

La protección de las bases de datos a la luz de la propiedad intelectual en el marco de la Cuarta Revolución Industrial¹

Ip protection of databases under the Fourth Industrial Revolution

JULIANA MARÍA BERNAL GUZMÁN²

RESUMEN

En la era de la cuarta revolución industrial y la economía del conocimiento, los datos han emergido como activos de contenido económico que impulsan la innovación y los desarrollos tecnológicos y la competitividad en diversas industrias. Las bases de datos desempeñan un rol fundamental en este ecosistema en tanto permiten almacenar, procesar y en general explotar grandes volúmenes de datos. Sin embargo, el régimen actual de protección legal de las bases de datos, principalmente basado en el derecho de autor, fue concebido en un contexto tecnológico y económico diferente al actual. Este marco desactualizado puede no abordar plenamente las complejidades y desafíos que plantea el entorno digital moderno, donde los datos se generan,

1 Fecha de recepción: 31 de mayo de 2024. Fecha de aceptación: 20 de noviembre de 2024. Received: May 31, 2024. Accepted: November 20, 2024. Artículo de revisión. DOI: <https://doi.org/10.18601/16923960.v24n1.01>

2 Abogada de la Universidad de los Andes con Opción Académica en Gestión para las Artes y la Cultura, y Magíster en Derecho Internacional de la misma universidad. Ha trabajado en la Oficina TIC del Ministerio de Salud y Protección Social, la Dirección de Gobierno Digital, y en del Despacho del Ministro de Tecnologías de la Información y las Comunicaciones. También ha asesorado emprendimientos en asuntos relacionados con propiedad intelectual y comercio electrónico. Ha sido par evaluadora de revistas especializadas en propiedad intelectual, ciencia política y relaciones internacionales. Contacto: julianabernalguzman@gmail.com. Orcid Id: <https://orcid.org/0009-0007-4706-9805>

comparten y utilizan de manera rápida. Es necesario reevaluar adaptar las regulaciones actuales para garantizar una protección efectiva en un entorno digital dinámico, asegurando al mismo tiempo el acceso y la reutilización ética de los datos.

Palabras claves: Propiedad Intelectual, Bases de Datos, Modelos de Datos, Derecho de Autor, Inteligencia Artificial.

ABSTRACT

In the era of the fourth industrial revolution and the knowledge economy, data has emerged a critical economic asset, driving innovation, technological advancements, and competitive advantages across industries. Databases play a fundamental role in this ecosystem, allowing large volumes of data to be stored, processed and in general exploited. However, the current legal protection regime for databases, primarily grounded in copyright law, was conceived in a vastly different technological and economic context. This outdated framework may not fully address the complexities and challenges posed by modern digital environments, where data is rapidly generated, shared, and utilized. It is necessary to re-evaluate and adapt current regulations to guarantee effective protection in a dynamic digital environment, while ensuring access and ethical reuse of data.

Key words: Intellectual Property, Databases, Data Models, Copyright, Artificial Intelligence.

INTRODUCCIÓN

La economía global está siendo transformada por la especialización y diversificación de las Tecnologías de la Información y las Comunicaciones (TIC). En particular, los efectos que las tecnologías de la cuarta revolución industrial están teniendo en la estructura y funcionamiento de la economía debido a los avances de la Industria 4.0, están generado cambios estructurales y tecnológicos sin precedentes en diversos sectores. Este fenómeno está dando lugar a una economía donde los factores del conocimiento y la tecnología son determinantes para el crecimiento y la competitividad a nivel empresarial y nacional. De este modo los activos intangibles, o el "capital de conocimiento",

se han convertido en el epicentro de la investigación económica, destacándose como factores clave para el crecimiento y la competitividad³.

Bajo este contexto, los datos empiezan a ser concebidos como un capital intangible que impulsa las industrias tradicionales, nuevos modelos de negocio, y la innovación. Los datos tienen una capacidad transformadora de la sociedad, cuyo uso y aprovechamiento impulsa la toma de decisiones tanto en lo privado como en lo público, y habilita la creación de nuevos productos y servicios, así como el mejoramiento de aquellos que actualmente se encuentran circulando en el mercado. Piénsese, por ejemplo, en el uso de tecnologías para impulsar el agro (AgroTech), en el uso de internet de las cosas en dispositivos y aplicaciones médicas (Internet of Medical Things, IoMT), en el uso de la Inteligencia Artificial (IA) para impulsar los procesos de aprendizaje, o en la creación de modelos predictivos para el sector financiero, entre innumerables casos en los que el acceso, intercambio, reutilización y explotación de los datos juega un rol fundamental.

Las tecnologías emergentes se ven impulsadas por la aceleración de la digitalización, dependen de los datos y, a menudo, generan más datos. En el caso de la IA, los avances actuales están siendo impulsados por tres factores principales: avances en la potencia de cómputo, mejores algoritmos y la disponibilidad de datos⁴. A su vez, los grandes volúmenes de datos que se generan implican la necesidad de desarrollar tecnologías que puedan gestionar y analizar esas grandes cantidades de datos de manera eficiente. Sobre el particular, Martínez Devia señala:

“El uso continuo que se hace de distintas tecnologías (webs, aplicaciones, servicios, sensores incorporados en dispositivos, las búsquedas en internet, las redes sociales, computadores portátiles, teléfonos inteligentes, dispositivos GPS, entre otros) ha incrementado la cantidad de información que se almacena cada día.

El volumen de datos que se recolecta actualmente no podría ser analizado por los métodos tradicionales de almacenamiento, acceso y análisis, por esta razón se han desarrollado nuevas herramientas de IA, la implementación de algoritmos y estadísticas mediante las cuales se pueden obtener resultados, tales como el comportamiento de las personas, sus gustos, la toma de decisiones, el reconocimiento de voz, la identificación de objetos, el diagnóstico de enfermedades, el ahorro de energía, la elaboración de perfiles para analizar o predecir aspectos relativos al rendimiento profesional, la situación económica, la salud, los intereses, la

3 Yannis Caloghirou, et. Alia, “Tracking Intangibles in the EU: The GLOBALINTO Input-Output Intangibles Database”. Presentación para la 17ª Conferencia Internacional de Globelics, Heredia, Costa Rica, 3-5 de noviembre de 2021

4 Organización Mundial de la Propiedad Intelectual (OMPI), “Diálogo de la OMPI sobre Propiedad Intelectual (PI) y Tecnologías de Vanguardia: Cuarta sesión – Resumen de la Cuarta Sesión”, Ginebra, 22-23 de septiembre de 2021

fiabilidad, el comportamiento o la ubicación, datos que se utilizan con diversos fines y se encuentran al alcance de empresas, Estados e incluso particulares⁵.

Así pues, los volúmenes de datos que circulan en esta economía son cada vez mayores, lo cual ha impulsado el desarrollo del big data y una transformación de los métodos tradicionales de almacenamiento de datos en bases de datos convencionales⁶. Esto ha llevado a la adopción de nuevas formas de almacenamiento de datos, como data lakes⁷, data warehouses⁸ y bases de datos NoSQL⁹.

Ante este panorama, la forma en que son almacenados los datos plantean cuestiones frente a la Propiedad Intelectual (PI) y su ámbito de protección. Por una parte, es claro que los datos son explotados en aplicaciones y sistemas que son protegidos bajo la PI. Por otro lado, las bases de datos como tal, donde reposa la información también podrían estar protegidas bajo la PI. Desde este ámbito de protección, se busca proteger la originalidad en la selección y disposición de los datos.

Esta originalidad constituye una suerte de know-how de contenido económico, ya que una base de datos puede permitir agilizar las validaciones. Este aspecto, en la opinión de la autora, se trata de un ámbito de la propiedad intelectual que se ha dejado a un lado, a pesar del importante potencial que se tiene. De acuerdo con la Dirección Nacional de Derecho de Autor, desde el 1 de enero del año 2000 hasta el 4 de abril de 2024, sólo reposaban

5 Andrea Martínez Devia, "La inteligencia artificial, el big data y la era digital: ¿una amenaza para los datos personales?" *Revista La Propiedad Inmateria*. 27, (2019), 7.

6 Las bases de datos convencionales, o relacionales, se basan en un modelo en el que los datos se organizan en tablas, filas y columnas, con el fin de facilitar su visualización y comprensión frente cómo se relacionan las diferentes estructuras de datos entre sí. "¿Qué es una base de datos relacional?", Google Cloud, consultado el 10 de junio de 2024, <https://cloud.google.com/learn/what-is-a-relational-database?hl=es-419>

7 Los data lakes o "lago de datos" son repositorios que permiten centralizar tanto datos estructurados como no estructurados a cualquier escala. "¿Qué es un lago de datos?", Amazon Web Services, consultado el 10 de junio de 2024, <https://aws.amazon.com/es/what-is/data-lake/>

8 Los data warehouses o almacenes de datos son sistemas de gestión de datos que centraliza y fusiona grandes cantidades de datos de múltiples fuentes, con el fin de habilitar y dar soporte a las tareas de inteligencia empresarial (BI), especialmente las analíticas. "¿Qué es un almacén de datos?", Oracle, consultado el 10 de junio de 2024, <https://www.oracle.com/co/database/what-is-a-data-warehouse/>

9 "NoSQL, también conocido como "no sólo SQL" o "no SQL", es un enfoque utilizado en el diseño de bases de datos que permite el almacenamiento y consulta de datos fuera de las estructuras tradicionales que se encuentran en las bases de datos relacionales.". "¿Qué es una base de datos NoSQL?", IBM, consultado el 10 de junio de 2024, <https://www.ibm.com/es-es/topics/nosql-databases>

29 inscripciones de obras en las cuales se hace referencia a que son una base de datos.¹⁰

Así las cosas, el presente escrito tiene por objetivo poner de presente que la protección de las bases de datos bajo el derecho de autor se fundamenta en normas concebidas antes del surgimiento de tecnologías que hoy impulsan la economía basada en datos. Estas normas se basan, en muchos casos, en analogías con la protección que el derecho de autor otorga a las compilaciones y antologías, y que pueden estar dejando de lado otras creaciones que también comprenden un esfuerzo intelectual y que resultan fundamentales para el desarrollo de las tecnologías emergentes. De esta manera, la normativa actual se queda corta frente a las necesidades del mercado y la evolución de la industria, que cada vez pone más su atención en el uso de tecnologías digitales. En consecuencia, la protección que hoy otorga la propiedad intelectual a las bases de datos en Colombia resulta ser una figura atemporal y obsoleta.

GENERALIDADES SOBRE LA PROTECCIÓN DE LAS BASES DE DATOS EN EL MARCO DE LA PROPIEDAD INTELECTUAL

Tradicionalmente, las bases de datos han sido protegidas por el derecho de autor, bajo el entendido de que son el resultado del esfuerzo de su creador de sistematizar un conjunto de datos, lo cual da lugar a la creación de una obra. Así las cosas, se parte de la idea de que un dato aislado no es una creación intelectual en sí misma, ni constituye conocimiento *per se* sino cuando

10 Dirección Nacional de Derecho de Autor, respuesta del 4 de abril de 2024 al derecho de petición con Radicado 1-2024-33658. Sobre el particular, la DNDA hace las siguientes precisiones: "Respecto a la inscripción de una base de datos en el Registro Nacional del Derecho de Autor, conforme a lo dispuesto en el artículo 21 de la Resolución 303 del 5 de noviembre de 2010: '*Son registrables ante la DNDA las compilaciones o colecciones que hayan sido formadas gracias a un método o sistema de selección que constituya una creación original. Del mismo modo, serán registrables, bien sea en la forma de un soporte lógico (software) o como una obra escrita (literaria), las bases de datos cuya selección o disposición de las materias que la conforman, constituyan una creación intelectual*'. Entonces, nuestro sistema no permite filtrar la búsqueda de bases de datos *per se*, y la mayoría de bases de datos se encuentran inscritas bajo la categoría de obra literaria y obra de soporte lógico. Teniendo en cuenta eso, se realizó la búsqueda en el sistema del Registro Nacional del Derecho de Autor, buscando "BASE DE DATOS", en título de la obra y en casilla de observaciones, la cual arrojó los siguientes resultados: Bajo el parámetro de búsqueda por título: se encontraron dieciocho (18) inscripciones desde el 1 de enero del año 2000 hasta el 4 de abril de 2024, nueve correspondientes a la categoría de soporte lógico y nueve correspondientes a la categoría de obra literaria inédita. Bajo el parámetro de búsqueda por casilla de observaciones: se encontraron once (11) inscripciones desde el 1 de enero del año 2000 hasta el 4 de abril de 2024, tres correspondientes a la categoría de soporte lógico y ocho correspondientes a la categoría de obra literaria inédita.

es analizado en conjunto con otros datos. En consecuencia, la propiedad intelectual entra a proteger el esfuerzo creativo y organizativo invertido en la recopilación, ordenamiento y presentación de esos datos.

De acuerdo con la Oficina del Derecho de Autor de Estados Unidos, en la terminología del derecho de autor, una base de datos se considera una 'compilación', es decir, una obra formada por la recopilación y ensamblaje de materiales preexistentes o de datos¹¹. Así pues, señala esta oficina que las compilaciones constituyen una de las formas más antiguas de autoría protegidas por la ley estadounidense, remontándose al siglo XVIII, cuando eran protegidas como una suerte de 'libros' bajo el primer estatuto federal de derechos de autor¹².

Durante el Siglo XIX se desarrollaron dos corrientes para la protección de las bases de datos bajo el derecho de autor. Por una parte, la doctrina del "sudor de la frente" ("sweat-of-the-brow doctrine" en inglés), planteaba que el mero esfuerzo en que incurre una persona para crear una base de datos genera derechos de propiedad intelectual, independientemente de que haya o no creatividad u originalidad en la creación de la misma. Según esta doctrina, el esfuerzo económico, técnico e investigativo en el que incurre el compilador, es una contribución que justifica la protección¹³.

Por otra parte, la doctrina de la originalidad fue surgiendo como un contrapeso a la doctrina del sudor de la frente, especialmente en casos que involucran compilaciones de materiales textuales, como por ejemplo libros de derecho. En estos casos, se aduce que existe en efecto una contribución crítica por parte del autor en la selección y organización de los componentes que conforman la compilación y, en consecuencia, dicha contribución es la que justifica la protección a la luz del derecho de la propiedad intelectual.

Como se explicará más adelante con mayor profundidad, para la protección de las bases de datos a la luz del derecho de autor, particularmente en Colombia, finalmente se acogió la doctrina de la originalidad, protegiendo sólo aquellas bases de datos que cumplieran con este requisito, y, por consiguiente, quedando por fuera de este régimen aquellas creaciones en las que no existe como tal el elemento de creatividad.

11 Ver art. 101 del Título 17 del Código de los Estados Unidos: "A "compilation" is a work formed by the collection and assembling of preexisting materials or of data that are selected, coordinated, or arranged in such a way that the resulting work as a whole constitutes an original work of authorship. The term "compilation" includes collective works."

12 Oficina del Derecho de Autor de Estados Unidos, "Report on legal protection for databases", agosto de 1997. Disponible en: <https://www.copyright.gov/reports/db4.pdf> p. 3

13 Op. Cit. p. 3

Existen varios instrumentos internacionales que respaldan la protección de las bases de datos a través del Derecho de Autor. En primer lugar, el artículo 2.5 del Convenio de Berna para la Protección de las Obras Literarias y Artísticas protege las bases de datos como compilaciones de obras literarias o artísticas, en función de la selección o disposición de las materias, que dan lugar a la creación de una obra independiente. Sobra mencionar que este tratado fue adoptado en 1886, cuando no existía lo que hoy se conoce como bases de datos:

"Las colecciones de obras literarias o artísticas tales como las enciclopedias y antologías que, por la selección o disposición de las materias, constituyan creaciones intelectuales estarán protegidas como tales, sin perjuicio de los derechos de los autores sobre cada una de las obras que forman parte de estas colecciones"¹⁴

Posteriormente, en 1994 se celebró el Acuerdo de la OMC sobre los Aspectos de los Derechos de Propiedad Intelectual relacionados con el Comercio (ADPIC), reconociendo una protección tanto a los programas de ordenador (*software*) como a las compilaciones de datos, a la luz de la propiedad intelectual:

"Artículo 10. Programas de ordenador y compilaciones de datos

1. Los programas de ordenador, sean programas fuente o programas objeto, serán protegidos como obras literarias en virtud del Convenio de Berna (1971).
2. Las compilaciones de datos o de otros materiales, en forma legible por máquina o en otra forma, que por razones de la selección o disposición de sus contenidos constituyan creaciones de carácter intelectual, serán protegidas como tales. Esa protección, que no abarcará los datos o materiales en sí mismos, se entenderá sin perjuicio de cualquier derecho de autor que subsista respecto de los datos o materiales en sí mismos"¹⁵.

Esta norma es relevante no solo por la protección que otorga a las bases de datos, sino porque el artículo reconoce la relación existente entre el software y las bases de datos como dos objetos independientes (pero correlacionados) de protección según el derecho de autor. En efecto, dado que en la práctica el software y las bases de datos están fuertemente interrelacionados, especialmente en lo que respecta a la interoperabilidad entre sistemas de

14 Convenio De Berna para la Protección de las Obras Literarias y Artísticas. Art. 2, núm. 5.

15 Acuerdo de la OMC sobre los Aspectos de los Derechos de Propiedad Intelectual relacionados con el Comercio (ADPIC), Art. 10.

información. Las bases de datos son fundamentales para el intercambio de información, ya que almacenan los datos que los sistemas de información comparten.¹⁶

Posteriormente, a través del Tratado de la OMPI sobre Derecho de Autor (TODA o WTC por sus siglas en inglés) de 1996, se desarrolló y complementó los principios de la Convención de Berna, con el fin de ofrecer una protección adecuada del derecho de autor en el entorno digital, ocupándose de los programas de ordenador y las compilaciones de datos u otros materiales (bases de datos)¹⁷. En particular, el artículo 5 señala:

“Las compilaciones de datos o de otros materiales, en cualquier forma, que por razones de la selección o disposición de sus contenidos constituyan creaciones de carácter intelectual, están protegidas como tales. Esa protección no abarca los datos o materiales en sí mismos y se entiende sin perjuicio de cualquier derecho de autor que subsista respecto de los datos o materiales contenidos en la compilación”¹⁸.

Ahora bien, en el marco de la Comunidad Andina de Naciones, el literal II) del artículo 4 de la Decisión Andina 351 de 1993 establece la protección de las bases de datos como obras protegibles a la luz del derecho de autor, en los siguientes términos:

16 Sobre el particular, la Directiva 2009/24/CE del Parlamento Europeo y del Consejo sobre la Protección Jurídica de los programas de ordenador dispone: “La función de un programa de ordenador es comunicarse y trabajar con otros componentes del sistema de ordenador y con sus usuarios y, a tal fin, se exige contar con un sistema lógico y, cuando sea conveniente, físico de interconexión e interacción para permitir a los elementos de los soportes físicos y lógicos trabajar con otros soportes físicos y lógicos y con usuarios, en la forma prevista. Las partes del programa que establecen dicha interconexión e interacción entre los elementos de software y hardware suelen denominarse «interfaces». La interconexión e interacción funcional suele conocerse como «interoperabilidad»; dicha interoperabilidad puede ser definida como la capacidad de los programas de ordenador para intercambiar información y utilizar mutuamente la información así intercambiada”. En concordancia con lo anterior, en Colombia, el Decreto 620 de 2020, establece el servicio de interoperabilidad como uno de los servicios ciudadanos digitales base para la transformación digital del Estado, definiéndolo en los siguientes términos: “Servicio de interoperabilidad: Es el servicio que brinda las capacidades necesarias para garantizar el adecuado flujo de información e interacción entre los sistemas de información de las entidades, permitiendo el intercambio, la integración y la compartición de la información, con el propósito de facilitar el ejercicio de sus funciones constitucionales y legales, acorde con los lineamientos del marco de interoperabilidad”.

17 “Tratado de la OMPI sobre Derecho de Autor”, Organización Mundial de la Propiedad Intelectual OMPI, consultado el 10 de junio de 2024, <https://www.wipo.int/treaties/es/ip/wct/>

18 Tratado de la OMPI sobre Derecho de Autor. Art. 5.

"Artículo 4.- La protección reconocida por la presente Decisión recae sobre todas las obras literarias, artísticas y científicas que puedan reproducirse o divulgarse por cualquier forma o medio conocido o por conocer, y que incluye, entre otras, las siguientes:

(...)

II) Las antologías o compilaciones de obras diversas y las **bases de datos**, que por la **selección o disposición de las materias** constituyan **creaciones personales**." (negrillas por fuera del texto original)¹⁹

En concordancia, el artículo 28 *ibidem* señala:

"Artículo 28.- Las bases de datos son protegidas siempre que **la selección o disposición de las materias constituyan una creación intelectual**. La protección concedida **no se hará extensiva a los datos o información compilados**, pero no afectará los derechos que pudieran subsistir sobre las obras o materiales que la conforman". (negrillas por fuera del texto original)²⁰

Como ya se había indicado, en Colombia, el artículo 5 de la Ley 23 de 1982 protege como obras independientes, "las obras colectivas, tales como publicaciones periódicas, antologías, diccionarios y similares, cuando el método o sistema de selección de organización de las distintas partes u obras que en ellas intervienen, constituye una creación original (...)"²¹.

Asimismo, mediante Circular Conjunta firmada por la Procuraduría General de la Nación y la Dirección Nacional de Derecho de Autor del 24 de abril de 2006, se indicó que una base de datos puede ser definida como "(...) una recopilación de datos puntuales u obras (preexistentes o no, originales o derivadas) que pueden ser utilizados manualmente o por medios electrónicos, hecha de forma organizada de tal manera que permite la recuperación de la información por los usuarios"²², y que la selección y disposición de materias es la que le otorga el carácter de original.

De acuerdo con lo anterior, la protección de bases de datos bajo el derecho de autor en Colombia depende de su carácter original. Sobre el particular, la Dirección Nacional de Derecho de Autor señala que el criterio

19 Decisión Andina 351 de 1993. Art. 4.

20 Decisión Andina 351 de 1993. Art. 28.

21 Ley 23 de 1982. Art. 5.

22 Circular Conjunta 01 de 2006. [Procuraduría General de la Nación & Dirección Nacional de Derecho de Autor]. Orientaciones para el cumplimiento de normas de Derecho de Autor y Derechos Conexos, en lo pertinente a la contratación estatal de obras y prestaciones protegidas y sus buenas prácticas. 24 de abril de 2006. (Colombia.)

de originalidad comprende una recolección de datos que, en función de la selección o disposición de los datos:

"(...) la selección, disposición y orden sistemático de la información debe constituir en si misma una creación original. Por oposición, la simple ordenación mecánica o acumulación de datos sin ningún criterio de selección, o la simple ordenación alfabética, numérica o cronológica de éstos, por muy dispendiosa que sea, no reviste altura creativa pues constituye una labor mecánica que bien puede ser realizada por una máquina. La originalidad debe hallarse en el particular método o sistema de selección o de organización de los distintos datos."²³

Ahora bien, es importante señalar que en la Circular Conjunta antes mencionada, se indica que el examen de originalidad de las bases de datos no puede ser realizado de manera abstracta, quedando sujeto a un análisis caso a caso:

"No es posible de manera general y abstracta determinar a-priori cuáles bases de datos se constituyen como creaciones originales y cuáles no. Este análisis debe ser adelantado para cada caso en concreto atendiendo a la destreza, el esfuerzo y la contribución del compilador en la selección, distribución y orden sistemático de la información"²⁴.

Ahora bien, en lo que corresponde a la protección de las bases de datos a la luz del derecho norteamericano, la Ley de Copyright de Estados Unidos de 1976, en su sección 101, define una base de datos como una "compilación", es decir, una obra que es formada a partir de la colección y ensamblaje de materiales preexistentes de datos que son seleccionados, coordinados u organizados de forma tal que el resultado de dicho trabajo, considerado como un "todo", constituye una obra original de autoría²⁵. Más adelante, en la sección 103, se indica:

23 Dirección Nacional de Derecho de Autor, respuesta del 8 de noviembre de 2005 al derecho de petición con Radicado 2-2005-10632. Disponible en: <https://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?dt=S&i=20039>

24 Circular Conjunta 01 de 2006. [Procuraduría General de la Nación & Dirección Nacional de Derecho de Autor]. Orientaciones para el cumplimiento de normas de Derecho de Autor y Derechos Conexos, en lo pertinente a la contratación estatal de obras y prestaciones protegidas y sus buenas prácticas. 24 de abril de 2006. (Colom.)

25 "A "compilation" is a work formed by the collection and assembling of preexisting materials or of data that are selected, coordinated, or arranged in such a way that the resulting work as a whole constitutes an original work of authorship. The term "compilation" includes collective works". Ley de Derechos de Autor de 1976 (Copyright Act of 1976), incorporada en el Título 17 del Código de los Estados Unidos. Sección 101 (EE.UU).

"The copyright in a compilation or derivative work extends only to the material contributed by the author of such work, as distinguished from the preexisting material employed in the work, and does not imply any exclusive right in the preexisting material. The copyright in such work is independent of, and does not affect or enlarge the scope, duration, ownership, or subsistence of, any copyright protection in the preexisting material"²⁶.

Si bien podría inferirse que la inclusión del término "original" dentro de la definición de compilación, implica que la Ley de Copyright de Estados Unidos de 1976 había optado por la doctrina de la originalidad, no fue sino hasta el año 1991 que la Corte Suprema de los Estados Unidos rechazó explícitamente la doctrina del sudor de la frente en el caso *Feist Publications vs. Rural Telephone Service Co.* Antes de que se profiriera dicha decisión, en varios casos se había discutido la viabilidad de la doctrina del sudor de la frente²⁷. Así pues, en la sentencia del caso *Feist Publications*, la Corte hizo una revisión de la historia de la protección de las compilaciones a la luz del Copyright y del desarrollo de la doctrina del sudor de la frente, indicando lo siguiente:

"The key to resolving the tension lies in understanding why facts are not copyrightable. The *sine qua non* of copyright is originality. To qualify for copyright protection, a work must be original to the author (...). Original, as the term is used in copyright, means only that the work was independently created by the author (as opposed to copied from other works), and that it possesses at least some minimal degree of creativity (...). To be sure, the requisite level of creativity is extremely low; even a slight amount will suffice. The vast majority of works make the grade quite easily, as they possess some creative spark, "no matter how crude, humble or obvious" it might be. (...) Originality does not signify novelty; a work may be original even though it closely resembles other works so long as the similarity is fortuitous, not the result of copying. To illustrate, assume that two poets, each ignorant of the other, compose identical poems. Neither work is novel, yet both are original and, hence, copyrightable (...)"²⁸.

En este caso, Rural Telephone Service Co había demandado a Feist Publications por una presunta infracción al Copyright al haber utilizado la información de sus directorios telefónicos para crear su propio. La Corte Suprema de los Estados Unidos indicó que los directorios carecían de la "chispa creativa"

26 Ley de Derechos de Autor de 1976 (Copyright Act of 1976), incorporada en el Título 17 del Código de los Estados Unidos. Sección 103.

27 Ver casos: *Jeweler's Circular Publishing Co. v. Keystone Publishing Co.* (1922); *Leon v. Pacific Telephone & Telegraph Co.* (1933); *Southern Bell Telephone & Telegraph Co. v. Associated Telephone Directory Publishers* (1951); *Merritt Forbes & Co. v. Newman Investment Securities, Inc.* (1982),

28 Corte Suprema de los Estados Unidos [C.S.E.U.] [U.S. Supreme Court], marzo, 27, 1991 (EEUU.)

(*creative spark*) en la selección de los contenidos. Así las cosas, por tratarse de una mera recopilación de datos, ausente de creatividad, los directorios telefónicos no pueden ser objeto de protección por parte del Copyright, rechazando la doctrina del sudor de la frente.

Por su parte, durante la década de los 90 en la Comunidad Europea también se puso de presente la necesidad de contar con un régimen claro, armonizado y uniforme para la protección de bases de datos por parte de los Estados miembros. Es así como surge la Directiva 96/9/CE del Parlamento Europeo y del Consejo, de 11 de marzo de 1996, sobre la protección jurídica de las bases de datos. Para efectos del presente escrito, resulta pertinente indicar que, dentro de las motivaciones para la expedición de este instrumento normativo, se puso de presente la importancia de contar con este régimen de cara a las implicaciones negativas que podían generarse en el mercado a falta de seguridad y protección jurídica:

“(.. 2) Considerando que las diferencias de protección jurídica de las bases de datos en las legislaciones de los Estados miembros inciden de forma directa y negativa en el funcionamiento del mercado interior en lo que se refiere a las bases de datos y, en particular, en la libertad de las personas físicas y jurídicas de suministrar bienes y prestar servicios en el sector de las bases de datos de acceso en línea conforme a un fundamento jurídico armonizado en toda la Comunidad; que dichas diferencias pueden agudizarse a medida que los Estados miembros adopten nuevas disposiciones en un sector que está cobrando una dimensión cada vez más internacional;

3) Considerando que deben suprimirse las diferencias que tienen un efecto distorsionados sobre el funcionamiento del mercado interior y que debe prevenirse la aparición de otras nuevas; que no es preciso eliminar las diferencias que en la actualidad no afectan negativamente al funcionamiento del mercado interior o al desarrollo de un mercado de la información en la Comunidad; (...)²⁹
(Negrillas por fuera del texto original)

Es importante mencionar que con la Directiva 96/9/EC sobre la protección jurídica de las bases de datos, la Unión Europea optó por contemplar, además del régimen establecido en el marco de derechos de autor para las bases de datos originales, con un régimen paralelo que permitiera proteger el esfuerzo invertido en la construcción de aquellas bases de datos que no contaran con ese elemento de creatividad para su construcción, denominado derecho sui generis. Este derecho otorga un término de protección de 15 años a partir del 1 de enero del año siguiente a la fecha de finalización de la creación de

29 Directiva 96/9/CE. [Parlamento Europeo y Consejo Europeo]. Protección jurídica de las bases de datos. 11 de marzo de 1996. (Unión Europ.)

la base de datos. Este término puede extenderse si se realizan inversiones sustanciales en la base de datos que resulten en una nueva inversión sustancial, cualitativa o cuantitativa.³⁰

DERECHO SUI GENERIS SOBRE LAS BASES DE DATOS

Teniendo en cuenta que la protección de las bases de datos otorgada por el derecho de autor se terminó construyendo bajo la teoría de la originalidad, surgió una cuestión frente a cómo proteger aquellas bases de datos cuya elaboración representaba una inversión sustancial, en términos no sólo del tiempo y el esfuerzo que ha requerido su elaboración, sino en los recursos financieros requeridos para tal efecto.

Es importante anotar que la forma de protección del derecho sui generis es distinta a la del derecho de autor. Mientras que el derecho de autor protege la forma en que está diseñada la base de datos, el derecho sui generis protege el contenido de las bases de datos como tal³¹. Es decir, bajo el amparo de este derecho sui generis, se prohíbe la extracción de información o el uso de los datos sin consentimiento o licencia. No obstante, no existe impedimento alguno para acceder y explotar a los datos a través de otras fuentes, sin que constituya infracción alguna³².

Ahora bien, frente a la protección otorgada por el derecho sui generis para las bases de datos, el Tribunal de Justicia de la Unión Europea ha señalado que se requiere una inversión sustancial para que pueda aplicar este derecho. Por ejemplo, en el caso *Fotball Dataco Ltd, Scottish Premier League Ltd, Scottish Football League, PA Sport UK Ltd vs. Yahoo! UK Ltd*. del año 2012, se señaló que la simple creación de una base de datos para partidos de fútbol, como lo es una relación de horarios y resultados de los partidos, no constituye una base de datos original (y por ende no aplica la protección del derecho de autor), ni una inversión sustancial en la obtención y organización de datos, no resultando procedente la protección bajo el derecho sui generis.

Por otra parte, en el caso *CV-Online Latvia SIA vs. Melons* del año 2014, correspondió analizar si una base de datos que contenía ofertas de trabajo podía ser protegida por el derecho sui generis. CV-Online Latvia SIA, una empresa de reclutamiento había demandado a Melons por copiar contenido

30 Directiva 96/9/CE. [Parlamento Europeo y Consejo Europeo]. Protección jurídica de las bases de datos. 11 de marzo de 1996. (U.E.). Art. 10.

31 Eduardo De La Parra Trujillo, "La protección de las bases de datos no originales en la legislación autoral mexicana (comparación con la directiva de la Unión Europea sobre bases de datos)". *Revista De La Facultad De Derecho De México*, 54(241), (2017):211–237, doi: <https://doi.org/10.22201/fder.24488933e.2004.241.61594>

32 Sergio Cámara Lapuente, "Últimas orientaciones internacionales sobre la protección jurídica de las bases de datos". *Revista Chilena de Derecho*, 26(1), (1999):9-61

de su base de datos, a partir de un hipervínculo con un motor de búsqueda especializado que indexaba y copiaba en su propio servidor el contenido de los sitios de Internet con ofertas de empleo, entre otros, el sitio de Internet de CV-Online, proporcionando a sus usuarios acceso al contenido de otras bases de datos. Así pues, el Tribunal determinó que había una inversión sustancial en la obtención, verificación o presentación del contenido de la base de datos de CV-Online.

Por otro lado, en el caso *Ryanair Ltd v PR Aviation BV* del año 2015, Ryanair alegó que PR Aviation había vulnerado sus derechos sobre su recopilación de datos y que había infringido las condiciones generales de utilización de su sitio en Internet. PR Aviation disponía un sitio en Internet en el que los consumidores podían hacer búsquedas en los datos de vuelo de compañías aéreas de bajo coste, comparar los precios y reservar un vuelo a cambio del pago de una comisión. Ryanair alegaba que dicho sitio infringía sus derechos sobre sus bases de datos, más aun teniendo en cuenta que sus propios términos y condiciones indicaban que Ryanair no autorizaba la venta de tiquetes por parte de otros sitios de Internet, y prohibía el uso de sistemas automatizados o de software para la extracción de datos, salvo que se hubiera celebrado un contrato de licencia para tal efecto. Así las cosas, el Tribunal de Justicia de la Unión Europea estableció en instancia prejudicial lo siguiente:

“La Directiva 96/9/CE del Parlamento Europeo y del Consejo, de 11 de marzo de 1996, sobre la protección jurídica de las bases de datos, debe interpretarse en el sentido de que no es aplicable a una base de datos que no está protegida por los derechos de autor ni por el derecho sui generis en virtud de la propia Directiva, por lo que los artículos 6, apartado 1, 8 y 15 de ésta no se oponen a que el creador de dicha base de datos establezca limitaciones contractuales a su utilización por terceros, sin perjuicio del Derecho nacional aplicable”³³.

Si bien en el marco de la Unión Europea se definió un régimen de protección para las bases de datos no originales, en otros lugares del mundo, la viabilidad para contar con una regulación de esta naturaleza no era tan clara. Así pues, la decisión del caso *Feist Publications vs. Rural Telephone Service Co.*, así como los desarrollos tecnológicos adelantados en los principios de los años 90 para la recolección, organización, reproducción y difusión de información, y las acciones adelantadas por la Unión Europea en la armonización sobre sus normas relacionadas con la protección de bases de datos, y el derecho sui generis otorgado a estas, dieron como resultado, que hacia 1996 se empezaran a dar discusiones globales para una nueva forma de protección de bases

33 Tribunal de Justicia de la Unión Europea [T.J.U.E.] [Court of Justice of the European Union], Sala Segunda. 15 de enero de 2015. Asunto C-30/14.(U.E)

de datos, tanto en la Organización Mundial para la Propiedad Intelectual (OMPI), como en el Congreso de los Estados Unidos.³⁴

Así las cosas, en ese año la Unión Europea propuso, en el marco de las reuniones regulares de los Comités de Expertos de la OMPI un tratado sobre la protección jurídica de las bases de datos, que contemplara la protección otorgada por el derecho *sui generis*, similar a la proporcionada en la Directiva 96/9/EC, pero adaptada a un contexto internacional. De este modo, se buscaba proteger las inversiones que requieren la construcción y mantenimiento de las bases de datos, y de esta forma establecer un incentivo para invertir en la recopilación y organización de datos, promoviendo la innovación y el desarrollo económico. Es oportuno indicar que, para ese entonces, se estaba discutiendo el Tratado de la OMPI sobre Derecho de Autor (WCT) y el Tratado de la OMPI sobre Interpretación o Ejecución y Fonogramas (WPPT).

Si bien la propuesta de la Unión Europea no dio como resultado un tratado internacional sobre la protección de bases de datos, sí influyó en discusiones globales e internas para desarrollar su propia legislación sobre las bases de datos. En efecto, con el propósito de contar con una adecuada protección de las bases de datos con los desarrollos digitales, y reconociendo que estas constituían un elemento fundamental para el desarrollo de múltiples industrias, a finales de la década de los 90, el Congreso de los Estados Unidos solicitó a la Oficina del Derecho de Autor realizar un informe en el que se examinara el régimen de protección legal de las bases de datos bajo el derecho norteamericano.

Así pues, con el Reporte sobre la Protección Legal para las bases de datos de la Oficina del Derecho de Autor de Estados Unidos de 1997, se realizó una revisión de la normativa nacional e internacional sobre la materia, consultando además a expertos en propiedad intelectual, creadores de bases de datos, y representantes de la industria, entre otros.

El Reporte encontró que el régimen definido en el Copyright Act de 1976, así como el precedente establecido en el caso *Feist Publications*, dejaban por fuera del ámbito de protección a las bases de datos no originales. En este sentido, se encontró que la regulación no proveía incentivos adecuados para producir distintas formas de bases de datos, lo cual desincentivaba a los productores y afectaban las decisiones de inversión en este mercado, ya que la percepción del riesgo era muy alta. De igual modo, se puso de presente la necesidad de encontrar un equilibrio entre los derechos de los creadores de las bases de datos y los derechos de los usuarios, señalando la necesidad de proteger los derechos de acceso a la información, y que no se obstaculizara la investigación y la promoción del conocimiento.

34 Reporte sobre la Protección Legal para las bases de datos. [Oficina del Derecho de Autor de Estados Unidos] Agosto de 1997. (EE.UU) P. 1.

En términos de Eduardo De La Parra Trujillo, aunque efectivamente hubo intentos de iniciativas legislativas para proteger el contenido de las bases de datos de forma similar al modelo de la Unión Europea, las críticas de diversos sectores que exigían un sistema más flexible, hicieron que los proyectos legislativos posteriores se enfocaran más en un sistema de represión de la competencia desleal y hacia la apropiación indebida (*unfair competition and misappropriation* en inglés).³⁵ Así las cosas, en Estados Unidos la protección de bases de datos no originales se protege a través de un conjunto de normas, como lo es la Ley de Secretos Comerciales (*Uniform Trade Secrets Act*) y la Ley de Competencia Desleal (*Lanham Act*), así como la doctrina de la apropiación indebida (*misappropriation doctrine*).³⁶

Por otro lado, resulta oportuno indicar que en el año 2002 se llevó a cabo el Comité Permanente de Derecho de Autor y Derechos Conexos de la OMPI, en el que la Comunidad Europea y sus Estados miembros presentaron el documento "La protección Jurídica de las bases de datos", poniendo de presente nuevamente la necesidad de no sólo proteger las bases de datos que constituyen creaciones creativas, sino retomar la protección de aquellas denominadas "hechas con el sudor de la frente":

"A juicio de la Comunidad Europea y sus Estados miembros, las bases de datos "hechas con el sudor de la frente" deberían gozar asimismo, bajo ciertas condiciones, de protección en tanto que propiedad intelectual – siguiendo la misma lógica que requiere que se utilice el potencial de los derechos de propiedad intelectual para la creación de empleo, el crecimiento, la prosperidad y la divulgación de información y conocimientos técnicos en otras esferas. La protección de la propiedad intelectual es un mecanismo para la distribución de contenidos de calidad de acuerdo a términos apropiados y, como tal, ha dado pruebas de los beneficios que podía aportar a la música, la literatura, las películas, los programas informáticos y todo tipo de productos industriales en el entorno tradicional, tanto en el siglo pasado como en la actualidad. Resulta esencial que también

35 Eduardo De La Parra Trujillo, "La protección de las bases de datos no originales en la legislación autoral mexicana (comparación con la directiva de la Unión Europea sobre bases de datos)". *Revista De La Facultad De Derecho De México*, 54(241), (2017):211–237, doi: <https://doi.org/10.22201/fder.24488933e.2004.241.61594>. P. 217

36 La doctrina de la apropiación indebida se origina con el caso *International News Service v. Associated Press* de 1918, en donde *International News Service* copiaba las noticias de *Associated Press* sobre la evolución de la Primera Guerra Mundial en Europa, y las distribuía como propias. Así pues, La Corte Suprema de los Estados Unidos estableció que estableciendo que *International News Service* estaba cometiendo un acto de competencia desleal al apropiarse del trabajo de otro. La Corte sostuvo que, si bien las noticias son hechos no pueden ser protegidas por derechos de autor, el esfuerzo sustancial, tiempo y costo invertido, generando una ventaja competitiva indebida y un perjuicio económico, sí pueden ser protegidos bajo la doctrina de la apropiación indebida.

las bases de datos electrónicas se beneficien de este mecanismo. El comercio electrónico sería un contrasentido sin una protección equilibrada por propiedad intelectual de las bases de datos que son fundamentales para su funcionamiento.

A nuestro juicio, es necesaria una protección coherente de las bases de datos mediante propiedad intelectual, a fin de promover la innovación y la inversión en productos de la información. Esta protección proporciona el incentivo necesario para divulgar una gran variedad de nuevas compilaciones en línea y fuera de línea, muchas de las cuales comportan una importante dimensión cultural. La seguridad jurídica permite el desarrollo saludable del sector de las bases de datos, aportando beneficiosos efectos al conjunto de la sociedad. De hecho, únicamente la protección claramente definida de la propiedad intelectual puede proporcionar la suficiente seguridad jurídica a los inversores y usuarios y el acceso de acuerdo con términos apropiados.³⁷

No obstante, a la fecha no se cuenta con un tratado internacional en los términos propuestos por la Unión Europea. En Colombia tampoco existe como tal un derecho sui generis para las bases de datos. Al respecto, la Dirección Nacional de Derecho de Autor, mediante Concepto 10632 de 2005, se pronunció en los siguientes términos:

"La información puntual contenida en una base de datos puede llegar a constituirse en un elemento de gran valor económico para el desarrollo de los mercados a nivel mundial. En la mayoría de los casos su recopilación y estructuración requieren de gran inversión de recursos humanos, técnicos y económicos.

En el transcurso de la década de los 90's ha surgido a nivel mundial la preocupación por resguardar este tipo de información. De tal manera, algunas legislaciones, entre ellas las de países pertenecientes a la Comunidad Europea han optado por crear un derecho sui generis que protege los fabricantes de bases de datos no originales contra los actos de apropiación de los resultados obtenidos, pasando por alto las inversiones económicas y el trabajo realizado por quien buscó y recopiló la información.

En suma se concede una protección en razón a la inversión, en términos cuantitativos y cualitativos de tiempo, esfuerzo y dinero, a la producción de la base de datos no original, siendo irrelevante a éste fin las características de la información así recopilada.³⁸

37 Organización Mundial de la Propiedad Intelectual (OMPI), Comité Permanente de Derecho de Autor y Derechos Conexos "Protección de las bases de datos", Ginebra, 4-5 de noviembre de 2002. Documento presentado por la Comunidad Europea y sus Estados Miembros.

38 Dirección Nacional de Derecho de Autor, respuesta del 8 de noviembre de 2005 al

INVERSIÓN SUSTANCIAL, MACHINE-GENERATED DATA Y EL DATA ACT: REDEFINIENDO EL ALCANCE DEL DERECHO SUI GENERIS

El régimen sui generis de la UE, concebido para ampliar el espectro de protección de las bases de datos definida por el derecho de autor tradicional, enfrenta retos significativos en su implementación. Un aspecto particularmente controvertido versa sobre la definición de lo que constituye una "inversión sustancial", concepto que ha generado diversas interpretaciones que conllevan a una incertidumbre para los creadores de bases de datos. Esta incertidumbre puede constituir un desincentivo, en tanto no existe certeza frente a si una base de datos es susceptible de protección bajo el régimen sui generis.

A manera de ejemplo, en el caso *British Horseracing Board Ltd v. William Hill Organization Ltd* del 2004, correspondió analizar si constituía una infracción al régimen sui generis que una empresa de apuestas utilizara los datos de una organización encargada de recopilar y gestionar los datos de carreras de caballos en el Reino Unido. El Tribunal de Justicia de la Unión Europea señaló que los datos reproducidos por la empresa de apuestas solo representaban "sino una mínima proporción"³⁹ de la base de datos del demandante, concluyendo entonces que los datos extraídos por el demandado no constituían una parte sustancial del contenido de la base de datos sobre la cual se solicitaba la protección bajo el régimen sui generis.

Asimismo, en la providencia en comento, el Tribunal explicó que la protección sui generis definido en la Directiva 96/9/CE se circunscribía únicamente a la obtención, verificación o presentación de los datos, dejando por fuera la creación de datos como tal. Con base lo anterior, concluyó que los datos extraídos por el demandado, al ser datos generados en el marco de las operaciones normales del demandante (fechas, lugares, participantes, entre otros.), se consideraban parte de la creación de los datos, mas no de la estructuración de la base de datos en sí, quedando por fuera del ámbito de protección del régimen sui generis.

Posteriormente, en el año 2018 la Comisión Europea publicó una Evaluación de la Directiva 96/9/CE sobre la protección jurídica de las bases de datos. En este informe se puso de presente que, si bien la Directiva ofrece un marco normativo significativo para la protección de las bases de datos, y que con esta se esperaba que el régimen sui generis incentivara las creaciones

derecho de petición con Radicado 2-2005-10632. Disponible en: <https://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?dt=S&i=20039>

39 Tribunal de Justicia de la Unión Europea [T.J.U.E.] [Court of Justice of the European Union], Gran Sala. 9 de noviembre de 2004. Asunto C-203/02.(U.E). P. 28.

de bases de datos, la evidencia sobre su efectividad en la promoción de la innovación es limitada.

La inseguridad jurídica generada a partir de la dificultad en definir qué se entiende por "inversión sustancial" es un desincentivo para los creadores de bases de datos. Asimismo, se señala que la postura adoptada por los tribunales en casos como *British Horseracing Board v. William Hill*⁴⁰ han tenido como resultado una limitación en el ámbito de protección del régimen sui generis, dejando por fuera las bases de datos creadas en el marco de las operaciones normales de una organización. Por lo anterior, se recomendó revisar el texto de la Directiva con el propósito de aclarar de definición de "inversión sustancial", de modo tal que esté alineada con las necesidades actuales y en particular del entorno digital.

El informe también puso de presente que esta Directiva, al haber sido redactada en un contexto anterior a la explosión del internet y la economía digital, puede ser un instrumento que no responda a las necesidades actuales, especialmente de cara a las nuevas tecnologías y el Big Data. Por lo anterior, se recomendó evaluar la posibilidad de establecer nuevas formas de protección o incentivos para sectores basados en el uso de los datos y las tecnologías emergentes.

Así las cosas, esta evaluación a la Directiva 96/9/CE señaló que la aplicación de este instrumento normativo ha presentado importantes limitaciones en la práctica, y en este sentido, se hace necesario abordar los desafíos que la protección a las bases de datos demanda en el contexto de una economía basada en datos. Para ello, se sugiere una revisión integral de la Directiva, de modo tal que logre promover la innovación y la inversión en la economía digital de la UE. El informe igualmente subrayó la importancia de equilibrar la protección a los creadores de las bases de datos y el acceso a la información, de modo que se genere un ambiente adecuado para promover la innovación y el desarrollo de la economía digital.

Más adelante, para la construcción de la Ley Europea de Datos – Data Act, se realizó una revisión de aspectos relacionados con el régimen sui generis de las bases de datos. Es así como en el Informe de Evaluación de Impacto del Documento de Trabajo del Personal de la Comisión del año 2022, que acompañó el entonces proyecto regulatorio, se puso de presente la relevancia que ha tomado el derecho sui generis en el marco de la economía de datos y los avances de la tecnología.

En este sentido, se examinó cómo la protección establecida en la Directiva 96/9/CE para las bases de datos no originales podría llegar a impactar otros

40 Ver también caso *Fixtures Marketing v. Svenska Spel*. Tribunal de Justicia de la Unión Europea [T.J.U.E.] [Court of Justice of the European Union], Gran Sala. 9 de noviembre de 2004. Asunto C-338/02.(U.E.)

derechos como el acceso a la información, y eventualmente podría limitar la reutilización de los datos, lo cual podría inhibir la innovación y la libre competencia. Es decir, se señaló que un uso excesivo de este derecho podría dar como resultado un bloqueo a otros competidores para mantener el control exclusivo y monopolizado sobre ciertos datos que podrían utilizarse para la creación de nuevos productos o servicios.

El informe también abordó la generación automática de grandes volúmenes de datos mediante dispositivos IoT, robótica, y otros dispositivos inteligentes (machine-generated data), y si en estos escenarios aplicaría la protección sui generis. Tradicionalmente, el derecho sui generis se aplicaba a bases de datos donde la inversión estaba en la recopilación y organización de los datos. Sin embargo, en el caso de dispositivos inteligentes, la inversión se refleja en la tecnología que genera los datos, no en la recopilación de los mismos, lo que podría llevar a una sobreprotección y limitar el acceso a datos generados sin una inversión significativa en la organización de la base de datos.

A manera de ejemplo, el informe trajo a colación el caso *Autobahn*, en el cual se analizó si una base de datos generada por un sistema de peaje para autopistas en Alemania que recopilaba datos sobre el tráfico y las tarifas de peaje a través de unos sensores instalados en carreteras estaba protegida bajo el derecho sui generis. En este caso, el Tribunal de Federal de Justicia de Alemania reconoció este derecho, indicando que, si bien la recopilación de los datos era automática, se cumplía con el criterio de inversión sustancial. Una de las grandes preocupaciones que surgió a partir de este caso se encuentra relacionado con el riesgo de generar una sobreprotección que limite el acceso y uso de unos datos, sin que haya una inversión significativa en la organización de las bases de datos.

Teniendo en cuenta lo anterior, en el informe se exploraron posibles reformas para armonizar la protección a los creadores de las bases de datos con el acceso abierto a la información. Entre otras, se consideraron opciones como reformar la Directiva, redefinir su alcance para evitar la configuración de monopolios de datos que limiten el acceso a la información en beneficio de la ciudadanía. Incluso se llegó a contemplar dentro del análisis de alternativas dentro del mencionado informe, la eliminación del derecho sui generis para aquellos casos en que se obstaculice el acceso a la información pública o la libre competencia.

Finalmente, mediante el artículo 43 del Data Act, que entró en vigor el 11 de enero de 2024 y será aplicable en septiembre de 2025, se excluyó de la protección del régimen sui generis a las bases de datos compuestas por datos generados automáticamente por máquinas. Lo anterior, buscando facilitar el acceso y uso de los datos, promover la innovación, reducir costos de transacción, y a su vez generar efectos positivos en el mercado con el desarrollo de nuevos productos. Es así como el Considerando 112 de este

instrumento normativo precisamente explica que el derecho sui generis no se aplica a estas bases de datos:

"Con el fin de eliminar el riesgo de que titulares de datos que están en bases de datos, obtenidos o generados por medio de componentes físicos, como sensores, de un producto conectado y de un servicio relacionado u otros datos generados por máquinas, reclamen el derecho sui generis en virtud del artículo 7 de la Directiva 96/9/CE obstaculizando de este modo, en particular, el ejercicio efectivo del derecho de los usuarios a acceder y utilizar los datos y el derecho a compartir datos con terceros en virtud del presente Reglamento, el presente Reglamento debe aclarar que el derecho sui generis no es aplicable a dichas bases de datos, ya que no se cumplirían los requisitos de protección. Ello no obsta a la posible aplicación del derecho sui generis en virtud del artículo 7 de la Directiva 96/9/CE a bases de datos que contengan datos que no entren en el ámbito de aplicación del presente Reglamento, siempre que se cumplan los requisitos de protección en virtud del apartado 1 de dicho artículo."

BUSCANDO OTRAS FORMAS DE PROTECCIÓN AL RÉGIMEN SUI GENERIS: MODELOS DE DATOS, BASES DE DATOS, Y OTRAS FORMAS DE ALMACENAMIENTO Y SU IMPACTO DE CARA A LA IA Y EL BIG DATA

Teniendo en cuenta el marco normativo anteriormente señalado, resulta necesario hacer unas precisiones conceptuales en entre los datos, los modelos de datos y las bases de datos, en tanto sólo estas últimas son las que hoy en día son objeto de protección a la luz de la propiedad intelectual.

Como se ha indicado, en el marco de las tecnologías de la cuarta revolución industrial, que impulsan una sociedad del conocimiento donde los datos son, en sí mismos, activos estratégicos para las empresas, y en consecuencia los datos *per se* no corresponden a una creación del intelecto, por lo que, no son objeto de protección a la luz del sistema de la propiedad intelectual.

Por otro lado, las bases de datos son definidas por la doctrina como "(...) una recopilación de datos puntuales u obras (preexistentes o no, originales o derivadas) que pueden ser utilizados manualmente o por medios electrónicos, hecha de forma organizada de tal manera que permite la recuperación de la información por los usuarios. La selección o disposición de materias es el elemento creativo que le confiere categoría de obra protegida por el derecho de autor"⁴¹.

41 Marvilia Carracedo Gonzalez, "La Protección Jurídica de la Base Datos". Documento expuesto en el marco del curso académico regional de la OMPI sobre derecho de autor y derechos conexos para países de América Latina, LA Habana, 22 al 30 de junio de 1998. Tomado de Dirección Nacional de Derecho de Autor, respuesta del 4 de abril de 2024 al derecho de petición con Radicado 1-2024-33658.

Es importante precisar que el modelo de datos no es lo mismo que una base de datos, y como se verá posteriormente, no cuenta con la protección otorgada por el régimen de derecho de autor. El Centro de Datos, Análisis y Reportes (Center for Data, Analytics and Reporting–CeDAR) de la Universidad de Princeton define el modelo de datos en los siguientes términos:

“Un modelo de datos organiza elementos de datos y estandariza cómo se relacionan los elementos de datos entre sí. Dado que los elementos de datos documentan personas, lugares y cosas de la vida real y los eventos entre ellos, el modelo de datos representa la realidad. Por ejemplo, una casa tiene muchas ventanas o un gato tiene dos ojos.

Los modelos de datos se utilizan a menudo como ayuda para la comunicación entre los empresarios que definen los requerimientos de un sistema informático y los técnicos que definen el diseño en respuesta a tales requerimientos. Se utilizan para mostrar los datos necesarios y creados por los procesos de negocio.

Un modelo de datos determina explícitamente la estructura de los datos. Los modelos de datos se especifican en una notación de modelado de datos, que suele tener un formato gráfico.

A veces se puede hacer referencia a un modelo de datos como estructura de datos, especialmente en el contexto de los lenguajes de programación. Los modelos de datos suelen complementarse con modelos de funciones”.⁴²

Así pues, el modelo de datos corresponde a una representación conceptual de la estructura y relaciones de los datos almacenados en bases de datos convencionales u otras formas de almacenamiento, como archivos de datos, sistemas de archivos distribuidos, data lakes, entre otros. El modelo de datos describe la manera como se organizan los datos, la tipología de datos que se almacenan, y cómo se relacionan entre sí. En este sentido, los modelos de datos no están limitados a las bases de datos, pudiéndose aplicar a una variedad de sistemas de almacenamiento de datos.

Ahora bien, es importante tener en cuenta que, de acuerdo con la normativa actual los modelos de datos no se encuentran dentro del ámbito de protección que otorga el derecho de autor a las bases de datos. Lo anterior, teniendo en cuenta conceptualmente los modelos de datos no almacenan la información como tal, sino que conservan una representación abstracta de esta.⁴³

42 Traducción Libre. “What is a Data Model?”, Center for Data, Analytics and Reporting (CeDAR) – Princeton University, consultado el 2 de junio de 2024, <https://cedar.princeton.edu/understanding-data/what-data-model>

43 Cristina Mesa, “IA generativa: ¿pueden protegerse los modelos de datos?”, Garrigues Digital, 13

De acuerdo con lo anterior, y pensando en otras formas de protección establecidas en el sistema de propiedad intelectual, es viable preguntarse si los modelos de datos podrían ser protegidos como patentes. No obstante, hay que tener en cuenta que los modelos de datos son en sí mismos estructuras abstractas, por lo que podría aducirse que se trata de una idea abstracta, o de una teoría matemática que estaría excluida de protección. Si el modelo de datos fuera empleado en el marco de una solución técnica específica y novedosa que tiene un efecto técnico, podría ser patentable, siempre y cuando se diera cumplimiento a los requisitos previstos en la normativa específica aplicable. Para el caso colombiano, se tendría que cumplir con los criterios de novedad, actividad inventiva, aplicación industrial y suficiencia descriptiva, de conformidad con la Decisión Andina 486 de 200 y el Decreto 2591 de 2000.

Por otra parte, también se podría argumentar que, si bien el modelo de datos no es susceptible de protección a la luz del derecho de autor, el código fuente que lo implementa podría ser protegido como un programa de ordenador. No obstante, no todos los modelos de datos son implementados en código fuente, sino que también pueden ser implementados en documentación, esquemas, bases de datos, incluso configuraciones del sistema, de aplicaciones, de APIs y Servicios web entre otros. El ciclo de vida de los modelos de datos puede comenzar en un diseño conceptual, y después de diversas etapas, previo a ser implementado en un sistema de software, se refina a través de un modelo lógico, que pasa a ser un modelo físico⁴⁴. En consecuencia, puede que la protección que establece el derecho de autor a los programas de ordenador no se compadezca con el esfuerzo intelectual que implica un modelo de datos.

Es importante mencionar que los modelos de datos han tomado bastante relevancia, especialmente en lo que atañe a la IA, en tanto son el "cerebro" de estos sistemas⁴⁵. Sin embargo, como se ha indicado, la normativa actual no protege los modelos de datos a la luz de la propiedad intelectual, lo que pone aun más de presente la necesidad de pensar la forma de protección de estos activos:

"El modelo de datos es el resultado del entrenamiento y afinamiento mediante el uso masivo de contenidos o datos. Actúa como el "cerebro" del sistema, y es

de octubre de 2023, https://www.garrigues.com/es_ES/garrigues-digital/ia-generativa-pueden-protgerse-modelos-datos

44 Thomas Connolly y Carolyn Begg, *Database Systems: A Practical Approach to Design, Implementation, and Management*, 6ª ed. (Harlow, Inglaterra: Pearson, 2014). Pp. 354-359

45 Cristina Mesa, "IA generativa: ¿pueden protegerse los modelos de datos?", Garrigues Digital, 13 de octubre de 2023, https://www.garrigues.com/es_ES/garrigues-digital/ia-generativa-pueden-protgerse-modelos-datos

lo que permite generar contenido de calidad. La clave para que el sistema sea eficaz está en la ingesta masiva de datos. Es esta ingesta la que permite que el modelo identifique y aprenda patrones y características de la información y, a partir de este aprendizaje, sea capaz de hacer predicciones, clasificaciones o generar nuevo contenido

(...)

Si bien es cierto que los modelos de datos pueden alimentarse de bases de datos, no deben confundirse con ellas. Un modelo de datos no almacena la información de entrenamiento como tal, sino que conserva una representación abstracta de esa información. El modelo contiene un conjunto de pesos, parámetros y conexiones, pero no los datos originales en sí. Del mismo modo, cuando de niños aprendemos qué es un gato, no recordamos cada gato individual que hemos visto, sino una idea generalizada o concepto de "gato" basado en las características comunes que hemos observado a través de nuestra experiencia".⁴⁶

La eficiencia de los modelos de datos puede impactar el rendimiento de los sistemas de IA. Por ejemplo, si un modelo de datos no está optimizado para acceder de manera eficiente a los datos relevantes, el procesamiento de la información se puede retrasar, afectando la toma de decisiones. Lo mismo puede pasar si el modelo de datos es demasiado complejo, aumentando el tiempo de entrenamiento de la IA. Asimismo, si un modelo de datos no es escalable, la IA podría tener dificultades para manejar grandes volúmenes de datos, o incluso adaptarse a cambios a medida que va creciendo el sistema.⁴⁷

En este sentido, los modelos de datos pueden tener un valor económico significativo, no solo en el marco de sistemas de IA, sino en el contexto de industrias y aplicaciones, en tanto pueden mejorar la toma de decisiones, impactando positivamente la eficiencia operativa, la optimización de los procesos, así como la identificación de oportunidades de mejora de las organizaciones. También pueden personalizar sus productos, servicios y experiencias ofrecidas a los clientes, a partir del análisis de datos de comportamiento y preferencias. También pueden contribuir en la identificación de riesgos potenciales y anticiparse a problemas, así como generar *insights* para el diseño de nuevos productos y servicios a partir del análisis de mercado.

Teniendo en cuenta que el sistema de propiedad intelectual no establece un mecanismo que permita una protección adecuada para los modelos de

46 Cristina Mesa, "IA generativa: ¿pueden protegerse los modelos de datos?", Garrigues Digital, 13 de octubre de 2023, https://www.garrigues.com/es_ES/garrigues-digital/ia-generativa-pueden-protegerse-modelos-datos

47 Cabe mencionar, que existen otros factores importantes, como el diseño del algoritmo, la calidad de los datos de entrenamiento y el volumen de datos de entrenamiento, entre otros.

datos, el secreto industrial podría constituir una alternativa de protección. Cabe mencionar que, en este escenario, la forma de protección sería distinta, en tanto se protegería la confidencialidad más que la obra o creación en sí misma. En consecuencia, el modelo de datos quedaría expuestos a procesos de ingeniería inversa, a la divulgación del secreto, o que un tercero pueda desarrollar la misma solución sin necesidad de acceder a la información.⁴⁸

Por otro lado, paralelo a la discusión frente a si deberían ser protegidos los modelos de datos, también importante mencionar que la IA generativa también puede emplear bases de datos en el proceso de entrenamiento de sus modelos, especialmente si se basa en técnicas de *machine learning* y *deep learning*. Durante el entrenamiento, estos modelos suelen alimentarse con grandes conjuntos de datos para aprender patrones y generar contenido nuevo de manera más precisa. Estos conjuntos de datos pueden estar almacenados en bases de datos, o en otras formas de almacenamiento de datos, como ya se ha indicado en el presente escrito. Posteriormente, cuando el modelo ya termina su periodo de entrenamiento, la IA utiliza los patrones aprendidos, no siendo necesario acceder al conjunto de datos original.

En este sentido, el uso de bases de datos en la IA generativa también genera retos frente a la normativa actual que protege el uso de bases de datos originales y no originales. Entre ellos, surge la cuestión frente a las implicaciones sobre el uso sin autorización de bases de datos para el proceso de entrenamiento de modelos. Sobre el particular, en un reciente artículo, Kretschmer, Margoni y Oruç ponen de presente la necesidad de establecer políticas claras y armonizadas para el uso de datos que reposan en bases de datos protegidas por derechos de autor y derechos sui generis en el aprendizaje automático, así como de materiales como artículos científicos, imágenes, videos, que suelen constituir el insumo para los algoritmos de aprendizaje automático. De acuerdo con los autores, la ambigüedad actual que existe en el régimen legal de recolección y procesamiento de datos puede obstaculizar la innovación y la investigación científica, sugiriendo la necesidad de repensar las excepciones y el otorgamiento de licencias para estos fines⁴⁹. Si bien los autores circunscriben su análisis al régimen establecido en la UE, se considera que estos retos podrían ser aún mayores en legislaciones que ni siquiera cuentan con un régimen de protección para bases de datos no originales, como lo es el colombiano.

48 Cristina Mesa, "IA generativa: ¿pueden protegerse los modelos de datos?", Garrigues Digital, 13 de octubre de 2023, https://www.garrigues.com/es_ES/garrigues-digital/ia-generativa-pueden-protegerse-modelos-datos

49 Martin Kretschmer, Thomas Margoni and Pinar Oruç, "Copyright law, and the life-cycle of machine learning models". *International Review of Intellectual Property and Competition Law* (IIC). 55, (2024):110–138, doi: <https://doi.org/10.1007/s40319-023-01419-3>

De igual modo, existen otros sistemas de IA que requieren bases de datos como parte de su funcionamiento, tales como los Sistemas de Recomendación⁵⁰, que utilizan bases de datos de preferencias de usuario, historiales de compra entre otros; los Sistemas de Información de Gestión (MIS por sus siglas en inglés), que almacenan datos estructurados para su posterior explotación⁵¹; y los Chatbots y asistentes virtuales, que en ocasiones utilizan bases de datos para almacenar y recuperar información específica.

REFLEXIONES FINALES

La protección de las bases de datos surge en sus inicios como una suerte de analogía a las compilaciones, como lo puede ser una antología de cuentos o un directorio telefónico. Para ese entonces, el sistema de derecho de autor tuvo que adaptarse para establecer unas disposiciones más específicas sobre la materia, como lo es la protección definida en el ADPIC, así como el TODA,

50 "Los sistemas de recomendación son aplicaciones de software que ayudan a los usuarios a encontrar elementos de interés en situaciones de sobrecarga de información. Las investigaciones actuales a menudo asumen un paradigma de interacción de una sola vez, donde las preferencias de los usuarios se estiman en función del comportamiento observado en el pasado y donde la presentación de una lista clasificada de sugerencias es la forma principal y unidireccional de interacción del usuario. Los sistemas de recomendación conversacional (SRC) adoptan un enfoque diferente y admiten un conjunto más rico de interacciones. Estas interacciones pueden, por ejemplo, ayudar a mejorar el proceso de obtención de preferencias o permitir al usuario hacer preguntas sobre las recomendaciones y dar comentarios. El interés en SRC ha aumentado significativamente en los últimos años. Este desarrollo se debe principalmente al importante progreso en el área del procesamiento del lenguaje natural, la aparición de nuevos asistentes domésticos controlados por voz y el mayor uso de la tecnología chatbot." Traducción libre. Dietmar Jannach, Ahtsham Manzoor, Wanling Cai, Li Chen, "A Survey on Conversational Recommender Systems". *ACM Computing Surveys*. 54(5), (2021):1-36, doi: <https://doi.org/10.1145/3453154>

51 Los Sistemas de Información de Gestión ("MIS" por sus siglas en inglés) se utilizan para la toma de decisiones y la coordinación, control, análisis y visualización de la información en una organización. Un MIS debe tener los componentes necesarios para recopilar, procesar, almacenar y recuperar la información necesaria para entregarla a los líderes de la organización. Son componentes de un MIS: i) Sistema de información: una combinación de hardware, software, personal e infraestructura que ayuda en la recopilación de datos almacenados en el MIS; Sistema de gestión de bases de datos (DBMS por sus siglas en inglés): software que maneja el almacenamiento, recuperación y actualización de datos en un sistema informático; Sistema de inteligencia: procesa los datos recopilados y los presenta de forma comprensible – también denominada "inteligencia empresarial"; Sistema de investigación: Identifica los principales problemas de gestión dentro de la organización y proporciona alternativas de solución, ayudando a que se analicen todas las opciones posibles y se tome la decisión adecuada. Vangie Beal, "What is a Management Information System?". *Webopedia*, consultado el 2 de junio de 2024, <https://www.webopedia.com/definiciones/mis/>

y la Decisión Andina 351 de 1993 para el sistema comunitario, lo cual se vio reflejado en la ley colombiana de derechos de autor. Tal como se indica en la sentencia del caso *Feist Publications*, al ser la originalidad un elemento *sine qua non* del derecho de autor, la protección sobre una base de datos quedó circunscrita a este elemento. En contraste, a través de la Directiva 96/9/CE, la Unión Europea adoptó un régimen sui generis para la protección de aquellas bases de datos que, aunque carecen del criterio de originalidad, involucran una inversión sustancial en la obtención, verificación o presentación de los datos, protegiendo así los derechos en su extracción y uso.

Los retos que trae la cuarta revolución industrial y las denominadas industrias 4.0, hacen que los datos adquieran un rol preponderante en tanto mueven la economía, y se convierten en un habilitador para la innovación. Sobre el particular, el Informe mundial sobre la Propiedad Intelectual de 2024 de la OMPI destaca la incidencia positiva de la innovación y de las políticas industriales que promuevan el desarrollo tecnológico, para impulsar la economía de los países. Para ello, es apenas natural contar con un régimen de protección que incentive a los creadores, y a la vez brinde seguridad jurídica frente a su aplicación⁵².

Así pues, las bases de datos son en la práctica una herramienta necesaria en esta economía basada en datos, ya que permiten almacenar la información, y de esta forma pueden permitir el procesamiento y explotación de grandes volúmenes de datos. También son un insumo fundamental para la analítica de datos, permitiendo identificar patrones, correlaciones y tendencias en la información, lo cual puede impulsar la toma de decisiones tanto de las empresas y organizaciones, como incluso las políticas en el sector público. De igual forma, las bases de datos proporcionan la infraestructura para el desarrollo de nuevas tecnologías.

En este sentido, como se ha podido evidenciar en la revisión normativa en el presente escrito, el régimen de protección a las bases de datos fue concebido en un contexto muy distinto en el que nos encontramos ahora, pero que, sin duda alguna, se estaba anticipando a los avances tecnológicos y la inmersión de las TIC en diversas industrias. Sin embargo, muchas de las discusiones que se dieron a finales de la década de los 90 en torno a contar con un régimen especializado, no solo para las bases de datos que cumplen con ese elemento de originalidad, sino para aquellas bases de datos que surgen en contextos más operativos que estratégicos, resurgen a medida que la tecnología evoluciona y las necesidades de contar con una protección adecuada para este tipo de obras cada vez se vuelven relevantes.

52 Organización Mundial de la Propiedad Intelectual – OMPI. “World Intellectual Property Report 2024”. <https://www.wipo.int/web-publications/world-intellectual-property-report-2024/>

Ahora bien, tal como lo señaló el Reporte sobre la Protección Legal para las bases de datos de la Oficina del Derecho de Autor de Estados Unidos, realizado en 1997, las bases de datos no sólo deben ser entendidas como un simple activo de contenido económico, sino que debe ser entendido en un contexto en el que también es necesario promover el acceso a la información:

“Las bases de datos siempre han sido *commodities* de valor comercial y utilidad social, desde su encarnación temprana en el siglo XVIII como directorios compilados puerta a puerta hasta los compendios de millones de artículos en formato electrónico de finales del siglo XX. La cuestión de si las bases de datos deberían ser protegidas por la ley y cómo hacerlo nunca ha sido fácil, ya que implica necesariamente encontrar un equilibrio entre dos objetivos sociales potencialmente contradictorios: el objetivo de proporcionar incentivos adecuados para su producción continua y el objetivo de garantizar el acceso público a la información que contienen. En diferentes momentos y en diferentes sociedades, ese equilibrio se ha logrado de diferentes maneras.”⁵³

Sin duda alguna, los avances de la IA y el Deep Learning ponen de presente que esta discusión se mantiene vigente, en tanto el acceso a la información permite que los sistemas evolucionen y, en consecuencia, las restricciones frente al uso de bases de datos protegidas eventualmente podrían llegar a constituir un obstáculo. Sobre el particular, el desarrollo de iniciativas como la de los Open Data Commons (ODC)⁵⁴. Asimismo, los avances normativos sobre infraestructura de datos en Colombia⁵⁵, y particularmente para el sector público, también ponen de presente la necesidad de contar con herramientas jurídicas que proporcionen un ambiente adecuado para promover la apertura de datos, no sólo del sector público, sino del sector privado. En este sentido, el régimen de propiedad intelectual de cara a la economía basada en datos resulta fundamental.

53 Reporte sobre la Protección Legal para las bases de datos. [Oficina del Derecho de Autor de Estados Unidos] Agosto de 1997. (E.E.U.U) P. 1.

54 Los Open Data Commons son una iniciativa que promueve la apertura de datos, así como el uso de datos abiertos. Consisten en licencias que permiten el uso de conjuntos de datos, de acuerdo con los términos y condiciones definidos en la licencia. “Introduction to intellectual property rights in data management”, Cornell University, consultado el 16 de junio de 2024, <https://data.research.cornell.edu/data-management/sharing/intellectual-property/>

55 Sobre el particular, resulta oportuno tener en cuenta la siguiente normativa: Ley 1712 de 2014 sobre transparencia y acceso a la información pública; Resolución 460 de MinTIC, que establece el Plan Nacional de Infraestructura de Datos; Decreto 1389 de 2022, que establece los lineamientos generales para la gobernanza de infraestructura de datos y crea el Modelo de gobernanza de la infraestructura de datos para el país; y el Decreto 767 de 2022, que define el Estado Abierto como uno de las líneas de acción de la Política de Gobierno Digital.

Por otro lado, de la revisión de doctrina y jurisprudencia realizada, los debates más recientes en materia de protección jurídica a las bases de datos se han dado en el contexto de las bases de datos no originales. En este sentido, la falta de mecanismos jurídicos para la protección de bases de datos no originales en Colombia constituye una circunstancia que pone en riesgo los esfuerzos de los creadores de bases de datos de esta naturaleza, y en consecuencia desincentiva este sector de la industria. Para ello, la revisión de la experiencia de la Unión Europea, y el análisis de la evolución que ha tenido la implementación de este derecho, podría proporcionar valiosas lecciones para el caso colombiano, creando un ambiente que brinde seguridad jurídica a los creadores de las bases de datos, y que a su vez promueva la innovación y el acceso y explotación de datos para el desarrollo de la economía basada en datos. En particular, será esencial el seguimiento a la implementación del Data Act y el análisis de impacto de la excepción del derecho *sui generis* contemplada en el artículo 43.

La protección *sui generis* puede ser un incentivo valioso para la creación de bases de datos, lo que podría tener efectos positivos para el desarrollo de diversos sectores. No obstante, la experiencia de la Unión Europea demuestra que existen desafíos significativos en esta materia, especialmente con la aparición del IoT, y en general aquellas tecnologías que automatizan la recopilación de datos. Estas tecnologías generan cuestiones frente a lo que constituye una "inversión sustancial", y plantean el riesgo de generar una protección excesiva a los creadores, que dé lugar a la configuración monopolios de datos, que restrinjan la tanto la innovación y el desarrollo de nuevos productos y servicios, como el acceso a la información, contraviniendo los objetivos de un marco regulatorio equilibrado.

Por otra parte, es importante subrayar que las regulaciones deben evolucionar para seguir el ritmo de los avances tecnológicos y las necesidades económicas contemporáneas. En este sentido, en el presente escrito también se ha puesto de presente que, en el contexto de la infraestructura de datos, los modelos de datos también pueden constituir un activo de contenido económico, que en el caso de algunos sistemas de IA pueden resultar incluso más relevantes que las bases de datos en sí mismas. También se ha puesto de presente que la definición de bases de datos está concebida bajo una suerte de analogía con lo que eran las compilaciones y antologías, por lo que también habrá que revisar si la definición de base datos es compatible y brindará una adecuada protección a las nuevas formas de almacenamiento de datos, tales como los data lakes, data warehouses y bases de datos NoSQL, entre otras.

En este sentido, se hace necesario reflexionar si el régimen de derecho de autor tradicional es suficiente para establecer los incentivos necesarios para la innovación que demanda una economía basada en datos, o si deberían protegerse otros activos de contenido económico, como los modelos de

datos, que preceden a la creación de las bases de datos e implican un esfuerzo intelectual (y podría decirse que "original") por parte de su creador.

BIBLIOGRAFÍA

Amazon Web Services (AWS). "¿Qué es un lago de datos?". Consultado el 3 de junio de 2024, <https://aws.amazon.com/es/what-is/data-lake/>

Abrams, Howard. "Copyright, Misappropriation, and Preemption: Constitutional and Statutory Limits of State Law Protection". *The Supreme Court Review*. Vol. 1983 (1983): 509-581. <http://www.jstor.org/stable/3109483>

Martínez Devia, Andrea. "La Inteligencia Artificial, El Big Data Y La Era Digital: ¿una Amenaza Para Los Datos Personales?". *Revista la Propiedad Inmaterial*. 27(junio), (2019): 5-23. <https://doi.org/10.18601/16571959.n27.01>.

Caloghirou, Yannis., Dimas, Petros., Stamopoulos, Dimitrios.; Tsakanikas, Aggelos. "Tracking Intangibles in the EU: The GLOBALINTO Input-Output Intangibles Database". Submission for the 17th Globelics International Conference, Heredia, Costa Rica, 3-5 November 2021. Disponible en: https://www.researchgate.net/publication/355757356_Tracking_Intangibles_in_the_EU_The_GLOBALINTO_Input-Output_Intangibles_Database

Cámara Lapuente, Sergio . "Últimas orientaciones internacionales sobre la protección jurídica de las bases de datos". *Revista Chilena de Derecho*, 26(1), (1999):9-61

Carracedo Gonzalez, Marvilia. "La Protección Jurídica de la Base Datos". Documento expuesto en el marco del curso académico regional de la OMPI sobre derecho de autor y derechos conexos para países de América Latina, LA Habana, 22 al 30 de junio de 1998.

Center for Data, Analytics and Reporting (CeDAR) – Princeton University. "What is a Data Model?", consultado el 2 de junio de 2024, <https://cedar.princeton.edu/understanding-data/what-data-model>

Circular Conjunta 01 de 2006. [Procuraduría General de la Nación & Dirección Nacional de Derecho de Autor]. Orientaciones para el cumplimiento de normas de Derecho de Autor y Derechos Conexos, en lo pertinente a la contratación estatal de obras y prestaciones protegidas y sus buenas prácticas. 24 de abril de 2006. (Colom.)}

Comisión Europea [European Commission]. *Commission Staff Working Document: Evaluation of Directive 96/9/EC on the Legal Protection of Databases*. Bruselas, 25 de abril de 2018.

Comisión Europea [European Commission]. *Commission Staff Working Document Impact Assessment Report Accompanying the Document Proposal for a Regulation of the European*

Parliament and of the Council on Harmonised Rules on Fair Access to and Use of Data (Data Act). Bruselas, 23 de febrero de 2022.

Comisión Europea, "Ley de Datos", consultado el 23 de agosto de 2024, <https://digital-strategy.ec.europa.eu/es/policies/data-act#:~:text=Entrada%20en%20aplicaci%C3%B3n,aplicable%20en%20septiembre%20de%202025>.

Cornell University, "Introduction to intellectual property rights in data management", consultado el 16 de junio de 2024, <https://data.research.cornell.edu/data-management/sharing/intellectual-property/>

Corte de Justicia del Reino Unido [C.J.U.K.] [Court of Justice of the United Kingdom], marzo, 1, 2012. Case C-604/10 (U.K.)

Corte Suprema de los Estados Unidos [C.S.E.U.] [U.S. Supreme Court], marzo, 27, 1991. No. 89-1909. (EE.UU.).

Dietmar Jannach, Ahtsham Manzoor, Wanling Cai, Li Chen, "A Survey on Conversational Recommender Systems". *ACM Computing Surveys*. 54(5), (2021):1-36, doi: <https://doi.org/10.1145/3453154>

Dirección Nacional de Derecho de Autor, respuesta del 8 de noviembre de 2005 al derecho de petición con Radicado 2-2005-10632. Disponible en: <https://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?dt=S&i=20039>

Dirección Nacional de Derecho de Autor, respuesta del 4 de abril de 2024 al derecho de petición con Radicado 1-2024-33658.

Directiva 96/9/CE. [Parlamento Europeo y Consejo Europeo]. Protección jurídica de las bases de datos. 11 de marzo de 1996. (Unión Europ.)

De La Parra Trujillo, Eduardo. "La protección de las bases de datos no originales en la legislación autoral mexicana (comparación con la directiva de la Unión Europea sobre bases de datos)". *Revista De La Facultad De Derecho De México*, 54(241), (2017):211–237, doi: <https://doi.org/10.22201/fder.24488933e.2004.241.61594>

Google Cloud. "¿Qué es una base de datos relacional?", consultado el 10 de junio de 2024, <https://cloud.google.com/learn/what-is-a-relational-database?hl=es-419>

IBM. "¿Qué es una base de datos NoSQL?", consultado el 10 de junio de 2024, <https://www.ibm.com/es-es/topics/nosql-databases>

Kretschmer, Martin, Margoni Thomas, y Oruç, Pinar. "Copyright law, and the lifecycle of machine learning models". *International Review of Intellectual Property and Competition Law (IIC)*. 55, (2024):110–138, doi: <https://doi.org/10.1007/s40319-023-01419-3>

- Ley de Derechos de Autor de 1976 (Copyright Act of 1976), incorporada en el Título 17 del Código de los Estados Unidos. (EE.UU).
- López Alzaga, Ricardo, "*Hot news: ¿se protegen las noticias?*", Blog Propiedad Intelectual e Industrial Garrigues, 20 de diciembre de 2022, <https://blogip.garrigues.com/propiedad-intelectual/hot-news-se-protegen-las-noticias>
- Mesa, Cristina "*IA generativa: ¿pueden protegerse los modelos de datos?*", Garrigues Digital, 13 de octubre de 2023, https://www.garrigues.com/es_ES/garrigues-digital/ia-generativa-pueden-protegerse-modelos-datos
- Nurton, James, "*Los datos son el combustible de la transformación de la economía mundial*", Revista OMPI. Marzo de 2022, https://www.wipo.int/wipo_magazine_digital/es/2022/article_0002.html.
- Oracle. "*¿Qué es un almacén de datos?*", consultado el 10 de junio de 2024, <https://www.oracle.com/co/database/what-is-a-data-warehouse/>
- Organización Mundial de la Propiedad Intelectual (OMPI), "*Diálogo de la OMPI sobre Propiedad Intelectual (PI) y Tecnologías de Vanguardia: Cuarta sesión – Resumen de la Cuarta Sesión*", Ginebra, 22-23 de septiembre de 2021
- Organización Mundial de la Propiedad Intelectual – OMPI. "*Tratado de la OMPI sobre Derecho de Autor*, consultado el 10 de junio de 2024, <https://www.wipo.int/treaties/es/ip/wct/>
- Organización Mundial de la Propiedad Intelectual -OMPI, Comité Permanente de Derecho de Autor y Derechos Conexos "*Protección de las bases de datos*", Ginebra, 4-5 de noviembre de 2002. Documento presentado por la Comunidad Europea y sus Estados Miembros
- Organización Mundial de la Propiedad Intelectual – OMPI. "*World Intellectual Property Report 2024*". <https://www.wipo.int/web-publications/world-intellectual-property-report-2024/>
- Oficina del Derecho de Autor de Estados Unidos, "*Report on legal protection for databases*", agosto de 1997. Disponible en: <https://www.copyright.gov/reports/db4.pdf>
- Peña Valenzuela, Daniel. "*Big Data y Propiedad Intelectual*", Blog de Derecho de los Negocios Universidad Externado, 10 de octubre de 2020, <https://dernegocios.uexternado.edu.co/big-data-y-propiedad-intelectual/>
- Rephen, Matthew. "*Publicly Available Data and the Misappropriation Doctrine*", Fordham Intellectual Property Institute. Sf. <https://fordhamipinstitute.com/publicly-available-data-and-the-misappropriation-doctrine/>

Sanabria de Luque, Jhohan Julian. "Sector Privado y Libre Competencia: Implicaciones Jurídicas del Web Scraping", (tesis de maestría), Universidad Externado de Colombia, 2021, <https://bdigital.uexternado.edu.co/server/api/core/bitstreams/5abd583d-ae25-453d-b1d8-2a4803049694/content>

Thomas Connolly y Carolyn Begg, *Database Systems: A Practical Approach to Design, Implementation, and Management*, 6ª ed. (Harlow, Inglaterra: Pearson, 2014).

Tribunal de Justicia de la Unión Europea [T.J.U.E] [Court of Justice of the European Union], Gran Sala- Noviembre, 9, 2004. Asunto C-203/02 (U.E)

Tribunal de Justicia de la Unión Europea [T.J.U.E] [Court of Justice of the European Union], Gran Sala. 9 de noviembre de 2004. Asunto C-338/02.(U.E).

Tribunal de Justicia de la Unión Europea [T.J.U.E] [Court of Justice of the European Union], Sala Segunda. Enero, 15, 2015. Asunto C-30/14.(U.E)

Tribunal de Justicia de la Unión Europea [T.J.U.E] [Court of Justice of the European Union], Sala Quinta. Junio, 3, 2021. Asunto C-762/19.(U.E)

Tribunal de Justicia de la Unión Europea [T.J.U.E] [Court of Justice of the European Union], Sala Tercera. Octubre, 18, 2012. Asunto C-173/11 (U.E)

Tribunal Federal de Justicia de Alemania [Bundesgerichtshof, BGH], Primera Sala Civil. Marzo, 25, 2010. Expediente 47/08

Vangie Beal, "What is a Management Information System?". Webopedia, consultado el 2 de junio de 2024, <https://www.webopedia.com/definiciones/mis/>

