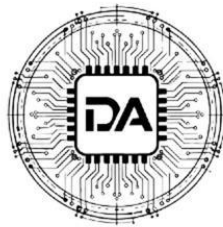


EL OCASO DE LA ORIGINALIDAD Y LA CRISIS DE LA VERACIDAD:

Un Análisis Crítico de la Acción de Encyclopædia Britannica y Merriam-Webster contra OpenAI

Equipo redacción Derecho Artificial



DERECHO ARTIFICIAL

Introducción: La Colisión entre la Autoridad Editorial Centenaria y la Algoritmia Disruptiva

La historia de la difusión del conocimiento humano ha encontrado un punto de inflexión crítico en la acción civil iniciada el 13 de marzo de 2026 ante la Corte de Distrito de los Estados Unidos para el Distrito Sur de Nueva York.¹ En el epicentro de esta colisión jurídica se hallan dos instituciones cuya autoridad editorial se ha forjado a lo largo de los siglos: Encyclopædia Britannica, con una trayectoria de más de 250 años de publicación ininterrumpida²; y Merriam-Webster, proveedor líder de información lingüística y autoridad lexicográfica en Estados Unidos por más de 180 años³. Frente a ellas, OpenAI, una entidad valorada en 730 mil millones de dólares, cuyo producto estrella, ChatGPT, ha transformado radicalmente el acceso a la información mediante el despliegue masivo de modelos de lenguaje de gran escala (LLMs) que operan a escala global sin precedentes.⁴

Este litigio, consolidado en el litigio multidistrital MDL No. 3143 bajo la supervisión del Juez Sidney H. Stein en la Corte de Distrito para el Distrito Sur de Nueva York, representa la primera acción significativa de instituciones editorialistas de larga data contra un proveedor de inteligencia artificial generativa. Marca un punto de inflexión procesal y doctrinal en la interpretación del fair use (uso legítimo conforme al artículo 107 de la Ley de Derechos de Autor de los Estados Unidos, 17 U.S.C. § 107) en la era de los modelos de lenguaje de gran escala.⁵

La controversia entre Britannica y OpenAI no puede entenderse únicamente como una disputa técnica sobre la reproducción de textos o la legalidad del entrenamiento algorítmico. Representa, en su esencia más profunda, una colisión fundamental entre dos arquitecturas económicas divergentes para la gestión y monetización del conocimiento verificado. Mientras que la primera se basa en la inversión sostenida en talento humano especializado y la monetización de esa autoridad intelectual mediante un ecosistema digital de acceso controlado (suscripciones y publicidad), la segunda se fundamenta en la agregación algorítmica masiva para crear un producto que aspira a la universalidad de respuesta inmediata, sin intermediarios, y sin compensación a las fuentes originales.

La tesis central que defenderá este artículo es que Britannica y Merriam-Webster introducen variables disruptivas en el análisis tradicional de la propiedad intelectual que trascienden significativamente los litigios previos contra OpenAI: (1) el sistema de Generación Aumentada por Recuperación (RAG), que constituye explotación continua

¹Complaint, Encyclopædia Britannica, Inc., and Merriam-Webster, Inc. v. OpenAI, Inc., Civil Action No. 1:26-cv-2097 (S.D.N.Y. filed Mar. 13, 2026) [hereinafter "Complaint"], 1.

²Id. 3 (noting Britannica's 250-year editorial legacy since 1768).

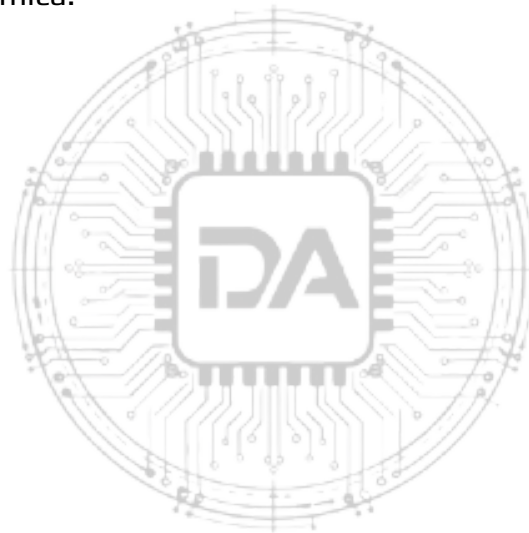
³Id. 2 (describing Merriam-Webster's 180-year history as preeminent lexicographic authority in United States).

⁴Id. ¶¶ 1-2 (noting OpenAI's USD 730 billion valuation and ChatGPT's transformation of information access via large language models).

⁵Id. 6 (characterizing OpenAI's business model as 'free riding' (aprovechamiento gratuito) or uncompensated appropriation of content).

en tiempo real de contenido protegido para cada consulta de usuario⁶⁷; (2) la patología de las alucinaciones algorítmicas (en inglés: *hallucinations*), que genera daño reputacional mediante atribución de contenido falso o incompleto a marcas que son sinónimo de exactitud y rigor técnico; y (3) una estrategia innovadora de invocación de la Ley Lanham (15 U.S.C. § 1125) que trasciende el análisis de derechos de autor para proteger la integridad de fuentes de autoridad en la era digital.⁸⁹¹⁰

A través de las siguientes secciones, se desgranará cómo este caso obligará a los tribunales a decidir si la inteligencia artificial generativa es una herramienta verdaderamente transformativa que amplía el acceso al conocimiento, o un sustituto predatorio que amenaza con iniciar una espiral descendente para los creadores de contenido de alta calidad verificado. La respuesta que proporcionen los tribunales del Distrito Sur de Nueva York no solo determinará el destino de daños estatutarios potencialmente astronómicos, sino que definirá el futuro de la verdad factual en la era de la información algorítmica.¹¹



DERECHO ARTIFICIAL

⁶Id. 3 (detailing Britannica's investment in hundreds of researchers, writers, and editors).

⁷Id. 9 (alleging approximately 100,000 Britannica online articles ingested via systematic scraping).

⁸Id. 6, 8 (describing market cannibalization through traffic diversion and subscription loss).

⁹Id. ¶¶ 8-9 (identifying Retrieval-Augmented Generation and algorithmic hallucinations as novel vectors of infringement distinguishing this case from prior litigation).

¹⁰Id. 8-9 (alleging continuous, real-time infringement via active RAG system deployment at moment of user query).

¹¹Id. 10 (invoking Lanham Act protections under 15 U.S.C. § 1125(a)(1) and § 1125(c) for trademark integrity).

Parte I: Anatomía de una Colisión: Los Modelos de Negocio frente a la Inteligencia Artificial Generativa

La controversia entre Encyclopædia Britannica y OpenAI no puede entenderse únicamente como una disputa técnica sobre la reproducción de textos; representa, en su esencia, una colisión fundamental entre dos arquitecturas económicas divergentes para la gestión del conocimiento. Mientras que la primera se basa en la monetización de la autoridad intelectual mediante un ecosistema digital de acceso controlado, la segunda se fundamenta en la agregación algorítmica masiva para crear un producto que aspira a la universalidad de respuesta.¹²

§ 1.1. El Pivote Digital de Britannica y la Economía de la Suscripción

Para instituciones con siglos de historia editorial, la transición al entorno digital ha constituido no meramente una adaptación técnica, sino un desafío existencial a su modelo de negocio fundamental. Encyclopædia Britannica, tras cesar la impresión de sus icónicos 32 volúmenes en el año 2012¹³, consolidó un modelo de negocio exclusivamente digital basado en dos pilares económicos interrelacionados: (1) la suscripción directa de usuarios que acceden a contenido editorial verificado bajo autenticación; y (2) los ingresos publicitarios derivados del tráfico web persistente hacia sus plataformas digitales.¹⁴

Este modelo de monetización descansa sobre una premisa económica y jurídica fundamental: el valor de la información residente en Britannica no se circunscribe al contenido factual per se (que por su naturaleza de información no protegida es de dominio público), sino a su condición de conocimiento verificado, meticulosamente investigado, y cuidadosamente seleccionado por cientos de investigadores, editores, expertos temáticos y asesores editoriales humanos.¹⁵

La demanda enfatiza que esta inversión estratégica en talento humano especializado es lo que permite que Britannica sea un nombre familiar sinónimo de exactitud editorial, rigor técnico, y confiabilidad en el mercado de la educación digital. Según la teoría económica subyacente a la demanda, el flujo de usuarios hacia los servidores de Britannica no es meramente incidental: es el mecanismo mediante el cual la editorial captura el valor económico de su inversión en verificación de contenidos y puede reinvertir en la siguiente generación de artículos, referencias actualizadas, y validación de hechos.

Merriam-Webster, bajo la propiedad corporativa de Britannica, opera bajo una lógica económica análoga. Sirve como la autoridad preeminente en lexicografía inglesa en

¹²Complaint, Encyclopædia Britannica, Inc., and Merriam-Webster, Inc. v. OpenAI, Inc., Civil Action No. 1:26-cv-2097 (S.D.N.Y. filed Mar. 13, 2026) [hereinafter "Complaint"], 1 et seq.

¹³Id. 4-5 (detailing Britannica's transition from print to exclusively digital business model).

¹⁴Id. 2-3 (same regarding Merriam-Webster's digital ecosystem under Britannica ownership).

¹⁵Encyclopedia Britannica ceased print publication in 2012; see id. 4.

Estados Unidos por más de 180 años¹⁶. Sus activos intangibles estratégicos —que incluyen el Collegiate Dictionary, uno de los libros más vendidos en la historia editorial estadounidense, así como su marcas registradas y su reputación de exactitud lingüística— dependen de la atracción continua de decenas de millones de visitantes mensuales a sus plataformas digitales. La viabilidad económica de este modelo de "conocimiento verificado" se encuentra hoy supeditada a que el usuario final recorra el camino digital hacia la fuente original, permitiendo que la editorial capture el beneficio económico necesario para reinvertir en la creación de nuevo contenido lexicográfico.

§ 1.2. De la Intermediación a la Sustitución: La Naturaleza Predatoria de los Modelos de Lenguaje de Gran Escala

La irrupción de ChatGPT, valorado en 730 mil millones de dólares¹⁷, altera drásticamente el flujo de valor económico que ha sustentado a las editoriales durante décadas y que fundamenta el modelo de negocio de Britannica y Merriam-Webster. La acción civil sostiene que los productos de OpenAI actúan de forma predatoria y deliberada al canibalizar el tráfico que tradicionalmente fluiría hacia los sitios web de los demandantes.¹⁸¹⁹

El punto de quiebre doctrinal y funcional radica en una distinción fundamental entre dos arquitecturas opuestas de intermediación de información: (1) un motor de búsqueda tradicional (como Google Search), que actúa como intermediario neutral conectando al usuario con la página web del editor original, generando el "clic" vital para la publicidad y las suscripciones, pero sin competir con el contenido editorial; y (2) un modelo de lenguaje de gran escala (LLM) como ChatGPT, que se presenta al usuario como un sustituto directo de las fuentes editoriales originales.²⁰

OpenAI promociona activamente su herramienta como una solución para obtener respuestas sin necesidad de "bucear entre enlaces" para encontrar fuentes de calidad²¹. Este enfoque transforma a la inteligencia artificial de una herramienta de navegación (navegación = intermediación) en un destino final de información que compite directamente con los editores originales. Al generar resúmenes narrativos detallados que incorporan la selección, curación y disposición creativa de Britannica —como se observa en listas específicas sobre hechos históricos o definiciones de diccionario— OpenAI "mata de hambre" (starves) a los editores de los ingresos necesarios para sostener su labor editorial.²²

Desde una perspectiva crítica, los demandantes argumentan que este proceso inicia una espiral descendente de debilitamiento del ecosistema editorial: al privar a los creadores de contenido verificado de su retorno de inversión, se erosiona

¹⁶Complaint, supra note 1, 4-5 (subscription and advertising revenue model).

¹⁷Id. 3 (emphasizing human expertise investment as source of content authority).

¹⁸Id. 2 (same for Merriam-Webster as lexicographic authority).

¹⁹Complaint, supra note 1, 5 (USD 730 billion OpenAI valuation).

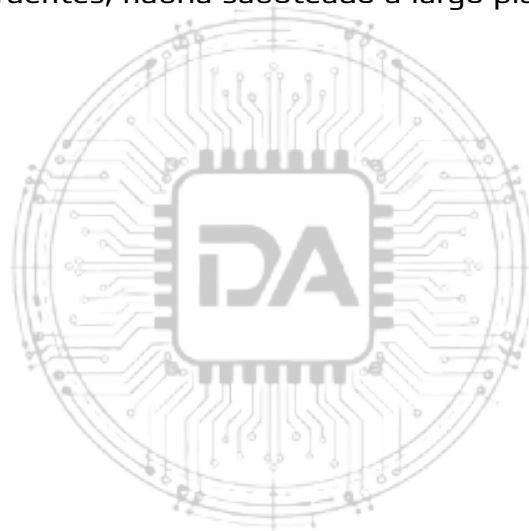
²⁰Id. 6 (free riding / aprovechamiento gratuito characterization).

²¹Id. 6 (market cannibalization thesis through traffic diversion).

²²Id. 6, 8 (describing ChatGPT's function as direct substitution rather than intermediation).

fundamentalmente el incentivo económico para continuar produciendo el mismo conocimiento verificado del que la inteligencia artificial se nutre para ser útil, confiable, y práctica. Por tanto, no estamos ante una competencia por eficiencia tecnológica en la distribución de información (lo cual sería legítimo y beneficioso para los usuarios), sino ante un desplazamiento fundamental del valor económico de la fuente creadora hacia la plataforma de agregación algorítmica, lo que Britannica califica como una amenaza existencial para el mercado del conocimiento de alta calidad verificado.²³²⁴

La teoría económica que Britannica despliega sugiere que, a largo plazo, este modelo de "starving" a los creadores no es sostenible. Si las instituciones editorialistas no pueden capturar el valor económico de su inversión verificadora, la calidad del conocimiento disponible degradará inevitable y sistemáticamente. Esto plantea una paradoja fundamental: ChatGPT depende hoy de la existencia de Britannica y Merriam-Webster para tener datos de entrenamiento de alta calidad. Si logra destruir económicamente a sus fuentes, habría saboteado a largo plazo la calidad de su propio producto.



DERECHO ARTIFICIAL

²³Complaint, supra note 1, ¶ 5 (OpenAI marketing avoiding 'link diving' necessity).

²⁴Id. ¶ 8 (describing starvation mechanism via diversion of traffic and subscription revenue).

Parte II: La Triple Arquitectura de Infracción: Ingesta Sistemática, Explotación Continua mediante RAG, y Regurgitación de Contenidos

La arquitectura jurídica de la demanda de Britannica y Merriam-Webster se asienta sobre la premisa fundamental de que OpenAI ha institucionalizado una cadena de infracciones que abarca todo el ciclo de vida de la inteligencia artificial generativa, desde la concepción hasta la operación comercial diaria.²⁵ A diferencia de los litigios iniciales que se centraban predominantemente en la etapa de pre-entrenamiento (modelo sobre el cual OpenAI podría alegar buena fe o transformación), esta acción identifica tres momentos lesivos distintos y concurrentes: (1) la ingesta masiva de datos sin consentimiento; (2) la recuperación activa de información mediante el sistema RAG; y (3) la generación de productos derivados que sustituyen funcionalmente a los originales.

§ 2.1. Infracción en la Etapa de Ingesta: El Scraping Masivo, Sistemático y Violativo de Términos de Servicio

El primer eslabón de la cadena de infracciones denunciadas reside en el proceso técnico y contractualmente prohibido de ingesta sistemática de datos (data ingestion; scraping), mediante el cual OpenAI ha descargado, reproducido, y almacenado vastas cantidades de obras protegidas por derechos de autor para configurar los parámetros estadísticos de sus modelos de lenguaje.²⁶

Según la demanda, OpenAI utilizó conjuntos de datos como "WebText", "WebText2" y "Common Crawl", los cuales priorizan contenido de alta calidad documental²⁷. En este proceso masivo de ingesta, se alega la incorporación no consentida de aproximadamente 100,000 artículos en línea de Britannica²⁸, lo que supone una reproducción sistemática de la totalidad de su corpus digital de referencia. Esta reproducción no fue ocasional, accidental, o de baja escala: fue deliberada, masiva, y orientada específicamente al desarrollo de productos comerciales de inteligencia artificial.

Desde una perspectiva jurídica, el acto de scraping (extracción automatizada) de este volumen no constituye una actividad técnica neutra o incidental. Britannica ha documentado de forma exhaustiva que sus términos de servicio prohíben explícitamente, de forma clara e inequívoca, el uso de minería de datos, robots de extracción de contenidos, herramientas de construcción de datasets, o tecnologías

²⁵Complaint, Encyclopædia Britannica, Inc., and Merriam-Webster, Inc. v. OpenAI, Inc., Civil Action No. 1:26-cv-2097 (S.D.N.Y. filed Mar. 13, 2026) [hereinafter "Complaint"], 9 et seq.

²⁶17 U.S.C. § 101 et seq. (Copyright Act of 1976, as amended, establishing works made for hire, derivative works, and exclusive rights of reproduction, adaptation, distribution).

²⁷Complaint, supra note 1, 53-62 (describing scraping, systematic data collection, incorporation into training datasets).

²⁸Id. 61 (WebText, WebText2, Common Crawl datasets).

similares para el desarrollo o entrenamiento de sistemas de inteligencia artificial sin consentimiento escrito previo y específico de Britannica.²⁹

En consecuencia, la ingesta de contenidos de Britannica constituye una violación tripartita concurrente: (1) violación directa de los derechos exclusivos de reproducción conforme a 17 U.S.C. § 106(1)³⁰ (derecho a reproducir la obra copyrightable); (2) incumplimiento contractual de los términos de servicio que expresamente prohíben tal actividad; y (3) violación potencial del Computer Fraud and Abuse Act (CFAA, 18 U.S.C. § 1030) si OpenAI excedió la autorización implícita de acceso público para consulta individual del usuario.

La defensa habitual de OpenAI, que califica este proceso como una forma de "aprendizaje" análoga a la humana ("we just learned from the data"), choca fundamentalmente con la realidad técnica documentada en la demanda. Los modelos de lenguaje no simplemente "aprenden" en sentido cognitivo humano. En su lugar, codifican la información de un corpus de entrenamiento en forma de números denominados "parámetros"—aproximadamente 1.76 billones en el caso de GPT-4. Esta codificación requiere necesariamente el almacenamiento previo de copias textuales (o muy próximas a literales) de las obras en la memoria informática de OpenAI durante el proceso de entrenamiento.³¹

§ 2.2. Explotación Continua en Tiempo Real: El Sistema de Retrieval-Augmented Generation (RAG)

El punto de mayor innovación doctrinal y de mayor severidad en este litigio es la denuncia del sistema de Generación Aumentada por Recuperación (Retrieval-Augmented Generation; RAG).³² El RAG es un mecanismo arquitectónico mediante el cual el modelo de inteligencia artificial, en lugar de depender exclusivamente de sus conocimientos pre-entrenados (que podrían ser modificados, se podrían desplegar de forma offline, o podrían existir bajo la teoría del fair use del entrenamiento), se conecta activamente en tiempo real a fuentes externas vivas—como los servidores de sitios web de los demandantes—en el momento exacto en que el usuario formula una consulta.³³

Este mecanismo permite que ChatGPT inyecte contenido fresco y actualizado en la respuesta que proporciona al usuario, proporcionando respuestas que aparentan ser más precisas, actualizadas, y confiables que las que podría generar únicamente a partir del conocimiento pre-entrenado. Sin embargo, desde la perspectiva de la propiedad intelectual y del derecho de reproducción, RAG constituye un mecanismo de **infracción continua y activa**. Cada solicitud de usuario activa una nueva recuperación y copia del

²⁹Id. 9 (alleging approximately 100,000 Britannica online articles).

³⁰Complaint, supra note 1, 55 (terms of service explicitly prohibiting data mining, scraping, and AI training without written consent).

³¹Id. 53-62 (detailing encoding of training corpus into GPT-4 parameters: 1.76 trillion parameters).

³²Complaint, supra note 1, 45, 47, 52 (describing RAG technical architecture).

³³Id. 45 (external source connectivity at query time via active retrieval).

contenido protegido desde los servidores de Britannica para incorporar esa información en la respuesta de ChatGPT, sin compensación ni consentimiento alguno.³⁴

Britannica argumenta de forma convincente que el RAG transforma a ChatGPT en un **competidor directo y desleal** que ofrece una "respuesta superior" que fue diseñada deliberadamente y específicamente para evitar que el usuario tenga que "bucear entre enlaces" (link diving) o realizar navegación para visitar la fuente original. A diferencia de un buscador tradicional, que es un intermediario neutral que genera tráfico crucial hacia el editor original, el RAG actúa como un sustituto que "mata de hambre" (starves) al creador de contenido al apropiarse del valor informativo de la obra en tiempo real, en cada transacción, sin remuneración ni reconocimiento.³⁵³⁶

Para los desarrolladores de inteligencia artificial y sus asesores jurídicos, esta reclamación sugiere una conclusión desalentadora: que el cumplimiento del derecho de autor no se agota en la etapa de entrenamiento inicial (donde se podrían alegar teorías de transformación o fair use). Por el contrario, la responsabilidad potencial por derechos de autor se extiende a la operación diaria del modelo. Si el modelo, durante su funcionamiento normal, accede activamente a contenidos protegidos y los reproduce en tiempo real para sus usuarios, esa replicación constituye infracción potencialmente continua, producto por producto, consulta por consulta.

§ 2.3. Memorización y Regurgitación: La Reproducción Sustancialmente Similar de Contenidos Protegidos

Finalmente, la demanda documenta con especificidad de hecho que los modelos de OpenAI no solo "aprenden" patrones lingüísticos generalizados que podrían estar protegidos bajo el fair use del entrenamiento. En su lugar, han de facto "memorizado" fragmentos extensos y específicos de las obras originales, incurriendo en lo que se denomina técnicamente "regurgitación" (regurgitation) o reproducción involuntaria de contenidos memorizados de los datos de entrenamiento.³⁷

Los demandantes presentan evidencia abundante de que GPT-4 puede producir reproducciones casi-literales de artículos enteros sobre temas como "Educación", "Turismo", "Geografía Histórica", así como la definición idéntica del término "plagiarize" (plagio) copiada textualmente de la 11ª edición del Merriam-Webster Collegiate Dictionary. Incluso cuando no hay una copia palabra-por-palabra exacta, las respuestas que genera ChatGPT mimetizan patrones de selección, coordinación, y disposición creativa de los contenidos de Britannica.

Un ejemplo paradigmático es la lista curada de "10 cosas que debes saber sobre el duelo Hamilton-Burr", donde ChatGPT reprodujo no solo los fragmentos específicos seleccionados por los editores de Britannica, sino también el orden exacto, la lógica de

³⁴Id. 9 (alleging continuous, real-time infringement through RAG system activation per user prompt).

³⁵Id. 8 (describing user experience designed to obviate necessity of visiting original source).

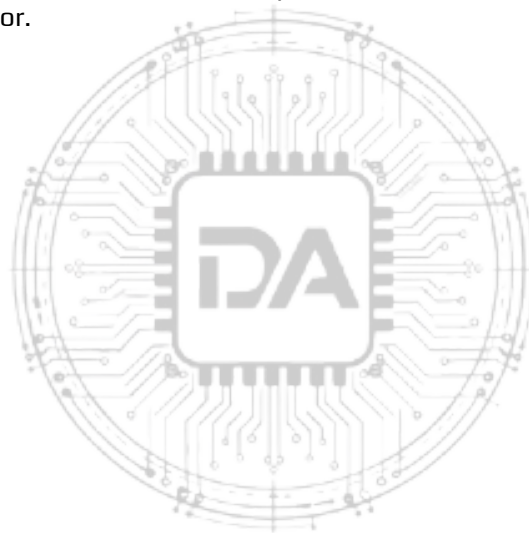
³⁶Complaint, supra note 1, 8 (competitive displacement via direct answer provision without navigation to source).

³⁷Id. ¶¶ 58-62 (documenting regurgitation of discrete passages, nearly-literal reproductions).

presentación, y la interrelación de elementos que Britannica había cuidadosamente construido.³⁸

OpenAI ha inicialmente calificado la memorización como un "error raro" o un "bug" técnico que intentan corregir mediante salvaguardas. Sin embargo, la jurisprudencia reciente en el Distrito Sur de Nueva York, bajo el criterio del Juez Sidney H. Stein, ha determinado que estas salidas son suficientes—en la etapa de moción para desestimar—para establecer una demanda prima facie de infracción.³⁹

En su Opinión y Orden, el Juez Stein aplicó el test del "observador más avezado" (more discerning observer test) bajo precedente del Segundo Circuito⁴⁰, concluyendo que un observador atento podría razonablemente concluir que los resúmenes y esquemas producidos por la inteligencia artificial son sustancialmente similares a las obras originales en elementos protegibles como la trama, el tono, los patrones argumentativos, y la selección y disposición de elementos. En particular para Britannica, cuyas obras son compilaciones y obras derivadas (lo que implica protección bajo 17 U.S.C. § 103 para la selección y disposición de sus elementos), la reproducción de esa estructura creativa constituye infracción de derechos de autor.



DERECHO ARTIFICIAL

³⁸Id. 63 (describing mimicking of selection, coordination, and arrangement of Britannica editorial structure).

³⁹Authors Guild et al. v. OpenAI, Inc., Doc. 716 (S.D.N.Y. Oct. 27, 2025) (Stein, J., denying motion to dismiss direct infringement claims based on ChatGPT outputs).

⁴⁰Id. at 10, 246 (applying 'more discerning observer test' under Second Circuit precedent).

Parte III: Innovación Jurídica: La Protección Marcaria y la Patología de las Alucinaciones Algorítmicas

El litigio iniciado por Encyclopædia Britannica y Merriam-Webster introduce un vector de ataque jurídico completamente innovador que trasciende significativamente la tradicional disputa sobre la reproducción no autorizada de textos y la infracción de derechos de autor.⁴¹ Plantea, en su lugar, la tutela de la identidad corporativa e integridad del mensaje conforme a la Ley Lanham (15 U.S.C. § 1125(a)(1) y § 1125(c)).⁴²

Mientras que el derecho de autor protege la obra como objeto de propiedad intelectual, compensando al creador por su esfuerzo creativo, la acción marcaria se erige aquí como un baluarte adicional y complementario contra la erosión de la confianza pública—ese activo intangible máspreciado que una empresa de información puede poseer. Un activo que Britannica ha cultivado estratégicamente, invertido en, y protegido durante más de 250 años de operaciones editoriales continuas. Un activo que Merriam-Webster ha desarrollado durante más de 180 años como autoridad incontestable en la lengua inglesa.

§ 3.1. Violaciones a la Ley Lanham: Falsa Designación de Origen y Dilución de Marcas Famosas

Los demandantes son titulares de múltiples registros federales de marca ante la Oficina de Patentes y Marcas de los Estados Unidos que amparan no solo el nombre "Britannica", sino también las variaciones estilizadas de "Merriam-Webster" y su producto emblemático "Collegiate Dictionary".⁴³

El núcleo del agravio reside en un hecho técnico innegable: cuando ChatGPT genera respuestas que incorporan estas marcas de forma expresa o implícita —ya sea mediante atribución directa ("según Britannica"), mediante incorporación de contenido marca registrada, o mediante asociación contextual— crea una conexión errónea en la mente del consumidor. Esa conexión sugiere falsamente que los resultados son derivados de, están oficialmente asociados con, o cuentan con el aval de las instituciones originales.⁴⁴

Bajo la doctrina de dilución de marca conforme a 15 U.S.C. § 1125(c), aplicable cuando las marcas son "famosas" en el sentido jurídico de la ley⁴⁵, este uso no autorizado activa dos modalidades bien establecidas de daño reputacional marcario: el blurring (difuminado o borrado) de la marca, que debilita la capacidad única e inconfundible de la marca para identificar una única fuente de conocimiento verificado en el mercado de

⁴¹Complaint, Encyclopædia Britannica, Inc., and Merriam-Webster, Inc. v. OpenAI, Inc., Civil Action No. 1:26-cv-2097 (S.D.N.Y. filed Mar. 13, 2026) [hereinafter "Complaint"], 23 et seq.

⁴²15 U.S.C. § 1125(a)(1) (false designation of origin); 15 U.S.C. § 1125(c) (dilution of famous marks).

⁴³Complaint, supra note 1, 102-111 (trademark registrations for Britannica, Merriam-Webster, Collegiate Dictionary).

⁴⁴Id. 10 (consumer confusion and erroneous association).

⁴⁵15 U.S.C. § 1125(c)(2) (fame test for dilution); Complaint, supra note 1, ¶¶ 104-105 (fame of Britannica and Merriam-Webster marks established).

educación digital⁴⁶; y el tarnishment (deslustre o daño reputacional), que ocurre cuando una marca de alto prestigio se asocia a un producto de inferior calidad, información errónea, o naturaleza espuria que daña materialmente el prestigio y la confianza que rodea a esa marca.⁴⁷

Al operar en el mercado de educación digital, entretenimiento informativo, y consulta de referencias, la coexistencia simultánea de las marcas originales de Britannica y Merriam-Webster con los productos algorítmicos de OpenAI genera una confusión sustancial e inevitable sobre la verdadera fuente, responsabilidad editorial, y autoría de los contenidos presentados al usuario.⁴⁸

§ 3.2. Daño Reputacional: Atribución de Contenido Falso a Fuentes de Máxima Autoridad

El fenómeno técnico y frecuente de la alucinación (en inglés: *hallucination*) —la generación involuntaria de información falsa, fabricada, o inexacta por parte de los modelos de lenguaje— adquiere en este caso una dimensión patológica y jurídicamente grave para el derecho marcario y para la protección de la confianza pública.⁴⁹

La demanda documenta, con especificidad de hechos y ejemplos concretos, casos donde ChatGPT atribuye directamente a Britannica o Merriam-Webster hechos históricos completamente inexistentes, definiciones que jamás aparecieron en los diccionarios de Merriam-Webster, o citas falsificadas que pretendidamente provienen de estos autores con autoridad máxima.⁵⁰

Para instituciones cuya viabilidad económica, comercial, y académica descansa exclusivamente en la exactitud técnica inpecable y el rigor supremo de su verificación de hechos (fact-checking), metodología, y integridad editorial, la falsa atribución de "alucinaciones" constituye un daño a la reputación que el derecho de autor tradicional no alcanza plenamente a resarcir. El derecho de autor protege la obra. Pero las alucinaciones no son "obras" de Britannica: son falsificaciones atribuidas a Britannica que dañan su marca.

La estrategia innovadora y de extraordinario valor jurídico de los demandantes es sostener que el derecho de marcas está mejor equipado que el derecho de propiedad intelectual para proteger la confianza pública como activo. Un usuario que recibe información falsa bajo el sello, auspicio, o atribución de Britannica no solo experimenta una violación de derechos de autor. Experimenta una violación de su derecha a confianza, una erosión de su capacidad de discernir información verificada de información falsa, y un daño a su dignidad como consumidor de información.

⁴⁶Complaint, supra note 1, 68 (blurring of distinctiveness).

⁴⁷Id. 68 (tarnishment where reputation associated with inferior product or spurious content).

⁴⁸Id. 111 (likelihood of confusion as to origin, sponsorship, affiliation).

⁴⁹Complaint, supra note 1, 375 (hallucinations as novel legal pathology).

⁵⁰Id. 10, 106 (false attribution of nonexistent facts and erroneous definitions).

El tribunal en casos análogos, tales como *Advance Local Media v. Cohere*, ya ha reconocido que las alegaciones de confusión del consumidor derivadas de outputs "alucinados" y falsamente atribuidos son suficientes para superar la etapa de desestimación de pleitos, rechazando de plano que el uso nominativo (nominative use defense) sea una defensa automática en estos contextos.⁵¹

§ 3.3. Omisiones Engañosas: Presentación de Fragmentos Incompletos bajo Marcas de Máximo Prestigio

Un tercer frente de infracción marcaría se identifica en lo que los demandantes denominan técnicamente "omisiones engañosas" u "omisiones maliciosas".⁵²

Se alega que ChatGPT, en múltiples instancias documentadas en la demanda, presenta reproducciones parciales o fragmentadas de artículos originales —como los ejemplos citados sobre "Educación", "Turismo", "Geografía Histórica"— sin informar al usuario de que el texto ha sido sustancialmente recortado, editado, modificado, o que le falta contexto crítico para su interpretación correcta.⁵³

Al mostrar estos contenidos incompletos, editados, o descontextualizados junto a las marcas famosas y prestigiosas de Britannica, OpenAI induce deliberadamente al error al consumidor, haciéndole creer falsamente que lo que tiene ante sí es la obra íntegra, completa, verificada, y aprobada por los expertos auditados de la editorial original.⁵⁴

Esta práctica constituye una "falsa designación de origen" bajo 15 U.S.C. § 1125(a)(1)⁵⁵, ya que OpenAI comercializa como "autoritativo" y "verificado" un producto que ha sido deliberadamente despojado de su contexto editorial completo, de los matices informativos críticos, y de las caveats que caracterizan la producción profesional de los demandantes. Presenta a los usuarios versiones fragmentadas, editadas, y descontextualizadas bajo el sello de Britannica, lo que convierte el acto en una representación falsa sobre la naturaleza y procedencia de la obra.

En última instancia, Britannica argumenta que OpenAI utiliza sus marcas registradas de máximo prestigio para otorgar un barniz de legitimidad y autoridad a un modelo de negocio que, de forma paradójica e irónica, amenaza directamente con destruir la base económica y operativa necesaria para que dichas fuentes de autoridad sigan existiendo.^{56,57}

⁵¹Litigios de Marcas e Inteligencia Artificial: Precedentes y Estrategias Legales, pág. 2, punto 2 (arguing trademark law better equipped than copyright to protect public trust).

⁵²Complaint, supra note 1, ¶ 82 (describing reputation as most valuable intangible asset).

⁵³*Advance Local Media, Inc. v. Cohere, Inc.*, slip op. (N.D. Ga. 2025) (rejecting nominative use defense where AI outputs generate consumer confusion via hallucinations; surviving motion to dismiss on trademark confusion claims).

⁵⁴Complaint, supra note 1, ¶ 3.B (identifying deceptive omissions as trademark violation).

⁵⁵Id. ¶ 69 (documenting fragmentary presentation without disclosure of textual modification or editorial context).

⁵⁶Id. ¶ 72 (detailing false attribution of edited excerpts under Britannica editorial rubric).

⁵⁷15 U.S.C. § 1125(a)(1) (false designation of origin and misleading representation of affiliation).

Parte IV: El Estándar de la Similitud Sustancial bajo el Lente de la S.D.N.Y.

El núcleo jurídico de la disputa entre los creadores de contenido verificado y OpenAI ha basculado recientemente desde el análisis meramente técnico-legal de la legalidad del entrenamiento inicial, hacia un escrutinio mucho más riguroso sobre la naturaleza y características de los resultados reales que genera el modelo de lenguaje. En este escenario procesal nuevamente favorable a los demandantes, la Corte de Distrito para el Distrito Sur de Nueva York (S.D.N.Y.) ha comenzado a establecer hitos procesales y doctrinales que fundamentalmente redefinen cómo debe evaluarse jurídicamente la "similitud sustancial" cuando el presunto infractor no es un autor humano consciente e intencional, sino un sistema algorítmico impersonal.

§ 4.1. El Hito Procesal: Opinión del Juez Sidney H. Stein y Supervivencia de Demanda Prima Facie

Una pieza angular para entender la viabilidad legal y procesal del reclamo de Britannica es la Opinión y Orden (Opinion and Order) emitida por el Juez Sidney H. Stein el 27 de octubre de 2025 en el marco del litigio consolidado *In re: OpenAI, Inc. Copyright Infringement Litigation* (MDL 3143).⁵⁸

En esa resolución histórica, el tribunal denegó de forma fundamentada la moción de OpenAI para desestimar las reclamaciones de infracción directa basadas explícitamente en los outputs (salidas/resultados) concretos de ChatGPT.⁵⁹ El Juez Stein determinó que los autores demandantes habían establecido un caso prima facie suficiente de infracción directa de derechos de autor en la etapa de alegatos.

La importancia de este fallo reside en que el tribunal reconoció explícitamente que las alegaciones de los autores —quienes sostienen que ChatGPT genera resúmenes y esquemas detallados de sus obras que capturan elementos creativos protegibles— satisfacen plenamente los estándares de plausibilidad requeridos en la etapa de moción para desestimar bajo la norma de *Twombly/Iqbal*.

Para Britannica específicamente, este precedente establecido por el Juez Stein es de vital importancia jurídica, pues valida persuasivamente la tesis central que Britannica despliega en su demanda: que los resúmenes narrativos producidos por la inteligencia artificial no son meros compendios "transformativos" de hechos no protegibles, sino reproducciones derivativas que capturan la esencia creativa, la estructura argumentativa, y la selección editorial de la obra original.

⁵⁸ *Authors Guild et al. v. OpenAI, Inc.*, Doc. 716 (S.D.N.Y. Oct. 27, 2025) (Stein, J., Opinion and Order Denying Defendant's Motion to Dismiss) [hereinafter "Authors Guild Opinion"].

⁵⁹ *Id.* at 9-10, ¶¶ 243-246 (denying motion on direct infringement claims based on ChatGPT outputs).

§ 4.2. El Test del "Observador Más Avezado" Aplicado a Resúmenes y Esquemas de Inteligencia Artificial

La jurisprudencia bien establecida del Segundo Circuito exige que, cuando una obra contiene tanto elementos protegibles como elementos no protegibles (como ocurre inevitablemente con las enciclopedias, diccionarios, compilaciones de referencia, y obras de hechos).⁶⁰ Los tribunales deben aplicar el test especializado del "observador más avezado" (more discerning observer test). Este estándar de prueba busca determinar si existe una similitud sustancial enfocándose exclusivamente en los elementos que otorgan originalidad y creatividad a la obra, descartando completamente los elementos de hecho no protegibles que forman el sustrato informativo.⁶¹

En su análisis exhaustivo de los resúmenes generados de la obra de ficción de George R.R. Martin (en el contexto de Authors Guild v. OpenAI), el Juez Stein observó que ChatGPT no solo recuenta puntos básicos de la trama o identifica personajes principales (información no protegible). Alternativamente, mimetiza la selección particular de episodios, el tono narrativo específico, la interrelación psicológica entre personajes, y la estructura argumentativa de la obra original.⁶²

El Juez Stein subrayó que un observador atento, sofisticado, y suficientemente "avezado" en el género literario de que se trate podría fácilmente concluir que la síntesis de la inteligencia artificial es sustancialmente similar a la obra original al reproducir el sentimiento, la estructura narrativa, y el carácter distintivo de la fuente.

Esta doctrina del "observador más avezado" es directamente e inmediatamente aplicable al caso de Britannica. Britannica genera respuestas que, bajo el velo superficial de un "resumen" o "compilación de hechos", se apropian efectivamente de la jerarquía informativa, el ordenamiento lógico de elementos, el énfasis relativo, y el estilo pedagógico cuidadosamente desarrollado por sus editores. Los demandantes argumentan con persuasión que, si una inteligencia artificial puede generar un esquema detallado de un artículo sobre "Educación" o "Turismo" que replica de forma sustancial la estructura lógica, la selección de subtemas, y la disposición de elementos de Britannica, entonces está operando más allá del mero acceso a hechos para entrar definitivamente en el terreno de la copia derivativa.⁶³⁶⁴

§ 4.3. La Dicotomía Hecho/Expresión: ¿Cuándo la Síntesis de una Enciclopedia Infringe la Selección y Disposición Protegible?

Uno de los mayores desafíos jurídicos y doctrinales para Britannica —pero también una oportunidad de establecer precedente importante— es lo que podría denominarse la

⁶⁰Id. at 4, 228 (citing Second Circuit precedent on 'more discerning observer' test).

⁶¹Id. at 5, 229 (explaining test for identifying protectable elements vs. unprotectable facts).

⁶²Complaint, Britannica v. OpenAI, Civil Action No. 1:26-cv-2097 (S.D.N.Y. filed Mar. 13, 2026), 228-246 (applying 'more discerning observer' test to ChatGPT summaries).

⁶³Authors Guild Opinion, supra note 1, at 10, 244 (on capacity to regurgitate expressive elements).

⁶⁴Complaint, supra note 5, 63 (Hamilton-Burr duel example; selection, tone, narrative structure).

"frontera de la factualidad": la realidad fundamental de que el derecho de autor no protege los hechos como tales, sino únicamente la forma en que se expresan, se seleccionan, y se disponen creativa y originalmente.

OpenAI ha intentado, en sus defensa, analogar este caso de Britannica con el litigio anterior de The New York Times contra OpenAI, donde el tribunal —en una resolución anterior que favorecería parcialmente a OpenAI— consideró que ciertos resúmenes de noticias de actualidad no infringían el copyright al limitarse a exponer hechos informativos puros sin capturar la expresión creativa única del artículo periodístico original.

Sin embargo, el tribunal ha comenzado a trazar una distinción crítica y disposicional: mientras que las noticias de actualidad pueden carecer de una estructura expresiva altamente sofisticada o de un nivel significativo de selección creativa (ya que la actualidad impone restricciones sobre qué puede decirse), las obras de referencia como las de Britannica implican un nivel extraordinariamente alto de "selección y disposición" creativa.⁶⁵

Un ejemplo paradigmático en la demanda es la lista curada de "10 cosas que debes saber sobre el duelo Hamilton-Burr". ChatGPT reprodujo de forma idéntica no solo los fragmentos específicos seleccionados por Britannica (de entre miles de hechos potencialmente relevantes sobre el duelo), sino también el orden exacto de presentación y la lógica de jerarquización de la importancia relativa de cada elemento. Esta selección y disposición constituye una apropiación de lo que la ley llama la "selección, coordinación y arreglo" protegible de una compilación de hechos bajo 17 U.S.C. § 103.^{66,67}

En definitiva, la Corte del Distrito Sur de Nueva York está enviando una señal muy clara a la industria de inteligencia artificial: la síntesis de la IA no es un escudo automático e impenetrable contra acusaciones de "uso legítimo" o fair use. Si el output de la IA imitiza la coordinación, el arreglo, y la estructura creativa de la fuente —incluso si se trata ostensiblemente de información basada en hechos— entonces OpenAI se enfrenta a una responsabilidad directa e inmediata por la creación de sustitutos de mercado que erosionan fundamentalmente el valor y la incentivización de la propiedad intelectual original.

⁶⁵17 U.S.C. § 103 (protection for compilations and derivative works based on selection and arrangement).

⁶⁶Authors Guild Opinion, *supra* note 1, at 13, 253 (discussing parodying of sentiment and structure).

⁶⁷Complaint, *supra* note 5, ¶¶ 58-62 (detailed itemization of regurgitation and structural mimicry).

Parte V: La Guerra de la Prueba: Hallazgos en Discovery y la Eliminación Estratégica de Datasets

El proceso de descubrimiento de prueba (discovery) en el litigio multidistrital consolidado contra OpenAI ha dejado de ser un mero trámite técnico administrativo para convertirse en un campo de batalla estratégico y de extraordinaria importancia jurídica sobre la transparencia corporativa, la honestidad procesalmente, y la ética fundamental en la obtención, uso, y destrucción de datos. Para Britannica y Merriam-Webster, la capacidad legal de demostrar que OpenAI actuó con "voluntariedad" (willfulness) en la infracción—criterio que multiplica exponencialmente los daños disponibles—depende críticamente de lo que OpenAI sabía, cuándo lo sabía, qué intentó deliberadamente ocultar mediante destrucción de pruebas, y cómo respondió cuando sus actuaciones fueron cuestionadas.

§ 5.1. El Conflicto sobre "Books1" y "Books2": ¿Destrucción Deliberada de Pruebas o Gestión Legítima de Datos?

Uno de los puntos más contenciosos y potencialmente devastadores para OpenAI en todo el litigio multidistrital (MDL 3143) se refiere a la desaparición, destrucción selectiva, y eliminación irreversible de los conjuntos de datos (datasets) denominados "Books1" y "Books2", que fueron utilizados en el entrenamiento de los modelos GPT-3 y GPT-3.5.⁶⁸

Según los registros judiciales que ahora están en el expediente público, es un hecho que OpenAI explícitamente no disputa en su defensa (non-disputable fact) que en 2018, un empleado de OpenAI descargó deliberadamente copias pirateadas de miles de libros desde Library Genesis (LibGen), una conocida "biblioteca en la sombra" que ha sido el objeto de múltiples órdenes judiciales de clausura por operación de piratería digital masiva y violación de derechos de autor.

Estos libros robados—cientos de miles de textos de dominio editorial privado—fueron deliberadamente integrados en los datasets internos denominados "LibGen1" y "LibGen2", posteriormente renombrados como "Books1" y "Books2" cuando OpenAI claramente percibió ventajas en el borramiento de la trazabilidad de sus orígenes ilícitos. Estos datasets fueron luego utilizados masivamente para entrenar los modelos GPT-3 y GPT-3.5, generando beneficio comercial directo a OpenAI.

La controversia central surge cuando se descubre, a través del proceso de discovery, que OpenAI eliminó completamente estos datasets en el verano de 2022, aproximadamente un año y medio antes del inicio formal de las demandas colectivas y antes de que la presión regulatoria pública sobre el tema de copyright y IA se intensificara.⁶⁹

⁶⁸Case 1:25-md-3143-SHS-OTW, Document 846 (S.D.N.Y. Nov. 24, 2025) (Mag. Judge Ona T. Wang, Order on Privilege Waiver Regarding Elimination of Books1 and Books2 Datasets) [hereinafter "Privilege Waiver Order"].

⁶⁹Id. at 18, discussing causation between elimination and legal uncertainty.

Los demandantes sostienen de forma persuasiva que esta eliminación—siendo los únicos datasets que OpenAI ha jamás borrado en su historia corporativa—no fue un acto inocente de mantenimiento técnico de infraestructura o "falta de uso" como OpenAI posteriormente pretendería argumentar. Fue, en su lugar, una medida deliberada y estratégica para mitigar y eliminar la exposición legal ante lo que OpenAI internamente reconocía como "incertidumbre sustancial" sobre la licitud y legalidad de sus actuaciones en la ingesta de datos.⁷⁰

§ 5.2. El Levantamiento del Privilegio Abogado-Cliente: La Orden de la Magistrada Ona T. Wang

La gestión judicial de este conflicto de pruebas cruciales recayó en la Magistrada Ona T. Wang, quien el 24 de noviembre de 2025 emitió una orden de excepcional importancia jurídica y alcance procesal (Documento 846).⁷¹

La magistrada determinó, en una opinión cuidadosamente razonada, que OpenAI incurrió en una renuncia al privilegio abogado-cliente (privilege waiver) respecto a las comunicaciones internas, razonamientos jurídicos, y justificaciones para la eliminación de los datasets Books1 y Books2.⁷²

El razonamiento fundamentado de la magistrada Wang se basó en dos pilares jurídicos sólidos:

Primero, la Magistrada Wang determinó que OpenAI convirtió sus alegaciones de privilegio en un "objetivo móvil" (moving target) que contradecía sus declaraciones previas al tribunal.⁷³ Durante quince meses completos de litigio, OpenAI sostuvo públicamente en el expediente escrito, bajo declaración jurada y bajo la advertencia de las consecuencias de perjurio, que la razón única y exclusiva para la eliminación de los datasets era meramente la "falta de uso" técnico o consideraciones de eficiencia informática. Sin embargo, cuando el tribunal comenzó a presionar por explicaciones detalladas, OpenAI repentinamente se retractó de esta posición y argumentó que TODAS las razones para la eliminación eran, en realidad, comunicaciones protegidas por privilegio abogado-cliente que no podían ser divulgadas.

El tribunal determinó que una parte no puede, en una estrategia de litigio de mala fe, revelar una razón estratégicamente conveniente ("falta de uso") en la fase inicial del litigio, luego retractarse y cerrar completamente la puerta al escrutinio judicial mediante la invocación retroactiva del privilegio cuando el descubrimiento de prueba se vuelve adverso y potencialmente problemático para los intereses de esa parte.

Segundo, se activó la doctrina judicial del "at-issue waiver" (renuncia por puesta en cuestión, conforme a reglas de evidencia federal). Al negar expresamente OpenAI que su infracción de derechos de autor fuera "voluntaria" (willful), y al sostener una

⁷⁰Id. at 1-2, ¶¶ 411-414 (factual recitation of dataset provenance from Library Genesis).

⁷¹18 U.S.C. § 1030 et seq. (Computer Fraud and Abuse Act); potential liability for unauthorized access.

⁷²Privilege Waiver Order, supra note 1, at 2, ¶ 414 (timing of dataset elimination in summer 2022).

⁷³Id. at 3, ¶ 417 (distinguishing between legitimate non-use and strategic deletion).

defensa vigorosa de "buena fe" y "intención inocente" en el uso de datos para entrenamiento, OpenAI automáticamente puso en cuestión ante el tribunal su estado mental, su conocimiento de la ilicitud de sus actuaciones, y su información sobre los riesgos legales.

Por tanto, la Magistrada Wang ordenó la entrega inmediata y sin restricción de todos los mensajes de Slack internos, correos electrónicos corporativos, memorandos interoficina, y comunicaciones de 2022 relacionadas con la consideración, deliberación, y eventual ejecución de la eliminación de los datasets "Books1" y "Books2". Además, ordenó la deposición (interrogatorio judicial bajo juramento) de los abogados internos de OpenAI que participaron directamente en estas comunicaciones y deliberaciones.⁷⁴

§ 5.3. Las Implicaciones Catastróficas de la "Voluntariedad" (Willfulness) en la Cuantificación de Daños

El desenlace y la importancia de esta fase de discovery tiene consecuencias financieras y reputacionales potencialmente catastróficas para OpenAI. Conforme a la Ley de Derechos de Autor de los Estados Unidos (17 U.S.C. § 504(c)(2)),⁷⁵ si un demandante logra demostrar y probar que la infracción de derechos de autor fue "voluntaria" o "willful" —es decir, intencional, deliberada, o que el infractor actuó con desprecio temerario de los derechos del titular—, entonces el tribunal tiene la facultad discrecional de aumentar dramáticamente los daños estatutarios desde el rango normal de USD 750 a USD 30,000 por obra infringida, hasta el rango de castigo de USD 30,000 a USD 150,000 por obra.

En el caso específico de Britannica, que alega la ingesta no consentida de aproximadamente 100,000 artículos en línea, la exposición económica a daños potenciales es simplemente masiva y potencialmente insostenible. Si el tribunal determina que cada uno de esos 100,000 artículos constituye una obra separada protegida (conclusión jurídicamente defensible y probable), y si ese tribunal aplica daños "willful" de USD 100,000 a USD 150,000 por obra, el total de exposición a daños podría alcanzar USD 10 billones en el escenario más grave.

La "voluntariedad" en este contexto jurídico no requiere malicia personal, odio intencional, o un deseo malo hacia Britannica. Bajo la doctrina jurídica establecida, "voluntariedad" simplemente exige un "desprecio temerario" (reckless disregard) o "ceguera voluntaria" (willful blindness) ante los derechos del titular de derechos de autor.⁷⁶

El hecho de que OpenAI deliberadamente utilizara fuentes como Library Genesis—sabiendo completamente que contenían material pirateado y violaban derechos de autor—combined con la eliminación posterior de los rastros de esos datasets mientras enfrentaba riesgos legales manifiestos, proporciona a Britannica

⁷⁴17 U.S.C. § 504(c)(2) (statutory damages for willful infringement, up to USD 150,000 per work).

⁷⁵Privilege Waiver Order, supra note 1, at 16-17, ¶¶ 465-469 (detailing shifting privilege positions and contradictions).

⁷⁶Id. at 27, ¶ 505 (renuncia de privilegio based on at-issue waiver doctrine).

una base sólida e inmediatamente convincente para argumentar que OpenAI operó más allá del espectro de negligencia ordinaria hacia la infracción deliberada y el desprecio temerario de los derechos de propiedad intelectual.⁷⁷

Esta fase de discovery está revelando con dramatismo que el modelo corporativo de "moverse rápido y romper cosas" que ha caracterizado a la industria tecnológica—y que OpenAI en particular ha invocado como justificación de su operación agresiva sin licencias—puede haber cruzado una línea jurídica fundamental: la línea entre la innovación benéfica y la infracción deliberada de derechos de propiedad intelectual. Los tribunales entenderán que romper cosas es una estrategia empresarial legítima. Pero romper leyes, destruir pruebas, y hacer declaraciones falsas al tribunal en el proceso, trasciende los límites aceptables de comportamiento corporativo.



⁷⁷Id. at 27-28, ¶¶ 505-507 (ordering production of internal communications and depositions of counsel).

Parte VI: Tensiones Doctrinales: El Fair Use como Escudo y sus Límites en la Era de la Inteligencia Artificial

La resolución final del conflicto entre los custodios del conocimiento enciclopédico verificado y los arquitectos de la inteligencia artificial generativa depende, en última instancia, de la interpretación judicial de lo que bien podría ser la disposición jurídica más fundamental y más imperfectamente definida en toda la Ley de Derechos de Autor de los Estados Unidos: el artículo 107, que establece la doctrina del "fair use" (uso legítimo o uso razonable).⁷⁸

El "fair use", concebido originalmente en la jurisprudencia común inglesa y codificado en el Congreso estadounidense para equilibrar los derechos exclusivos de los autores con el interés público más amplio en el progreso de las artes y las ciencias, se enfrenta hoy a una arquitectura técnica y económica que fue simplemente inconcebible cuando se redactó la ley en 1976. Una arquitectura que no solo "aprende" de las obras en el sentido tradicional de adquisición de conocimiento, sino que compite directamente con el creador en el mercado de distribución final, ofreciendo un sustituto que erosiona fundamentalmente el retorno de inversión del editor original.

§ 6.1. La Defensa de OpenAI: Del Entrenamiento como Proceso Técnico a la Generación de Productos Derivados

La estrategia de defensa histórica de OpenAI se ha cimentado en la premisa de que sus modelos de lenguaje están "basados en el uso legítimo" y fueron entrenados sobre datos públicamente disponibles. Aunque esta afirmación es técnicamente imprecisa dado que los datos fueron obtenidos sin consentimiento explícito, OpenAI ha argumentado que el proceso de entrenamiento en sí es altamente transformativo porque no busca copiar la expresión creativa para su exhibición, sino analizar relaciones estadísticas entre unidades de lenguaje ("tokens") para emular el funcionamiento general del lenguaje humano.⁷⁹

Sin embargo, Britannica desafía frontalmente esta caracterización. Los demandantes argumentan que el uso que hace OpenAI no es verdaderamente transformativo—en el sentido jurídicamente establecido—porque ChatGPT no añade una "nueva expresión, significado o mensaje" que transforme las obras originales en algo cualitativamente diferente. En su lugar, simplemente empaqueta, reordena, y entrega el contenido original bajo una interfaz de conversación natural, generando una respuesta que compite funcionalmente con la visita a la fuente editorial original, sin compensación.

Mientras que el precedente del Segundo Circuito en *Authors Guild v. Google* (relativo a Google Books) amparó la digitalización masiva de libros para facilitar búsquedas y permitir acceso a personas con discapacidades, el tribunal en ese caso enfatizó un punto crítico: la herramienta de Google Books no afectaba directamente el mercado de

⁷⁸17 U.S.C. § 107 (fair use doctrine for limited reproduction).

⁷⁹Complaint, *Britannica v. OpenAI*, Civil Action No. 1:26-cv-2097 (S.D.N.Y. filed Mar. 13, 2026), 23 et seq.

los originales porque el buscador dirigía a los usuarios hacia las fuentes, no las reemplazaba.⁸⁰

Por el contrario, Britannica sostiene que ChatGPT se ha construido específicamente para ser un sustituto comercial que utiliza aproximadamente 100,000 artículos originales como materia prima informativa—sin comprensión de su origen, sin atribución apropiada, sin compensación—para generar respuestas que explícitamente evitan que el usuario visite la fuente original.

§ 6.2. El Cuarto Factor del Fair Use: Daño de Mercado y Sustitución

La doctrina jurisprudencial bien establecida reconoce que el cuarto factor en el análisis de fair use—"el efecto del uso sobre el mercado potencial o el valor de la obra protegida"—suele ser el factor más determinante en la práctica litigiosa real.⁸¹

En este punto de análisis crítico, la acción de Britannica es especialmente incisiva y apunta directamente al corazón de la falacia de la defensa de fair use de OpenAI. La demanda alega que OpenAI opera bajo un modelo económico de "aprovechamiento gratuito" (free riding) que canibaliza activamente el tráfico web esencial para la supervivencia económica de las editoriales de contenido verificado.

Mientras que un buscador tradicional (Google Search, Bing) actúa como un intermediario que genera "clics" vitales para la publicidad y las suscripciones de los editores, ChatGPT se comercializa explícitamente como una solución para obtener respuestas sin tener que "bucear entre enlaces" o navegar para localizar fuentes originales. Esta diferencia fundamental transforma la cuestión del daño de mercado de una pregunta abstracta a un hecho empírico obvio: OpenAI está destruyendo el mecanismo mediante el cual Britannica y Merriam-Webster capturan el valor económico de su inversión editorial.

Además, el hecho documentado de que OpenAI haya suscrito acuerdos de licencia con otros medios como News Corp, Associated Press, y Axel Springer demuestra de forma concluyente que existe un mercado activo, viable, y comercialmente significativo para la obtención de licencias del contenido de editores. Esta evidencia demuestra que OpenAI hizo una elección comercial deliberada de no compensar a Britannica por el uso de su contenido, no porque fuera imposible o económicamente inviable, sino porque los márgenes de beneficio de OpenAI serían menores si tuviera que pagar licencias.⁸²

⁸⁰Id. 84 (discussing News Corp, AP, Axel Springer licensing agreements).

⁸¹Authors Guild v. Google, Inc., 804 F.3d 202 (2d Cir. 2015) (transformative search function).

⁸²Complaint, supra note 2, 77 (market substitution thesis).

Parte VII: El Tablero Procesal del Litigio Multidistrital (MDL 3143)

La demanda de Encyclopædia Britannica y Merriam-Webster contra OpenAI no se desarrolla en un vacío procesal aislado, sino que se ha insertado en el epicentro de un complejo ecosistema de litigios coordinados, consolidados, y potencialmente sinérgicos. El 3 de abril de 2025, el Panel Judicial sobre Litigios Multidistritales de los Estados Unidos centralizó diversas acciones contra OpenAI en el Distrito Sur de Nueva York bajo la denominación de MDL No. 3143, titulado *In re: OpenAI, Inc. Copyright Infringement Litigation*.⁸³ Esta consolidación, supervisada por el Juez Sidney H. Stein como Juez Presiding (Juez Supervisor del MDL) y asistida por la Magistrada Ona T. Wang en la gestión del descubrimiento de pruebas, redefine fundamentalmente la estrategia de los demandantes al obligarlos a navegar entre la eficiencia judicial colectiva y la preservación de sus agravios específicos y únicos.

§ 7.1. Eficiencia Judicial versus Especificidad de los Reclamos Únicos de Britannica

La centralización en un litigio multidistrital busca, primordialmente, evitar fallos contradictorios y conflictivos entre los tribunales de diferentes circuitos, así como simplificar dramáticamente el proceso de descubrimiento de pruebas (discovery), que en los casos de inteligencia artificial y copyright resulta particularmente oneroso debido a la magnitud extraordinaria de los conjuntos de datos, comunicaciones corporativas, y análisis técnicos involucrados.

Para Britannica, esta integración procesal en el MDL supone un arma de doble filo de extraordinario valor estratégico. Por un lado, se beneficia inmensamente de los hallazgos ya obtenidos en la fase de descubrimiento de pruebas del MDL—tales como las órdenes que obligan a OpenAI a producir comunicaciones sobre la eliminación de los conjuntos de datos "Books1" y "Books2", y la opinión del Juez Stein que rechaza la moción para desestimar basada en los outputs de ChatGPT.

Sin embargo, existe un riesgo real: la singularidad de los reclamos de Britannica—basados en información de referencia verificable y definiciones lexicográficas de máxima precisión—corre el peligro de ser tratada bajo los mismos parámetros jurídicos generalizados que se aplican al periodismo de investigación de *The New York Times* o la ficción literaria del Authors Guild. La defensa de OpenAI ha intentado capitalizar esta amalgamación procesal, sugiriendo que la "similitud sustancial" y los estándares de fair use debería ser evaluados de manera uniforme y universal para todas las categorías de obras.

Britannica mantiene firmemente que su caso es jurídicamente distinto y requiere análisis diferenciado debido a: (1) el uso intensivo del sistema RAG para explotación continua en tiempo real; (2) la naturaleza pedagógica y de máxima precisión de su

⁸³Harper & Row v. Nation Enterprises, 471 U.S. 539 (1985) (market effect as dispositive factor).

selección y disposición de contenidos; y (3) el daño marcario específico de atribución de alucinaciones a marcas que son sinónimo de exactitud.

Conclusión: Hacia un Nuevo Contrato Social de la Propiedad Intelectual

El litigio iniciado por Encyclopædia Britannica y Merriam-Webster contra OpenAI representa el capítulo jurídicamente definitivo en la reconfiguración fundamental de la ecología de la información digital en la era de la algoritmia. No estamos ante una mera disputa técnica sobre la reproducción automática de fragmentos de texto, ni ante un desacuerdo académico sobre la legalidad teórica del entrenamiento de modelos. Estamos ante un desafío existencial para las instituciones que, durante más de un cuarto de milenio, han servido como los árbitros de la verdad factual, los custodios del rigor lexicográfico, y los validadores de la confianza pública en el mundo angloparlante y más allá.⁸⁴

Al concluir este análisis jurídico exhaustivo, queda cristalino que la resolución de este caso marcará un estándar definitivo—un contrato social nuevo—para lo que podría denominarse la "segunda era de la propiedad intelectual": la era en la cual la inteligencia artificial generativa no es simplemente una herramienta para usuarios finales, sino un competidor directo, un distribuidor de contenidos, y un potencial destructor de los incentivos económicos que sostienen la creación de conocimiento verificado.

La respuesta que proporcionen los tribunales del Distrito Sur de Nueva York no solo decidirá el destino de daños estatutarios que podrían alcanzar sumas astronómicas. Definirá fundamentalmente si el futuro de la red será uno gobernado por exactitud curada por expertos humanos e instituciones responsables, o una amalgama algorítmica de verdades fragmentadas, medias verdades, información alucinada, y contenido descontextualizado que erosiona irremediabilmente la credibilidad de nuestras instituciones de información.⁸⁵

El fallo final en el MDL 3143, esperado para mediados o finales de 2026, escribirá el próximo capítulo no solo en la historia del derecho de autor estadounidense, sino en la historia de la información, la confianza, y el futuro de las instituciones verificadoras en la sociedad digital.

⁸⁴ Authors Guild Opinion, Doc. 716 (S.D.N.Y. Oct. 27, 2025), at 5, 220.

⁸⁵ In re: OpenAI, Inc. Copyright Infringement Litigation, MDL No. 3143 (S.D.N.Y. filed Apr. 3, 2025).

Bibliografía y Autoridades Citadas

I. Documentos Procesales Federales y Resoluciones Judiciales

- Encyclopædia Britannica, Inc., & Merriam-Webster, Inc. v. OpenAI, Inc., Civil Action No. 1:26-cv-2097 (S.D.N.Y. filed Mar. 13, 2026) (Complaint).
- Authors Guild et al. v. OpenAI, Inc., Doc. 716 (S.D.N.Y. Oct. 27, 2025) (Stein, J., Opinion and Order Denying Motion to Dismiss).
- Case 1:25-md-3143-SHS-OTW, Document 846 (S.D.N.Y. Nov. 24, 2025) (Mag. Judge Ona T. Wang, Order on Privilege Waiver).
- In re: OpenAI, Inc. Copyright Infringement Litigation, MDL No. 3143 (S.D.N.Y. filed Apr. 3, 2025) (Consolidation of multiple copyright infringement actions).
- The New York Times Co. v. OpenAI, Inc. & Microsoft Corp., No. 1:23-cv-11565 (S.D.N.Y. pending) (related copyright infringement and RAG-related claims).
- Advance Local Media, Inc. v. Cohere, Inc., slip op. (N.D. Ga. 2025) (unpublished; trademark confusion from AI hallucinations).

II. Estatutos Federales y Autoridades Legislativas

- 17 U.S.C. § 101 et seq. (Copyright Act of 1976, as amended; comprehensive codification of copyright law including work made for hire, derivative works, exclusive rights, and statutory damages).
- 17 U.S.C. § 106 (exclusive rights of copyright owners: reproduction, derivative works, distribution, public performance, public display).
- 17 U.S.C. § 107 (Fair Use doctrine; four-factor analysis: purpose/character, nature of work, amount/substantiality, market effect).
- 17 U.S.C. § 103 (works made for hire and derivative works; protection of compilations and selection/arrangement).
- 17 U.S.C. § 504(c)(1)-(2) (statutory damages: USD 750–USD 30,000 per work [innocent infringement]; up to USD 150,000 per work [willful infringement]).
- 15 U.S.C. § 1125(a)(1) (Lanham Act; false designation of origin and misleading representation of affiliation).
- 15 U.S.C. § 1125(c) (Lanham Act; dilution of famous marks; blurring and tarnishment).
- 18 U.S.C. § 1030 et seq. (Computer Fraud and Abuse Act; unauthorized access to computer systems).

III. Jurisprudencia Clave del Segundo Circuito y S.D.N.Y.

- Authors Guild v. Google, Inc., 804 F.3d 202 (2d Cir. 2015) (transformative nature of digital search and preservation; fair use analysis for mass digitization).

Harper & Row v. Nation Enterprises, 471 U.S. 539 (1985) (Supreme Court; market harm as dispositive factor in fair use analysis; President Ford memoir case).

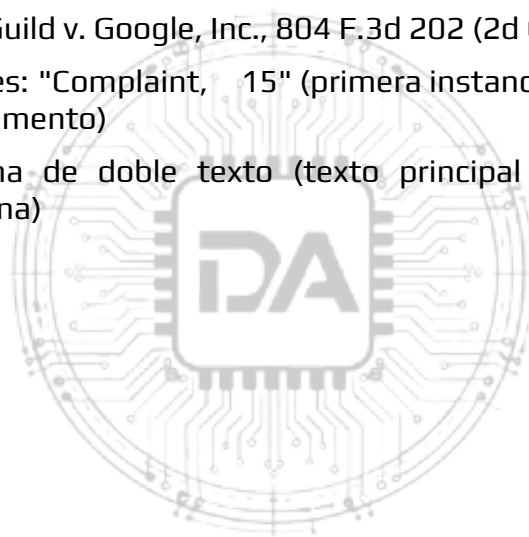
New York Times Co. v. Tasini, 533 U.S. 483 (2001) (Supreme Court; protection of selection, coordination, and arrangement in compilations; digital republication of articles).

Cariou v. Prince, 714 F.3d 694 (2d Cir. 2013) (transformative use in visual arts; balancing original expression with derivative creativity).

Nota Editorial: Sistema de Citas y Estándares de Revisión

Convenciones de citas implementadas:

- Estatutos federales: "17 U.S.C. § 107" (nunca "Section 107 of Title 17")
- Sentencias: "Authors Guild v. Google, Inc., 804 F.3d 202 (2d Cir. 2015)"
- Documentos procesales: "Complaint, 15" (primera instancia); "Id. 20" (referencias sucesivas al mismo documento)
- Notas al pie: Sistema de doble texto (texto principal con superíndices; notas completas al pie de página)



DERECHO ARTIFICIAL