

Derecho de autor



MINERÍA DE TEXTOS Y DATOS COMO (NUEVO) LÍMITE AL DERECHO DE AUTOR

Elena Vicente Domingo

Catedrática de Derecho Civil
Universidad de Burgos

Teresa Rodríguez Cachón

Profesora Ayudante Doctora
Universidad de Burgos



REUS
EDITORIAL

COLECCIÓN DERECHO DE AUTOR

TÍTULOS PUBLICADOS

1. *Obras colectivas y propiedad intelectual*, Carlos Rogel Vide, (2021).
2. *Obras de dominio público, digitalización y preservación natural*, Susana Navas Navarro, (2021)
3. *Minería de textos y datos como (nuevo) límite al Derecho de autor*, Elena Vicente Domingo y Teresa Rodríguez Cachón, (2021).

Colección Derecho de Autor
Director de la Colección: Carlos Rogel Vide
Catedrático emérito de Derecho Civil

**MINERÍA DE TEXTOS
Y DATOS COMO (NUEVO)
LÍMITE AL DERECHO DE
AUTOR**

Elena Vicente Domingo
Catedrática de Derecho Civil
Universidad de Burgos

Teresa Rodríguez Cachón
Profesora Ayudante Doctora
Universidad de Burgos



Madrid, 2021

© Elena Vicente Domingo y Teresa Rodríguez Cachón
© Editorial Reus, S. A., para la presente edición
C/ Rafael Calvo, 18, 2º C – 28010 Madrid
+34 91 521 36 19 – +34 91 522 30 54
Fax: +34 91 445 11 26
reus@editorialreus.es
www.editorialreus.es

1ª edición, REUS, S.A. (septiembre, 2021)
ISBN: 978-84-290-2541-5
Depósito Legal: M-25634-2021
Diseño de portada: María Lapor
Impreso en España
Printed in Spain

Imprime: *Estilo Estugraf Impresores S.L.*

Ni Editorial Reus ni sus directores de colección responden del contenido de los textos impresos, cuya originalidad garantizan sus propios autores. Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra sólo puede ser realizada con la autorización expresa de Editorial Reus, salvo excepción prevista por la ley.

Fotocopiar o reproducir ilegalmente la presente obra es un delito castigado con cárcel en el vigente Código penal español.

CAPÍTULO 1. MINERÍA DE TEXTOS Y DATOS: NOCIONES BÁSICAS

1.1. INTRODUCCIÓN

Un día cualquiera. Suena el despertador de tu *smartphone* y la persiana de tu habitación se levanta automáticamente. Coges tu teléfono y, gracias a su aplicación de reconocimiento facial, te haces un perfecto selfi con filtro. Revisas las notificaciones de redes sociales que han llegado a tu teléfono durante la noche; todas te resultan absolutamente interesantes. Te levantas y, antes de entrar en la ducha, le pides viva voz a tu altavoz inteligente que ponga la lista de reproducción mejor puntuada en tu aplicación de música en *streaming*. Antes de vestirte, consultas el tiempo que hará a lo largo del día y, al ver que lloverá toda la semana, entras en la aplicación de tu tienda favorita para comprar unas botas. Te agrada ver que los primeros resultados que te ofrece resultan completamente coincidentes con tus gustos. Mientras desayunas, con-

sultas tu bandeja de correo electrónico, en la que ya hace años que apenas tienes que lidiar con *spam*. Al tiempo, accedes a la aplicación de tu banco y preguntas directamente por chat una duda sobre tus últimas transacciones; recibes respuesta de forma instantánea. Al ojear los principales titulares del día de la prensa internacional, aparecen ciertas expresiones en inglés cuyo significado desconoces, ante lo cual consultas su traducción en una aplicación de traducción automática. Al salir de casa y coger el coche, consultas el tráfico en tiempo real y tu aplicación de navegación te sugiere una ruta alternativa para evitar atascos, a la vez que te propone una lista de *podcasts* para escuchar de camino al trabajo¹.

Esta situación describe de forma prototípica la mañana de una persona cualquiera en la actualidad. Durante la primera hora en que estamos despiertos, es frecuente que hagamos uso de no menos de 10 aplicaciones móviles o tecnologías aplicadas, cuyo funcionamiento descansa principalmente en inteligencia artificial.

A pesar de que su delimitación resulta un tanto controvertida, se entiende usualmente por inteligencia artificial aquel campo de las ciencias de la

¹ Esta situación ha sido recreada partiendo de la entrada del blog “10 ejemplos de que ya dependes de la IA en tu vida diaria”, de 6 de septiembre de 2019 del blog <https://cutt.ly/sbmLoSf>

computación centrado en la creación de programas y mecanismos capaces de mostrar comportamientos inteligentes, preparados para aprender y razonar como lo haría un ser humano, mediante la imitación de sus funciones cognitivas, entre las que destacan el pensamiento, el razonamiento, el aprendizaje, la predicción, la recomendación, la toma de decisiones y la resolución de problemas².

En consecuencia, la inteligencia artificial puede ser aplicada no solo al desarrollo de aplicaciones tecnológicas de uso generalizado, sino también –por nombrar algunos ejemplos– a la investigación científica en cualquier ámbito, como la investigación médica, antropológica o jurídica³, a la investigación criminal, a la prevención de la ciberdelincuencia y a una amplísima variedad de modelos de negocio, muchos de los cuales incluso han nacido en este contexto.

² FERNÁNDEZ CARBALLO-CALERO, Pablo, 2021, pp. 21-22; OECD, “Recommendation of the Council on Artificial Intelligence”, (disponible en: <https://cutt.ly/MbEImfU>); y COM(2018) 237 final, 25.4.2018.

³ Diane McDONALD y Ursula KELLY (2012), en su informe “The Value and Benefits of Text Mining. The Value and Benefit of Text Mining to UK Further and Higher Education” exponen cómo el TDM y la inteligencia artificial pueden eficientemente analizar el constante creciente volumen de textos legales complejos y jurisprudencia

Documento disponible en: <https://bit.ly/3wSseRq>

ÍNDICE

CAPÍTULO 1. MINERÍA DE TEXTOS Y DATOS: NOCIONES BÁSICAS.....	5
1.1. Introducción	5
1.2. Minería de textos y datos: estado de la cuestión.....	16
1.3. Diferencias entre minería de datos y minería de textos.....	26
1.4. Fases de la minería de textos y datos ...	31
1.4.1. Fase 1: rastreo y acceso al contenido.	32
1.4.2. Fase 2: extracción y copia del contenido, en caso de ser requerido	34
1.4.3. Fase 3: minería de datos y descubrimiento de conocimiento.....	36
CAPÍTULO 2. MINERÍA DE TEXTOS Y DATOS COMO LÍMITE AL DERECHO DE AUTOR	43
2.1. ¿Qué derechos de propiedad intelectual pueden colisionar con el TDM?... ..	43
2.1.1. Situaciones no problemáticas.....	43
A. Reproducciones provisionales	44

B. Extracción de hechos o datos y de partes no sustanciales.....	45
C. Límite de ilustración aplicable a actividades de TDM	48
2.1.2. Colisión entre TDM y derechos de propiedad intelectual.....	50
A. Contenidos a extraer/copiar incluidos en una base de datos.....	51
B. Contenidos a extraer/copiar no incluidos en una base de datos.....	54
2.1.3. Colisión entre TDM y otros derechos	57
2.2. Economía y Derecho de Autor	59
2.2.1. El Mercado único digital	59
2.2.2. De forma habitual, ¿quiénes ostentan el derecho de reproducción sobre las obras sometidas a TDM?.....	67
2.2.3. Nuevos modelos de negocio nacidos al calor del TDM. En particular, el caso Elsevier <i>vs</i> Tilburg University..	74
2.2.4. Situación del TDM en el Derecho comparado.....	81
2.3. Las dos versiones de la excepción en la Directiva.....	91
CAPÍTULO 3. TDM CON FINES DE INVESTIGACIÓN CIENTÍFICA	97
3.1. Introducción	97

3.2. Elemento subjetivo: quiénes pueden realizar TDM con fines de investigación científica.....	100
3.2.1. Doble requisito	100
3.2.2. Organismos de investigación e instituciones responsables del patrimonio cultural.....	101
3.2.3. ¿Qué es investigación científica?	110
3.3. Elemento objetivo: sobre qué obras u otras prestaciones se puede realizar TDM con fines de investigación científica	111
3.3.1. Bases de datos y derecho de reproducción en la sociedad de la información.....	111
3.3.2. Publicaciones de prensa en lo relativo a usos en línea.....	113
3.4. ¿Remuneración?, ¿compensación equitativa?	114
CAPÍTULO 4. TDM SIN FINES DE INVESTIGACIÓN CIENTÍFICA	123
4.1. Introducción	123
4.2. Elemento subjetivo: quiénes pueden realizar TDM sin fines de investigación científica.....	125
4.3. Elemento objetivo: sobre qué obras u otras prestaciones se puede realizar TDM sin fines de investigación científica.....	125

4.4. Posibilidad de excluir la aplicación de este límite.....	127
5. REFLEXIONES FINALES.....	131
6. BIBLIOGRAFÍA	133

Durante la primera hora en que estamos despiertos, es frecuente que hagamos uso de no menos de 10 aplicaciones móviles, cuyo funcionamiento descansa principalmente en inteligencia artificial, la cual, en un gran número de ocasiones, debe ser “alimentada” mediante minería de textos y datos (TDM).

Asumiendo la indudable necesidad de fomentar el desarrollo tecnológico, la aparición de una nueva excepción de TDM al Derecho de autor plantea una infinidad de incertidumbres, e incluso llega a poner en cuestión ciertas bases tradicionales de nuestro sistema del Derecho de autor. El presente libro trata de arrojar luz sobre algunas de estas incertidumbres.

Las autoras, **Teresa Rodríguez Cachón** y **Elena Vicente Domingo**, son profesoras de Derecho Civil en la Universidad de Burgos, donde la segunda es Catedrática y la primera, Ayudante Doctora. Existe entre ambas una estrecha colaboración, que ha dado como principal resultado la tesis doctoral titulada *Relaciones contractuales de la cadena alimentaria: estudio desde el análisis económico del Derecho*, recientemente premiada por su destacado impacto socioeconómico.

En la trayectoria de la profesora **VICENTE DOMINGO**, hay varias obras dedicadas a la propiedad intelectual, con una primera monografía: *El droit de suite de los artistas plásticos*, publicado en esta editorial. Además, ha participado en cursos y conferencias y ha organizado encuentros en los que se ha debatido sobre *Ingeniería y Derecho de autor y Estados civiles y Derecho de autor*, *El precio del libro*, y otras muchas, tanto en colaboración con el Gremio de editores de CyL como en el seno de la Asociación para la Enseñanza del Derecho de autor (ASEDA), a la que ambas pertenecen.

Por su parte, **RODRÍGUEZ CACHÓN** ha comenzado hace un tiempo a interesarse y a especializarse en el Derecho de autor y ha coordinado un congreso internacional dentro del proyecto *Propiedad intelectual y open data en la Universidad: intersección entre propiedad intelectual, reutilización de la información del sector público y protección de datos*. Una de las publicaciones derivadas de este proyecto es el embrión de presente trabajo. Es en este contexto en el que nace este libro para la nueva colección que dirige el Prof. Carlos **ROGEL VIDE**, el cual ha sido preparado con ilusión, en una interesante y positiva coautoría.