

El derecho al olvido en Internet en la era de la inteligencia artificial

The Right to Be Forgotten on the Internet in the Age of Artificial Intelligence

POR MARÍA FERNANDA SÁNCHEZ DÍAZ (*)

Palabras claves

derecho al olvido
inteligencia artificial
protección de datos
personales
machine unlearning
dignidad humana

Resumen

El presente artículo analiza los desafíos que enfrenta el derecho al olvido en la era de la inteligencia artificial, particularmente en un contexto en el que los modelos de inteligencia artificial son entrenados con grandes volúmenes de información obtenida tanto de Internet como de la interacción cotidiana de las personas con estos sistemas. El objetivo consiste en analizar los límites que presenta el modelo tradicional del derecho al olvido cuando la información ya ha sido incorporada al proceso de aprendizaje de la inteligencia artificial. Para ello, se desarrolla un análisis jurídico desde la perspectiva de los derechos humanos, la protección de datos personales y la gobernanza de la inteligencia artificial, tomando como referencia la evolución del derecho al olvido y los nuevos desafíos que plantea la memoria digital. El estudio concluye que la protección efectiva de este derecho humano exige replantear sus mecanismos tradicionales frente al aprendizaje permanente de los sistemas de inteligencia artificial.

Abstract

Keywords

right to be forgotten
artificial intelligence
personal data protection
machine unlearning
human dignity

This article examines the challenges faced by the right to be forgotten in the era of artificial intelligence, particularly in a context where artificial intelligence models are trained on vast amounts of information obtained both from the Internet and from the daily interactions of individuals with these systems. The objective of this study is to analyze the limitations of the traditional model of the right to be forgotten once personal information has been incorporated into the learning process of artificial intelligence systems. To this end, the article develops a legal analysis from the perspective of human rights, personal data protection, and artificial intelligence governance, taking as its starting point

(*) Licenciada, Maestra y Doctora en Derecho, Universidad Nacional Autónoma de México. Máster en Protección Internacional de los Derechos Humanos, Universidad de Alcalá, España. Investigadora Nacional del Sistema Nacional de Investigadores (SNII), nivel 1. Catedrática de Licenciatura y Posgrado, Universidad Nacional Autónoma de México y Universidad Internacional de La Rioja, España. Certificada en Inteligencia Artificial por American Business College, USA. Miembro del Comité Científico del *Annual International Congress on Computer Science*, Oxford, Reino Unido como experta en inteligencia artificial. ORCID: <https://orcid.org/0000-0002-7519-5258>

the evolution of the right to be forgotten and the challenges posed by digital memory. The study concludes that the effective protection of this human right requires rethinking its traditional mechanisms in light of the continuous learning processes of artificial intelligence systems.

I. Introducción

El presente artículo aborda un tema cuya relevancia resulta cada vez más evidente en el contexto del desarrollo de la inteligencia artificial: los desafíos que enfrenta el derecho al olvido frente a sistemas capaces de aprender, generar inferencias y construir memoria a partir de los datos con los que son entrenados. Se trata de un fenómeno que trasciende el ámbito de la protección de datos personales, pues involucra directamente el ejercicio de derechos como la autodeterminación informativa, el derecho al honor, la privacidad y, de manera particular, la dignidad humana.

El derecho al olvido no se encuentra reconocido de manera uniforme en todos los ordenamientos jurídicos. La Unión Europea ha sido de las regiones que mayor desarrollo ha alcanzado en esta materia, particularmente a partir de la sentencia dictada por el Tribunal de Justicia de la Unión Europea en el caso Mario Costeja González. En otros países, la protección de intereses similares se materializa mediante figuras como la cancelación o supresión de datos personales, lo que pone de manifiesto que aún no exista una respuesta homogénea frente a los desafíos que plantea la permanencia de la información en los entornos digitales.

No obstante, el desarrollo de la inteligencia artificial ha transformado profundamente este escenario. El principal desafío ya no consiste únicamente en eliminar información de Internet o impedir su indexación por los motores de búsqueda. Actualmente, resulta necesario reflexionar sobre la capacidad de los sistemas de inteligencia artificial para dejar de utilizar la información con la que fueron entrenados cuando una persona ejerce su derecho al olvido. En otras palabras, el reto consiste en analizar si estos sistemas pueden desaprender aquello que previamente incorporaron durante su proceso de aprendizaje, garantizando con ello una protección efectiva de los datos personales y de los derechos fundamentales de las personas.

Lo anterior obliga a replantear los mecanismos tradicionales de protección de datos personales y a reflexionar sobre la necesidad de contar con marcos regulatorios capaces de responder a los nuevos desafíos tecnológicos. En este contexto, la privacidad desde el diseño y por defecto debe trascender al ámbito de la inteligencia artificial, incorporando mecanismos que permitan materializar de manera

efectiva el derecho al olvido cuando la información ya forma parte del proceso de aprendizaje de un modelo y ya no es de interés público. Desde esta perspectiva, el presente artículo analiza la viabilidad del *machine unlearning* como una tecnología emergente que puede contribuir a la protección de la dignidad humana en la era de la inteligencia artificial, sin perder de vista que su implementación deberá integrarse dentro de un modelo de gobernanza sustentado en los derechos humanos y la ética tecnológica.

II. El derecho al olvido como mecanismo de protección de la dignidad humana en el entorno digital

Los datos personales constituyen, en la actualidad, uno de los activos más valiosos para gobiernos, empresas, medios de comunicación y un amplio número de actores públicos y privados. De una u otra manera, representan un elemento fundamental para la toma de decisiones. Sin embargo, su importancia no radica únicamente en el uso inmediato que pueda hacerse de ellos, sino también en su potencial utilización futura. Con frecuencia, la información es recabada y almacenada sin que exista una finalidad específica en ese momento, aunque posteriormente pueda ser analizada o utilizada conforme evolucionen las necesidades de quien la posee.

En este sentido, Carissa Véliz sostiene que: “El poder de las compañías tecnológicas se forma, por un lado, a partir de la posesión de un control exclusivo sobre nuestros datos y, por su capacidad de prever todos nuestros movimientos, lo que, a su vez, les brinda múltiples oportunidades de influir en nuestra conducta y de vender esa influencia a otros, gobiernos incluidos” (2021, p. 68).

Los modelos de negocio que utilizan tanto las grandes empresas como las pequeñas se sustentan en la explotación de datos personales. Con frecuencia, la justificación para el tratamiento masivo de esta información descansa en la personalización de los servicios que se ofrecen a los usuarios. Las personas tienen la percepción de que las empresas y la tecnología les faciliten la vida, y de que existe un interés por conocer mejor a sus clientes para brindarles un servicio más integral. Sin embargo, los datos personales que son objeto de tratamiento representan un activo económico de enorme valor, tanto a corto como a mediano y largo plazo. Actualmente, la economía digital se sustenta, en gran medida, en el tratamiento de datos personales, los hábitos de consumo y la perfilación de los usuarios.

En este contexto, resulta importante reflexionar sobre una consecuencia que durante muchos años pasó prácticamente inadvertida: la desaparición del olvido como un proceso natural de la vida de las personas. Durante mucho tiempo, el olvido formó parte natural de la vida de las personas. Permitía continuar con la vida sin que hechos ocurridos en el pasado afectaran permanentemente el presente y, sobre

todo, el futuro. Asimismo, permitía equivocarse, aprender de los errores cometidos y cambiar para seguir adelante. En este sentido, el olvido hacía posible que el tiempo cumpliera su función, particularmente en aquellos casos en los que determinada información únicamente resultaba relevante para quienes participaron en los hechos.

Sin embargo, con la llegada de Internet, este anhelo de dejar atrás el pasado dejó de ser una posibilidad real. La información que compartimos en la red se difunde a una velocidad de la que muchas personas no son plenamente conscientes. La realidad es que, una vez que una persona publica información en Internet, pierde, en gran medida, el control sobre ella.

Para comprender la dimensión de este problema, pensemos en una persona que cometió un delito hace quince o treinta años, cumplió con la sanción que le fue impuesta conforme a la ley, posteriormente se reinsertó en la sociedad, reconstruyó su proyecto de vida, obtuvo un empleo estable, formó una familia y maduró. Sin embargo, años después, una persona, un grupo de personas o incluso una empresa o un gobierno decide que el hecho por el cual fue sancionada debe seguir siendo objeto de difusión. En este sentido, resulta pertinente preguntarse: ¿debe una persona cargar de manera permanente con las consecuencias sociales de un hecho por el cual ya respondió jurídicamente?, ¿tiene derecho a dejar atrás su pasado?, o bien, ¿la sociedad tiene derecho a continuar estigmatizándola bajo el argumento de un supuesto interés nacional?

No obstante, este equilibrio se ha roto. Actualmente vivimos en un entorno en el que la información permanece almacenada tanto en servidores físicos como en sistemas de almacenamiento en la nube. Además, cada vez resulta más económico conservar grandes volúmenes de información, una parte importante de la cual también es utilizada para el entrenamiento de sistemas de inteligencia artificial.

Uno de los casos más emblemáticos en materia de derecho al olvido es el del ciudadano español Mario Costeja González, quien, en el año 2010, presentó una reclamación ante la Agencia Española de Protección de Datos contra Google España y el periódico La Vanguardia Ediciones, ante la negativa de ambas empresas de atender su solicitud de oposición al tratamiento y supresión de los datos personales derivados de una publicación periodística en la que se le identificaba como moroso.

El reclamo del Señor Mario Costeja se originó porque, al realizar una búsqueda de su nombre en Google, aparecía una antigua publicación periodística correspondiente a una digitalización de la versión impresa del periódico La Vanguardia Ediciones, en la que se le identificaba como deudor. En la misma página también aparecía una nota relacionada con el caso de la solicitud de eutanasia de Ramón Sampederro, asunto que, por sus implicaciones jurídicas y sociales, era considerado de interés público

en una época en la que la eutanasia no estaba permitida y en la que el señor Ramón Sampredo buscaba hacer valer lo que se reconoce como el derecho a una muerte digna. Cabe destacar que, aun cuando Mario Costeja obtuvo una resolución favorable para el ejercicio de su derecho al olvido, la publicación periodística continúa, al mes de julio de 2026, disponible en el sitio de Internet del periódico La Vanguardia.

En el año 2014, el Tribunal de Justicia de la Unión Europea dictó la sentencia que se convertiría en un referente en materia de derecho al olvido en Internet y protección de datos personales. En ella analizó la responsabilidad de los denominados motores de búsqueda, particularmente Google Spain, filial de Google Inc., respecto del tratamiento de los datos personales que eran indexados por dichos motores de búsqueda a partir de contenidos publicados por terceros.

Tras el análisis de las cuestiones, el Tribunal concluyó, en términos generales, que la normativa europea, tanto la Unión Europea como la legislación española que la desarrollaba, resultaba aplicable a Google Spain; que los proveedores de servicios de motores de búsqueda debían ser considerados responsables del tratamiento de datos personales y, en consecuencia, reconoció la posibilidad de que las personas titulares de los datos solicitaran que su nombre dejara de estar vinculado, como criterio de búsqueda, a determinadas páginas web. A ello se suma el énfasis del Tribunal al considerar que la información estructurada que proporciona un motor de búsqueda sobre una persona física puede generar una afectación significativa tanto a su vida privada como a la protección de sus datos personales. (Maqueo, 2016, p. 76)

La información hoy puede replicarse con enorme facilidad a escala global, generando impactos tanto positivos como negativos. Es decir, aquello que anteriormente desaparecía con el transcurso del tiempo hoy permanece almacenado en bases de datos, servidores, plataformas digitales y sistemas algorítmicos, los cuales son utilizados para el entrenamiento de sistemas de inteligencia artificial. Lo que antes se desvanecía gradualmente de la memoria social, hoy puede conservarse y recuperarse de manera prácticamente indefinida.

En este sentido, Guzmán señala que: “La conservación de grandes cantidades de información, la inmensa capacidad de almacenamiento de la red informática mundial y la posibilidad de acceder a través de esta a dicha información prácticamente en cualquier tiempo y lugar, así como las técnicas de búsqueda e indexación, hacen más sencillo el acceso a la misma, con consecuencias que en ocasiones resultan dañinas para las personas” (2024, s.p.).

En este contexto, resulta necesario analizar este fenómeno desde la perspectiva

de la privacidad y la protección de datos personales, sin perder de vista que el eje central de la discusión debe ser la dignidad humana. Lo que está en juego no es únicamente la protección de la información, sino la posibilidad de que las personas puedan evolucionar, cambiar, reconstruir su proyecto de vida y no quedar definidas permanentemente por un momento específico de su historia.

Lo anterior pone de manifiesto la necesidad de establecer límites al desarrollo tecnológico y al tratamiento de la información en los entornos digitales. Ello exige incorporar una perspectiva transversal de derechos humanos que oriente el diseño, desarrollo e implementación de las tecnologías, a fin de garantizar que la innovación tecnológica se encuentre al servicio de la persona y del respeto a su dignidad humana¹.

En este sentido, De León Batista advierte que la revolución digital ha transformado profundamente la manera en que se ejercen y protegen los derechos humanos. Entre los principales desafíos identifica el acopio masivo de datos por actores públicos y privados, conscientes del valor estratégico de la información; el uso de la inteligencia artificial para la toma de decisiones; la propagación de discursos de odio; el ciberacoso; la infodemia; y el empleo de algoritmos en actividades tradicionalmente reservadas al juicio humano (De León, 2020, p. 674).

Lo anterior resulta especialmente si se considera que una memoria digital sin límites jurídicamente definidos puede convertirse en una forma de sanción permanente de facto, impuesta sin la intervención de una autoridad jurisdiccional imparcial, sin un análisis del contexto, sin criterios de proporcionalidad, legalidad y seguridad jurídica, así como sin una posibilidad real de reintegración a la vida pública y privada.

Ahora bien, uno de los aspectos fundamentales que deben considerarse en materia de derecho al olvido es que la tecnología ya no se limita a almacenar información. Actualmente también la organiza, la jerarquiza, la recomienda y la hace visible, con el consecuente impacto en derechos como la autodeterminación informativa, el derecho a la propia imagen y, por supuesto, el derecho al olvido. Ello obedece a que la tecnología no solamente recuerda, sino que también decide qué se recuerda. Precisamente por ello, resulta necesario ampliar la discusión hacia un nuevo paradigma centrado en la gobernanza de la memoria digital.

El escenario se volvió aún más complejo con la llegada, en noviembre de 2022, de los sistemas de inteligencia artificial de acceso masivo, los cuales hoy se encuentran al alcance de prácticamente cualquier persona. En muchos casos, los usuarios interactúan con estos sistemas sin ser plenamente conscientes de ello y comparten información sin reflexionar sobre el tratamiento que recibirán sus datos personales,

(1) Cfr. Calera y Rosell, 2022.

la forma en que estos podrán ser utilizados para el entrenamiento de los modelos de inteligencia artificial, las inferencias que podrán generarse a partir de ellos o el tiempo durante el cual permanecerán siendo sujetos a tratamiento.

Uno de los principales retos radica en la facilidad y rapidez con que los sistemas de inteligencia artificial son capaces de localizar información contenida en millones de fuentes, identificar patrones y generar nuevas inferencias en cuestión de segundos. Se trata de una capacidad que, hasta hace algunos años, resultaba considerablemente más limitada, ya que, aun cuando existían sistemas computacionales capaces de procesar grandes volúmenes de información, se requería infraestructura tecnológica altamente especializada, a la que únicamente podían acceder un reducido número de empresas o instituciones debido a los elevados costos que ello implicaba.

Actualmente, estas herramientas son mucho más accesibles y pueden ser utilizadas por prácticamente cualquier persona, incluso por usuarios que desconocen las implicaciones jurídicas y éticas que puede tener compartir información propia o de terceros con sistemas de inteligencia artificial. En muchos casos, no existe conciencia sobre la forma en que esa información será tratada, las inferencias que podrán generarse a partir de ella o las posibles afectaciones que ello puede producir en derechos como la privacidad, la protección de datos personales y, en aquellos ordenamientos jurídicos en los que se encuentra reconocido, el derecho al olvido.

Lo anterior implica que, aun cuando una persona considere haber eliminado determinada información, esta puede continuar existiendo en otros sistemas accesibles para la inteligencia artificial. Ello resulta especialmente relevante si se considera que el desarrollo de esta tecnología avanza a un ritmo acelerado, mientras que los marcos regulatorios y los mecanismos de control ético no han evolucionado con la misma rapidez. Así, aun cuando la persona titular haya eliminado un determinado contenido, ya sea por iniciativa propia o mediante el ejercicio del derecho al olvido, basta con que dicha información haya sido replicada por terceros o almacenada en diferentes bases de datos para que continúe siendo accesible y susceptible de tratamiento. Lo anterior pone de manifiesto que la materialización del derecho al olvido enfrenta desafíos cada vez mayores frente a la capacidad de los algoritmos para localizar, relacionar y reutilizar información disponible en el entorno digital.

En este contexto, resulta necesario analizar jurídicamente los desafíos que enfrenta el modelo tradicional del derecho al olvido frente a la inteligencia artificial y la memoria digital. Ello permitirá determinar si los mecanismos jurídicos existentes continúan siendo suficientes para garantizar una protección efectiva de los derechos de las personas en el entorno digital y, particularmente, para salvaguardar la dignidad humana en un escenario caracterizado por la permanencia, reutilización e inferencia constante de la información.

III. Los límites del derecho al olvido frente a la inteligencia artificial y la memoria digital

Como se ha señalado anteriormente, los sistemas de inteligencia artificial requieren grandes volúmenes de información para su entrenamiento y el perfeccionamiento continuo de sus modelos. Sin embargo, pese a los avances alcanzados en los últimos años, estas tecnologías aún presentan importantes limitaciones. Una de las más relevantes es el fenómeno conocido como alucinación, término utilizado para describir aquellos supuestos en los que un sistema de inteligencia artificial genera respuestas que aparentan ser correctas, pero que carecen de sustento fáctico o jurídico.

Lo anterior obedece, entre otras razones, a que estos sistemas procesan información proveniente de un universo inmenso de fuentes disponibles en Internet, además de la información contenida en sus propios modelos. En consecuencia, todavía no son capaces de distinguir de manera automática y plenamente confiable entre información verdadera e información falsa. En el ámbito jurídico, por ejemplo, pueden generar referencias a sentencias, criterios jurisprudenciales o disposiciones normativas inexistentes. Esta situación hace indispensable la intervención del usuario para verificar la información generada por los sistemas de inteligencia artificial.

Lo anterior obliga a reflexionar sobre la necesidad de que quienes utilizan sistemas de inteligencia artificial cuenten con competencias mínimas en materia de protección de datos personales, ética, derechos humanos e inteligencia artificial, un escenario que, en las condiciones actuales, resulta particularmente complejo de alcanzar. El acceso masivo a estas tecnologías ha permitido que prácticamente cualquier persona pueda utilizar sin conocer plenamente las implicaciones jurídicas, éticas y sociales derivadas del tratamiento de la información que comparten con ellas. Como consecuencia, es cada vez más frecuente encontrar escenarios en los que se introducen datos personales de terceros sin aplicar procesos adecuados de anonimización² o seudonimización³, incrementando el riesgo de afectar su privacidad, su huella digital y su memoria digital. Ello adquiere especial relevancia si se considera que, dependiendo del sistema utilizado y de sus condiciones de funcionamiento, esa información puede ser objeto de almacenamiento, tratamiento o reutilización para el entrenamiento y mejora de los modelos de inteligencia artificial.

(2) "La anonimización de datos define la metodología y el conjunto de buenas prácticas y técnicas que reducen el riesgo de identificación de personas, la irreversibilidad del proceso de anonimización y la auditoría de la explotación de los datos anonimizados, monitorizando quién, cuándo y para qué se usan. Es decir, cubre tanto el objetivo de anonimización, como el de mitigación del riesgo de reidentificación, siendo este último un aspecto clave". (Ministerio para la Transformación Digital y de la Función Pública, 2021, s.p.).

(3) "La seudonimización se presenta como una técnica crucial en la gestión de datos, ya que requiere una capa adicional de seguridad para evitar la reidentificación de las personas. Un ejemplo de esta técnica es la creación de un "hash", un algoritmo matemático que transforma los datos en una serie de caracteres, o la utilización de claves adicionales para acceder a datos encriptados, intercambiando datos personales por tokens o números aleatorios, entre otras estrategias (Delgado, 2024, s.p.).

En este sentido, Andrés Delgado sostiene que uno de los principales desafíos en la implementación y uso de la inteligencia artificial se encuentra en la protección de los datos personales. Actualmente, una parte importante de nuestra interacción cotidiana ya se desarrolla en entornos digitales que incorporan sistemas de inteligencia artificial, aun cuando muchas personas no sean plenamente conscientes de ello. Basta pensar en el uso de dispositivos móviles o de plataformas de streaming como Youtube, Netflix o Prime Video, por mencionar algunos ejemplos, cuyos algoritmos analizan constantemente nuestros hábitos de consumo, preferencias e interacciones con el propósito de personalizar los contenidos que nos muestran. Lo mismo ocurre con las redes sociales, cuyos algoritmos buscan mantener la atención de las personas mediante la recomendación continua de contenidos acordes con sus intereses. Todo ello implica un tratamiento permanente de datos personales que va conformando nuestra huella digital, sin que los usuarios conozcan con certeza qué información se almacena, durante cuánto tiempo permanecerá disponible o cuál podrá ser su utilización en el futuro. En este contexto, resulta pertinente preguntarse si, frente a la permanencia de la información en los ecosistemas digitales y a la capacidad de los sistemas de inteligencia artificial para reutilizarla, el ejercicio del derecho al olvido será suficiente para garantizar una protección efectiva de los derechos de las personas.

En palabras de Andrés Delgado, el nudo gordiano consiste en:

¿Qué podemos hacer para prevenir el uso indebido o inapropiado de nuestros datos? En Argentina, la regulación de datos personales existe, plasmada en leyes tales como la 25.326 de Protección de Datos Personales, aunque fue pensada en una Web 1.0 y 2.0, en donde la interacción era limitada. En dichas interacciones de la red, el usuario tenía un perfil más bien pasivo, mientras que la Web 2.0 y 3.0 busca más un perfil del usuario más activo. (2024, s.p.)

Actualmente, algunas regiones del mundo, como la Unión Europea, y algunos países, como España, han comenzado a avanzar en la construcción de modelos de gobernanza de la inteligencia artificial con el propósito de establecer límites al desarrollo de esta tecnología. Ello resulta particularmente relevante si se considera que el desarrollo de la inteligencia artificial se ha convertido en una posición de liderazgo. Se trata de una competencia que demanda inversiones económicas sin precedentes, pero que, sobre todo, requiere cantidades masivas de datos para el entrenamiento y perfeccionamiento continuo de los algoritmos, tal como se ha señalado a lo largo del presente artículo.

El caso de España resulta especialmente relevante, ya que su legislación en materia de inteligencia artificial, la Ley Orgánica para el buen uso y la gobernanza de la inte-

ligencia artificial, incorpora la figura del delegado de inteligencia artificial, tomando como referencia el modelo previsto en el Reglamento General de Protección de Datos de la Unión Europea respecto del delegado de protección de datos. No obstante, aún queda por resolver un aspecto fundamental: determinar el perfil profesional que deberá reunir quien desempeñe dicha función. La discusión no debería centrarse únicamente en si cuenta con conocimientos tecnológicos suficientes, sino también en si posee una formación jurídica sólida, con perspectiva de derechos humanos, protección de datos personales y ética aplicada a la gobernanza de la inteligencia artificial. Ello resulta particularmente relevante si se considera que, frente a los nuevos desafíos tecnológicos, la protección de derechos como el derecho al olvido exigirá analizar si basta con eliminar la información de los sistemas tradicionales o si, además, es necesario desarrollar mecanismos que permitan a los modelos de inteligencia artificial dejar de utilizar la información con la que previamente fueron entrenados.

Otra de las aportaciones de la norma, que no procede de la adaptación del RIA, es impulsar el buen uso de la IA en el sector público estatal, medida que se incluye a petición de numerosas propuestas incluidas en la audiencia pública. Para ello, introduce la creación de un inventario de sistemas de IA utilizados en procedimientos administrativos, no solo para sistemas de alto riesgo, reforzando la transparencia. Además, establece la figura del delegado de IA, que estará encargado de coordinar la aplicación normativa y asesorar en proyectos y contratación pública. Tanto el inventario como el delegado se desarrollarán por Real Decreto. Adicionalmente, se dará impulso a la formación y concienciación de los empleados públicos en materia de inteligencia artificial. (Ministerio para la Transformación Digital y de la Función Pública, 2026, s.p.)

Las medidas adoptadas por España no permiten afirmar, por sí mismas, que se encuentre garantizado un desarrollo responsable y un uso adecuado de la inteligencia artificial. Sin embargo, representan un avance significativo frente a un fenómeno tecnológico que, en muchos países, todavía no ha sido asumido en toda su dimensión por los gobiernos. En buena medida, continúa prevaleciendo un modelo en el que la regulación avanza a un ritmo considerablemente más lento que el desarrollo tecnológico, dejando amplios espacios para la autorregulación de las empresas que diseñan, desarrollan e implementan sistemas de inteligencia artificial.

La complejidad que plantea este escenario es relevante, particularmente si tomamos en consideración que, hoy en día, los datos personales ya no permanecen únicamente como información almacenada, sino que son transformados en patrones, relaciones e inferencias que alimentan el funcionamiento de los sistemas de *machine learning*. Ello obliga a replantear la forma en que tradicionalmente se ha entendido el derecho al olvido, pues ya no basta con eliminar o desindexar determinada informa-

ción de la Internet cuando ésta ha sido utilizada para el entrenamiento de modelos de inteligencia artificial. En este contexto, resulta necesario analizar la viabilidad del *machine unlearning* como una tecnología emergente que permita hacer efectivo el derecho al olvido frente a la inteligencia artificial, mediante el desarrollo de mecanismos capaces de lograr que los modelos dejen de utilizar la información previamente incorporada durante su proceso de entrenamiento.

Para comprender la importancia del *machine unlearning*, resulta necesario partir del concepto de *machine learning* del ecosistema de la inteligencia artificial. En términos generales, el *machine learning* puede entenderse como “una disciplina del campo de la IA que se centra en el desarrollo de algoritmos y modelos estadísticos que permiten a los sistemas informáticos mejorar su rendimiento en una tarea específica a través de la experiencia (...)”, “(...) esta tecnología permite a las máquinas aprender de los datos sin ser programadas explícitamente para cada tarea. Esto significa que los sistemas pueden identificar patrones, hacer predicciones y tomar decisiones basadas en la información que reciben y procesan” (Universidad Alfonso X el Sabio, 2026, s.p.).

Por un lado, el *machine learning* permite que los sistemas de inteligencia artificial generen conocimiento a partir de los datos con los que interactúan. Esta capacidad resulta indispensable para mejorar continuamente su desempeño y se fortalece conforme aumenta la interacción de los sistemas con las personas y con la información disponible en Internet. Sin embargo, cuando este fenómeno se analiza desde la perspectiva de los derechos humanos, particularmente de la protección de datos personales, el derecho al honor, la autodeterminación informativa y, por supuesto, la dignidad humana, surge una cuestión fundamental: así como los sistemas de inteligencia artificial son capaces de aprender, también deben ser capaces de dejar de utilizar determinada información cuando ello resulte necesario para garantizar los derechos de las personas. Lo anterior cobra especial relevancia si se considera que una parte importante de la información con la que se entrenan estos sistemas proviene precisamente de los datos e interacciones generados por las personas en los entornos digitales. En este contexto, resulta necesario analizar la figura del *machine unlearning* como un instrumento que, desde una perspectiva garantista, podría contribuir a hacer efectivo el derecho al olvido frente a los sistemas de inteligencia artificial.

Los sistemas de inteligencia artificial, también conocidos como LLM (*Large Language Models*), ofrecen cada vez respuestas más precisas y una interacción más cercana con las personas, como resultado del entrenamiento continuo de sus modelos y del creciente número de usuarios que interactúan con estas tecnologías. Desde la liberación de *ChatGPT* para el público en general, en noviembre de 2022, quedó de manifiesto que nos encontrábamos frente a una tecnología profundamente disruptiva.

tiva, cuyo desarrollo y adopción avanzaban con mucha mayor rapidez que los marcos jurídicos destinados a regularla. Si bien en un inicio existían importantes cuestionamientos sobre los riesgos asociados a su utilización, hoy resulta evidente que cada vez son más las personas que interactúan con sistemas de inteligencia artificial y que, de manera consciente o inconsciente, proporcionan información personal e incluso datos de terceros durante dichas interacciones. Esta realidad ha obligado a la comunidad científica y jurídica a reflexionar sobre la necesidad de desarrollar mecanismos que permitan el desaprendizaje de la información por parte de los modelos de inteligencia artificial, ya sea para dar cumplimiento a las obligaciones derivadas de la normativa en materia de protección de datos personales o para reducir los riesgos asociados al tratamiento masivo de información personal.

En este sentido, un estudio realizado por la Universidad de Harvard explica una de las principales razones que justifican el desarrollo del *machine unlearning*:

(...) sin que el propietario del modelo lo sepa durante el entrenamiento, un pequeño subconjunto de los datos utilizados para entrenarlo podría ser privado. La preocupación radica en que, al implementar el modelo entrenado, este podría reproducir los datos privados textualmente, exponiéndolos a terceros que interactúan con él. Incluso si el modelo no memoriza los datos con exactitud, nuestro trabajo reciente demuestra que un atacante que acceda al modelo puede, en algunos casos, extraer los datos privados. Otra razón es que un usuario simplemente podría haber solicitado la eliminación de sus datos, un derecho que le asiste gracias a marcos normativos como el RGPD. (Layne, 2024, s.p.)

Ante este escenario, resulta evidente que el derecho al olvido enfrenta un desafío sin precedentes. Mientras que el modelo tradicional busca eliminar o suprimir la información, los sistemas de inteligencia artificial transforman esa información en patrones, relaciones e inferencias que pasan a formar parte de su proceso de aprendizaje. Ello obliga a replantear la manera en que debe materializarse el derecho al olvido en la era de la inteligencia artificial. En este contexto, tanto las empresas que desarrollan o implementan sistemas de inteligencia artificial como los gobiernos deberán impulsar mecanismos que permitan a estos modelos dejar de utilizar la información respecto de la cual se haya ejercido el derecho al olvido, evitando que continúe influyendo en las respuestas, inferencias o decisiones generadas por el sistema, sin que ello implique reconstruir completamente el modelo desde cero. En otras palabras, el verdadero desafío consiste en lograr que el sistema desaprenda únicamente aquello que se pretende eliminar, preservando el resto del conocimiento adquirido durante su proceso de entrenamiento.

Es importante dejar claro que el *machine unlearning* representa una tecnología prometedora, pero de ninguna manera constituye una solución mágica. Su implementación no sustituye la obligación de incorporar la privacidad desde el diseño y por defecto, prevista en el Reglamento General de Protección de Datos (RGPD); tampoco exime del cumplimiento de principios como la minimización de datos, ni de la realización de evaluaciones de impacto en materia de protección de datos personales. De igual forma, no releva a los Estados de su obligación de establecer modelos de gobernanza ética de la inteligencia artificial que contemplen reglas claras, mecanismos efectivos de supervisión y un régimen de responsabilidades y sanciones frente a aquellas prácticas que pongan en riesgo los derechos de las personas. En otras palabras, el *machine unlearning* no sustituye la responsabilidad jurídica ni ética de quienes desarrollan, implementan o utilizan sistemas de inteligencia artificial.

La dignidad humana no puede quedar supeditada a las capacidades de la tecnología. Por ello, el verdadero desafío no consiste únicamente en lograr que los sistemas de inteligencia artificial olviden aquello que deban olvidar, sino en diseñar y desarrollar sistemas que, desde su origen, respeten los derechos humanos y minimicen el riesgo de vulnerarlos.

Frente a este escenario, los desafíos son múltiples. Uno de los más importantes es el regulatorio. Actualmente persisten importantes zonas de incertidumbre jurídica respecto de lo que realmente significa eliminar información de un sistema de inteligencia artificial. Será necesario definir con claridad qué debe entenderse por el ejercicio al derecho al olvido cuando la información ya forma parte del proceso de aprendizaje de un modelo, cómo podrá verificarse que dicho modelo efectivamente dejó de utilizar esa información y qué mecanismos de supervisión permitirán generar certeza sobre el cumplimiento de este derecho. A ello se suma una cuestión fundamental: determinar quién deberá asumir la responsabilidad cuando la vulneración a los derechos de las personas ya se haya materializado.

Aunado a lo anterior, será indispensable realizar un ejercicio de ponderación entre los distintos derechos que convergen en este escenario, la protección de la privacidad y de los datos personales deberá armonizarse con la libertad de expresión, el derecho a la información, la seguridad jurídica y otros derechos fundamentales que pueden entrar en tensión cuando se pretende hacer efectivo el derecho al olvido. Ello exigirá el desarrollo de criterios jurídicos claros, sustentados en los principios de proporcionalidad, razonabilidad y en una visión democrática de los derechos humanos.

Otro desafío radica en la asimetría de poder que actualmente existe entre las grandes empresas tecnológicas, los gobiernos y las personas usuarias. En gran medida, las decisiones sobre qué información permanece visible, qué contenidos se priorizan

y qué información deja de ser relevante son adoptadas por las propias plataformas mediante mecanismos de autorregulación. La transparencia sobre el funcionamiento de los algoritmos continúa siendo limitada y los mecanismos de rendición de cuentas siguen siendo insuficientes. En muchos casos, los avances regulatorios han surgido únicamente después de que una persona ha sufrido una afectación en el ejercicio de sus derechos fundamentales.

Todo lo anterior permite advertir que el verdadero desafío ya no consiste únicamente en garantizar el derecho al olvido en Internet, sino en replantear su alcance frente a sistemas de inteligencia artificial capaces de aprender, generar inferencias y reconstruir información a partir de los datos con los que fueron entrenados. En este contexto, el *machine unlearning* representa una herramienta con un enorme potencial; sin embargo, su implementación deberá formar parte de una estrategia mucho más amplia, sustentada en la privacidad desde el diseño y por defecto, la gobernanza ética de la inteligencia artificial, la transparencia algorítmica y la protección efectiva de los derechos humanos.

La pregunta que debe orientar la discusión ya no es únicamente cómo hacer que la inteligencia artificial olvide determinada información. La verdadera pregunta es cómo proteger la dignidad humana en un entorno en el que la tecnología no solo recuerda, sino que es capaz de reorganizar la información, generar nuevas inferencias y reconstruir permanentemente el pasado de las personas. La respuesta a este interrogante definirá, en buena medida, el futuro de la protección de la dignidad humana en la era de la inteligencia artificial.

IV. Conclusiones

Todo lo analizado en el presente artículo permite arribar a las siguientes conclusiones.

El derecho al olvido ya no puede ni debe seguir siendo entendido únicamente como el derecho a eliminar o desindexar información disponible en Internet. El desarrollo de la inteligencia artificial ha modificado profundamente la naturaleza de este derecho, pues los datos personales dejan de existir únicamente como información aislada para convertirse en patrones, relaciones e inferencias que alimentan el proceso de aprendizaje de los sistemas de inteligencia artificial. En este nuevo escenario, el verdadero desafío consiste en garantizar que las personas puedan ejercer un control efectivo no solo sobre sus datos personales, sino también sobre la forma en que estos continúan siendo utilizados para generar respuestas, inferencias y decisiones algorítmicas sin supervisión humana.

El análisis realizado permite advertir que el ejercicio efectivo del derecho al olvido exige replantear los mecanismos tradicionales de protección de datos personales. La

privacidad desde el diseño y por defecto continúa siendo un principio indispensable; sin embargo, resulta insuficiente si no se acompaña de modelos de gobernanza de la inteligencia artificial sustentados en una perspectiva de derechos humanos, transparencia algorítmica, rendición de cuentas y responsabilidad ética de quienes diseñan, desarrollan e implementan estas tecnologías.

El *machine unlearning* representa una tecnología emergente con un enorme potencial para contribuir a la protección de derechos como el derecho al olvido. No obstante, su implementación no constituye una solución aislada ni sustituye las obligaciones jurídicas derivadas de la protección de datos personales, la evaluación de impacto, la minimización de datos o la privacidad desde el diseño. Su verdadero valor radica en abrir una nueva línea de reflexión sobre la necesidad de desarrollar mecanismos técnicos que permitan a los sistemas de inteligencia artificial dejar de utilizar la información previamente incorporada durante su entrenamiento cuando ello resulte necesario para garantizar los derechos de las personas.

Asimismo, corresponde al Estado asumir un papel mucho más activo en la regulación del diseño, desarrollo e implementación de la inteligencia artificial. La velocidad con la que evolucionan estas tecnologías no puede justificar la ausencia de marcos regulatorios efectivos ni la delegación de decisiones fundamentales en mecanismos de autorregulación diseñados por las propias empresas tecnológicas. La protección de los derechos humanos no puede depender exclusivamente de la buena voluntad de quienes desarrollan estas herramientas.

Finalmente, el desarrollo tecnológico no constituye, por sí mismo, el problema. La inteligencia artificial representa una de las herramientas con mayor capacidad para transformar positivamente la forma en que las personas trabajan, acceden al conocimiento y reciben servicios. El verdadero reto consiste en garantizar que dicha transformación ocurra dentro de un marco ético y jurídico que coloque a la dignidad humana como el eje rector de toda innovación tecnológica. Los datos personales, la memoria digital y la inteligencia artificial no deben ser entendidos como enemigos de los derechos humanos, sino como realidades que exigen nuevos mecanismos de protección, mayores niveles de responsabilidad y una gobernanza capaz de responder a los desafíos del siglo XXI.

En consecuencia, la discusión ya no debe centrarse únicamente en determinar si la inteligencia artificial puede olvidar determinada información. La verdadera pregunta consiste en cómo garantizar la dignidad humana en un entorno en el que los sistemas de inteligencia artificial no solo recuerdan, sino que son capaces de aprender, generar nuevas inferencias y reconstruir permanentemente el pasado de las personas.

V. Referencias

Caldera Ynfante, J. E. y Rosell Aiquel, R. (2022). Dignidad humana y mínimo vital digital. *Revista Internacional de Cultura Visual*, 12(4), 1-16. <https://doi.org/10.37467/rev-visual.v9.3754>

De León Batista, H. A. (2020). La dignidad humana en la era digital. *Anuario de Derecho Constitucional Latinoamericano*, 26, 671-695.

Delgado, A. (6 de mayo de 2024). Privacidad: anonimización y seudonimización frente a la IA. *Río Negro*. <https://www.rionegro.com.ar/opinion/privacidad-anonimizacion-y-seudonimizacion-frente-a-la-ia-3561808/>

El Gobierno aprueba el proyecto de ley que garantizará una supervisión humana y un uso confiable de la IA. (26 de mayo de 2026). *Ministerio para la Transformación Digital y de la Función Pública*. <https://digital.gob.es/comunicacion/notas-prensa/mtdfp/2026/05/el-gobierno-aprueba-el-proyecto-de-ley-que-garantizara-una-super>

Guzmán Camacho, J. J. (2024). El ejercicio del derecho al olvido en México. *Estudios en Derecho a la Información*, 1(16), 35-49. <https://doi.org/10.22201/ij.25940082e.2023.16.18070>

Layne, R. (22 de febrero de 2024). How to make AI 'forget' all the private data it shouldn't have. *Working Knowledge, Harvard Business School*. <https://www.library.hbs.edu/working-knowledge/qa-seth-neel-on-machine-unlearning-and-the-right-to-be-forgotten>

Maqueo Ramírez, M. S. (2016). Análisis comparativo de las resoluciones emitidas por el Tribunal de Justicia de la Unión Europea y el Instituto Federal de Acceso y Protección de Datos respecto del motor de búsqueda gestionado por Google y la Protección de Datos Personales. *Boletín Mexicano de Derecho Comparado*, 1(145), 75-100. <https://doi.org/10.22201/ij.24484873e.2016.145.4992>

Machine learning: Qué es, beneficios y aplicaciones. (8 de junio de 2026). *Universidad Alfonso X el Sabio*. <https://www.uax.com/blog/ingenieria/que-es-el-machine-learning>

Ministerio para la Transformación Digital y de la Función Pública. (28 de octubre de 2021). La importancia de la anonimización y la privacidad de datos. *Datos.gob.es* <https://datos.gob.es/es/blog/la-importancia-de-la-anonimizacion-y-la-privacidad-de-datos>

Véliz, C. (2021). *Privacidad es poder. Datos, vigilancia y libertad en la era digital*. Debate.

Fecha de recepción: 08-07-2026

Fecha de aceptación: 24-08-2026