionan como símbolos de

reconocimiento social y contribuyen a reforzar las dinámicas de pertenencia o diferenciación entre los individuos.

El comportamiento del consumidor dentro del espacio digital es el resultado, por lo tanto, de la interacción entre varias dimensiones: preferencias personales, influencias sociales y mecanismos tecnológicos de recomendación. Comprender al consumidor en estas dinámicas representa un desafío importante para las empresas, que buscan transformar los comportamientos observables en información útil (marketing y comunicación).

2.4 Métodos y herramientas para el análisis de datos

Los métodos de análisis tradicionales se basan en una base de datos declarativa. Eso significa que la información utilizada proviene de declaraciones a través de encuestas de satisfacción, cuestionarios o entrevistas. Hoy es posible tener una base de datos más amplia y también con mayor precisión. De hecho, no proviene de una encuesta con puntos de escala, sino de opiniones y declaraciones en directo en redes sociales, blogs y plataformas, lo que permite una mayor precisión (Mayer-Schönberger & Cukier, 2013).

Frente a esta limitación, los nuevos enfoques se basan en datos de comportamiento reales, obtenidos mediante el seguimiento de las acciones efectivas de los usuarios en los entornos digitales. Su carácter masivo y no mediado los hace tanto más precisos como más representativas que los datos declarativos.

En este contexto, la nueva base de datos que puede utilizar no está estructurada. De hecho, no se organizan en líneas o columnas predefinidas como puede ser los textos, las imágenes o las publicaciones. Así, los investigadores han desarrollado nuevas herramientas capaces de tratar este tipo de formato que no puede ser analizados directamente por los métodos estadísticos convencionales.

Las herramientas presentadas a continuación (análisis de sentimientos, minería de textos y Big Data) tienen sus raíces metodológicas en las humanidades digitales, campo que las desarrolló inicialmente para el análisis de corpus literarios, históricos y culturales a gran escala. Su aplicación al marketing constituye precisamente el objeto de este trabajo. Esta combinación permite detectar tendencias a gran escala mientras se comprende su significado.

Hoy en día, el desafío no está únicamente en el análisis de una base de datos para extraer conclusiones, sino también en comprender la manera en que se realizan esos análisis. En efecto, las decisiones metodológicas condicionan los resultados, los diferentes algoritmos pueden introducir sesgos y los resultados dependen del marco interpretativo elegido. En otras palabras, como en cualquier investigación, la reflexión sobre la metodología utilizada es un requisito fundamental. En el ámbito de las humanidades digitales, esta reflexión resulta especialmente importante debido al uso de herramientas algorítmicas y al tratamiento de grandes volúmenes de datos.

2.4.1 Análisis de sentimientos

El análisis de sentimientos consiste en identificar las emociones gracias a patrones que son expresados a través de las opiniones de los clientes o los comentarios disponible en línea. Concretamente, este análisis se basa en cortar los textos, identificar cada palabra clave para atribuirles un tono que puede ser positivo, negativo o neutro. Los principios de esta técnica fueron formalizados por Liu (2012) y Pang & Lee (2008).

Por ejemplo, una empresa como Apple puede ser analizada en pocos segundos a través de experiencias de clientes encontradas en línea. Las palabras asociadas a emociones positivas pueden ser “rápido” o “eficaz”. Al contrario "lento" o "decepcionante" pueden ser conectados con emociones negativas. Como muestran Hu & Liu (2004) en su análisis de las opiniones de los consumidores sobre productos, este método permite identificar rápidamente los puntos de satisfacción e insatisfacción relacionados con un producto o una marca.

Sin embargo, este enfoque presenta limitaciones importantes. Cambria y White (2014) muestran que las palabras extraídas de su contexto pueden recibir una interpretación radicalmente diferente de la formulada. La ironía, la ambigüedad o ciertas construcciones sintácticas resisten al análisis algorítmico: un comentario como «super producto... otra avería» corre el riesgo de ser clasificado como positivo cuando en realidad es una crítica. Este tipo de problema está en el centro del análisis del discurso que consiste a estudiar precisamente la conexión del sentido de un enunciado con su contexto, las intenciones del hablante o de las convenciones sociales que enmarcan su interpretación. Fairclough (1995) y van Dijk (1997), figuras destacadas del análisis crítico del discurso, han documentado ampliamente estos mecanismos.

El análisis de sentimientos constituye así una herramienta eficaz para extraer tendencias a gran escala. Estas tendencias deben, no obstante, ser sometidas a una interpretación humana para evitar conclusiones apresuradas: solo esta última puede integrar el contexto, la cultura de marca y las sutilezas del discurso de los consumidores.

2.4.2 Minería de textos

La minería de textos consiste en ordenar y estructurar datos a partir de una base de datos bruta, que significa un conjunto de textos no organizados. Este proceso permite extraer información importante como opiniones, comentarios o aun publicaciones en redes sociales en una forma adecuada para ser analizada.

En primer lugar, los textos son limpiados. Esta etapa consiste en suprimir las palabras inútiles. Una palabra inútil, en este marco, es una palabra que es gramaticalmente necesaria pero sin valor para el análisis. Se trata entonces de los artículos, las preposiciones o las conjunciones. Así, otras herramientas pueden tomar el relevo para identificar la recurrencia de elementos y la asociación de palabras utilizadas. En esta etapa, el contexto en el que aparecen estas palabras no se tiene en cuenta. El objetivo es obtener la frecuencia, lo que ya constituye en sí mismo una limitación del método.

Por ejemplo, una empresa es capaz, gracias a este método, de identificar las asociaciones hechas con su marca o su producto. De hecho, si los términos «fácil de usar», «práctico» o incluso «buena relación calidad-precio» aparecen con frecuencia, indican puntos fuertes percibidos por los consumidores. Por el contrario, si hay una alta frecuencia de términos como «mala relación calidad-precio», «frágil» o incluso «decepcionado» revela problemas recurrentes en la experiencia del usuario.

Sin embargo, este enfoque también tiene sus límites. Como muestra Hearst en (1999), para que el lenguaje humano sea explotable por los algoritmos, la lengua debe ser simplificada al extremo, lo que es fuente de error. El ejemplo de la palabra «rápido» lo ilustra bien: este término será interpretado positivamente en la expresión «entrega rápida», pero negativamente si califica la duración de vida de una batería de teléfono.

Esto plantea una cuestión más amplia: si el algoritmo no puede interpretar el contexto, ¿es necesario que un ser humano verifique constantemente la relevancia de sus resultados? La respuesta es sí, y es precisamente esto lo que fundamenta el valor del enfoque mixto mencionado anteriormente. La minería de textos es una herramienta eficaz para la detección de tendencias a gran escala, pero no sustituye la interpretación humana. Para las empresas, esto implica una intervención cualitativa regular para validar que los patrones identificados corresponden bien a los comportamientos reales de los consumidores, y no a artefactos producidos por la simplificación algorítmica del lenguaje.

2.4.3 Datos masivos y visualización

El big data es un término bastante amplio que hace referencia a la capacidad de recopilar, almacenar y analizar grandes bases de datos. Estos datos pueden provenir de diferentes fuentes, como la navegación por Internet, las transacciones bancarias, las redes sociales, los dispositivos conectados... La idea es procesar un gran volumen de información para identificar tendencias o correlaciones. Este campo ha sido teorizado, en particular, por Mayer-Schönberger y Cukier (2013), quienes destacan que el valor del Big Data reside menos en los datos en sí mismos que en los usos analíticos que estos hacen posibles.

Primero, la información es recolectada y almacenada en bases de datos adaptadas como los centros de datos. Luego es analizada mediante diferentes algoritmos según la fuente de los datos y el objetivo de la empresa. Este enfoque permite detectar patrones, correlaciones, similitudes o incluso opiniones.

Por ejemplo, Amazon utiliza el Big Data en las sugerencias personalizadas hechas a los consumidores. El historial de compras y la navegación de los usuarios se analizan para proponer productos que correspondan a sus necesidades. Si una persona acaba de comprar un iPhone, proponer una funda y un cargador adaptado es más pertinente que proponer un Samsung.

Estos datos en bruto se vuelven difíciles de explotar sin un formato adecuado. Es aquí donde interviene la visualización de datos, que consiste en representar la información en forma de gráficos, tableros de control o matrices para hacerla legible y accionable. Un tablero de control

puede así mostrar en tiempo real el valor medio del carrito o el número de pasos en la caja, permitiendo a un responsable orientar rápidamente las decisiones de su equipo. Esta presentación facilita la lectura, pero también conlleva un riesgo, al simplificar datos complejos, puede ocultar matices o inducir interpretaciones demasiado rápidas.

Según Boyd y Crawford (2012), estas herramientas abren oportunidades inéditas pero con límites. De hecho, la cantidad masiva de datos puede crear una ilusión de objetividad porque los números parecen hablar por sí mismos, se olvida que son el resultado de elecciones metodológicas, qué recoger, cómo tratar, qué mostrar. Sin embargo, una correlación no es una causalidad.

Así, el Big Data y la visualización son herramientas poderosas para analizar comportamientos a gran escala, siempre y cuando no se limiten a leer los datos, sino que se interpreten en su contexto para construir estrategias pertinentes. Este trabajo de interpretación exige una toma de distancia metodológica, pero también una reflexión sobre las condiciones en las que se han recopilado estos datos. Ante la magnitud de lo que estas tecnologías permiten, surge una pregunta: ¿en qué medida el uso de estos datos sigue siendo ético, legítimo y conforme a la ley?

2.5 Ética, privacidad y enfoque interdisciplinario en la evaluación de datos

El uso masivo de los datos digitales también plantea un problema ético en torno a la dudable consciencia de los consumidores sobre el uso de sus datos personales. De hecho, la recopilación y el tratamiento de estos datos pueden hacer que nos preguntemos si siempre se respeta nuestra privacidad y cuál es el verdadero nivel de consentimiento informado de los consumidores. Como muestran Acquisti, Brandimarte y Loewenstein (2015), el consentimiento de los usuarios es a menudo formal en lugar de real: los individuos aceptan las condiciones de uso sin medir las implicaciones concretas en su privacidad.

Hoy en día, en la práctica, algunas empresas se especializan en recopilar y revender información sobre el comportamiento de los consumidores. Lo que se revende son historiales de navegación, listas de compras en línea e interacciones en las redes sociales. Cuando un usuario acepta las cookies en un sitio web, generalmente ignora la extensión exacta de los datos a los que la empresa accede y a quién se los transmite. Esta falta de transparencia está en el centro de las críticas dirigidas al modelo económico dominante de las plataformas digitales. Zuboff (2019) analiza en detalle este fenómeno en su obra de referencia sobre el capitalismo de la vigilancia. Entonces, la autora muestra que las grandes empresas especializadas como Google, Amazon, Facebook, Apple, Microsoft, no se contentan con ofrecer servicios. Más bien, basan gran parte de su modelo de negocio en la recopilación y explotación de datos de comportamiento con fines de predicción e influencia. Los clics, los "me gusta" y los historiales de navegación se convierten así en activos monetizables, explotados para ofrecer contenidos publicitarios dirigidos.

Para proteger a los consumidores de estos problemas, la Unión Europea adoptó en 2018 el Reglamento General de Protección de Datos (RGPD). El texto obliga a las empresas a cumplir cuatro cosas principales. Primero, las empresas deben pedir el consentimiento claro de los usuarios antes de recoger cualquier dato. Segundo, las empresas deben usar los datos solo para

la finalidad que se declaró. Tercero, las empresas deben dar a los usuarios el derecho a acceder a los datos, a corregir los datos y a borrar los datos. Cuarto, las empresas deben avisar a las autoridades cuando haya una violación (Parlamento Europeo, 2016). Esta es la información proporcionada por el usuario: En Francia, la Comisión Nacional de Informática y Libertades (CNIL) cuida que se cumplan los principios (CNIL, 2023). Pero el marco jurídico no basta para asegurar el uso ético. La complejidad de las tecnologías y la rapidez de la innovación hacen que la práctica sea difícil de detectar. La práctica también resulta difícil de sancionar.

En este contexto, las humanidades digitales tienen un papel para integrar una reflexión sobre los usos de los datos. El desafío tiene dos partes. Por un lado es usar los datos para mejorar las estrategias de marketing y para entender los comportamientos de los consumidores. Los métodos que aparecen en la sección responden a la necesidad: análisis de sentimientos, minería de textos, Big Data. Por otro lado, hay que asegurar el respeto de los derechos de los individuos frente a la recopilación de datos que a veces son tan grandes que los individuos no los comprenden. La tensión entre el rendimiento del análisis y la ética de los datos no solo ocurre en el marketing. La tensión también está en todas las ciencias sociales actuales. Por eso, el enfoque mixto que combina el tratamiento algorítmico y la interpretación humana se vuelve necesario como método (Creswell, 2014).

3. Contexto: las tácticas de marketing y la estrategia corporativa

Como se ha visto anteriormente, el marco teórico ha permitido determinar los fundamentos del análisis de las huellas digitales relacionadas con el comportamiento del consumidor. En esta parte, es pertinente ponerlos en contexto para estudiar la aplicación concreta en el mundo empresarial. Esta revolución tecnológica orientada hacia lo digital ha cambiado completamente las estrategias de marketing, comunicación y la toma de decisiones dentro de las empresas. Hoy en día, ya no se trata solo de tener una idea vaga de las necesidades de los clientes. Deben conocerlos y apoyarse en datos en tiempo real para optimizar sus estrategias y ganar competitividad. Esta parte tiene como objetivo examinar cómo el marketing digital ha evolucionado hacia un modelo basado en cifras. La idea es, por lo tanto, saber cómo aprovechar las huellas digitales y cuáles son los desafíos que esta evolución impone en la toma de decisiones.

3.1 Vista general de la estrategia de marketing digital y los negocios

Hoy en día, las empresas interactúan de una manera nueva con la aparición del marketing digital. De hecho, Chaffey y Ellis Chadwick (2019) lo ilustran a través del hecho de que no se trata solo de cambiar de entorno y utilizar las técnicas tradicionales en lo digital. El conjunto de la cadena de valor de la empresa se ve afectado, desde la identificación de oportunidades de mercado hasta la fidelización de la clientela.

El modelo tradicional consiste en difundir un mensaje a través de un anuncio de televisión, un cartel o incluso un catálogo. El consumidor estaba entonces expuesto sin poder tener una

interacción inmediata. Con la llegada de lo digital, el consumidor es a la vez el receptor y el emisor de la información. A través de una publicidad en las redes sociales, puede comentar, compartir o incluso evolucionar el producto. Así es como las empresas han dejado de utilizar el "broadcasting" que consiste en tocar al mayor número de persona posible por una estrategia de "narrowcasting" que consiste en personalizar el mensaje para que sea relevante a cada segmento de consumidores.

Es a través del cambio que Kotler, Kartajaya y Setiawan (2021) explican el concepto de marketing 5.0. Este último permite integrar la evolución del marketing donde la tecnología es una herramienta al servicio de una experiencia del cliente humana y personalizada. En otras palabras, la empresa es capaz de anticipar las necesidades de un cliente, de adaptar su oferta en tiempo real sin que el cliente tenga que manifestarlo.

Para comprender los engranajes de este modelo, es útil examinar los tres mecanismos que lo caracterizan.

El primero es el algoritmo de recomendación. El ejemplo más destacado es el de Amazon con su algoritmo llamado A9. De hecho, este último analiza los comportamientos de compra y las búsquedas (incluyendo el tiempo pasado en cada búsqueda) de cada usuario para sugerirle otros productos. Así, cada sugerencia es individualizada. Este mecanismo transforma totalmente la búsqueda en línea, pero también el posicionamiento de las marcas, un aspecto crucial para estas últimas.

El segundo mecanismo es la publicidad programática. En el modelo tradicional, una marca negociaba directamente con un medio la compra de un espacio publicitario, a precio fijo y con varias semanas de antelación, sin un control preciso sobre la audiencia expuesta.

La publicidad programática automatiza completamente este proceso. A cada nueva carga de página web, se desencadena una subasta entre varios anunciantes. En más o menos 100 milisegundos (el tiempo de carga de una página web), un postor gana así la subasta, lo que le permite el derecho de mostrar su anuncio a un usuario específico. Este usuario se determina en función de su perfil hecho a partir de sus huellas digitales. El anunciante que ofrece el precio más alto gana la subasta. Todo esto ocurre de manera invisible, antes incluso de que el usuario haya terminado de ver la página cargarse.

La principal ventaja de esta herramienta es la precisión. De hecho, si tomamos el ejemplo de Spotify Ad Studio. Una marca es capaz de dirigir sus anuncios a los usuarios en función de su perfil de edad, ubicación, pero también según el estado de ánimo de la persona basado en las últimas canciones escuchadas. Así, un usuario que escucha música para trabajar no tendrá los mismos anuncios que un usuario que escucha música festiva, incluso si tienen el mismo perfil sociodemográfico.

El tercero y último es la personalización de los contenidos. Por ejemplo, LVMH ha invertido masivamente en su plataforma 24S para poder personalizar la experiencia de compra en línea, como ya lo han hecho Dior o Givenchy, sin sacrificar la percepción de exclusividad que define la posición misma de LVMH. Su solución, adaptar de manera dinámica el contenido mostrado, ya sean las imágenes, la descripción o incluso los productos destacados, y esto en función del

perfil de cada cliente. El objetivo es crear una experiencia similar a la de una tienda física pero a mayor escala.

Sin embargo, queda un límite. El objetivo mismo de todos estos mecanismos es maximizar el tiempo pasado, ya sea en la plataforma o en un sitio web. Por ejemplo, dentro de TikTok reside una paradoja. Si su propio algoritmo tiene como objetivo mantener al usuario el mayor tiempo posible, esto reduce fuertemente la eficacia publicitaria de las marcas que pagan por ser mostradas en la plataforma. Para las empresas, el principal desafío sigue siendo la comprensión de los mecanismos que rigen todas estas herramientas para utilizarlas de manera óptima y alineada con sus objetivos estratégicos.

Aunque el marketing digital permite un marco estratégico general. Es importante examinar de manera precisa cómo se explotan concretamente las huellas digitales y qué aportan en términos de comprensión del consumidor.

3.2 Usos de la explotación de las huellas digitales para entender la conducta y cultura digital

Actualmente, analizar las huellas digitales se ha convertido en un desafío estratégico importante para las empresas contemporáneas. De hecho, las empresas pueden no solo comprender los comportamientos de compra, sino también identificar las tendencias culturales y anticipar la evolución del mercado. Como muestran Boyd y Crawford (2012), estos datos permiten alcanzar un nivel de precisión sin precedentes en las prácticas sociales a gran escala, abriendo así la caja de Pandora para las estrategias de las empresas que saben interpretarlas. Para comprender cómo funciona esto, es interesante estudiar tres mecanismos concretos que ilustran tanto su diversidad como su complejidad.

El primero es la segmentación comportamental. La segmentación se basaba tradicionalmente en criterios «fijos» como la edad, el género, los ingresos... La segmentación comportamental, por su parte, se basa en las acciones del usuario, búsqueda en internet, producto consultado, lista de espera... Como muestran Wedel y Kannan (2016), este tipo de segmentación permite identificar patrones de comportamiento mucho más predictivos de la intención de compra que los criterios demográficos clásicos. Es un instrumento aún más poderoso que gigantes del comercio electrónico como Amazon lo han vinculado con su sistema de "logística preventiva". De hecho, el algoritmo anticipará la demanda en ciertas zonas bajo restricciones de costo logístico y ubicación, permitiendo una mejor disponibilidad para los clientes y reduciendo así los plazos de entrega. En otras palabras, las huellas digitales no solo son útiles para la personalización de la experiencia del cliente, sino también para optimizar toda la cadena logística.

El segundo mecanismo utilizado se llama Retargeting, que describe el hecho de redirigir mensajes publicitarios a usuarios que han visitado un sitio web sin haber finalizado su compra (por ejemplo, añadir un iPhone al carrito). Matz et al (2017) demuestran que un buen retargeting permite una tasa de conversión notablemente superior a la media de las generadas por la publicidad tradicional. La compañía Meta ha construido su modelo económico alrededor de

este principio gracias al píxel de Meta. Concretamente, es una pequeña etiqueta de código instalada en los sitios web que se recopila posteriormente cuando un usuario se reconecta en la plataforma Meta (Instagram, Facebook...). Una marca que lo utiliza diariamente es Sephora. De hecho, cuando un cliente ha puesto un producto en su carrito sin finalizar su compra, ese mismo artículo aparecerá unas horas más tarde en su feed de noticias, a veces acompañado de una oferta promocional para atraer nuevamente al consumidor y empujarlo a la compra. Es gracias a este nivel de precisión que una publicidad digital se transforma en una verdadera conversación continua entre marca, consumidor y gran plataforma de red social.

El tercer mecanismo es el marketing basado en datos, un verdadero activo de marca. Como muestra Kozinets (2010), las comunidades digitales crean un flujo constante que las empresas pueden analizar posteriormente, pero también es una poderosa ventaja de marketing. Varias marcas de renombre como Netflix o Spotify lo utilizan. De hecho, Spotify, a través de su funcionalidad "Wrapped", permite cada año en diciembre republicar en versión resumida las músicas que el usuario ha escuchado más. Resultado, millones de usuarios comparten al mismo tiempo el resumen de su mejor artista o álbum en las redes sociales, transformando así los datos de comportamiento en una campaña de marketing a gran escala y a bajo coste.

Sin embargo, Pariser (2011) muestra que este tipo de explotación excesiva de datos puede crear una visión distorsionada del consumidor. Al analizar únicamente los comportamientos pasados del consumidor. El algoritmo puede categorizarlo solo en un único tipo de gusto o personalidad, creando así un bucle y una imagen estática del consumidor. En realidad, los gustos evolucionan constantemente. Mayer-Schönberger y Cukier (2013) refuerzan esta advertencia recordando que una correlación no es sinónimo de causalidad. Zalando fue uno de los primeros en darse cuenta de esto porque su algoritmo estereotipaba demasiado a los clientes, limitando así el potencial de gasto en su plataforma. Esta limitación muestra bien que las huellas digitales, aunque poderosas, siguen siendo una herramienta y no reemplazan totalmente la importancia de los estudios cualitativos que aportan más contexto y sutileza en el análisis.

3.3. Retos y corrientes actuales en la toma de decisiones estratégicas

La toma de decisiones en las empresas ha evolucionado profundamente con el uso masivo de datos y la explotación de las huellas digitales. Sin embargo, no solo resulta en aspectos positivos como oportunidades, sino también en importantes desafíos que lo acompañan. De hecho, el libre acceso a tal cantidad de información no garantiza necesariamente tomar la decisión más apropiada. Así, el verdadero desafío reside en la capacidad de cada uno para interpretar correctamente estos datos e integrarlos de la mejor manera posible en la estrategia de la empresa.

Como señalan Mayer-Schönberger y Cukier (2013), paradójicamente, el primer desafío es la sobrecarga de información. De hecho, el big data se caracteriza por un volumen y una velocidad de información sin precedentes. Esta abundancia de información permite así tener numerosos indicadores a veces contradictorios o muy contextuales. Estas últimas pueden generar confusión e inducir decisiones poco coherentes. Por ejemplo, una empresa puede enfrentarse a tasas de

clics altas en una campaña publicitaria con una baja tasa de conversión. Comprender la causa de esta divergencia es la clave.

El segundo desafío importante es el de los sesgos algorítmicos. De hecho, como muestra Cathy O'Neil (2016), los algoritmos dependen en gran medida de los datos sobre los que han sido entrenados, lo que da lugar a algoritmos que no son neutros. Si estos datos contienen sesgos, entonces el algoritmo los reproducirá a su vez e incluso los amplificará. Por ejemplo, un algoritmo puede detectar que un perfil de consumidor es estratégico para una empresa y así tendrá tendencia a reforzar esta posición a través de estrategias de marketing. Sin embargo, el hecho de elegir este eje estratégico tenderá a excluir otros perfiles que podrían tener un mayor potencial. Esto limita a las empresas a explorar nuevos mercados, reforzando las desigualdades existentes.

Como ya hemos mostrado anteriormente, la interpretación errónea entre correlación y causalidad puede llevar a estrategias ineficaces, e incluso contraproducentes. Ante estos desafíos, surgen nuevos modelos llamados híbridos o interdisciplinarios. Como propone John W. Creswell (2014), un enfoque mixto tiene la virtud de combinar tanto un análisis cuantitativo como la interpretación cualitativa. En el mundo empresarial, esto se traduce a través de la integración de múltiples perfiles, tales como: el analista de datos, expertos en marketing, sociólogos... El objetivo es no solo analizar los datos, sino comprender su significado.

Así, el hecho de depender cada vez un poco más de los datos puede constituir un freno en el crecimiento de las empresas. De hecho, como lo muestra Shoshana Zuboff (2019), las empresas tienden la tendencia a tomar sus decisiones basándose en datos cuantificables en detrimento del análisis cualitativo. Sin embargo, no todo es medible. Los factores emocionales, culturales o incluso las motivaciones profundas del consumidor no pueden ser retranscritos de manera fiel a través de análisis cuantitativos. Por ejemplo, una disminución de las ventas relacionada con la percepción de la marca, como fue el caso de Tesla en 2024, no aparece directamente en los indicadores cuantitativos.

Finalmente, otro desafío que no es menor es la temporalidad de la base de datos. De hecho, el entorno produce constantemente nuevos datos, el comportamiento de los consumidores evoluciona en tiempo real. Una estrategia basada en datos que no son frescos puede volverse obsoleta muy rápidamente. Por ejemplo, una tendencia estudiada en redes sociales como TikTok puede desaparecer en unas pocas semanas o meses. Esto obliga a las empresas a adaptarse diariamente, donde la toma de decisiones es un proceso dinámico.

Es en este contexto que la toma de decisiones estratégicas no puede basarse únicamente en la recopilación de datos. El análisis de estos datos debe basarse en un espíritu crítico y una interpretación contextualizada. El verdadero desafío para las empresas radica en disponer de datos recientes, pero también en poder transformarlos en decisiones pertinentes, sostenibles y adecuadas.

4. Investigación empírica: aplicación de las huellas digitales

La transformación digital también impacta el sector del lujo. En 2024, LVMH ya dedicaba una parte importante de su presupuesto de marketing a los canales digitales (11,5%, The robin Report). Las proyecciones indican que el comercio electrónico representará entre el 28% y el 30% del sector del lujo al plazo 2025-2030 (Bain & Company). Para empresas como Louis Vuitton, comprender lo que los consumidores piensan, sienten y comparten es fundamental. Cada comentario, cada publicación, cada mención de la marca en las redes sociales que constituye un fragmento de datos puede ser utilizado a través de las humanidades digitales.

Así, esta parte tiene como objetivo utilizar las humanidades digitales y sus marcos teóricos desarrollados anteriormente en un caso concreto, Louis Vuitton. Se articula en cinco puntos : la selección del terreno de estudio (4.1), la recopilación de datos (4.2), el procedimiento de análisis (4.3), los resultados y su discusión (4.4), y finalmente las limitaciones y oportunidades identificadas (4.5).

4.1. Selección de empresas o sectores

El hecho de elegir Louis Vuitton por esta investigación no es casual. De hecho, fundada en 1854 es hoy un buque insignia de LVMH, el primer grupo mundial sobre los productos de lujo, cuyo volumen de negocio alcanzó los 80,8 mil millones de euros en 2025 (LVMH 2026 *Rapport annuel 2025*). Louis Vuitton representa un caso particular con la coexistencia entre un legado cultural secular y una presencia digital masiva. La marca deja en el mismo momento huellas pasivas (Datos de navegación, comportamiento de compra en línea y geolocalización) y las huellas activas (publicaciones en redes sociales, comentarios, avisos), formando un ecosistema de datos importantes. Cada interacción digital de un consumidor con Louis Vuitton constituye una parte de esta huella colectiva que revela no solo los comportamientos individuales sino también tendencias culturales con la percepción del lujo en la era digital.

La pertinencia de esta empresa también se basa en la riqueza de los datos publicados. De hecho, Louis Vuitton adopta un enfoque multiplataforma (55,5 millones de suscriptores en Instagram, 25,5 millones en Facebook, 14 millones en TikTok), la marca genera así un volumen considerable de datos sociales a explotar. Esta presencia digital multiplataforma convierte a LV en un observatorio ideal para analizar cómo los consumidores de lujo construyen, comparten y negocian su identidad en el espacio digital.

Así, la dimensión intercultural de LV la convierte en un terreno de estudio particularmente adecuado para examinar las tendencias culturales globales en torno al lujo digital.

4.2. Recolección de datos: desde las plataformas de redes sociales, la opinión de los clientes

Para recoger los datos relevantes sobre el comportamiento digital de los consumidores de lujo es importante movilizar fuentes diversas que se pueden complementar. En el caso de LV, se han identificado dos grandes categorías : las plataformas de redes sociales como TikTok y las opiniones de los clientes a través de Trustpilot o Google review.

En primer lugar, sobre los datos procedentes de TikTok, los datos fueron recopilados directamente en la plataforma utilizando la herramienta de extracción Octoparse a partir de publicaciones asociadas a los hashtags #LouisVuitton, #LVBag y #LVunboxing. Así, este proceso permitió constituir un corpus de 575 comentarios de fuentes diversas, desempaqueado (unboxing), críticas de productos, contenidos aspiracionales o, por el contrario, contestatarios. Concretamente, los datos fueron recopilados a partir de tres perfiles de videos distintos y complementarios.

El primer perfil corresponde a videos de “unboxing” publicados por usuarios particulares cuyos comentarios reflejan un discurso dominado por el deseo aspiracional y la comparación de modelos. El segundo perfil es el de un asesor de ventas de Louis Vuitton (client advisor) cuya sección de comentarios revela una comunidad de consumidores comprometidos, expertos y en busca de consejos. El tercero es un video de naturaleza contestataria, que plantea públicamente la cuestión de la fabricación de los bolsos Louis Vuitton en China. Además, se han integrado datos provenientes del perfil oficial de Louis Vuitton en TikTok (@louisvuitton), con publicaciones relacionadas con la dirección artística de Pharrell Williams.

En segundo lugar, las opiniones de los clientes recopiladas a través de plataformas como Trustpilot y Google Reviews (Paris, Madrid, Roma) también gracias a Octoparse. Al estar exentas de cualquier control editorial por parte de la marca, capturan la voz auténtica del consumidor: sus frustraciones, sus expectativas y sus experiencias vividas tanto en la tienda como en línea. Este tipo de fuente complementa lo que se obtiene en las redes sociales, donde la dimensión aspiracional puede a veces restar sinceridad al discurso.

4.3. Procedimiento de análisis

4.3.1. Preprocesamiento de datos (eliminación de datos incorrectos, tokenización, lematización).

El pretratamiento constituye la primera etapa fundamental de cualquier análisis textual automatizado. El objetivo es transformar los datos brutos que son por definición heterogéneos en un corpus normalizado que se puede utilizar por los algoritmos de análisis. El preprocesamiento se articuló en tres fases sucesivas: la limpieza de los datos en bruto, la tokenización y la lematización con eliminación de las palabras vacías (stopwords). Para ello, se utilizó la herramienta Claude IA para generar el código de análisis que luego ejecuté, permitiéndome automatizar esta tarea mientras mantenía una visibilidad completa sobre el conjunto de datos que fueron conservados o eliminados.

1. Preprocesamiento de datos que proviene de TikTok.

1) Ejemplo de limpieza de los datos en bruto.

<u>TIPO BRUTO</u>	<u>Ejemplo bruto</u>	<u>Después de la limpieza</u>
<u>URL y enlaces</u>	BellarteLuxShop.com	(eliminado)
<u>Menciones @usuario</u>	@Voguerunwaydubai	(eliminado)
<u>Emojis y caracteres no ASCII</u>	🥰💕🌟🥰	(eliminado)
<u>Caracteres especiales y puntuación</u>	Omg, just a few days ago!	omg just a few days ago
<u>Mayúsculas → minúsculas</u>	LOVE IT, So CUTE	love it so cute

2) Ejemplo de la tokenización.

Comentario limpio : "omg just a few days ago i got the exact same color"

Tokens generados: ["omg", "just", "a", "few", "days", "ago", "i", "got", "the", "exact", "same", "color"]

3) Ejemplo de Lematización.

La lematización reduce cada token a su forma canónica (lema), independientemente de su conjugación, declinación o número. Así:

- "sizes" → "size"
- "bought" → "buy"
- "feels" → "feel"

Esta normalización es crucial para evitar la fragmentación de los resultados durante los análisis de frecuencia y coocurrencia.

4) Ejemplos completos de la cadena de pretratamiento

Etapa	Contenido
Comentario bruto	Omg, just a few days ago I got the exact same color from my vendor and honestly I love it. The leather feels so soft! I got mine from @Voguerunwaydubai on Instagram ❤️
Limpieza	omg just a few days ago i got the exact same color from my vendor and honestly i love it the leather feels so soft
Tokenización	["omg", "just", "a", "few", "days", "ago", "i", "got", "the", "exact", "same", "color", ...]
Lematización	["omg", "day", "ago", "get", "exact", "color", "vendor", "honestly", "love", "leather", "feel", "soft"]
Observación	Eliminación de @mención, emoji y puntuación. "days" → "day", "feels" → "feel", "got" → "get".

2. Preprocesamiento de datos que proviene de Trustpilot y Google Review.

Ejemplo 1 : Aviso positivo (Trustpilot).

<u>Etapa</u>	<u>Resultado</u>
Comentario bruto	Love it. Purchased a weekend bag and we had a great experience.
Limpieza	love it purchased a weekend bag and we had a great experience
Tokenización	["love", "it", "purchased", "a", "weekend", "bag", "and", "we", "had", "a", "great", "experience"]
Lematización	["love", "purchase", "weekend", "bag", "great", "experience"]
Observación	"purchased" → "purchase" : retour à l'infinitif. Stopwords ("it", "a", "we", "had") supprimés.

Ejemplo 2 : Aviso negativo (Trustpilot)

<u>Etapa</u>	<u>Resultado</u>
Comentario bruto	I'm very disappointed with this company. The items aren't worth the price. The canvas print is peeling off the bag.
Limpieza	im very disappointed with this company the items arent worth the price the canvas print is peeling off the bag
Tokenización	["im", "very", "disappointed", "with", "this", "company", "the", "items", "arent", "worth" ...]
Lematización	["disappointed", "company", "item", "worth", "price", "canvas", "print", "peel", "bag"]
Observación	"items" → "item", "peeling" → "peel" : normalisation morphologique.

Ejemplo 3 : Aviso en frances (Google Reviews Madrid).

<u>Etapa</u>	<u>Resultado</u>
Comentario bruto	Très mauvaise expérience. Il y avait trop de monde. Le service exécrable. Aucun respect pour les clients.
Limpieza	trs mauvaise exprience il y avait trop de monde le service excrable aucun respect pour les clients"
Tokenización	["trs", "mauvaise", "exprience", "il", "avait", "trop", "monde", "service", "excrable" ...]
Lematización	["mauvaise", "exprience", "trop", "monde", "service", "excrable", "respect", "client"]
Observación	Los acentos franceses desaparecen durante la limpieza ASCII (limitación de la plantilla en inglés). Para un procesamiento óptimo, sería necesaria una plantilla específica en francés (spaCy fr).

4.3.2. Análisis de textos y de la sociedad (análisis de las emociones, minería textual)

Una vez completada la etapa de pretratamiento del corpus hecho, el análisis se basó en dos enfoques que son complementarios, el análisis de sentimientos y emociones, así como la minería textual temática. El análisis de sentimientos que fue realizado con el modelo VADER, muestra una diferencia notable entre las dos plataformas. De hecho, los comentarios procedentes de TikTok, presentan una tonalidad globalmente positiva (puntuación media de +0,208), con solo un 11,2 % de contenidos negativos. Al contrario,

los avisos procedentes de Google Reviews y Trustpilot presentan una tonalidad neutra (puntuación media de +0,077) y una proporción de comentarios negativos que alcanza el 30%. Esta diferencia se puede explicar por la naturaleza de las plataformas. Entonces, TikTok refleja más una relación aspiracional (significa aquí el deseo de pertenecer simbólicamente a un universo) con el lujo, basada en el deseo y la admiración mientras que las opiniones de los clientes a través de Trustpilot muestran la experiencia real después de la compra.

La minería textual temática revela también diferencias significativas entre los corpus. En TikTok, tres grandes temas estructuran las discusiones: el deseo de compra y el atractivo por los productos, los debates sobre los precios y la accesibilidad, así como las controversias sobre el origen de fabricación de los productos. Estos intercambios subrayan la importancia de la autenticidad en la percepción del lujo dentro de los espacios digitales.

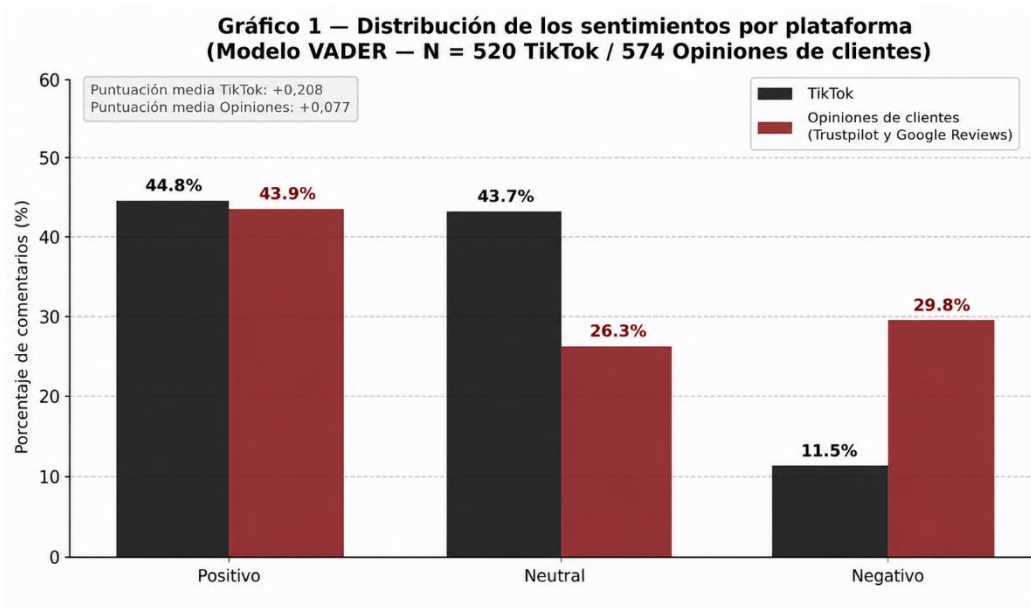
En las opiniones de los clientes, los temas negativos más frecuentes se refieren a la calidad del servicio al cliente por ejemplo cuando el volumen de clientes es elevado, las dificultades relacionadas con las reparaciones, la relación calidad-precio, los problemas de entrega, las experiencias de discriminación en la tienda y los defectos de calidad de los productos. Por su parte, las reseñas positivas se centran principalmente en la calidad de los productos, la experiencia en tienda, la fidelidad a la marca y la excelencia del servicio. Los resultados muestran que las expectativas de los consumidores de lujo superan con creces el producto en sí: la calidad del acompañamiento y la experiencia relacional juegan un papel central en la satisfacción. Los testimonios relativos a la discriminación en boutique también ponen de manifiesto cuestiones sociales y culturales que van más allá del simple marco comercial e ilustran el interés de las humanidades digitales para el estudio del sector del lujo.

En conjunto, este análisis pone de relieve la existencia de dos formas complementarias de huella digital: una huella aspiracional e identitaria en TikTok, centrada en el deseo y la representación del lujo, y una huella experiencial en las opiniones de los clientes, más vinculada a la evaluación concreta de los productos, los servicios y la experiencia de consumo.

4.3.3. Representación gráfica de datos (gráficos, tablas, tableros)

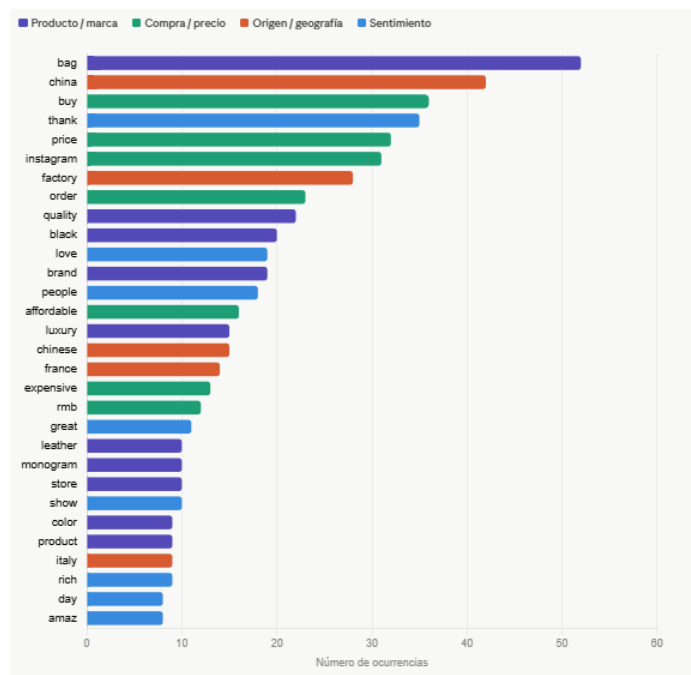
Una etapa clave de este análisis es la representación gráfica. En efecto, esta etapa permite transformar gran cantidad de datos e información de manera visual clara y accesible. Gracias a ella, entender rápidamente las tendencias y comparar los resultados más relevantes. Esta etapa además ayuda a facilitar la interpretación de los datos y así tener una mejor comprensión de los fenómenos observados. Los gráficos que se muestran en el presente estudio fueron realizados con herramientas Excel, WordCloud y Canva con la información previamente extraída y procesada. Utilizaron estas herramientas para mostrar claramente la información y, sobretodo, apelando a los objetivos del análisis, lucidamente visualizados.

1. Distribución de los sentimientos por plataforma.



El primer gráfico compara la distribución de sentimientos dentro de las plataformas analizadas. Tenemos en negro la plataforma TikTok y en rojo las opiniones de clientes de la plataforma Trustpilot y Google Reviews. La diferencia estructural entre ambas plataformas es inmediatamente visible. De hecho, de un lado tenemos TikTok que concentra 44,8 % de comentarios positivos y solo el 11,5 % de negativos, los avisos de clientes presentan una proporción de negativos que alcanza el 29,8 %, con un score medio cuatro veces inferior (+0,077 frente a +0,208). Este contraste ilustra también la diferencia fundamental entre una plataforma aspiracional y una plataforma de evaluación post-compra.

2. Análisis de las lemas por TikTok.



El análisis de los lemas mas frecuentes ha permitido identificar cuatro grandes clústeres temáticos que son presentes en los comentarios de los usuarios en TikTok sobre Louis Vuitton.

- **Cluster 1 – Producto y marca.**

Palabras claves : Quality, black, Brand, luxury, leather y monogram .

Estas palabras muestran la importancia que ponen los usuarios sobre las características de los productos. De hecho la estética, el tamaño y los materiales utilizados son muy importante por los clientes. Así, este clúster refleja una relación aspiracional con el producto, percibido no solo como un accesorio funcional, sino también como un símbolo de estatus y prestigio.

- **Cluster 2 – Compra y precio.**

Palabras claves : buy, Price, factory, order, affordable, expensive.

Estas palabras muestran que los usuarios son atentos sobre la relación calidad-precio del producto. También en numerosos comentarios muchos usuarios comparaban los precios oficiales de Louis Vuitton con los precios propuestos por fabricante o revendedores especialmente chino. Este fenómeno refleja una creciente discusión sobre la accesibilidad de los productos de lujo y la existencia de circuitos paralelos de compra promovidos a través de las redes sociales.

- **Cluster 3 – Origen geográfico**

Palabras claves : China, chinese, France, Italy.

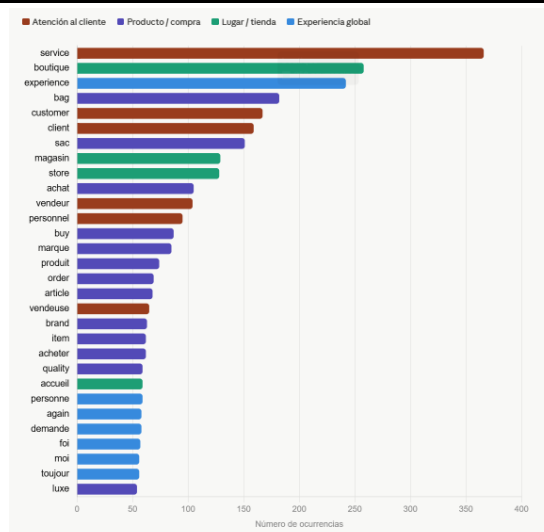
Los usuarios de TikTok debaten de manera activa sobre el país de fabricación de los productos así como la deslocalización masiva y la autenticidad de los productos. Comprar un producto de lujo es también comprar un saber hacer y una reputación.

- **Cluster 4 – Sentimiento, compromiso**

Palabras claves : thank (gracias), love , people, great, amaz (incredible), show.

Este tipo de palabra muestra la importancia de pertenecer a una marca. Hay un tono globalmente positivo y aspiracional, típico de los comentarios en TikTok.

3. **Análisis de las lemas por Trustpilot y Google reviews.**



- **Cluster 1 : Atención al cliente**

Palabras claves : service (servicio), customer, client (cliente), vendeur (vendedor).

Estas palabras representan las expectativas altas que tienen los compradores sobre a la atención que se les presta especialmente en el sector del lujo.

- **Cluster 2 : Lugar de venta**

Palabras claves : boutique (tienda), store, accueil (el punto de atención).

Estas palabras tienen similitudes con el atención al cliente pero relevan la presencia de reseñas en ambos idiomas. La experiencia en tienda física predomina claramente sobre la compra online.

- **Cluster 3 : Producto y compra**

Palabras claves : (bag, sac (bolsa), achat (compra), order, quality, Brand)

Estas palabras muestran la importancia que ponen los usuarios sobre las características de los productos. Además, los compradores tienen expectativas muy altas por el precio y la reputación de la marca.

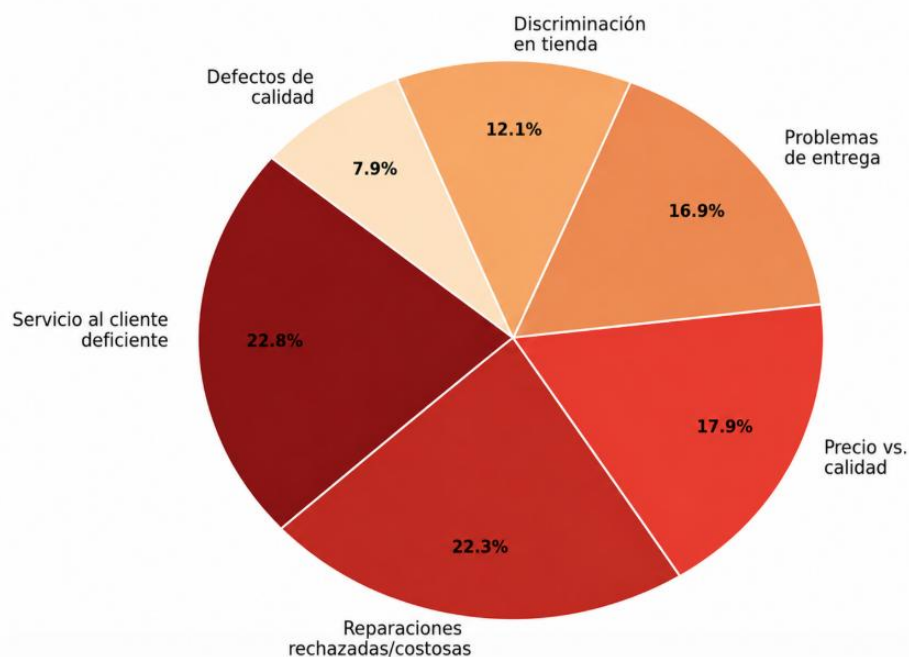
- **Cluster 4 : Experience global**

Palabras claves : experience (experiencia), again, toujours (siempre)

Clúster evaluativo donde los clientes emiten un juicio global sobre su recorrido de compra. La presencia de again y toujours (siempre) señala recurrencia, ya sea de clientes fieles o de problemas que se repiten.

4. Análisis de insatisfacción de los opiniones de clientes.

Gráfico 4 — Distribución de los temas negativos en las opiniones de clientes (Trustpilot y Google Reviews)



El gráfico muestra la distribución de los principales motivos de insatisfacción expresados por los clientes en Trustpilot y Google Reviews. Las críticas se dirigen principalmente al deficiente servicio al cliente (22,8 %) y a las reparaciones rechazadas o consideradas demasiado caras (22,3 %), que suman casi la mitad de las opiniones negativas. Destacan también las quejas sobre la relación calidad-precio (17,9 %) y problemas de entrega (16,9 %), lo que apunta a una percepción de valor insuficiente y dificultades en la experiencia del cliente. En menor medida, los usuarios hacen referencia a situaciones de discriminación en tienda (12,1 %) y defectos de calidad (7,9%).

4.4. Resultados y discusión: patrones de comportamiento y tendencias culturales

Los resultados y conclusión del análisis empírico son obtenidos gracias a un corpus total de 1 109 comentarios procedentes de TikTok (N=534) y de las plataformas de opinión de clientes Trustpilot y Google Reviews (N=575). Esta base de dato real ha permitido identificar cinco patrones de comportamiento y tendencias culturales en relación con los consumidores con Louis Vuitton en el espacio digital. Así, estos patrones ilustran de manera concreta en que dimensión las huellas digitales relevan dinámicas que trascienden el simple acto de compra para inscribirse en fenómenos sociológicos y culturales más amplios.

En primer lugar, los resultados muestran una diferencia importante entre TikTok y las plataformas de opiniones de clientes (Trustpilot y Google Reviews). De hecho, de un lado hay TikTok que presenta una predominancia de comentarios positivos. De otro lado, hay Trustpilot y Google Reviews concentran una proporción mucho más elevada de opiniones negativas. Esta diferencia puede explicarse por la naturaleza de cada plataforma. TikTok funciona principalmente como un espacio digital de inspiración y construcción de imagen. Los usuarios interactúan con contenidos aspiracionales asociados al lujo, la moda y el prestigio social. Al contrario, las plataformas de reseñas son utilizadas mayoritariamente por consumidores que desean compartir una experiencia concreta de compra, especialmente cuando esta ha sido insatisfactoria.

En segundo lugar, el producto es un elemento central de la conversación sobre Louis Vuitton. Tanto en TikTok como en las plataformas de opiniones aparecen términos relacionados con los bolsos, la calidad, la marca y el lujo. Esto confirma que en el mente de los consumidores hay una asociación clara dentro Louis Vuitton y los atributos tradicionales de exclusividad y prestigio. Sin embargo, este tipo de expectativa eleva mucho las exigencias en torno a la relación calidad-precio. La presencia recurrente de términos como “price”, “expensive” y “affordable” en el corpus sugiere una creciente sensibilidad económica, incluso dentro del mercado del lujo. Los usuarios muestran interés para buscar alternativas más accesibles con una calidad similar, lo que refleja un acceso fácil a la información sobre productos de lujo.

Otro aspecto relevante es la importancia del país de fabricación y la percepción de autenticidad. Hay una frecuencia de palabras como China, Francia o Italia que demuestra la preocupación de los consumidores y el valor que ponen al origen geográfico de los productos. Para muchos consumidores, el lujo no se limita únicamente al producto final sino como el producto final es

así. Los símbolos como saber hacer artesanal, la tradición y la procedencia son elementos muy importantes por los consumidores de lujo. El debate constante sobre la producción en china ponen de manifiesto las tensiones existentes entre la globalización, la cadena de valor mundial, y las expectativas de autenticidad asociadas a las marcas de lujo.

Asimismo, gracias a el análisis de las opiniones negativas, una gran parte de insatisfacción de los consumidores no están relacionadas directamente con el producto. De hecho, la insatisfacción está asociada con los servicios asociados a la experiencia de compra de un producto de lujo. El servicio al cliente, las reparaciones y los problemas logísticos representan la mayoría de las críticas observadas. Dado el elevado precio de los artículos, las expectativas respecto al servicio son especialmente altas, y cualquier incumplimiento genera una fuerte insatisfacción.

Los redes sociales han cambiado el papel de los consumidores, ya no son sólo compradores, sino también creador y difusor masivo de contenidos. En la plataforma TikTok, los usuarios publican una gran cantidad de comentarios (debates, opiniones y recomendaciones), lo que puede influir en la imagen de marca, que ya está bien consolidada. Las plataformas de reseñas también se convierten en espacios de expresión y evaluación colectiva donde los consumidores influyen directamente en la reputación de la empresa. Esta dualidad muestra que hoy en día la percepción de una marca de lujo se construye a través de múltiples canales digitales y depende tanto de la comunicación corporativa como de las experiencias compartidas por los propios consumidores.

En conjunto, los resultados sugieren que Louis Vuitton cuenta con una imagen aspiracional muy sólida en las redes sociales, pero enfrenta retos relacionados con la gestión de la experiencia de sus clientes. La mezcla de admiración por la marca y las críticas a ciertos aspectos del servicio muestra las nuevas demandas de los consumidores digitales, que esperan una perfecta coherencia entre la promesa de lujo comunicada por la marca y la experiencia recibida.

4.5. Obstáculos, limitaciones y oportunidades en la recolección

Obstáculos encontrados

El primer obstáculo encontrado fue de orden técnico. La herramienta Octoparse, utilizada para la extracción de datos en TikTok, Trustpilot y Google Reviews, presentó limitaciones importantes en términos de volumen de datos accesibles en su versión gratuita. Las plataformas como TikTok restringen activamente el acceso automatizado a sus datos mediante sistemas anti-scraping, lo que obligó a realizar extracciones manuales y parciales. Esta restricción limita la representatividad del corpus y reduce la posibilidad de análisis a gran escala que permitirían conclusiones más sólidas.

Limitaciones del análisis

La primera limitación se el sesgo de selección. Los datos recopiladas en TikTok solo proviene de cuatro videos específicos elegido también por los temas tratados. Por ejemplo el debate sobre la fabricación en china esta sobre representado porque estaba un video viral con 2,4 millones de visualizaciones. Un corpus más diversificado, construido a partir de más videos produciría resultados más equilibrados.

La segunda limitación es de orden lingüístico. El modelo principal utilizado por el analisis de sentimientos, VADER, estaba hecho al inicio por el inglés. Aunque el corpus contiene avis en francés, español e incluso mandarin. Así pierde un poco de precisión. Como esta posible ver en el 4.3.1, los acentos de los comentarios franceses desaparecen durante la limpieza ASCII, generando así lemas incompletos como "exprience" en lugar de "expérience". Un análisis verdaderamente multilingüe requeriría modelos específicos por idioma, como spaCy en francés por ejemplo.

Finalmente, el análisis presenta una limitación temporal : los datos fueron recopilados entre octubre de 2025 y junio de 2026, lo que constituye una fotografía puntual de un ecosistema digital en constante evolución. Los patrones identificados podrían variar significativamente en función de los eventos de marca, las campañas publicitarias o las controversias surgidas fuera de este período.

Oportunidades identificadas

A pesar de las limitaciones presentadas, la recolección de huellas digitales abre perspectivas analíticas considerable especialmente por los comercios de proximidad. De hecho, con la gratuidad y la accesibilidad de las fuentes utilizadas tal como TikTok, Trustpilot y Google reviews es posible construir un corpus rico sin necesidad de grandes recursos financieros.

En segundo lugar, el hecho de cruzar las fuentes ha resultado ser una fortaleza metodológica real. En efecto, captar la dimensión aspiracional pre compra a través de TikTok y la dimensión evaluativa post compra a través de Trustpilot y Google Review ha permitido obtener una vision completa del comportamiento del consumidor de lujo.

Finalmente, la presencia del debate sobre la fabricación en china ha permitida tener datos espontáneos que ninguna encuesta dicha tradicional podría captar. Este tipo de resultado emergente, ilustra la capacidad de las huellas digitales para revelar tendencias culturales latentes que escapan a los métodos de investigación clásicos.

5. Conclusión

5.1 Resultados principales.

El objetivo de este trabajo es examinar de que manera las huellas digitales y las humanidades digital puede ser utilizado para comprender el comportamiento de los consumidores y las tendencias culturales en el sector del lujo, tomando como caso de estudio la marca Louis Vuitton para ilustrarlo. A través de un recorrido que ha articulado fundamentos teóricos, metodología empírica y análisis de datos reales, es posible formular un balance en tres dimensiones.

En el plano teórico, este trabajo ha permitido demostrar que las humanidades digitales ofrecen un marco conceptual interesante y pertinente para el análisis del comportamiento del consumidor. De hecho, la noción de huella digital se ha relevado como una herramienta analítica mayor y útil. Además, las aportaciones de autores reconocido en esta área como Kotler y Keller (2016) sobre la evolución del comportamiento del consumidor, cruzadas con los marcos de las humanidades digitales, han permitido construir una base teórica sólida.

En el plano empírico, gracias a el análisis de los 1 109 comentarios procedentes de TikTok, Trustpilot y Google Reviews ha producido resultados concretos y originales. Los cinco patrones de comportamiento y tendencias culturales identificados así como la diferencia emocional entre plataformas, la controversia sobre el origen de fabricación de los productos, el servicio de atención al cliente y por fin la dualidad entre huella aspiracional y huella experiencial ilustran perfectamente la complejidad de la relación entre los consumidores de lujo y la marca Louis Vuitton en la era digital.

En el plano metodológico, este trabajo ha demostrado la viabilidad de una investigación empírica sobre las huellas digitales con recursos accesibles. La combinación de Octoparse para la extracción de datos, Claude IA, Python y NLTK para el preprocesamiento y la lematización, y VADER para el análisis de sentimientos, constituye una cadena metodológica reproducible y transferible a otros contextos de investigación. Hay solo que tomar en cuenta las limitaciones identificadas tal como el sesgo lingüístico o la representatividad parcial del corpus.

En última instancia, este trabajo invita a considerar las huellas digitales como un objeto de estudio a pleno derecho en el campo de las humanidades digitales. En un mundo donde cada interacción en línea deja una traza, la capacidad de leer e interpretar estas huellas se convierte en una competencia esencial, tanto para los investigadores como para los profesionales del sector del lujo.

5.2 Consecuencias prácticas para la comprensión cultural y estrategias digitales

Como se ha visto, la sección anterior ha presentado los aportes teóricos y metodológicos de este trabajo. El objetivo de esta última sección se centra en sus consecuencias concretas para los actores del sector del lujo. Los resultados empíricos no tienen solo un interés académico. Generan implicaciones concretas en prácticas directas tanto para la comprensión cultural del consumidor como para el diseño de estrategias digitales.

Comprender al consumidor de lujo en la era digital. El primer aporte práctico de este trabajo es demostrar que el consumidor de lujo contemporáneo no puede ser comprendido a través de una sola fuente de datos. El análisis comparado de TikTok y de las plataformas de avis ha revelado que un mismo consumidor puede expresar entusiasmo aspiracional en las redes sociales y frustración profunda en sus opiniones post-compra. Las marcas que limiten su escucha digital a los canales que ellas mismas controlan (sus propias cuentas, sus campañas) se privarán de la parte más auténtica e informativa de las huellas digitales de sus clientes. La recolección sistemática y el análisis de datos multicanal deben convertirse en una práctica estándar de inteligencia de mercado en el sector del lujo.

Gestionar la reputación en tiempo real. El caso del vídeo viral sobre la fabricación en China ilustra la rapidez con la que una controversia puede propagarse y acumularse en forma de huella reputacional colectiva. Una estrategia de social “listening” activa, capaz de detectar estos fenómenos en sus fases iniciales, permitiría a las marcas responder con mayor agilidad, ya sea desmintiendo informaciones incorrectas, aportando transparencia sobre sus procesos de fabricación, o movilizando a sus comunidades de consumidores fieles como vectores de contra narrativa.

Priorizar la experiencia humana como ventaja competitiva. Los resultados muestran que la fuente principal de satisfacción como de la insatisfacción es en relación con el servicio al cliente. Esta conclusión esta hecha gracias a la clasificación temática de los comentarios siguiendo el método de codificación descrito por Creswell (2014). En un contexto de creciente digitalización de los puntos de venta y de desarrollo del comercio electrónico, este resultado invita a las marcas de lujo a no sacrificar el factor humano en nombre de la eficiencia digital. La formación y la fidelización del personal de boutique constituyen, a la luz de estos datos, una inversión estratégica de primer orden.

Anticipar las tensiones culturales emergentes. El análisis de las huellas digitales ha puesto de manifiesto dos fenómenos culturales que trascienden el marco comercial : el escepticismo creciente sobre la autenticidad de los productos de lujo, alimentado por la viralidad de los debates sobre el origen de fabricación, y las experiencias de discriminación en boutique relatadas por una proporción significativa de clientes. Estos fenómenos no son marginales en el corpus analizado, que se basan en solo una marca y no permiten por si solas de generalizar : son señales débiles de transformaciones culturales profundas que afectan a la percepción del lujo en sociedades cada vez más sensibles a las cuestiones de transparencia, inclusividad y justicia social. Las marcas que sepan leer estas señales en sus huellas digitales dispondrán de una ventaja considerable para adaptar su posicionamiento antes de que estas tensiones se conviertan en crisis abiertas.

6. Limitaciones de la investigación y caminos futuros para las indagaciones

Hay diversas limitaciones en esta investigación que deben ser consideradas para la buena interpretación de sus conclusiones. En primer lugar, el estudio se centró exclusivamente en una única marca del lujo (Louis Vuitton). Aunque esta empresa representa un caso de estudio interesante no representa todo el sector. Así, los resultados obtenidos no pueden ser generalizados automáticamente al conjunto de la industria del lujo en su globalidad. Otras marcas pueden así, en función de su posicionamiento, estrategia de comunicación, estrategia digital o su relación con los consumidores tener conclusiones diferentes.

En segundo lugar, el análisis se basa a partir de datos textuales procedentes de TikTok, Trustpilot y Google reviews. Aunque esta base de dato es rico, no reflejan las percepciones de todos los consumidores. Los usuarios que publican comentarios o reseñas representan solo una parte de los consumidores de la marca, lo que puede generar ciertos sesgos de representación. Además, generalmente las personas que expresan opiniones en línea suele ser asociadas a una particularmente positiva o negativa experiencia lo que pone en un fondo las experiencias mas neutras.

Otra limitación importante a tener en consideración es sobre las herramientas utilizadas para el tratamiento de los dato. Aunque los métodos de análisis de sentimiento y minería texto estaban respetados la dificultad para interpretar elementos como la ironía, el sarcasmos son difíciles a captar. Es por esta razón que los resultados deben ser interpretado como una tendencia general y no como una representación absoluta de las opiniones de los consumidores.

A partir de estas dichas limitaciones, se abren varias perspectivas para futuras investigaciones. En primer lugar, sería interesante ampliar el análisis a otras marcas importante del sector del lujo tal como Chanel, Hermes, Gucci o Dior. El objetivo es de comparar los patrones observados y determinar las específicas de cada uno para explicar el mercado del lujo en su globalidad.

En segundo lugar, las futuras investigaciones podrían incorporar fuentes diferentes como Instagram, X (antes Twitter) o WeChat (una red social asiáticos). La combinación de estas diversas ecosistema permitiría obtener una visión más amplia de las conversaciones y así más representativo de los consumidores.

Finalmente, la combinación de métodos cuantitativos y cualitativos constituye una vía especialmente prometedora. El hecho de integrar entrevistas o grupos de discusión con el análisis automatizado de huellas digitales permitiría enriquecer la interpretación de los resultados.

En definitiva, esta investigación demuestra el potencial de las humanidades digitales para el análisis del comportamiento del consumidor, al tiempo que pone de manifiesto la necesidad de seguir desarrollando enfoques interdisciplinarios capaces de combinar la potencia analítica de los datos masivos con la riqueza interpretativa de las ciencias humanas y sociales.

7. Fuentes bibliográficas

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). *Privacy and human behavior in the age of information*. *Science*, 347(6221), 509–514.

Berry, D. M. (2012). *Understanding digital humanities*. Palgrave Macmillan.

Boyd, D., & Crawford, K. (2012). *Critical questions for Big Data*. *Information, Communication & Society*, 15(5), 662–679.

Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste*. Harvard University Press.

Cambria, E., & White, B. (2014). *Jumping NLP curves: A review of natural language processing research*. *IEEE Computational Intelligence Magazine*, 9(2), 48–57.

Cambridge Analytica. *A case study*. Cambridge Analytica.

Chaffey, D., & Ellis-Chadwick, F. (2019). *Digital marketing: Strategy, implementation and practice* (7e éd.). Pearson.

CNIL. (2023). *Comprendre le RGPD*. Commission nationale de l'informatique et des libertés. <https://www.cnil.fr>

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4e éd.). SAGE Publications.

Drucker, J. (2011). *Humanities approaches to graphical display*. *Digital Humanities Quarterly*, 5(1).

eMarketer / Statista. (2024). *Share of e-commerce in total retail sales worldwide, 2021–2027*. Statista.

Fairclough, N., & Wodak, R. (1997). *Critical discourse analysis*. In T. A. van Dijk (Ed.), *Discourse as social interaction* (pp. 258–284). SAGE Publications.

Hearst, M. A. (1999). *Untangling text data mining*. *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics*, 3–10.

Hu, M., & Liu, B. (2004). *Mining and summarizing customer reviews*. *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 168–177.

Kaplan, A. M., & Haenlein, M. (2010). *Users of the world, unite! The challenges and opportunities of social media*. *Business Horizons*, 53(1), 59–68.

Kotler, P., Kartajaya, H., & Setiawan, I. (2021). *Marketing 5.0: Technology for humanity*. Wiley.

Kotler, P., & Keller, K. L. (2016). *Marketing management* (15e éd.). Pearson.

Kozinets, R. V. (2010). *Netnography: Doing ethnographic research online*. SAGE Publications.

- Liu, B.** (2012). *Sentiment analysis and opinion mining*. Morgan & Claypool Publishers.
- Marshall, A.** (1890). *Principles of economics*. Macmillan.
- Matz, S. C., Kosinski, M., Nave, G., & Stillwell, D. J.** (2017). *Psychological targeting as an effective approach to digital mass persuasion*. *Proceedings of the National Academy of Sciences*, 114(48), 12714–12719.
- Mayer, J., & Mitchell, J.** (2012). *Third-party web tracking: Policy and technology*. *Proceedings of the 2012 IEEE Symposium on Security and Privacy*, 413–427.
- Mayer-Schönberger, V., & Cukier, K.** (2013). *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.
- LVMH.** (2026). *Rapport annuel 2025*. LVMH Moët Hennessy Louis Vuitton.
- O'Neil, C.** (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishers.
- Pang, B., & Lee, L.** (2008). *Opinion mining and sentiment analysis*. *Foundations and Trends in Information Retrieval*.
- Pariser, E.** (2011). *The filter bubble: What the internet is hiding from you*. Penguin Press.
- Parlement européen.** (2016). *Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel (RGPD)*.
- Schreibman, S., Siemens, R., & Unsworth, J.** (Éds.). (2004). *A companion to digital humanities*. Blackwell Publishing.
- Sunstein, C. R.** (2017). *#Republic: Divided democracy in the age of social media*. Princeton University Press.
- The Robin Report.** (2024). *Luxury Brands Continue to Increase Digital Marketing Investment*
- van Dijk, T. A.** (Éd.). (1997). *Discourse as social interaction*. SAGE Publications.
- Wedel, M., & Kannan, P. K.** (2016). *Marketing analytics for data-rich environments*. *Journal of Marketing*, 80(6), 97–121.
- Zuboff, S.** (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.

8. Apéndices

Anexo aviso Trustpilot y Google review. [Pincha Aquí](#).

Anexo comentarios TikTok. [Pincha Aquí](#).

Declaración de Uso de Herramientas de Inteligencia Artificial Generativa en Trabajos Fin de Grado

Por la presente, yo Bastien YEFESS, estudiante de María Eugenia Ramos Fernández de la Universidad Pontificia Comillas al presentar mi Trabajo Fin de Grado titulado “humanidades digitales” declaro que he utilizado la herramienta de Inteligencia Artificial Generativa ChatGPT, Claude y DeepL sólo en el contexto de las actividades descritas a continuación.

1. **Búsqueda** : Identificación de pistas e ideas relevantes al inicio del trabajo, con el fin de orientar la reflexión en una dirección determinada. (ChatGPT, OpenAI)
2. **Reestructuración de párrafos** : Para ello, se utilizó Claude (Anthropic) para reorganizar algunos párrafos para mejorar la coherencia argumentativa de ciertas secciones, a partir de contenidos previamente redactados.
3. **Preprocesamiento de los datos** : Para ello, se utilizó la herramienta Claude (Anthropic) para generar el código de análisis que luego ejecuté, permitiéndome automatizar esta tarea mientras mantenía una visibilidad completa sobre el conjunto de datos que fueron conservados o eliminados.
4. **Traducción** : Para ello, se utilizó DeepL y ChatGPT (openAI) para generar traducción de algunas expresiones francés al español y vocabulario adecuado.
5. **Corrección y revisión de la sintaxis**. Es decir, ayuda para incorporar las correcciones y sugerencias formuladas por la directora del trabajo, reformulando los pasajes señalados de manera coherente con el resto del texto.

En todos los casos, los contenidos que fueron generados por la herramienta de la inteligencia artificial fueron revisados, adaptados y validados. La inteligencia artificial no ha sustituido en ningún caso el pensamiento crítico ni el trabajo intelectual, sino que ha actuado como una herramienta de apoyo.

Afirmo que toda la información y contenido presentados en este trabajo son producto de mi investigación y esfuerzo individual, excepto donde se ha indicado lo contrario y se han dado los créditos correspondientes (he incluido las referencias adecuadas en el TFG y he explicitado para que se ha usado ChatGPT u otras herramientas similares). Soy consciente de las implicaciones académicas y éticas de presentar un trabajo no original y acepto las consecuencias de cualquier violación a esta declaración.

Fecha: 26/06/2026

Firma:

