JUSTICIA
ISSN impreso 0124-7441
ISSN digital 2590-4566

Instrumentos colaborativos en la regulación de la responsabilidad civil frente a daños ocasionados por sistemas de inteligencia artificial: sandboxes y evaluación de impacto algorítmica[1]

Collaborative instruments in the regulation of civil liability against damages caused by artificial intelligence systems: sandboxes and algorithmic impact evaluation

Sebastián Alexander Scheffer Lagos
Universidad Alberto Hurtado, Santiago, Chile
sebahhin90@gmail.com
https://orcid.org/0009-0001-5507-7895

Recibido: 28 de diciembre de 2023 / Aceptado: 29 de enero de 2024
https://doi.org/10.17081/just.29.45.7170

Resumen

Objetivo: evidenciar los aportes significativos de los instrumentos colaborativos, los sandboxes y las evaluaciones de impacto algorítmica, en los problemas jurídicos originados por las deficiencias de la regulación tradicional de la responsabilidad civil extracontractual aplicada a los supuestos de daño ocasionado por sistemas de inteligencia artificial (en adelante, IA). Método: para abordar el objetivo señalado, se ha utilizado una metodología de análisis documental a través de una revisión crítica de una amplia gama de libros, capítulos de libros, sitios web, artículos de revistas académicas destacadas, normas jurídicas, informes y recomendaciones internacionales sobre la regulación de los sistemas de IA. Resultados: la investigación evidencia que estos instrumentos contribuyen a resolver las deficiencias del régimen tradicional de responsabilidad civil extracontractual en relación con los sistemas de IA, en cuanto facilita la comprensión sobre el diseño y funcionamiento de éstos, reduce la asimetría de conocimiento entre los sujetos involucrados y otorga certeza y flexibilidad jurídica. Conclusiones: los sistemas de IA representan un avance para la sociedad, pero también conlleva riesgos, respecto de los cuales la regulación tradicional de la responsabilidad civil extracontractual tiene grandes deficiencias. Por ello, además de la incorporación de definiciones, criterios y procedimientos sobre los sistemas de IA en la regulación, ésta requiere de la aplicación de los instrumentos colaborativos de los sandboxes y evaluaciones de impacto algorítmica, en cuanto éstos permiten que las regulaciones entreguen respuestas acordes a la naturaleza dinámica y compleja

1 Este trabajo corresponde a una profundización de mi tesina para obtener el grado académico de Licenciado en Ciencias Jurídicas y Sociales de la Universidad Alberto Hurtado, Santiago de Chile, titulado *"Responsabilidad civil e inteligencia artificial: desafíos para una regulación civil adecuada frente a daños causados por sistemas de IA"*, la cual fue entregada con fecha 9 de septiembre del 2023 y evaluada con la nota máxima con fecha 20 de noviembre del mismo año, a cargo de la profesora guía doña María José Arancibia Obrador. Sin perjuicio de lo anterior, es menester aclarar que mi tesina no fue publicada.

de la IA, con un equilibrio en la protección de las personas y la innovación.

Palabras clave: evaluación de impacto algorítmica, innovación, inteligencia artificial, regulación, responsabilidad civil extracontractual, sandboxes.

Abstract

Objective: to demonstrate the significant contributions of collaborative instruments, sandboxes and algorithmic impact evaluations, in the legal problems caused by the deficiencies of the traditional regulation of non-contractual civil liability applied to cases of damage caused by artificial intelligence systems (hereinafter, IA). Method: to address the stated objective, a documentary analysis methodology has been used through a critical review of a wide range of books, book chapters, websites, articles from prominent academic journals, legal regulations, reports and international recommendations on the regulation of AI systems. Results: the research shows that these instruments contribute to resolving the deficiencies of the traditional non-contractual civil liability regime in relation to AI systems, in that they facilitate understanding of their design and operation, reducing the asymmetry of knowledge between the subjects involved and provides legal certainty and flexibility. Conclusions: AI systems represent an advance for society, but they also entail risks, with respect to which the traditional regulation of non-contractual civil liability has major deficiencies. Therefore, in addition to the incorporation of definitions, criteria and procedures on AI systems in the regulation, this requires the application of collaborative instruments of sandboxes and algorithmic impact evaluations, as these allow the regulations to provide appropriate responses to the dynamic and complex nature of AI, with a balance between protecting people and innovation.

Keywords: algorithmic impact assessment, artificial intelligence, innovation, non-contractual civil liability, regulation, sandboxes.

Como Citar:

Scheffer, S. (2024). Instrumentos colaborativos en la regulación de la responsabilidad civil frente a daños ocasionados por sistemas de inteligencia artificial: sandboxes y evaluación de impacto algorítmica. *Justicia*, *29* (45), 1-17. https://doi.org/10.17081/just.29.45.7170

I. INTRODUCCIÓN

En el contexto de la cuarta revolución industrial con el uso de nuevas tecnologías, los sistemas de inteligencia artificial (en adelante, IA) representan un avance fundamental en la humanidad. Esto es posible de observar en áreas tales como la salud, con sistemas de IA que permiten una detección temprana de enfermedades (Neumann, 2020; Hernández, 2023); en el transporte, con el uso de vehículos autónomos (Gana, 2023); en la construcción, con el uso de drones autónomos que ayudan a construir edificios (CNN Chile, 2022); entre otras áreas.

Sin embargo, también existen riesgos de que los sistemas de IA causen daño a las personas tales como que un vehículo autónomo choque (Gordo et al, 2019), que un dron autónomo ataque a personas (Vega, 2021), que un sistema de IA sesgado que produzca discriminación arbitraria en la elección de personas en un puesto de trabajo (Dastin, 2018), etc. Por otro lado, también podría ocurrir que el daño hacia las personas no derive de alguna falla o defecto del sistema de IA, sino que este sistema llegue a un grado de complejidad en que sea imposible comprenderlo, por ende, se pierde el control humano sobre el sistema de IA (Grupo de Diarios de América, 2020).

Lo anterior, justifica la necesidad de regulación por el Derecho, con el objeto de que las regulaciones jurídicas aprovechen los beneficios y disminuir los riesgos por el uso de sistemas de IA. En efecto, la postura

defendida en este trabajo es que ello es posible a través de la siguiente solución: la incorporación de los instrumentos colaborativos del sandbox y las evaluaciones de impacto algorítmica en las regulaciones jurídicas y, en particular, en la responsabilidad civil extracontractual subjetiva y objetiva, lo que permite una retroalimentación real entre las regulaciones y el funcionamiento de los sistemas de IA.

Por ello, el análisis de este trabajo contiene la siguiente estructura: 1) la comprensión de la naturaleza de los sistemas de IA mediante sus elementos centrales; 2) examinar las propiedades generales del régimen de responsabilidad civil extracontractual y los problemas derivados de su aplicación a los sistemas de IA en relación con los criterios de imputación; 3) conocer el funcionamiento y aporte de los instrumentos colaborativos del sandbox y las evaluaciones de impacto algorítmica a la regulación de la responsabilidad civil extracontractual, con el objeto de comprender la naturaleza de los sistemas de IA y el fomento de regulaciones flexibles en relación con estos sistemas.

II. MÉTODO

Para esta investigación, se ha utilizado una metodología de análisis documental mediante la cual hubo una revisión crítica de diversos documentos, tales como libros, capítulos de libros, sitios web, artículos de revistas académicas destacadas, normas jurídicas, informes y recomendaciones internacionales sobre la regulación de los sistemas de IA. En efecto, esta revisión crítica refiere a la recopilación, cotejo y un proceso lógico-racional de análisis, síntesis e interpretación sistemática de los postulados doctrinarios y normativos sobre la regulación tradicional de la responsabilidad civil extracontractual frente al supuesto de daño ocasionado por los sistemas de IA.

En específico, esta metodología es empleada bajo los siguientes términos estructurales. En primer término, el análisis de diversas fuentes para comprender la naturaleza de la IA, ya que ello es fundamental para entender los problemas de la IA en relación con la regulación civil frente a daños ocasionados por ésta y el aporte de los instrumentos colaborativos. En segundo término, el análisis de diversas fuentes para comprender las propiedades y problemas de la responsabilidad civil extracontractual tradicional en supuestos de daño ocasionado por sistemas de IA. Por último, el análisis de diversas fuentes para comprender el funcionamiento y aportes de los instrumentos colaborativos a la regulación tradicional de la responsabilidad civil extracontractual e IA.

III. RESULTADOS

Sobre la comprensión de la naturaleza de la IA

Sobre la comprensión de la naturaleza de la IA, autores como Aznar y Domingues (2022) plantean que no hay una definición unívoca sobre qué es lA. Por ello, este trabajo no busca ofrecer una definición única, sino delimitar el concepto a partir de sus elementos centrales. La ventaja metodológica de esta alternativa es otorgar certeza sobre la naturaleza de los sistemas de IA y, con ello, su relación con la responsabilidad civil y los instrumentos colaborativos a ésta.

Por lo tanto, a partir de los postulados de diversos autores, informes y recomendaciones sobre la naturaleza de la IA, este trabajo plantea que los elementos centrales de todo sistema de IA consisten en: 1) la capacidad de imitar el comportamiento humano; 2) la naturaleza dinámica; 3) la autonomía y 4) la falibilidad.

Sobre el primer elemento, Gill Press (2017) refiere que éste es observado desde el origen de la noción de inteligencia artificial, donde John McCarthy y otros expresan que el problema de la inteligencia artificial “es el de hacer que una máquina se comporte de una manera que se llamaría inteligente si un ser humano se comportara así”. Según Araya (2020), Morales (2021), Chui y McCarthy (2018), esto alude a que los sistemas de IA buscan imitar las funciones cognitivas humanas, tales como percibir, razonar, decidir, aprender, interactuar con el entorno, entre otras.

De esta forma, Calo (2015) entiende que los sistemas de IA, mediante algoritmos, perciben la información del entorno, la asocia con otra ya almacenada-sea por el ser humano o por el aprendizaje autónomo del sistema de IA-, lo que permite finalmente que tome una decisión reflejada en una acción concreta.

Con ello, es relevante hacer una primera precisión, la cual es que no todos los sistemas de IA son iguales, lo que es posible de observar en distinciones tales como los sistemas de IA débil y fuerte (Ataz, 2020; Mender, 2020; Morales, 2021; Aznar y Domingues, 2022; Mansilla et al, 2020); sistemas expertos, Machine y Deep Learning (Morales, 2021; Castello, 2020; Amunategüi, 2023), entre otras. En consecuencia, estas distinciones expresan un grado de complejidad diverso en el diseño y funcionamiento de los sistemas de IA, lo que impacta en el grado de comprensión y control humano sobre las mismas y, con ello, en la acreditación de los criterios de imputación de la responsabilidad civil extracontractual.

Sobre el segundo elemento, Llamas et al (2022) refieren que éste deriva de la capacidad de los sistemas de IA de interactuar con el entorno y su capacidad adaptativa. Según estos autores, el primer aspecto consiste en "la capacidad de un agente de percibir e interactuar con otros agentes, sean humanos o artificiales, con sus propias metas y capacidades" (p.32). Sobre el segundo aspecto, estos autores afirman que consiste en "la capacidad de aprender de las propias experiencias, sensaciones e interacciones para reaccionar con flexibilidad a los cambios del entorno" (p.32).

Por lo tanto, este elemento de todo sistema de IA apunta a la función cognitiva del aprendizaje. Sobre ello, Morales (2021) alude a que los sistemas de IA han evolucionado desde los sistemas expertos hasta los sistemas de IA con Machine Learning (aprendizaje automático) y Deep Learning (aprendizaje profundo). Por un lado, este autor plantea que el funcionamiento de los sistemas expertos son actos que se basan en reglas prestablecidas por el ser humano; por otro lado, este autor afirma que los sistemas de IA con Machine y Deep Learning no necesariamente se basan en reglas asignadas por el ser humano, sino por datos obtenidos por el propio sistema tras su interacción con el entorno. En consecuencia, los sistemas de IA han evolucionado su aprendizaje, desde los actos limitados completamente por las reglas humanas hasta actos realizados no sólo por una decisión humana, sino por decisiones tomadas por el mismo sistema de IA, a partir de la interacción y adaptación a las situaciones del entorno, lo cual aumenta su grado de autonomía.

Este elemento también tiene repercusiones jurídicas, ya que el cambio constante de los sistemas de IA por su naturaleza dinámica produce dificultades para la regulación civil tradicional (por su rigidez), lo que impacta en el grado de control humano y la acreditación adecuada de los criterios de imputación de la responsabilidad civil.

Sobre el tercer elemento, Araya (2020), Aznar y Domingues (2022), Llamas et al (2022) y Ataz (2020) expresan que la autonomía consiste en la capacidad de efectuar actos con una mayor o menor complejidad, a través de una intervención mínima o nula del ser humano. Como los sistemas de IA son distintos, cada sistema de IA tiene un grado de autonomía diferente, lo que es reconocido por la Organización para la Cooperación y Desarrollo Económicos (2019), la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (2021) y en las recomendaciones a la Comisión Europea (2019). Según Ataz (2020), el grado de autonomía está determinado por la existencia e intensidad de aspectos tales como la complejidad, opacidad, adaptabilidad, conectividad y aprendizaje del sistema de IA.

Ahora bien, lo anterior no sólo determina el grado de autonomía, sino que produce la siguiente consecuencia: la imprevisibilidad del comportamiento de los sistemas de IA. Para Ataz (2020) y Araya (2020), la imprevisibilidad consiste en que el fundamento de las decisiones de estos sistemas no es posible de prever para el ser humano, sea porque le es desconocido, sea porque es una decisión derivada de la interacción con el entorno, sea porque ha tenido en cuenta una cantidad de variables imprevisibles para el ser humano-por la limitación cognitiva del cerebro humano. Un ejemplo en relación con la opacidad es el fenómeno de la caja negra, donde Oliveira (2020) señala, en síntesis, que es posible conocer el acto ejecutado por el sistema, pero no su fundamento.

Por lo tanto, Araya (2020) plantea que este elemento también tiene relevancia jurídica, dado que el comportamiento autónomo e imprevisible de los sistemas de IA implica una posibilidad eventual de que éstos dañen al ser humano o a sus bienes a través de un acto imprevisible para el ser humano, lo que impacta en el grado de control humano y en los criterios de imputación de la responsabilidad civil.

Sobre el cuarto elemento, según Mender (2020), la falibilidad de los sistemas de IA consiste en que éstos pueden errar, ya que no son sistemas perfectos. Este autor plantea que el origen del error puede derivar de dos aspectos: 1) porque son una creación humana, y el ser humano puede errar; 2) por una obtención de datos erróneos y/o defectuosos del sistema de IA que causen una toma de decisión errada. Este autor y Gordo et al (2019), señalan que esto ocurre, por ejemplo, en el uso de vehículos autónomos y los sistemas biométricos.

Este elemento, al igual que los demás, tiene relevancia jurídica, en cuanto los sistemas de IA no sólo pueden dañar por una decisión imprevisible, sino también por una decisión errada.

Por lo tanto, todos los elementos señalados tienen un grado de repercusión jurídica en el control humano sobre los sistemas de IA y, con ello, en los criterios de imputación de la responsabilidad civil. Esto es relevante, dado que el uso masivo de estos sistemas y la posibilidad de que éstos dañen al ser humano o sus bienes (sea por un comportamiento imprevisible o errado), justifica la necesidad de que el Derecho regule los sistemas de IA y, en particular, la regulación civil frente a daños causados por éstos.

Por ello, la finalidad de toda regulación jurídica de los sistemas de IA-incluida la regulación de la responsabilidad civil- debe buscar el equilibrio entre dos aspectos esenciales: 1) la protección de la víctima frente a daños causados por sistemas de IA y 2) el fomento de la innovación, mediante el incentivo de desarrollar nuevos sistemas de IA.

En el siguiente apartado veremos, de forma general, los problemas de la aplicación del régimen tradicional de responsabilidad civil extracontractual a los sistemas de IA.

Sobre los problemas de la aplicación del régimen tradicional de responsabilidad civil extracontractual a los sistemas de IA

El análisis de las fuentes está orientado entender qué significa este régimen jurídico, por qué es conveniente su aplicación a los sistemas de IA y, con ello, comprender los problemas actuales de su aplicación a los sistemas de IA.

Según autores como Corral (2003) y Barros (2006), el régimen de responsabilidad civil supone el daño sufrido por una persona y la imputación a otra persona de la obligación de reparar íntegramente el daño. Por ello, la Resolución del Parlamento Europeo (2020) ha expresado que este régimen tiene dos grandes funciones, una esencialmente reparatoria y, residualmente, preventiva. En consecuencia, Araya (2020) explica que la aplicación de este régimen a los sistemas de IA es adecuada, ya que su objetivo esencial es la reparación íntegra del daño sufrido por la víctima y, de forma residual, prevenir su ocurrencia, lo que es pertinente respecto de estos sistemas, cuyo uso e impacto es masivo y tienen la posibilidad potencial de dañar a las personas o sus bienes.

Ahora bien, una interrogante a resolver en la investigación es la siguiente: ¿por qué este trabajo está enfocado en la responsabilidad civil extracontractual y no la contractual? Según Araya (2020), esto se debe a que la aplicación del régimen de responsabilidad civil contractual supone el reconocimiento de personalidad jurídica hacia los sistemas de IA, lo que no está establecido en ningún ordenamiento jurídico hasta la fecha, como es reconocido en la Resolución del Parlamento Europeo (2020). En efecto, Araya (2020) refiere que la responsabilidad civil contractual supone la capacidad de las partes involucradas y que posean la calidad de personas, y los sistemas de IA no son reconocidos como personas ni como sujetos capaces de ejercer derechos y contraer obligaciones. En cambio, Corral (2003) plantea que la responsabilidad civil extracontractual comprende una variedad y amplitud de situaciones jurídicas mayores que la contractual, ya que su origen no requiere un vínculo contractual, lo que conlleva una mayor cantidad de sujetos y relaciones jurídicas involucradas. Por ello, este trabajo tendrá un enfoque basado en el régimen de la responsabilidad civil extracontractual.

Según Hernán Corral (2003), define la responsabilidad civil extracontractual en los siguientes términos:

La responsabilidad civil extracontractual es aquella que proviene de un hecho ilícito perpetrado por una persona en perjuicio de la otra, que no constituye la violación de un deber contractual. El deber de reparar surge de la transgresión, no de una obligación propiamente tal, sino de un deber genérico de no dañar a otro (alterum non laedere), que es un principio general de todo ordenamiento jurídico (p.24).

Esta responsabilidad se clasifica en dos tipos, subjetiva y objetiva. Por un lado, según Barros (2006), el fundamento de la responsabilidad subjetiva recae en que “el daño ha sido causado por un hecho negligente, esto es, realizado con infracción a un deber de cuidado” (p.28). Por otro lado, este autor expresa que la responsabilidad objetiva no exige un comportamiento negligente del autor del daño, sino que “la mera relación causal entre el hecho del demandado y el daño sufrido por el demandante” (p.29), cuyo fundamento “es el riesgo creado por quien desarrolla la actividad respectiva y no la omisión de deberes de cuidado” (p.29), por lo que la atribución de responsabilidad no requiere de “un juicio de valor respecto de la conducta del demandado” (p.29). Además, para la doctrina mayoritaria, tales como Barros (2006), Araya (2020) y Amunategüi (2023), la regla general en los ordenamientos jurídicos es la aplicación del régimen de

responsabilidad subjetivo, donde el primer autor señalado expresa que la responsabilidad objetiva "es un régimen especial de derecho estricto, que rige ciertos ámbitos de conducta o tipos de riesgos definidos por el legislador" (p.30), por el peligro que conlleva algunos de sus criterios de imputación, que se verán más adelante.

Dicho lo anterior, ahora es necesario analizar en términos generales los problemas de la aplicación del régimen tradicional de responsabilidad civil extracontractual subjetivo y objetivo a los sistemas de IA.

Por un lado, los problemas de la aplicación del régimen de responsabilidad civil extracontractual subjetivo a los sistemas de IA radican en la excesiva carga probatoria hacia la víctima y el grado de desconocimiento de ambas partes (aunque, en general, será mayor el desconocimiento de la víctima que la del demandado). El fundamento de la carga probatoria elevada radica en la naturaleza dinámica, autónoma e imprevisible de los sistemas de IA, lo que produce una indeterminación del grado de control exigible y/o debida diligencia al sujeto responsable (culpa) y de la causa concreta generadora de la acción dañina del sistema de IA (relación de causalidad).

Por otro lado, el problema principal de la aplicación del régimen de responsabilidad civil extracontractual objetivo es que no considera el comportamiento del usuario del sistema de IA para la atribución de responsabilidad, sino sólo el del demandado, el cual, por el sólo hecho de introducir una actividad riesgosa en la sociedad (en este caso, un sistema de IA), deberá responder, aun cuando sea un riesgo fuera de su esfera de control. Esto genera injusticia y desincentivo en el uso de los sistemas de IA.

Teniendo en cuenta los problemas mencionados anteriormente, ahora haré una profundización breve sobre cada uno de ellos, tanto en la responsabilidad subjetiva como objetiva.

Sobre la responsabilidad civil extracontractual subjetiva

Según autores como Araya (2020), los criterios de imputación de este régimen de responsabilidad son los siguientes: "i) acción u omisión (hecho voluntario); ii) culpa (negligencia) o dolo; iii) el daño; y iv) relación de causalidad entre la acción u omisión dolosa o culpable y el daño" (p.270). En relación con los sistemas de IA, los criterios de imputación que producen una serie de problemas aluden a la capacidad, la culpa y la relación de causalidad, no así el daño. Por ello, sólo serán analizados estos tres primeros elementos.

Tal como fue mencionado con anterioridad, un primer problema alude a la capacidad, en cuanto los sistemas de IA no tienen personalidad jurídica, no son capaces y, por ello, no es posible que se les impute responsabilidad. Para Amunategüi (2023) esto es un problema, dado que este régimen está diseñado para que el sujeto a quien se le imputa responsabilidad sea el causante directo del daño, es decir, la persona que ha producido el daño debe responder por su comportamiento negligente o doloso. Sin embargo, en este caso, aunque los sistemas de IA son los que producen el daño, no es posible atribuir responsabilidad a estos sistemas y, con ello, la imposibilidad de que la víctima obtenga una reparación del daño por parte de los sistemas de IA.

Ahora bien, para que la víctima no quede en una desprotección total y tenga la posibilidad de obtener una reparación mediante este régimen, autores como Fossaceca y Moreyra (2020), Tapia (2021) y en regulaciones como la Resolución del Parlamento Europeo (2020), la OCDE (2019) y la UNESCO (2021), la solución jurídica consiste en que la responsabilidad por daños causados por sistemas de IA siempre debe recaer en una persona que sea capaz, y no en el sistema de IA. En consecuencia, si la persona es capaz, se puede continuar con el análisis de los otros criterios de imputación.

Un segundo problema por discutir dice relación con la culpa. Según Corral (2003), la culpa consiste en "la falta de aquella diligencia o cuidado que los hombres prudentes emplean ordinariamente en sus actos y negocios propios" (p.172), por lo que, además de la ocurrencia del daño, requiere que éste sea producido por un comportamiento negligente del agente. Este autor y Barros (2006) expresan que la forma de apreciación de la culpa en varios ordenamientos jurídicos es la culpa objetiva o en abstracto, la cual, en términos simples, refiere a una comparación del comportamiento del agente con un parámetro normativo de un sujeto prudente en el mismo contexto (persona prudente, buen padre de familia, etc). Por lo tanto, el análisis se puede resumir en lo siguiente: si el demandado tiene un comportamiento equivalente a la persona prudente, no hay culpa; de lo contrario, hay culpa.

El problema, en general, radica en la prueba de la culpa, donde es la víctima quien debe acreditar la existencia de un comportamiento negligente, lo cual es complejo en el contexto actual, donde las relaciones sociales son muy variadas. Según Aznar y Domingues (2022), lo anterior tiene un grado de complejidad mayor

en el caso de los sistemas de IA, por un lado, debido a la naturaleza dinámica, autónoma e imprevisible de dichos sistemas, que muchas veces impide conocer el fundamento de los actos y, con ello, que sean imprevisibles; por otro lado, debido al grado de desconocimiento de la víctima sobre el funcionamiento de los sistemas de IA. En consecuencia, estos autores concluyen que ello genera una pérdida de control humano sobre las acciones de éstos, lo que impide la acreditación de la culpa, ya que, si no es posible prever el acto dañino del sistema de IA-ni su fundamento- para un sujeto prudente en el contexto de la persona a quién se atribuye responsabilidad, no se puede afirmar que ha sido negligente, por lo que no habría culpa ni responsabilidad ni reparación del daño causado por el sistema de IA.

Con el objeto de reducir la carga probatoria de la víctima en la culpa, Amunategüi (2023) expresa que surgen las presunciones legales de culpa, donde la víctima sólo tendrá que acreditar los presupuestos de aplicación de la presunción, cuyo efecto será invertir la carga probatoria, esto es, que sea la persona a quién se imputa responsabilidad quién deba acreditar la ausencia de culpa, y no la víctima quién tenga que acreditar la existencia de culpa. Estas presunciones son tres: 1) presunción por el hecho propio; 2) presunción por el hecho ajeno; 3) presunción por el hecho de las cosas.

Ahora bien, ¿es posible aplicar las presunciones legales de culpa a los sistemas de IA? Para responder esta pregunta, el análisis estará centrado en la presunción legal de culpa por el hecho de las cosas, que es la presunción que ha causado mayor debate dentro de la doctrina y la comunidad en general (no así las otras dos presunciones).

Según Araya (2020), la presunción legal de culpa por el hecho de las cosas "hace responsable al dueño de las cosas, quien debe vigilarlas y mantenerlas en el estado que no cause daño" (p.273), por lo que supone un deber de vigilancia correspondiente a la persona dueña de las cosas. En particular, el análisis está centrado en la presunción legal por el hecho de los animales, donde Corral (2003) expresa que esta presunción también supone un deber de vigilancia que le corresponde al dueño del animal o "a todo el que se sirva de un animal ajeno" (p.247), con el objeto de evitar que el animal cause daño a otro. Por ello, Amunategüi (2023) plantea que, si el animal causa daño, se presume que el dueño o quien tenga a cargo el cuidado del animal no emplearon la debida vigilancia y, por lo tanto, la víctima no tendría que acreditar la culpa, sino que se invierte la carga probatoria a éstos, los cuales tendrán que acreditar la ausencia de culpa.

Ahora bien, ¿es posible aplicar esta presunción frente a daños causados por los sistemas de IA? Para ello, hay dos posturas.

Por un lado, la postura mayoritaria de la doctrina moderna, tales como Amunategüi (2023), Llamas et al (2022) y Ataz (2020), plantean que sí es posible aplicar esta presunción frente a daños ocasionados por los sistemas de IA. En primer lugar, según autores como Llamas et al (2022) y Díaz y Flórez (2022), esta postura expresa que es posible establecer una semejanza entre los sistemas de IA y los animales, debido a su naturaleza de objetos o cosas. En segundo lugar, hay sistemas de IA que pueden dañar a las personas o sus bienes, al igual que los animales, tales como los vehículos autónomos, drones autónomos y sistemas biométricos. En tercer lugar, tanto en los animales como en los sistemas de IA, es posible examinar el deber de vigilancia de la persona a cargo mediante el grado de control razonable a exigir, que es el de una persona prudente en el mismo contexto, donde para Barros (2006) en este contexto debe ser considerado el elemento del riesgo del animal concreto.

Las críticas a esta postura, principalmente, son dos: 1) la presunción, en vez de disminuir la carga probatoria, la aumenta, debido a que la naturaleza compleja de muchos sistemas de IA y el desconocimiento de la víctima producen una alta dificultad para que ésta pueda acreditar la falta de vigilancia y control razonable de las personas a cargo del sistema de IA; 2) Aznar y Domingues (2022) mencionan que, en algunos casos, hay problemas para delimitar y determinar la naturaleza mueble o inmueble del sistema de IA, tales como el caso de las viviendas inteligentes.

Por otro lado, otra postura niega la posibilidad de aplicar esta presunción, planteada por Araya (2020), cuyo fundamento radica principalmente en tres razones. En primer lugar, este autor y Čerka et al (2015) expresan que no es posible asimilar la naturaleza de los sistemas de IA con la de los animales, por las diferencias en su funcionamiento y grado de interacción con el entorno. En segundo lugar, expresa que, a veces, es complejo determinar la naturaleza mueble o inmueble de los sistemas de IA, cuya dificultad no se presenta en el caso de los animales. En consecuencia, esto produce que el grado de control exigible en cada caso no sea posible de asemejar, todo lo cual impide la aplicación de la presunción. En tercer lugar, si el sistema de IA es demasiado complejo, la aplicación de la presunción implica un aumento de la carga probatoria de la víctima, por la naturaleza compleja del sistema de IA y el desconocimiento de la víctima,

que impedirán que ésta acredite los presupuestos de aplicación de la presunción.

La crítica principal a esta postura es que olvida que los sistemas de IA son diferentes. En consecuencia, en los sistemas de IA de menor complejidad o donde es posible acreditar la naturaleza mueble y/o inmueble y el grado de vigilancia exigible, no hay impedimento para aplicar la presunción. Además, si el sistema de IA no tiene mayor complejidad en estos presupuestos, no implicará un aumento de la carga probatoria para la víctima.

Por lo tanto, la postura defendida en este trabajo es la mayoritaria, esto es, que es posible aplicar la presunción legal por el hecho de las cosas-en particular, por el hecho de los animales- frente a daños ocasionados por los sistemas de IA. En efecto, es posible asemejar los sistemas de IA a la naturaleza de objeto, y la presunción se podrá aplicar en la medida en que dichos sistemas no tengan un alto grado de complejidad, con el objeto de que la víctima tenga la posibilidad de determinar la naturaleza concreta del sistema y el grado de vigilancia exigible en el caso concreto, sin que implique un aumento de la carga probatoria, sino una disminución.

Por último, un tercer problema a discutir trata sobre la complejidad para acreditar la relación de causalidad entre la acción u omisión culpable o dolosa y el daño.

En general, Corral (2003) plantea que la acreditación de este criterio de imputación es compleja, y Amunategüi (2023) expresa que ello es aún más complejo respecto de los daños ocasionados por sistemas de IA. En efecto, este último autor dice que la relación de causalidad supone “un modelo linear de acciones que conducen a un resultado” (p.518), lo que, para Barros (2006), implica “el más general fundamento de justicia de la responsabilidad civil, porque la exigencia mínima para hacer a alguien responsable es que exista una conexión entre su hecho y el daño” (p.373).

El problema general de acreditación de la relación de causalidad es la multiplicidad de causas productoras del daño concreto, lo que aumenta en el caso de los sistemas de IA. Aun cuando Corral (2003) expresa que han surgido diversas teorías para determinar la causa concreta que funda la atribución de responsabilidad, éstas no resuelven los problemas de indeterminación de la causa concreta del daño ocasionado por sistemas de IA.

Estos problemas derivan, por un lado, de la naturaleza compleja e imprevisible de los sistemas de IA y, por otro lado, del grado de desconocimiento y limitación humana para comprender y determinar los motivos que fundan ciertas acciones de los sistemas de IA complejos. En efecto, autores como Araya (2020), Aznar y Domingues (2022) y Ataz (2020) mencionan que, tanto para el diseñador del sistema de IA como para el usuario, es difícil o imposible determinar la causa concreta del daño ocasionado por sistemas de IA con alto grado de complejidad, ya que muchas veces el motivo o fundamento de la acción dañina de dicho sistema es desconocido e imprevisto para el ser humano. Además, es necesario considerar la limitación cognitiva del cerebro humano para comprender la cantidad de variables que puede llegar a tener en cuenta un sistema de IA complejo para llevar a cabo una acción determinada.

Todo lo anterior, según estos autores y que adhiero en este trabajo, implica una carga probatoria excesiva para la víctima, sea por la complejidad para comprender el funcionamiento y fundamento de las acciones de un sistema de IA-que ni siquiera un experto a veces puede llegar a comprender-, sea por la limitación cognitiva del cerebro humano para entender lo anterior, sea por el desconocimiento de la víctima para acreditar la acción dañina del sistema de IA.

Por los motivos antes expuestos, ello produce la imposibilidad de imputar responsabilidad y la obligación de indemnizar perjuicios. En consecuencia, este es el principal criterio que dificulta la indemnización de perjuicios por daños ocasionados por sistemas de IA.

Sobre la responsabilidad civil extracontractual objetiva

En general, Barros (2006) y Corral (2003) expresan que este es un régimen excepcional-cuya regla general es el régimen subjetivo-, de fuente legal.

Según Barros (2006), el primer presupuesto de aplicación consiste en que este régimen no requiere de un comportamiento culpable o doloso, sino que sólo requiere la acreditación de la causalidad entre la acción y el daño producido por dicha acción riesgosa, siendo este el elemento determinante de la responsabilidad objetiva (p.448). El segundo presupuesto consiste en que el daño debe ser aquel que materializa el ámbito concreto de riesgo que la ley ha designado a las personas que desarrollan una determinada actividad, y no

otro tipo de riesgo (p.445). Por último, el tercer presupuesto consiste en que este riesgo debe estar bajo el control del responsable (p.445).

Si bien este régimen, en comparación al régimen subjetivo, tiene una mayor facilidad para acreditar el vínculo de causalidad y, por ello, la atribución de responsabilidad a un sujeto concreto, ya que no examina el comportamiento culpable o doloso del agente, sino si se ha producido un daño causado por la actividad riesgosa que éste ha introducido a la sociedad, tiene un carácter excepcional para actividades consideradas riesgosas, porque puede desincentivar dos aspectos relevantes.

En primer lugar, el comportamiento diligente de las personas, puesto que, si aún con un alto grado de diligencia se ha concretado un daño derivado de una actividad riesgosa, estas personas siempre deberán pagar una indemnización, por lo que el incentivo de las personas no será actuar diligentemente, sino el de evitar el pago de la indemnización-por ejemplo, con la contratación de un seguro de responsabilidad civil-.

En segundo lugar, el desarrollo de actividades económicas de emprendimiento innovadoras que impliquen riesgos, puesto que, si la persona siempre tiene que pagar una indemnización en caso de la producción de un daño derivado de una actividad riesgosa, ésta puede preferir no iniciar dicha actividad, para evitar el pago de una indemnización, lo que puede afectar a la economía y avance tecnológico de la sociedad.

De los aspectos mencionados anteriormente, Araya (2020) explica que el problema principal de la aplicación del régimen tradicional de responsabilidad civil objetivo a los sistemas de IA: este régimen ignora el grado de responsabilidad del usuario del sistema de IA. En efecto, muchas veces el usuario es el responsable y generador del riesgo concreto de la actividad del sistema IA y no quien introduce la actividad riesgosa.

Por lo tanto, bajo la aplicación de este régimen objetivo a los sistemas de IA, el diseñador, fabricante y/o la empresa desarrolladora de IA, por el sólo hecho de haber introducido una actividad riesgosa-el sistema de IA-, deberá responder, pese a que su proyecto sea innovador y genere grandes aportes a la sociedad (como podría ser el transporte público y privado mediante vehículos autónomos; maquinaria para el sector de la construcción, minería, etc). Todo ello, aun cuando se trate de un riesgo concreto fuera de la esfera de control de los agentes que introducen el sistema de IA al mercado, lo cual genera injusticia y desincentivo en el uso de los sistemas de IA.

En conclusión, ambos regímenes requieren de modificaciones para una aplicación adecuada frente a daños ocasionados por sistemas de IA, con el objeto de proteger a la víctima mediante una reparación integral del daño y, a la vez, fomentar la innovación con el desarrollo de nuevos sistemas de IA. En efecto, hay medidas propuestas por varias regulaciones, tales como la Resolución del Parlamento Europeo (2020), la Ley Act sobre IA (2021), la propuesta de directiva del Parlamento Europeo y del Consejo (2022), la UNESCO (2021), la orden ejecutiva de EEUU sobre IA segura (2023), la OCDE (2019) y el Reglamento sobre el uso de servicios de información de internet (2022). Estas regulaciones incluyen algunas de estas modificaciones, tales como la distinción de los sistemas de IA según el tipo de riesgo y aplicar el régimen subjetivo a los sistemas de IA menos complejos y el objetivo a los más complejos; establecer un catálogo de definiciones de criterios, tales como “alto riesgo”, “control”; creación de órganos que supervisen la incorporación y funcionamiento de los sistemas de IA; exigir la revelación de prueba, entre otras. Sin embargo, estas medidas son insuficientes, ya que, si bien otorgan mayor certeza jurídica-con ello, confianza en el uso de los sistemas de IA-, faltan otros instrumentos jurídicos que permitan la disminución de la carga probatoria, el aumento de conocimiento de los sujetos involucrados en el ciclo de vida de los sistemas de IA y, a la vez, el fomento de la innovación.

Debido a lo anterior, en el siguiente acápite se hará un análisis sobre algunos instrumentos colaborativos para una regulación adecuada de la responsabilidad civil extracontractual frente a daños ocasionados por sistemas de IA, cuyo objeto será abordar las insuficiencias jurídicas vigentes mencionadas anteriormente.

Instrumentos colaborativos para la regulación de la responsabilidad civil extracontractual por uso de sistemas de IA

Como fue expresado anteriormente, estos instrumentos tienen por finalidad la de colaborar en la delimitación, acreditación y/o exoneración de los criterios de imputación de la responsabilidad civil extracontractual, mediante la disminución de la carga probatoria, facilitar el aumento de conocimiento de los sujetos relacionados con los sistemas de IA y el fomento de la innovación. El mecanismo de regulación

para establecer estos instrumentos debe ser mixto, es decir, que el origen del uso de estos instrumentos no puede radicar sólo en la autorregulación de las empresas, sino que también deben derivar de leyes con un carácter imperativo, ya que, tal como ha expresado Barrio (2021), si el uso de estos instrumentos es meramente voluntario y, con ello, con la posibilidad de ignorar su uso, no habrá una mejora sustantiva en el régimen de responsabilidad civil extracontractual, la protección de la víctima ni la innovación. Por último, estos instrumentos permitirán a las regulaciones jurídicas distinguir de forma adecuada entre los sistemas de IA según el tipo de riesgo concreto, definir con precisión los parámetros sobre los sistemas de IA, actualizar de forma constante el catálogo de sistemas de IA y sus propiedades, entre otros aspectos.

Los instrumentos colaborativos por analizar en este trabajo son los sandboxes y la evaluación de impacto algorítmica.

Sobre los sandboxes

Yolanda Bustos Moreno (2022) define los sandboxes en los siguientes términos:

Son marcos concretos que proporcionan un contexto estructurado para la experimentación, y así permiten ensayar cuando procede, en situaciones reales, tecnologías, productos, servicios o enfoques innovadores —por el momento, sobre todo en el contexto de la digitalización—durante un periodo limitado y en una parte limitada de un sector o ámbito bajo supervisión regulatoria, y garantizar la existencia de salvaguardias adecuadas (p.325).

A esta definición, el trabajo de Guío (2021) sobre los sandbox regulatorio de IA en Chile, agrega que estas zonas de prueba segura conllevan que se pueden probar soluciones innovadoras sin que tengan las consecuencias normales de la regulación por la ejecución de dichos proyectos, lo cual es un beneficio para las empresas desarrolladoras de sistemas de IA, puesto que facilita la entrada de nuevos sistemas de IA al mercado, reduce las barreras de entrada a éste y mitiga el riesgo para la sociedad.

De la definición anterior, es posible establecer una serie de elementos.

El primer elemento consiste en que son un espacio seguro de experimentación. En efecto, para Herrera y Vadillo (2018), el control y seguridad del espacio de experimentación es posible de observar en el cumplimiento de una serie de requisitos para participar del sandbox, durante el período de pruebas y después de la obtención de los resultados del proyecto innovador. Por lo tanto, estos requisitos implican una garantía tanto para los usuarios de los sistemas de IA como para las empresas desarrolladoras de estos sistemas, en cuanto los proyectos que se busca obtener son aquellos que sean realmente viables, innovadores, seguros y que faciliten el acceso a la información mediante el uso de un lenguaje simple que puedan entender todos los sujetos involucrados.

El segundo elemento consiste en que los sandboxes tienen una duración limitada. Las empresas de soluciones innovadoras no pueden tener sus proyectos en período de prueba de forma indefinida, dado que, según Herrera y Vadillo (2018), estos espacios tienen por objeto "facilitar la validación y entender el funcionamiento de productos, servicios, soluciones tecnológicas o modelos de negocio innovadores, antes de ofrecerlos en un mercado mundial" (p.7). Por ello, el ente regulador fija un período de tiempo en el caso concreto que sea razonable para cumplir dicho objetivo.

El tercer elemento consiste en que es un entorno colaborativo de las partes interesadas y autoridades competentes. Según Herrera y Vadillo (2018), esto es posible de observar en aspectos tales como el establecimiento de mecanismos informativos que permitan una retroalimentación entre las empresas innovadoras, los usuarios y los entes reguladores; la búsqueda de medidas conforme al proyecto específico (las características de las pruebas, la información obtenida, etc) y el enfoque flexible para los sujetos involucrados.

El cuarto elemento consiste en que tienen un carácter excepcional. Si bien los sujetos involucrados buscan la máxima mitigación de riesgos durante el período de pruebas, esto no elimina el hecho de que es un espacio de experimentación sujeto a diversos riesgos-algunos conocidos y otros no-. Por ello, Herrera y Vadillo (2018) plantean que no es un espacio otorgado a cualquier empresa, sino sólo a aquellas que acrediten tener un proyecto innovador viable que implica un valor relevante para las personas.

El quinto elemento consiste en la presencia de un ente regulador, que supervisa y colabora en el espacio de experimentación.

El sexto elemento consiste en la flexibilidad normativa. Según Guío (2021), Herrera y Vadillo (2018) y Bustos (2022), esto es posible de observar en diversas medidas, tales como la posibilidad de que no haya una sanción inmediata por el cumplimiento de una norma jurídica tradicional, autorizaciones temporales durante el período de prueba que permitan a los agentes innovadores ejercer medidas para comprender el funcionamiento de su proyecto-lo que también ayuda a una mejor regulación jurídica-, las cláusulas de experimentación, entre otras.

Tenemos diversos ejemplos de experiencias de aplicación de los sandboxes. Sin embargo, la aplicación de sandboxes en inteligencia artificial más reciente es aquella realizada en España mediante el Real Decreto 817 de 2023, que entra en vigor el día 10 de noviembre del año 2023 (siendo este país el primero de los Estados miembros de la Unión Europea que pone en funcionamiento un sandbox en materia de IA). El objetivo es examinar la concreción de los requisitos establecidos en la Ley Act de la Unión Europea del año 2021 sobre IA, en relación con el funcionamiento de estos sistemas antes, durante y después de su implementación en el mercado. Esta regulación considera en diversos artículos los elementos antes mencionados sobre los sandboxes, tales como su carácter excepcional (art.1 y 5), el entorno colaborativo entre los sujetos involucrados (arts. 2, 4, 5, 6 a 15, 26 y siguientes), los requisitos de documentación y planes de gestión de riesgos que deben cumplir los proyectos de sistemas de IA innovadores antes, durante y después de su participación en los sandboxes (arts.6 a 15), entre otros.

Finalmente, se verán los aportes generales del sandbox a las regulaciones jurídicas y los aportes específicos de este instrumento a la regulación civil.

En primer término, los aportes generales del sandbox a las regulaciones jurídicas dicen relación con el establecimiento de regulaciones con un mayor grado de flexibilidad y, a la vez, mayor certeza jurídica, minimiza los riesgos de elaborar regulaciones rígidas y defectuosas y fomenta la innovación de las empresas y la protección a las personas.

Según Bustos (2022), la flexibilidad está dada en que, sin una alteración sustantiva de las regulaciones vigentes-en cuanto las modificaciones jurídicas en los sandboxes están limitadas a un espacio y período determinado, acorde a ciertas características-, permite al ente regulador obtener un alto grado de conocimiento y aprendizaje sobre el funcionamiento de soluciones innovadoras que tienen un aporte relevante a la sociedad en una etapa temprana de las mismas. Dicho conocimiento y aprendizaje otorga la posibilidad de que los reguladores puedan evaluar el estado actual de las regulaciones, aplicar modificaciones jurídicas sustantivas positivas-para un ingreso adecuado de los proyectos innovadores al mercado-, de acuerdo con los resultados reales de las pruebas hechas con las innovaciones.

Lo anterior aumenta la certeza jurídica, ya que, una mejor comprensión del funcionamiento, objetivos y riesgos de la innovación sometida a prueba, permite una elaboración adecuada de regulaciones jurídicas, en el sentido del establecimiento de requisitos equilibrados para los sujetos involucrados (protección a las personas y la innovación), mitigar los riesgos, crear órganos que supervisen antes, durante y después el funcionamiento del proyecto innovador. En efecto, esto es lo que autoras como Bustos (2022) han catalogado como “aprendizaje normativo proactivo”, es decir, fomentar “que los reguladores adquieran un mayor conocimiento normativo y detecten los mejores medios para regular las innovaciones, a partir de ensayos con datos reales, especialmente en su fase más incipiente” (p.342). Esta autora afirma que esto implica un aprendizaje para todos los sujetos involucrados, en los siguientes términos:

El agente innovador o promotor del proyecto aprende porque puede ver en funcionamiento el producto antes de lanzarlo a la realidad; la Administración experimenta aprendizaje porque puede entender los obstáculos o vacíos normativos que impiden el desarrollo de los proyectos; y la sociedad, en general, acumula experiencia asimismo al anticiparse a los cambios que la tecnología, el progreso y la innovación exigen a los ordenamientos jurídicos (p.343).

En segundo término, los aportes específicos de los sandboxes a la regulación de la responsabilidad civil extracontractual radican en la disminución de la carga probatoria, mitigar la brecha de la asimetría de conocimiento entre víctima y demandado, y fomenta la innovación. Por un lado, en la responsabilidad subjetiva, la información otorgada por los sandboxes sobre los proyectos innovadores-en este caso, sistemas de IA- facilita la acreditación o exoneración de la culpa y la relación de causalidad, en cuanto permite una mejor delimitación de la debida diligencia y ayuda a encontrar la posible causa concreta del daño que puedan ocasionar estos sistemas de IA. Por otro lado, en la responsabilidad objetiva, también la información entregada por los sandboxes sobre los proyectos de innovación permite delimitar con mayor precisión los riesgos concretos de la actividad desarrollada por el sistema de IA, lo cual permitirá distinguir

qué riesgos están incluidos y cuáles no y, por ello, si el sujeto es o no responsable y deba indemnizar o no a la víctima.

Por lo tanto, este instrumento facilita la acreditación o exoneración de los criterios de imputación de responsabilidad civil extracontractual, mitiga la brecha de conocimiento entre la víctima y el demandado, mitiga los riesgos; pero, también, es un instrumento que fomenta la innovación mediante el desarrollo de nuevos sistemas de IA.

Sobre la evaluación de impacto algorítmica

Según Llamas et al (2022), la evaluación de impacto algorítmica "es un modelo prometedor de gobernanza algorítmica, debido a que incluye una descripción de los daños potenciales y reales de un sistema para identificar quién es responsable de su reparación" (p.54).

En términos generales, estos autores plantean que este instrumento aborda tres elementos: "qué hace un sistema; quién puede hacer algo sobre lo que hace ese sistema; y quién debería tomar decisiones sobre lo que se le permite hacer al sistema" (p.54). En efecto, un programa promovido por el Ministerio para la transformación digital y de la función pública de España, llamado Digital Future Society (2024), profundiza en una serie de elementos que concretan los aspectos mencionados anteriormente.

Hay una serie de elementos expresados por esta iniciativa. El primer elemento es el foco, que permite comprender el funcionamiento del algoritmo y el contexto en el que se encuentra (p.17). El segundo elemento son los sujetos involucrados, lo que incluye al agente innovador, el ente regulador, los usuarios, entre otros (pp.17-18). El tercer elemento es el momento de realización de la evaluación. Si la evaluación es realizada ex ante, se evaluará sobre el diseño del sistema de IA y los posibles riesgos. Si la evaluación es realizada ex post, se evaluará los impactos reales del sistema de IA (p.19). El cuarto elemento son las orientaciones que puede aportar a la normativa vigente, dependiendo de si busca el cumplimiento legal, de buenas prácticas o de una certificación (p.19). El quinto elemento es el objeto de la evaluación (p.20). El sexto elemento es el nivel de acceso a la información sobre el algoritmo y su contexto (p.21). El último elemento mencionado es la metodología. Hay auditorías algorítmicas y evaluaciones de impacto algorítmicas, y este trabajo está centrado en la segunda metodología, cuya diferencia principal es que las primeras analizan el algoritmo en base a criterios específicos, y las segundas tienen un enfoque más amplio, donde analizan el algoritmo, su contexto, los riesgos ex ante, durante y ex post (p.21).

En efecto, en Digital Future Society (2024) afirman que las evaluaciones de impacto algorítmica:

Pueden medir los riesgos que entraña un sistema entre determinados colectivos de personas, antes o durante la implementación (análisis de riesgos o algorithmic risk assessment) o los impactos que se han generado después de su implementación (análisis de impacto o algorithmic impact evaluation) (p.21).

Si bien Digital Future Society (2024) plantea que la regulación en relación con las evaluaciones de impacto algorítmica aún es escasa, hay diversos instrumentos de evaluación del algoritmo y su contexto con diversos enfoques. En efecto, esta iniciativa expresa que hay evaluaciones que hacen un examen del código fuente del algoritmo y su funcionamiento interno, tales como las auditorías de código y el scraping; otras, evalúan los sistemas de IA y sus algoritmos desde la perspectiva de diversos sujetos, tales como los Sock puppet, Carrier puppet y auditorías colaborativas. Sin perjuicio de lo anterior, han surgido intentos de regulación mediante guías y directrices con parámetros generales para las evaluaciones de impacto algorítmica, tales como el trabajo del Banco de Desarrollo de América Latina y la OCDE (2022) donde es posible observar ejemplos de regulaciones de países como el caso de Canadá, México y Uruguay. En estas experiencias de regulación, han considerado todos los elementos señalados anteriormente, y el objetivo principal es la mitigación de riesgos en el uso de sistemas de IA mediante una evaluación que permita conocer el diseño del sistema concreto, su algoritmo y contexto, lo que contribuye a otorgar garantía de su calidad. Además, la evaluación es efectuada antes, durante y después de la implementación del sistema de IA al mercado, con la exigencia de una actualización constante de la información sobre el sistema y su evaluación.

Finalmente, en relación con los aportes de este instrumento a la regulación jurídica en general-y, en particular, a la regulación de la responsabilidad civil extracontractual frente a daños causados por sistemas de IA-, éstos son similares a aquellos obtenidos de los sandboxes. En efecto, tal como fue expresado con anterioridad, este instrumento permite un mayor grado de flexibilización normativa, mayor certeza jurídica, mitiga los riesgos de daño en los sistemas de IA, otorga la oportunidad de mejorar regulaciones rígidas y

defectuosas y fomenta la innovación de las empresas y la protección a las personas. En relación con la responsabilidad civil extracontractual, este instrumento también contribuye a una disminución sustantiva de la dificultad probatoria para acreditar los criterios de imputación, debido a que las evaluaciones del algoritmo del sistema de IA y su contexto permiten que todos los sujetos involucrados tengan una mejor comprensión respecto del diseño y funcionamiento del sistema durante todo su ciclo de vida. En consecuencia, una mayor comprensión reduce la asimetría de conocimiento entre las partes, en cuanto habrá una determinación más precisa de la debida diligencia (culpa), la causa concreta del daño del sistema de IA (relación de causalidad) y del riesgo concreto de la actividad (responsabilidad objetiva), con un equilibrio entre ambas partes, sin que ello implique desfavorecer a ninguna de éstas.

En síntesis, tanto los sandboxes como las evaluaciones de impacto algorítmica contribuyen a la protección de las personas y de la innovación, debido a que facilitan la comprensión del diseño, funcionamiento y riesgos de los sistemas de IA para todos los sujetos relacionados con éstos (usuarios, empresas desarrolladoras de IA y los agentes reguladores). Esto, permite que las regulaciones establezcan conceptos precisos, lo que favorece la certeza jurídica, sin que ello implique la rigidez tradicional de las normas jurídicas del Derecho. Por último, estos instrumentos colaboran en la incorporación o exclusión de sistemas de IA, de acuerdo con criterios determinados que buscan proyectos innovadores y, a la vez, que entreguen garantías de calidad y seguridad sobre sus productos.

IV. DISCUSIÓN

Como es posible observar en los resultados de esta investigación, las regulaciones civiles tradicionales de la responsabilidad extracontractual requieren de modificaciones sustantivas en aquellos supuestos de daño ocasionado por los sistemas de IA, debido a la naturaleza compleja de éstos, el desconocimiento de las personas sobre su diseño y funcionamiento y, con ello, la carga probatoria excesiva en este tipo de conflictos jurídicos.

Tanto para la responsabilidad civil subjetiva como objetiva, los autores y regulaciones mencionadas expresan modificaciones tales como la distinción entre los sistemas de IA según el riesgo, el establecimiento de definiciones y criterios sobre dichos sistemas y la creación de órganos para la actualización de los sistemas de IA, definiciones y criterios. Además, se ha propuesto la aplicación de las presunciones legales de culpa por el hecho de las cosas-en particular, por el hecho de los animales- respecto de los sistemas de IA sometidos al régimen de responsabilidad subjetivo.

Respecto de estas modificaciones, en este trabajo se ha expresado la adhesión a dichas modificaciones, por su preocupación de elaborar regulaciones que aumenten la certeza jurídica, sean flexibles y acordes a la naturaleza de los sistemas de IA; lo que se observa en el reconocimiento de que los sistemas de IA son distintos, dinámicos y complejos, con la posibilidad de cambio a través de órganos especializados. Sin embargo, también es menester expresar una crítica relevante: estas modificaciones son insuficientes, ya que, si bien es un buen primer paso regulatorio en el aumento de certeza jurídica, flexibilidad y confianza de las personas en el uso de los sistemas de IA, para una retroalimentación efectiva entre las regulaciones y los sistemas de IA es necesario incorporar otros instrumentos jurídicos colaborativos que permitan la disminución de la carga probatoria, el aumento de conocimiento de los sujetos involucrados en el ciclo de vida de los sistemas de IA y, a la vez, el fomento de la innovación. La necesidad de retroalimentación es reconocida por la Resolución de la Asamblea General de la Organización de las Naciones Unidas (2024), toda vez que fomenta la implementación de mecanismos de retroalimentación con el objeto de identificar y mitigar vulnerabilidades técnicas, usos indebidos de los sistemas de IA e incidentes, sea en la fase de desarrollo, puesta a prueba y despliegue en la sociedad (punto n°6, letra c).

Debido a lo anterior, la investigación hace un examen crítico sobre los instrumentos colaborativos de los sandboxes y las evaluaciones de impacto algorítmica. Como se puede observar de los resultados de la investigación, ambos instrumentos son una oportunidad fundamental de aprendizaje sobre el diseño y funcionamiento sobre los sistemas de IA. En efecto, estos instrumentos aumentan el conocimiento de los sujetos involucrados en estos sistemas, lo que produce mayor certeza jurídica, flexibilidad y disminuye la carga probatoria.

Además, el aporte no se reduce a un mayor acceso y comprensión de la información sobre los sistemas de IA, sino también a su seguridad: debido al aumento de conocimiento, ello constituye una oportunidad para crear sistemas de IA seguros, donde es posible incorporar medidas preventivas en el diseño y

funcionamiento de éstos, lo que finalmente aumenta la confianza y, con ello, el fomento de la innovación-por la reputación de las empresas de crear sistemas seguros, con cargas adecuadas en sus obligaciones-. En efecto, la protección de la seguridad en el diseño de los sistemas de IA se ha visto reflejada en diversas propuestas regulatorias de EEUU, tales como el documento propuesto por Microsoft titulado Governing AI: A Blueprint for the Future (2023) donde se propone la implementación de frenos de seguridad en el diseño de sistemas de IA que se ocupen de la infraestructura crítica; el índice de la Universidad de Stanford titulado The AI 2024 Index Annual Report AI (2024) y el plan de la Oficina de Política Científica y Tecnológica de la Casa Blanca titulado Blueprint for an AI Bill of Rights (2022), plantean como principio general para la evaluación responsable de un sistema de IA es que éstos sean seguros y eficaces, mediante el seguimiento constante de sus acciones en la sociedad. Por lo tanto, estos instrumentos permiten una retroalimentación efectiva entre las regulaciones jurídicas y los sistemas de IA.

Por último, si bien de las regulaciones e iniciativas vistas, tales como Real Decreto 817/2023 en los sandboxes, el trabajo del Banco de Desarrollo de América Latina y la OCDE (2022) y Digital Future Society (2024) en las evaluaciones de impacto algorítmica, es posible observar que la aplicación de los sandboxes y evaluaciones de impacto algorítmica en los sistemas de IA es algo reciente, esto no implica un riesgo relevante en relación con la efectividad de su aplicación a estos sistemas. En efecto, sobre los sandboxes, el trabajo de Herrera y Vadillo (2018) evidencia los buenos resultados de la aplicación de estos instrumentos en áreas de características similares a las de la IA (complejidad y dinamismo), como lo es el sector financiero. Lo mismo ocurre en el caso de las evaluaciones de impacto algorítmica, como es posible observar de los trabajos del Banco de Desarrollo de América Latina y la OCDE (2022) y de la iniciativa Digital Future Society (2024), que muestran la contribución de este instrumento en diversos países para un uso responsable de la IA.

V. CONCLUSIÓN

Actualmente, los sistemas de IA tienen un uso masivo en la sociedad, lo que conlleva beneficios y riesgos. Esto, en cuanto los sistemas de IA traen un alto grado de innovación y aportes en diversos sectores de la sociedad (transporte, construcción, minería, educación, etc), pero también éstos pueden dañar a las personas.

Por lo anterior, es necesario que el Derecho regule los sistemas de IA, con el objeto de proteger a las personas y, a la vez, fomentar la innovación. Sin embargo, ello no es posible si no se logra comprender la naturaleza de los sistemas de IA. Para ello, como no hay un concepto universal sobre la IA, a partir del trabajo doctrinario, informes, recomendaciones y normas jurídicas, este trabajo propone algunos elementos centrales de estos sistemas, que son la capacidad de imitar el comportamiento humano, su naturaleza dinámica, autónoma, imprevisible y falible.

Teniendo en cuenta la naturaleza de los sistemas de IA, que no son iguales y existe el riesgo de que causen daño, la responsabilidad civil extracontractual puede servir para establecer parámetros adecuados que permitan una protección de las personas-mediante una reparación integral del daño sufrido por sistemas de IA- y, a la vez, fomentar la innovación, a través de cargas adecuadas a las empresas desarrolladoras de sistemas de IA. Además, este régimen es compatible con otros instrumentos que contribuyan a un equilibrio entre todos los sujetos involucrados, facilitar la acreditación o exoneración de los criterios de imputación mediante una mejor comprensión de la naturaleza compleja de los sistemas de IA.

Sin embargo, el régimen tradicional de la responsabilidad civil extracontractual en los diversos ordenamientos jurídicos es insuficiente para regular adecuadamente los sistemas de IA. Esto, en cuanto es una regulación rígida, con una carga probatoria excesiva para la víctima y un alto grado de incertidumbre jurídica para las empresas desarrolladoras de IA, lo que implica un desincentivo de la innovación. En efecto, debido a la naturaleza compleja de los sistemas de IA (por su naturaleza cambiante, autónoma, imprevisible y falible) y un alto grado de asimetría entre la víctima y el demandado-y, muchas veces, éste último tampoco tiene un nivel de comprensión suficiente del sistema de IA-, la labor de acreditar los criterios de imputación de la responsabilidad subjetiva y objetiva son difíciles o imposibles, por la imposibilidad de determinar la debida diligencia, la relación de causalidad y el riesgo concreto generador del daño en la actividad realizada por el sistema de IA.

Por ello, esta investigación expresa que han surgido diversas propuestas de regulación en la Unión Europea, EEUU, Canadá, China y América Latina; las cuales dicen relación con propuestas de reglamentos

sobre la IA, recomendaciones, principios, entre otras. En estas regulaciones, se proponen soluciones para aumentar la seguridad jurídica, disminuir la asimetría de conocimiento y la carga probatoria, tales como aplicar presunciones legales de culpa por el hecho de las cosas (Resolución del Parlamento Europeo, 2020, p.9 y 20), distinguir los sistemas de IA por riesgos (Ley Act, 2021, arts.5, 6 y siguientes), exigir la revelación de prueba (Propuesta de Directiva del Parlamento Europeo y del Consejo, 2022, art.3 apartado 1), establecer un catálogo de definiciones (Ley Act, 2021, art.3), crear órganos especializados en sistemas de IA (Ley Act, 2021, art.3), entre otras.

El problema de las soluciones propuestas radica en que éstas no tienen una retroalimentación adecuada con la realidad del diseño, funcionamiento, impacto y riesgos de los sistemas de IA, debido a que éstos interactúan constantemente con el entorno y cambian sus decisiones conforme a ello. Esto puede generar rigidez en los sujetos involucrados y las regulaciones para adaptarse a los cambios de los sistemas de IA.

Por lo anterior, este trabajo busca contribuir a los avances de las regulaciones mediante el énfasis y aportes que pueden hacer los instrumentos colaborativos, que son los sandboxes y las evaluaciones de impacto algorítmica. Ambos instrumentos tienen por objeto facilitar la comprensión del diseño, funcionamiento, impacto y riesgos de los sistemas de IA; sea mediante pruebas del sistema de IA en un entorno real controlado, sea mediante diversos tipos de evaluaciones sobre el algoritmo, el contexto en que se encuentra, su impacto tras haber sido implementado en la sociedad, etc. En consecuencia, una mejor comprensión reduce la asimetría de conocimiento, aumenta la certeza jurídica con un grado de flexibilidad razonable en las regulaciones y, con ello, permite una delimitación más precisa de los criterios de imputación y las herramientas utilizadas como prueba para acreditar la negligencia y/o debida diligencia, la relación de causalidad y el riesgo concreto de la actividad realizada-lo que permite distinguir qué riesgo está dentro de la actividad y cuál no-.

Por lo tanto, estos instrumentos contribuyen a una regulación adecuada de la responsabilidad civil extracontractual-tanto subjetiva como objetiva-, mediante una delimitación precisa de los criterios de imputación, un mayor acceso a la información-mediante una comprensión más sencilla para los sujetos involucrados-, el fomento de elaborar sistemas de IA seguros, y la flexibilidad normativa a través de la retroalimentación constante y actualizada del diseño, funcionamiento e impacto de los sistemas de IA. Esto es congruente con el objetivo de toda regulación en relación con los sistemas de IA: proteger a las personas y la innovación; permitir la posibilidad de una reparación integral mediante una indemnización de perjuicios, pero sin una carga excesiva de obligaciones a los demandados; no impedir el desarrollo de nuevos sistemas de IA, sino hacer una regulación inteligente, que sea flexible y equilibrada para los agentes involucrados en todo el ciclo de vida de los sistemas de IA.

VI. REFERENCIAS

Amunategüi, C. (2023). Responsabilidad civil extracontractual e inteligencia artificial. En C. Céspedes (director) et al, *responsabilidad civil extracontractual: instrumentos de derecho comparado, proyectos de reforma y derecho chileno. Doctrina, jurisprudencia y derecho comparado*, 513-525. Tirant lo Blanch.

Araya, C. (2020). Desafíos legales de la inteligencia artificial en Chile. *Revista Chilena de Derecho y Tecnología, 9*(2), 257-290. http://dx.doi.org/10.5354/0719-2584.2020.54489

Aznar, A., & Domingues, M. (25 de julio de 2022). La responsabilidad civil derivada del uso de inteligencia artificial. *Diario constitucional*. https://www.diarioconstitucional.cl/2022/07/25/la-responsabilidad-civil-derivada-del-uso-de-inteligencia-artificial-por-antonio-aznar-domingo-y-maria-patrizia-domingues-villarroel/

Banco de Desarrollo de América Latina y Organización para la Cooperación y Desarrollo Económicos, (14 de septiembre de 2022). *Uso estratégico y responsable de la inteligencia artificial en el sector público de América Latina y el Caribe, Estudios de la OCDE sobre Gobernanza Pública*. https://doi.org/10.1787/5b189cb4-es

Barrio, M. (2021). Towards legal regulation of artificial intelligence. *Revista IUS, 15*(48), 35-53. https://doi.org/10.35487/rius.v15i48.2021.661

Barros, E. (2006). *Tratado de responsabilidad extracontractual*. Editorial Jurídica de Chile.

Bini Mender, S. (2020). Sistemas biométricos y machine learning: sus desafíos. En H. Granero (director académico)., J. Veltani y R. Lozano (Coords) et al, *Inteligencia artificial, un reto social*, 165-186. Editorial Albremática S.A.

Brynjolfsson, E., Clark, J., Etchemendy, J., Fattorini, L., Ligett, K., Lyons, T., Manyika, J., Maslej, N., Niebles, J., Parli, V., Perrault, R., Reuel, A., Shoham, Y., & Wald, R. (2024). *The AI index 2024 Annual report AI*. Institute for Human-Centered AI, Stanford University, Stanford, CA. https://aiindex.stanford.edu/report/

Bustos, Y. (2022). Análisis sobre las medidas de apoyo legal a la experimentación en tecnologías innovadoras. *Revista española de Derecho Aeronáutico y Espacial*, (2), 319-346. https://aedae-aeroespacial.org/wp-content/uploads/2022/09/Revista-AEDAE_2022_digital-27-09.pdf

Calo, R. (2015). Robotics and the lessons of Cyberlaw. *California Law Review, 103*(3), 513-563. https://digitalcommons.law.uw.edu/faculty-articles/23/

Castello, C. (2020). Inteligencia artificial, marco legal argentino y su gobernanza. En H. Granero (director académico)., J. Veltani y R. Lozano (Coords) et al, *Inteligencia artificial, un reto social*, 355-366. Editorial Albremática S.A.

Čerka, P., Grigiene, J., & Sirbikytė, G. (2015). Liability for damages caused by artificial intelligence. *Computer Law & Security Review, 31*(3), 376-389. https://doi.org/10.1016/j.clsr.2015.03.008

Chui, M., Kamalnath, V., & McCarthy, B. (2018). An executive's guide to AI. *McKinsey & Company*. https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Analytics/Our%20Insights/An%20executives%20guide%20to%20AI/Executives-guide-to-AI

Comisión Europea, Dirección General de Justicia y Consumidores. (2019). *Responsabilidad por la inteligencia artificial y otras tecnologías digitales emergentes*. Oficina de Publicaciones de la Unión Europea. https://data.europa.eu/doi/10.2838/573689

Corral, H. (2003). *Lecciones de responsabilidad civil extracontractual*. Editorial Jurídica de Chile.

Dastin, J. (14 de octubre de 2018). Amazon abandona un proyecto de IA para la contratación por su sesgo sexista. *Discover Thomson Reuters*. https://www.reuters.com/article/amazon-com-contratacion-ia-idESKCN1MO0M4

Desde los aires y emulando el trabajo en equipo de las abejas: este es el drone que puede imprimir un edificio en 3D (28 de septiembre de 2022). *CNN Chile, sección Futuro 360*. https://www.futuro360.com/ciudades-del-manana/drone-imprimir-3d_20220928/?fbclid=IwAR1oIjD7_XRF0JHbXUaDGDOqxm6AfARYD7QvlYZVVtXN9Dekx82uKiWN1aA

Díaz, C., & Flórez, J. (23 de junio de 2022). *Imputación de daños causados por robots con inteligencia artificial. Conceptos aplicables de la responsabilidad civil y del Estado*. Centro de Estudios Regulatorios. https://www.cerlatam.com/publicaciones/imputacion-de-danos-causados-por-robots-con-inteligencia-artificial-vigencia-de-los-presupuestos-tradicionales-de-la-responsabilidad-civil-y-del-estado/

Digital Future Society. (febrero de 2024). *Hacia un uso responsable de los algoritmos: métodos y herramientas para su auditoría y evaluación*. https://digitalfuturesociety.com/es/report/towards_accountable_algorithms/

Fossaceca, C., & Moreyra, P. (2020). Reflexiones sobre la inteligencia artificial desde la perspectiva jurídica. En H. Granero (director académico)., J. Veltani y R. Lozano (Coords) et al, *Inteligencia Artificial, un reto social*, 343-354. Editorial Albremática S.A.

Gana, A. (31 de mayo de 2023). El futuro de los vehículos autónomos. Hacia una revolución en la industria automotriz. *Publimetro*. https://www.publimetro.cl/tacometro/2023/05/31/el-futuro-de-los-vehiculos-autonomos/

Gordo, J., Malvaso, A., Mazzarella, C., Saldivio, A., & Sangineto, C. (2019). *Accidentes producidos por vehículos autónomos*. Universidad Tecnológica Nacional, Facultad Regional Buenos Aires. https://grupogemis.com.ar/wp-content/uploads/2019/05/AdS_M_AccidentesVehiculosAutonomos.pdf

Graff, M., Llamas, J., & Mendoza, O. (2022). Enfoques regulatorios para la inteligencia artificial (IA). *Revista chilena de derecho y tecnología, 49*(3), 31-62. https://doi.org/10.7764/R.493.2

Grupo de Diarios de América. (12 de mayo de 2020). Facebook tuvo que "apagar" inteligencia artificial que desarrolló su idioma. *El Economista*. https://www.eleconomista.net/tendencias/Facebook-tuvo-que-apagar-inteligencia-artificial-que-desarrollo-su-idioma-20200512-0017.html

Guío, A. (2021). *Sandbox regulatorio de inteligencia artificial en Chile*. Ministerio de Economía, fomento y turismo del Gobierno de Chile, FAST (futuro y adopción social de la tecnología) y CAF (Banco de Desarrollo de América Latina). https://www.economia.gob.cl/wp-content/uploads/2021/09/PaperSandboxIA.pdf

Hernández, T. (19 de mayo de 2023). Así es la Inteligencia Artificial que podría diagnosticar ataques cardíacos. *El Tiempo*. https://www.eltiempo.com/tecnosfera/novedades-tecnologia/asi-es-la-inteligencia-artificial-que-podria-diagnosticar-ataques-cardiacos-770132

Herrera, D., & Vadillo, S. (2018). *Sandbox regulatorio en América Latina y el Caribe para el ecosistema FinTech y el sistema financiero*. Banco Interamericano de Desarrollo. https://publications.iadb.org/es/publicacion/17483/sandbox-regulatorio-en-america-latina-el-caribe-para-el-ecosistema-fintech-y-

el

López Ataz, J. (agosto de 2020). Daños causados por las cosas: una nueva visión a raíz de la robótica y de la inteligencia artificial. *Cátedra Jean Monnet de Derecho privado Europeo*, 1-58. http://hdl.handle.net/2445/169850

Mansilla, A., Mansilla, E., & Ortiz, J. (2020). ¿Es necesario regular los algoritmos o la inteligencia artificial? En H. Granero (director académico)., J. Veltani y R. Lozano (Coords) et al, *Inteligencia artificial, un reto social*, 329-342. Editorial Albremática S.A.

Microsoft. (mayo de 2023). *Governing AI: a blueprint for the future*. https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RW14Gtw

Morales, A. (2021). El impacto de la inteligencia artificial en el Derecho. *Advocatus*, (39), 39-71. https://doi.org/10.26439/advocatus2021.n39.5117

Moyano, R. (2020). Decisiones generadas por inteligencia artificial para la resolución de conflictos y Estado de Derecho (de lege ferenda). En H. Granero (director académico)., J. Veltani y R. Lozano (Coords) et al, *Inteligencia artificial, un reto social*, 77-88. Editorial Albremática S.A.

Neumann, S. (18 de marzo de 2020). Empresa chilena desarrolla solución digital para detectar y monitorear coronavirus. *Diario financiero*. https://www.df.cl/df-lab/transformacion-digital/empresa-chilena-desarrolla-solucion-digital-para-detectar-y-monitorear

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. (23 de noviembre de 2021). *Recomendación de la UNESCO sobre la ética de la inteligencia artificial*. https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

Organización para la Cooperación y Desarrollo Económicos. (21 de mayo de 2019). *Recomendación del Consejo sobre inteligencia artificial*. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449#mainText

Press, G. (27 de agosto de 2017). Artificial intelligence (AI) defined. *Forbes*. https://bit.ly/2ApDKM2

Propuesta de Directiva del Parlamento Europeo y del Consejo, relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial (Directiva sobre responsabilidad en materia de IA) (2022/0303 (COD)), del 28 de septiembre del año 2022. en: https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52022PC0496

Propuesta de reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de Inteligencia Artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión. 21 de abril del año 2021. https://eurlex.europa.eu/legal-content/ES/TXT/?uri=celex%3A52021PC0206

Real Decreto 817 de 2023. Que establece un entorno controlado de pruebas para el ensayo del cumplimiento de la propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial. 9 de noviembre de 2023. BOE. No. 268. https://www.boe.es/eli/es/rd/2023/11/08/817

Reglamento 12 de 2023 [Administración del Ciberespacio de China Ministerio de Industria y Tecnología de la Información Ministerio de Seguridad Pública]. Sobre la gestión de síntesis en profundidad de los servicios de información de internet. 10 de enero de 2023. https://www.gov.cn/zhengce/zhengceku/2022-12/12/content_5731431.htm

Resolución 78/265 aprobada por la Asamblea General de la Organización de las Naciones Unidas, sobre aprovechar las oportunidades de sistemas seguros y fiables de inteligencia artificial para el desarrollo sostenible. 21 de marzo de 2024. https://digitallibrary.un.org/record/4043244/files/A_RES_78_265-ES.pdf?ln=en

Resolución del Parlamento Europeo con recomendaciones destinadas a la Comisión sobre un régimen de responsabilidad civil en materia de inteligencia artificial. 20 de octubre de 2020. https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52020IP0276

Tapia, A. (2021). La responsabilidad civil derivada del uso de la inteligencia artificial y su aseguramiento. *Revista Ibero-Latinoamericana de seguros, 30*(54), 107-146. https://doi.org/10.11144/Javeriana.ris54.rcdu

The White House. (octubre de 2022). *blueprint for an ai bill of rights making automated systems work for the american people*. https://www.whitehouse.gov/ostp/ai-bill-of-rights

The White House. (30 de octubre de 2023). *FACT SHEET: President Biden issues executive order on safe, secure, and trustworthy artificial intelligence*. https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/

Vega, G. (28 de mayo de 2021). La ONU informa del primer ataque de drones autónomos a personas. *El País*. https://elpais.com/tecnologia/2021-05-28/la-onu-informa-del-primer-ataque-de-drones-autonomos-a-personas.html