UNIVERSIDAD DE CHILE

Facultad de Derecho

Departamento de Derecho Privado

INTELIGENCIA ARTIFICIAL Y RESPONSABILIDAD EXTRACONTRACTUAL: DESAFÍOS FRENTE A LA NORMATIVA CHILENA ACTUAL.

VALERIA SCHNAKE MUÑOZ

Memoria de prueba para optar al grado de Licenciada en Ciencias Jurídicas y Sociales

Profesora guía: Dra. MARÍA MAGDALENA BUSTOS DÍAZ

Santiago

2024

"The original question, "Can machines think?" I believe to be too meaningless to deserve discussion. Nevertheless, I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted.[1]".

(Alan Turing, 1950).

[1]"La pregunta original: "¿Pueden pensar las máquinas?" Creo que carece de sentido como para merecer discusión. Dicho esto, creo que a finales del siglo el uso de las palabras y la opinión educada general se habrán alterado tanto que uno podrá hablar de máquinas pensantes sin esperar ser contradicho."

AGRADECIMIENTOS

A mi familia: mis padres, mis abuelos y mi hermana, por su compañía infaltable.

A mis amigos: especialmente a J. y L. por su perspectiva desde las ciencias de la computación.

A mi profesora guía: por su infinita paciencia, estímulo intelectual y rigor académico.

A J. C., por todo.

ÍNDICE

RESUMEN

La presente memoria de prueba presenta y analiza los desafíos que surgen a partir de los avances actuales que involucran sistemas de inteligencia artificial, particularmente respecto de la responsabilidad extracontractual en el orden jurídico chileno. Para ello, se analiza (i) qué son la inteligencia artificial y los sistemas de inteligencia artificial, (ii) cuáles son las dificultades que presentan dichos sistemas en general, en comparación con el resto de las tecnologías de la información y (iii) cómo dichos sistemas pueden presentar desafíos novedosos a la responsabilidad civil, a través de un análisis pormenorizado de la incidencia de dichos desafíos en los distintos elementos de la responsabilidad extracontractual.

ABSTRACT

This thesis presents and analyzes the challenges arising from current developments involving artificial intelligence systems, particularly with respect to tort liability in the Chilean legal system. To this end, it analyzes (i) what are artificial intelligence and artificial intelligence systems, (ii) what are the difficulties presented by such systems in general, in comparison with the rest of the information technologies and (iii) how such systems may present novel challenges to civil liability, through a detailed analysis of the impact of such challenges on the different elements of tort liability.

INTRODUCCIÓN

Ad portas de lo que algunos autores han calificado como 'la cuarta revolución industrial[2]', la inteligencia artificial (en adelante "IA") se ha convertido en la tecnología protagonista del avance tecnológico de las últimas tres décadas, expandiendo el horizonte de las tareas que creíamos podían realizar las máquinas y el software. A través de algoritmos y modelos como el aprendizaje automatizado (en inglés *Machine Learning*, en lo sucesivo "ML"), los sistemas de IA tienen la capacidad de aprender a partir de la información que reciben de su entorno[3], siendo adaptables, y en algunos casos, autónomos[4]. El emergente campo de investigación acerca de la IA concibe dichos sistemas como agentes inteligentes[5] no humanos capaces de procesar el lenguaje natural[6], aprender, percibir e inferir conocimiento[7].

Este tipo de tecnologías ya poseen una amplia aplicación en nuestra vida diaria: los ejemplos más cercanos van desde el asistente virtual presente en la mayoría de nuestros teléfonos inteligentes, los algoritmos que nos recomiendan contenido en redes sociales o en servicios de *streaming*, hasta la tecnología que hace posible que ciertos vehículos puedan ser piloteados de forma autónoma. La IA también está comenzando a ser implementada en contextos médicos, financieros[8], e incluso en la industria legal[9], existiendo sistemas de IA capaces de analizar contratos[10] o de realizar otras tareas legales normalmente consideradas como repetitivas.

Sin perjuicio de su ubicua posición en el mundo del siglo XXI, la IA puede parecernos una novedad confusa y difícil de entender en su aspecto técnico o en sus características esenciales. Los sistemas de IA, si bien son técnicamente complejos y presentan, a su vez, desafíos únicos en su aplicación,

---

[2] ROSS y MAYNARD (2021): 159-161.
[3] DOUGLAS WILL. (2021). MIT Technology Review. En: https://www.technologyreview.com/2021/05/27/1025453/artificial-intelligence-learning-create-itself-agi/.
[4] Véase: https://www.gov.uk/government/news/us-and-uk-research-labs-collaborate-on-autonomy-and-ai.
[5] Debe precisarse que lo que entendemos por *agencia* de la IA (o IA como "agente racional") se relaciona a aspectos técnicos que veremos con detalle más adelante.
[6] BROOKS. (2020). University of York. "The Role of natural language processing in AI". En: https://online.york.ac.uk/the-role-of-natural-language-processing-in-ai/
[7] IBM. Think. (2024) "What is artificial intelligence?". En: https://www.ibm.com/topics/artificial-intelligence.
[8] OCDE. (2021) "Artificial Intelligence, Machine Learning and Big Finance". En: https://www.oecd.org/finance/artificial-intelligence-machine-learning-big-data-in-finance.htm.
[9] KABIRI. M. Forbes. (2021) "How AI is being used in the legal industry". En: https://www.forbes.com/sites/forbesbusinesscouncil/2021/01/19/how-ai-is-being-used-in-the-legal-industry/?sh=2e3d953c50c6.
[10] ASLAN E. (2019) ABA Technology Review. "AI boosts contract analysis". En: https://www.americanbar.org/news/abanews/publications/youraba/2019/april-2019/ai-helps-take-automated-contract-analysis-to-the-next-level/.

han traído consigo importantes cambios científicos y tecnológicos que, incluso, han hecho posible resolver problemas que hasta el día de hoy no habían sido solucionados por el intelecto humano[11]. Es este tipo de tecnología que, según algunos, nos llevará a poner los pies en el planeta Marte, y que, según otros, podría llevarnos a la extinción[12], o a una "condición humana superior[13]", temas que dada su novedad parecieran hablarnos de un futuro propio de una novela de ciencia ficción.

Más allá de lo especulativo, ya es posible hablar y prever problemas concretos y actuales relacionados con la IA y, en particular, el derecho: desde la pérdida potencial de trabajos debido a la automatización de labores intelectuales[14], hasta la posible afectación de derechos fundamentales por aplicaciones de IA (tales como el derecho a la privacidad, los relacionados con el tratamiento de datos personales, el derecho a la vida, la integridad física, la libertad de movimiento, entre otros[15]), además de los daños que podrían ser causados por aplicaciones o *robots* de IA autónomos en el ámbito civil-patrimonial[16], en el ámbito penal[17] o en conflictos bélicos[18].

La novedad de este tipo de tecnología, su complejidad, el rápido avance en la investigación relativo a la misma que ha surgido en la última década, la posibilidad de que existan aplicaciones o usos completamente autónomos, así como la gran cantidad de actores involucrados en el ciclo de vida de la IA (entre ellos operadores, programadores, consumidores, usuarios, distribuidores, etc.); hacen necesario plantear la siguiente gran pregunta: ¿Cómo responde el derecho civil frente a posibles daños y responsabilidades que surjan a partir del uso de sistemas de IA? A partir de dicha

---

[11] LU (2019). New Scientist. "AI is helping tackle one of the biggest unsolved problems in maths". En: https://www.newscientist.com/article/2226493-ai-is-helping-tackle-one-of-the-biggest-unsolved-problems-in-maths/ y J.H. HEULE M. (2017) Communications of the ACM. "The Science of Brute Force". En: https://cacm.acm.org/magazines/2017/8/219606-the-science-of-brute-force/fulltext

[12] V. gr. las opiniones expresadas por científicos como Stephen Hawking, existiendo incluso un campo de estudio llamado 'X-Risk' al cual nos referiremos más adelante.

[13] V. gr. las ideas relacionadas con el "transhumanismo", que pretenden establecer una conexión entre humano y máquina, o aumentar la capacidad cognitiva del ser humano a través de aplicaciones tecnológicas.

[14] OCDE (2018).

[15] El 2021, la alta Comisionada de Derechos Humanos de la Unión Europea emitió un comunicado de prensa al respecto, destacando un reporte elaborado por la Oficina de Derecho Humanos de la UE referido al tema, el cual puede encontrarse en el siguiente enlace: https://www.ohchr.org/sites/default/files/HRBodies/HRC/RegularSessions/Session48/Documents/A_HRC_48_31_AdvanceEditedVersion.docx.

[16] Un ejemplo son los accidentes causados por vehículos autónomos, véase SCHELLEKENS (2015) y JUHÁSZ-REKA (2016).

[17] Se ha presentado la pregunta de si un sistema de IA puede ser criminalmente responsable: véase LIMA (2017) y HALLEVY (2013). Dicho esto, también existen una multiplicidad de cuestiones relacionadas con el uso de la IA en el proceso penal: a modo de ejemplo, en el estado de Florida (EE. UU.) se utilizan algoritmos para ayudar a calcular el monto de la fianza necesaria para otorgar la libertad condicional entre otros. Para esto último, véase KLEIBERG (2018) y ZAVRSNIK (2020).

[18] SAUER (2022): 237.

interrogante, surgen forzosamente muchas otras: ¿Cómo y a quién debemos atribuir responsabilidad en el caso de que un sistema de IA cause daños? ¿Cuál es el deber de cuidado que debemos asignar a los operadores de estas tecnologías? ¿Puede nuestro sistema general de responsabilidad civil responder de forma *satisfactoria* a dicha contingencia?

Frente a este escenario de cambio, el poder legislativo, en cumplimiento de su deber, cuenta con la tarea de decidir qué tipo de regulación deberá establecerse respecto de los sistemas de IA y de la responsabilidad civil ocasionada por ellos o por su uso. En ausencia de esta intervención, y debido a la novedad del fenómeno mismo, el jurista (y el juez) deben preguntarse de qué forma las normas comunes y supletorias de nuestro régimen jurídico responden a los problemas prácticos que podrían surgir a raíz de la novedad que esta tecnología presenta. Yendo más allá, debe también el jurista preguntarse si aquellas normas ya existentes *pueden* o *son suficientes para* resolver dichos problemas, teniendo en cuenta los fines de justicia propios de la regulación relativa al derecho de daños. Dada esta situación y dentro del marco general del derecho patrimonial, el caso específico que se pretende analizar en la presente memoria es el de la aplicación del régimen general del derecho de daños o responsabilidad civil extracontractual frente a casos en que los daños hayan sido causados por el uso de sistemas de IA.

Es por estos factores (novedad, falta de una legislación en nuestro país, proximidad y diversidad de aplicaciones de la IA en nuestra vida diaria) que la presente memoria pretende responder las siguientes grandes interrogantes: ¿*de qué forma* aplica el estatuto general de responsabilidad civil (o derecho de daños) frente a hipótesis de daños causados por el uso de sistemas de inteligencia artificial? ¿Resulta *satisfactoria* dicha aplicación?

Para abordar las interrogantes planteadas, la presente memoria propone que la aplicación de las reglas generales de la responsabilidad civil extracontractual contenidas en los artículos 2314 y siguientes del Código Civil resultan, en general, satisfactorias a la hora de imputar responsabilidad frente a *determinadas* hipótesis de daños causados por el uso de sistemas de IA. Sin embargo, debido a las características particulares de este tipo de tecnologías, la aplicación del estatuto general de responsabilidad trae consigo ciertos problemas, especialmente relacionados con la determinación de la diligencia exigible y la prueba de la causalidad. Frente a ello, resulta necesario (ante la falta de una ley que regule los daños causados por sistemas de IA) recurrir a los criterios doctrinarios más aceptados en materia de responsabilidad civil y a las normativas supranacionales que orientan y disponen los principios éticos que han de seguirse respecto al desarrollo y la utilización de sistemas

de IA (como lo son los principios relativos a la IA de la OCDE o las directrices sobre ética aplicable a los sistemas de IA de la Unesco).

Con el fin de responder las preguntas de investigación aquí planteadas, y desarrollar la propuesta previamente señalada, la memoria se dividirá en tres capítulos. El Capítulo 1 presentará un breve resumen de la historia del estudio de la inteligencia artificial, así como diversos acercamientos a la delimitación de su concepto. Particularmente, se propondrá desde las ciencias de la computación el concepto que identifica a la inteligencia artificial como un "agente racional", destacando que existe una distinción relevante entre la IA como un campo de estudio en y su aplicación práctica y concreta en sistemas de IA. Asimismo, se expondrán los nacientes acercamientos normativos referidos a la materia, que demuestran la relevancia que tanto en el ámbito internacional como nacional se le ha dado al tema. Lo anterior, con el propósito de desmitificar el concepto y funcionamiento de los sistemas de IA.

En el Capítulo 2, se identificarán y expondrán los principales desafíos vinculados con las características intrínsecas de los sistemas de IA. Para ello, nos centraremos en aquellos desafíos que, desde la óptica del presente trabajo, resultan relevantes para la responsabilidad civil y que darán lugar a las problemáticas que luego se examinarán en el Capítulo 3. Los desafíos que particularmente se analizarán son: (i) la complejidad técnica y la existencia de una pluralidad de actores, (ii) la opacidad algorítmica y el problema de la caja negra, (iii) el sesgo algorítmico, (iv) la autonomía y, finalmente, (v) la imprevisibilidad.

En el Capítulo 3 se expondrá someramente la regulación relativa a la responsabilidad civil, tanto contractual como extracontractual, en el derecho chileno. Se describirán los elementos que, de acuerdo con la doctrina nacional y el Código Civil, se constituyen como requisitos necesarios para llevar a cabo exitosamente un juicio de responsabilidad civil extracontractual. Respecto de cada uno de estos elementos (capacidad, culpa o dolo, daño y relación de causalidad), se identificarán y analizarán una serie de desafíos relacionados con el uso de sistemas de IA, los cuales se ejemplificarán a través de casos hipotéticos. A través del análisis de dichas problemáticas, así como de la aplicación del derecho a los casos de ejemplo, se justificará que la aplicación de las normas que componen el sistema general de responsabilidad civil resulta factible, pero que, como ya se señaló, presenta dificultades relevantes tanto respecto del elemento de la culpa, así como respecto del nexo causal; resultando necesario recurrir en complemento a las normativas supranacionales que orientan

y disponen los principios éticos que han de seguirse respecto al desarrollo y la utilización de sistemas de IA.

Para finalizar, se presentarán conclusiones acerca de lo expuesto a lo largo del trabajo. Estas conclusiones se enfocarán en sintetizar las ideas principales acerca de la forma en que el sistema de responsabilidad resulta aplicable frente a daños causados por IA, y señalar la posición que esta memoria tiene respecto de aquellos casos en que dicha aplicación no resulte satisfactoria.

CAPÍTULO 1: ¿QUÉ ES LA INTELIGENCIA ARTIFICIAL? UNA APROXIMACIÓN A SU CONCEPTO Y FUNCIONAMIENTO

Antes de entrar a analizar la aplicación del estatuto jurídico de la responsabilidad civil frente a la IA, lo primero que debemos hacer es definir el concepto de sistemas de IA. ¿Qué es la IA y por qué presenta desafíos tan diferentes a los que la tecnología ha presentado hasta ahora? ¿Cómo funcionan los sistemas de IA[19]?

No resulta extraño pensar que la tecnología, conforme a su avance, ha presentado desafíos impensados para la humanidad (y en consecuencia, para el derecho), y no es necesario utilizar un ejemplo exótico, solamente basta pensar en las innumerables repercusiones sociales y normativas que trajeron consigo grandes hitos tecnológicos como la Primera Revolución Industrial o la invención de la electricidad.

La idea de la IA, tal como la conocemos, existe desde los años cincuenta[20]; sin embargo, solo ha sido dentro de la última década que se ha producido un creciente y acelerado interés económico, político y legal en la materia. Esto se debe, principalmente, al paso desde sistemas de IA de primera generación a sistemas de IA de segunda generación, la cual es capaz de procesar grandes cantidades de información y fabricar de forma autónoma algoritmos refinados basados en lo que se conoce como ML*,* avance que ha llevado a que la IA se convierta en una gran industria de más de $66.8 mil millones de dólares[21].

Puede parecer que el concepto de la IA es amplio, elusivo y complejo, por lo que es necesario delimitarlo y desmitificarlo. Con este fin, repasaremos su origen histórico como campo de estudio, para luego entrar a explicar qué es la IA como concepto técnico y por qué los sistemas de IA, al usarse como *software* o en determinados tipos de tecnologías (como *robots*/*hardware*) presentan desafíos complejos y novedosos para el Derecho Civil y particularmente para el régimen de responsabilidad civil extracontractual, que requieren un estudio y atención necesarios.

[19] La utilización de la expresión "sistemas de IA" versus IA implica una relación de género a especie, es decir, hablamos de un sistema de IA particular vs. la IA como concepto en su generalidad.
[20] MOOR JAMES (2006): 87.
[21] US COMMMERCIAL SERVICE (2021): 6. En: https://www.trade.gov/sites/default/files/2022-05/Top%20Global%20AI%20Markets%20Report%204.20%20%282%29%20%281%29.pdf.

Este panorama nos ayudará a comprender los posibles escenarios que pueden darse en el derecho de responsabilidad civil extracontractual en relación con los sistemas de IA, lo que contribuirá a facilitar la identificación de problemas que surjan derivado de las interrogantes aquí planteadas.

1.1 La IA como campo de estudio o disciplina.

El estudio de la IA es relativamente reciente. Muchos, al intentar fijar un momento de "inicio" de una disciplina enfocada precisamente a ello, apuntan al año 1956: fecha de una conferencia organizada por una variedad de intelectuales en la Universidad de Dartmouth[22] y donde por primera vez se acuñó el término "IA"[23]. Desde entonces, la historia de la disciplina se ha visto marcada por épocas de entusiasmo, decepción y resurgimiento[24]. A pesar de ello, la última década ha significado un crecimiento gigantesco para el estudio y la industria de la IA, con hitos tales como la fundación de *DeepMind, OpenAI*, la creación del autopiloto de Tesla, y escándalos como el de Cambridge Analytica[25], pero también un amplio impacto en múltiples industrias, tales como la medicina, la educación y el ámbito financiero[26].

En la actualidad, la IA como área de estudio abarca una gran variedad de subcampos, que van desde áreas de propósito general como el estudio, aprendizaje y la percepción mediante IA, hasta otras más específicas como la aplicación de sistemas de IA en juegos contra humanos como el ajedrez, la demostración de teoremas matemáticos, la escritura de poesía o texto que se asemeje lo más posible a un lenguaje natural humano, así como investigación orientada al campo médico (diagnóstico de enfermedades o cirugías llevadas a cabo por *robots)*. Es por esto que algunos expertos han llegado a catalogar el estudio de la IA como un campo "genuinamente universal[27]".

1.2 La Inteligencia Artificial como concepto: el enfoque del agente racional.

Por su parte, definir la IA como concepto ha sido siempre una tarea difícil, lo que para algunos ha sido algo beneficioso para su desarrollo[28]. Un intento de dar una definición concreta de IA la describe como:

---

[22] ABELIUK (2021): 1-2.
[23] *Ibid.*
[24] Para ver más sobre la historia, CORDESCHI (2017).
[25] ISAAK y HANNA (2018): 57.
[26] UNESCO (2021).
[27] RUSSELL y NORVIG (2014): 1.
[28] WANG (2019): 1-2.

> "(…) aquella actividad dedicada a hacer a las máquinas inteligentes, siendo la inteligencia la cualidad que permite a una entidad funcionar de manera apropiada y con previsión en su ambiente.[29]"

RUSSEL y NORVIG, por su parte, señalan que al intentar definir el objeto de estudio de la IA la opinión de los expertos se ha dividido en dos: por una parte, aquellos que intentan definir el campo como el estudio de aquellos sistemas que "piensan" o "actúan" como humanos, es decir, en torno a la cognición e inteligencia humanas[30], y por otra parte, quienes estudian estas tecnologías como "sistemas inteligentes" o "sistemas racionales", poniendo énfasis en la idea de "racionalidad"[31] sin hacer referencia a un estándar de imitación a la conducta o inteligencia humanas. Bajo este último entendimiento, un "sistema racional" es tal en referencia a un concepto "ideal" de racionalidad que, en términos de los autores, se traduce en hacer lo "correcto[32]" en función del conocimiento que el agente o sistema tiene a su disposición en un momento dado[33]. Estos dos enfoques son los que han dominado la percepción de los sistemas inteligentes y la IA a través de su historia, evolución y desarrollo en general. El primer enfoque presenta una serie de problemas, dado que definir la inteligencia desde el *pensar* como humano resulta filosófica y técnicamente complejo[34]. El segundo enfoque, en cambio, es el que predomina en el estudio técnico de la IA. Esta última perspectiva es la que seguirá esta memoria, esto es, aquella que entiende el estudio de la IA como el estudio de agentes inteligentes o agentes racionales. Siguiendo esta línea de pensamiento, para destacar la diferencia entre el estudio de los programas computacionales o tecnologías de la información comunes y la IA, nos referiremos brevemente a la definición que RUSSEL y NORVIG presentan al enfoque del agente racional[35]".

---

[29] NILSSON (2010). Traducción propia.

[30] Este acercamiento, si bien se toma en cuenta a la hora de desarrollar aplicaciones que, por ejemplo, interactúan con humanos (como es el caso de los procesadores de texto naturales o chatbots), no deja de ser problemático dadas las diversas definiciones y entendimiento en torno a lo que constituye la "inteligencia" humana, la cual no solo se puede medir desde parámetros o medidas como "eficiencia" o "racionalidad".

[31] RUSSEL y NORVIG (2014). Señalan a lo largo de su texto que, en total, se establecerían cuatro categorías de entendimiento/definición de la IA: (i) sistemas que piensan como humanos (enfoque cognitivo), (ii) sistemas que actúan como humanos, (iii) sistemas que piensan racionalmente y, (iv) sistemas que actúan racionalmente.

[32] Lo correcto, en estos términos, suele entenderse como la opción más racional en términos de eficiencia. Debemos entender que, usualmente, nos estamos refiriendo a modelos matemáticos.

[33] RUSSELL y NORVIG (2014): 2.

[34] Esto dado que respecto de la cognición humana existen, como dictan las máximas de la experiencia, muchísimas variables, experiencias, y diferencias. Eso sí, debe destacarse que el comportamiento similar a los seres humanos sigue siendo un parámetro importante respecto a las aplicaciones de IA que interactúan con seres humanos: por ejemplo, las aplicaciones de *chatbot* simulan ser personas reales. El no poder distinguir una máquina de un computador es un componente esencial de lo que se conoce como "la prueba de Turing."

[35] RUSSELL y NORVIG (2014): 5

> "se espera que estos últimos agentes informáticos tengan otros atributos que los programas convencionales, como que estén dotados de controles autónomos que perciban su entorno, que persistan durante un periodo de tiempo prolongado, que se adapten a los cambios y que sean capaces de alcanzar objetivos diferentes.[36]"

Así, según ellos: "un agente racional es aquel que actúa con la intención de alcanzar el mejor resultado posible o, cuando hay incertidumbre, el mejor resultado esperado o posible dadas sus circunstancias[37]." Este concepto de "agente racional" es el concepto que debemos relacionar al pensar en los conceptos de "inteligencia" y "máquina inteligente" a lo largo de esta memoria[38]. Un agente "no racional", como una calculadora, por ejemplo, únicamente se limita a efectuar un cálculo simple luego de que introducimos la suma "1+1". Difícilmente consideraríamos que está desarrollando una operación consistente con el enfoque de un agente racional si la pantalla nos revela el resultado: "2". Contrastemos dicho ejemplo con el de una aspiradora robot[39]. Si durante su funcionamiento la aspiradora robot se ve atascada frente a un mueble, su programa (basado en IA) buscará la forma más eficiente para volver a circular. Así, su programa buscará la solución más eficiente para dichas circunstancias, la que probablemente esté en relación con una función similar a "retroceder" o "moverse a la derecha" dependiendo de lo que le indiquen sus sensores de proximidad.

Puede parecer que detenernos en la idea de inteligencia dentro de la "IA" es algo baladí, sin embargo, nos resultará de utilidad para comprender cómo la IA presenta desafíos únicos dentro de la responsabilidad civil.

1.3. ¿Cómo funciona la IA?

Si bien hemos introducido el concepto de forma simple, históricamente se han definido "generaciones"[40] en el desarrollo de la IA, de acuerdo con su progreso[41]: (i) la "IA simbólica" o de

---

[36] RUSSELL y NORVIG (2014): 5

[37] *Ibid.*

[38] Como se verá más adelante, dejaremos de lado las posibles aplicaciones de la responsabilidad sobre máquinas e Inteligencia Artificial autoconsciente, lo que, en nuestra opinión, podría quedar relegado a materias relativas al Derecho Constitucional, en caso de que la agencia de la IA sea completamente autónoma y autorreflexiva, al modo que los teóricos plantean (así lo han especulado KURZWEIL y otros). Dichos debates y su posibilidad escapan de lo comprendido en este trabajo.

[39] Este ejemplo es utilizado también por RUSSELL y NORVIG, pero es alterado y parafraseado aquí para facilitar su entendimiento.

[40] BOUCHER (2020): 2.

[41] Debe tenerse presente que, si bien se establece una idea de uno como más moderno que el otro, no quiere decir que la IA de primera generación esté obsoleta.

"sistemas expertos", (ii) la "IA de redes neuronales" y *Machine Learning* y (iii) una etapa eventual o futura, englobada en el concepto de "IA General"[42]. Los ejemplos y problemáticas presentados en esta memoria se referirán solo a sistemas calificables bajo las primeras dos generaciones de IA, puesto a que son las únicas con las que convivimos actualmente.

(i) Primera generación: IA Simbólica o de sistemas expertos.

La primera generación de IA se caracteriza por ser reactiva, es decir, responder al *input data* basándose en un programa o modelo[43] previo o preestablecido. El enfoque de los desarrolladores en esta época se centró en crear máquinas "inteligentes" capaces de codificar la experiencia de "expertos" en un conjunto de reglas que podrían ser ejecutadas por una máquina. Se denomina simbólica porque hace uso de lógica simbólica (por ejemplo, si X=Y, e Y=Z, entonces X=Z) para representar y resolver problemas[44]. Frente a este tipo de sistemas de IA, es fácil para los seres humanos comprender cómo dicho sistema llega a sus decisiones, debido a que el esquema de razonamiento nos resulta inteligible[45]. A pesar de desarrollar una amplia gama de tareas en forma automática, este tipo de IA solo puede hacerlo a partir de datos preestablecidos, y únicamente pueden mejorar o incorporar información adicional a través de intervención humana directa[46]. De todas formas, y a pesar de estas limitaciones, este tipo de sistemas de IA no están en absoluto obsoletos, y siguen siendo útiles en áreas que involucran decisiones repetitivas, como en el control de determinadas maquinarias[47].

(ii) Segunda generación: IA de redes neuronales y Machine Learning.

La segunda generación se caracteriza por utilizar *Machine Learning* (de aquí en adelante también "ML"), un conjunto de técnicas que automatizan el proceso de aprendizaje de los algoritmos[48]. ML se refiere al proceso a través del cual la IA, mediante algoritmos, "aprende"[49] sin necesidad de que exista intervención humana directa y constante, a diferencia de la IA de la primera generación**.** En este contexto, el sistema de IA observa datos determinados, crea un modelo basado en dichos datos,

[42] BOUCHER (2020): 2.
[43] GEORGEFF y LANSKY en GLAUBITZ (2021): 7.
[44] BOUCHER (2020): 2.
[45] *Ibid.*: 3.
[46] *Ibid.*
[47] *Ibid.*
[48] *Ibid.*: 2.
[49] RUSSELL y NORVIG (2020): 1202. Traducción propia.

y usa tal modelo como una hipótesis sobre el mundo para resolver problemas[50]. De esta forma, estos sistemas emulan la inteligencia humana al aprender del entorno que las rodea[51]. La IA de segunda generación:

> "reacciona al *input data* al aplicar experiencias pasadas a contextos nuevos (…) recordando de forma activa escenarios pasados o conjuntos de estímulos relacionados con la determinación que se le pida al algoritmo que haga, así, el algoritmo hace un contrapeso de sus opciones y basado en su memoria, toma una decisión en un ambiente extraño[52]".

El ML se ha vuelto un estándar en la ingeniería del *software*[53], y es una de las tecnologías que ha hecho posible la acrecentada evolución de los sistemas de IA que utilizamos en nuestro día a día (los vehículos autónomos, por ejemplo, se caracterizan por utilizar algoritmos de este tipo). Revisaremos algunos ejemplos de ML:

> Ejemplo 1: Un algoritmo determina el precio de un bien inmueble (por ejemplo, la casa Y) tras haber sido entrenado evaluando las características de millones de casas y sus precios. Al hacer esta evaluación, considerará variables tales como: número de habitaciones, ubicación en la ciudad, comuna, estado de reparación, número de baños, etcétera.
>
> Ejemplo 2: un *chatbot* responde preguntas humanas de forma espontánea y en lenguaje natural luego de haber sido entrenada con toda los datos presentes en Internet.

Como dijimos y como se ve en los ejemplos, la principal característica de este tipo de algoritmos es que son entrenados con modelos que contienen grandes cantidades de datos. Este proceso de entrenamiento implica el procesamiento de millones de datos que "alimentan" al algoritmo. Dichos datos son utilizados por el sistema para generar una soluciones de forma autónoma frente a escenarios novedosos no contemplados en los datos o modelos originales. Como vimos en el ejemplo 1, el sistema pudo calcular el precio de la casa Y efectuado, a pesar de que la casa Y jamás entró en el conjunto de datos original.

La utilización de este tipo de algoritmos ha llevado al desarrollo de tecnologías muy refinadas capaces de procesar, comparar y "aprender" enormes cantidades de información en poco tiempo,

---

[50] RUSSELL y NORVIG (2020): 1202. Traducción propia.
[51] EL NAQA y MURPHY (2015): 2
[52] GLAUBITZ (2021): 7.
[53] RUSSELL y NORVIG (2020): 1203. Traducción propia.

siendo un ejemplo paradigmático de este tipo de tecnología las redes neuronales: sistemas complejos basados en la estructura de las neuronas[54], el aprendizaje profundo o *Deep Learning,* el uso de inteligencia de datos o *Big Data*, entre otros; lo que han traído consigo importantes avances en el área médica (específicamente, oncológica)[55], el análisis financiero, las ciencias biológicas y farmacéuticas, entre otras.

(iii) Tercera generación o "Inteligencia Artificial General".

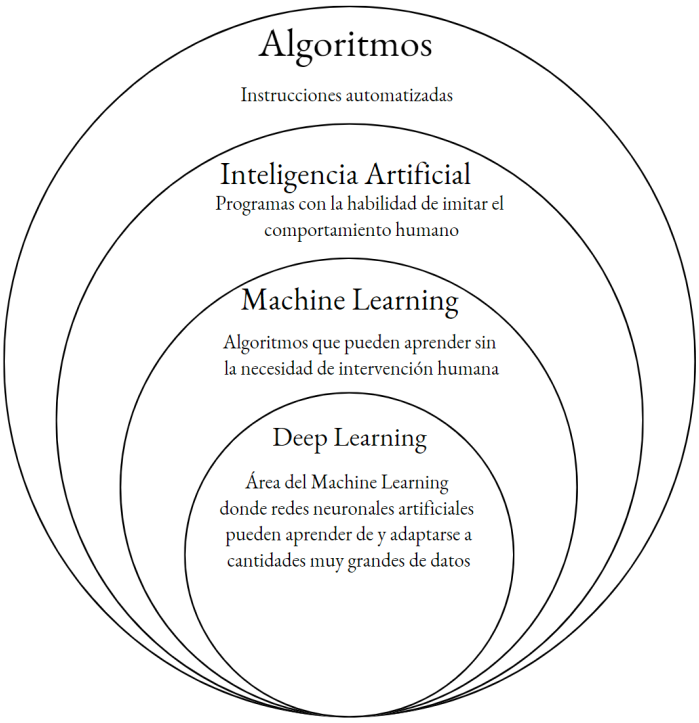
La tercera generación de la IA, propuesta por algunos teóricos, vendrá con un desarrollo de una teoría de la mente propia por parte de la IA, lo que se ha entendido por algunos como una IA con conciencia o autopercepción[56]. Legalmente, algunos han dicho que los sistemas de IA tendrán capacidad legal de la misma forma que una persona, por lo que podrán entrar en relaciones contractuales, ser demandados, etcétera[57]. Otros vaticinan escenarios apocalípticos en los que la IA alcanzará una capacidad intelectual tan grande que superará con creces a la humana, a tal punto que nos será imposible de comprender. Esto último sería la clave de nuestra evolución o extinción como especie, llegando a lo que algunos han llamado "la singularidad tecnológica"[58]. Otras perspectivas centran el análisis del riesgo existencial de la IA en la posible afectación que esta puede significar a los derechos humanos. Los sistemas de IA podría incidir en la estructura general de la sociedad, así como en las relaciones de poder entre Estados y sus ciudadanos[59]. Algunos científicos y filósofos más escépticos consideran que esto jamás será posible, aunque no han sido pocos los que han mostrado preocupación respecto al riesgo existencial de la IA, existiendo incluso un campo de estudio emergente al respecto[60].

1.4 Los algoritmos y el ciclo de vida de las IA.

Hasta ahora, nos hemos dedicado a describir, de forma general, lo que se entiende por IA, pero creemos pertinente detenernos en una pieza importante en el rompecabezas técnico de este fenómeno antes de continuar: los algoritmos. A lo largo de esta memoria y en muchos textos

[54] BISHOP (1994): 1804.
[55] Un ejemplo son los algoritmos diseñados para discernir sobre la malignidad de un tumor a partir del análisis de imágenes de diagnóstico, tales como rayos X y tomografías.
[56] Para definiciones relativas a la conciencia e IA, ver MCDERMOTT (2007).
[57] UMAR et al. (2023): 117.
[58] Es necesario señalar que, si bien es un término utilizado ampliamente en los medios de comunicación, no existe una definición técnica de lo que la "singularidad" realmente podría ser. Véase WANG, LIU y DOUGHERTY: 3-4.
[59] BUCKNALL, DORI-HACOHEN (2022): 121-122.
[60] *Ibid.*: 120.

referidos a la IA y los sistemas de IA, pareciera que "algoritmo" e "IA" a veces son términos aparentemente intercambiables, pero este no es el caso. Según DIAKOPOULOS: "un algoritmo es una serie de pasos que puede tomarse para resolver un problema particular u obtener un resultado definido. Esto puede llevarse a cabo por las personas, por la naturaleza, o por las máquinas[61]." De esto podemos concluir que un algoritmo es, en términos simples, un proceso de transformación de la información. Es importante dejar en claro, empero, que el algoritmo es la "herramienta" con la que se pueden llevar a cabo y resolver los problemas a través de IA y sistemas de IA. La relación entre IA y los algoritmos puede entenderse, por tanto, como una relación de género a especie. Nos sirve pensar, en este punto, en el concepto de agente racional, dado que en un solo sistema de IA, es decir, en una sola aplicación, máquina, código, API o página web que utilice IA, pueden operar una cantidad gigantesca de algoritmos simultáneamente. La relación se esclarece más con el siguiente esquema:



[62]

Por su parte, la expresión "ciclo de vida de la IA", utilizada tanto en la esfera técnica como normativa (por ejemplo, en los principios éticos de la OCDE), se refiere a las distintas etapas que,

[61] DIAKOPOULOUS (2014): 3.
[62] VRANA (2020). Traducción propia.

por regla general, se ven involucradas en el desarrollo y uso de los sistemas de IA. Las fases del ciclo de vida de la IA involucran[63]: (i) "diseño, data y modelos": la cual constituye la fase de planificación, diseño, recopilación de datos y procesamiento, así como la construcción de modelos, (ii) "verificación y validación", (iii) "despliegue" y "operación y monitoreo". Dichas fases suelen tener lugar de forma iterativa y no secuencial, ya que la decisión de retirar un sistema de IA de su operación puede ocurrir en cualquiera de estas[64]. Dicha diferenciación de etapas puede ser útil en la determinación de la responsabilidad, debido a que en cada una de ellas intervienen actores específicos. Así, a modo de ejemplo, si nos encontramos frente a un sistema de IA que ya ha sido desplegado en un determinado mercado, y que es utilizado por un usuario que, maliciosamente y a través de *hacking*, introduce modificaciones en el sistema de IA con el fin de producir daño, no habría mayor problema a la hora de establecer respecto del usuario la existencia de culpa o el nexo causal. Esto, a diferencia de otros casos hipotéticos en que el establecimiento, ya sea de la culpa o el nexo causal, resultará más dificultoso, problemáticas que revisaremos más adelante en la generalidad del Capítulo 3.

Habiéndonos detenido brevemente en la explicación de la IA y el ML, nos referiremos a los acercamientos normativos que este tipo de sistemas han tenido hasta ahora, para luego retomar y analizar pormenorizadamente algunas dificultades propias que la misma presenta desde un punto de vista jurídico.

1.5 Acercamientos normativos.

Los acercamientos normativos a la IA son bastante recientes. En el presente capítulo se revisarán algunos instrumentos normativos relevantes para nuestro país y la presente memoria. Los principales instrumentos que nos convocan son los siguientes:

(i) Legislación en la Unión Europea – Artificial Intelligence Act (AIA)

En marzo de 2024, el Parlamento Europeo adoptó oficialmente la Ley IA (también conocido como el *Artificial Intelligence Act* o AIA*).* Originalmente propuesto el 2021, el AIA pretende convertirse en un estándar normativo global relativo a la IA, tal como el Reglamento General de Protección de

[63] OCDE (2022): 7.
[64] OCDE (2022): 7.

Datos de la Unión Europea lo ha sido en temas de protección de datos personales[65]. El AIA consta de XIII capítulos, los cuales definen el objeto y ámbito de aplicación de la ley, así como las obligaciones, deberes de gobernanza, sanciones, entre otros.

El artículo 1 N.º 2 del AIA establece normas armonizadas para la comercialización, la puesta en servicio, y el uso de los sistemas de IA en la UE, entre las cuales se comprenden normas acerca de la prohibición de determinadas prácticas de IA, así como los requisitos y obligaciones específicos que los sistemas de IA de alto riesgo y sus operadores deben cumplir. Dicho artículo establece, además, normas de transparencia para determinados sistemas de IA según su categoría de riesgo, además de normas relativas para la comercialización de modelos de IA de uso general. Por último, el artículo contiene normas sobre supervisión, vigilancia, gobernanza y ejecución del sistema de IA en el mercado europeo, así como medidas de apoyo a la innovación con especial atención a las pymes.

En el AIA, los sistemas de IA se clasifican por su riesgo, existiendo sistemas de IA de riesgo inaceptable, alto riesgo, riesgo limitado y riesgo mínimo. Dicha clasificación se relega a anexos incorporados en la normativa, con la finalidad de que, frente a cambios tecnológicos, la adaptación de la normativa pueda ser flexible. Entre estos anexos se incluyen el anexo III, que se refiere a cuáles sistemas de IA se consideran como de alto riesgo, así como diversos otros relativos a la documentación técnica necesaria para poner en marcha un sistema de IA riesgoso o las sanciones penales a las que pueden ser sujetos aquellos que pongan comercialicen un sistema de IA de riesgo prohibido. En suma, el AIA establece una serie de prácticas de IA prohibidas y de alto riesgo.

El artículo 3 del AIA provee una serie de definiciones relevantes en materia de IA. En este sentido, define "sistema de IA" como:

> "un sistema basado en máquinas que está diseñado para funcionar con diversos niveles de autonomía y que puede mostrar capacidad de adaptación tras su despliegue, y que, para objetivos explícitos o implícitos, infiere, a partir de la entrada que recibe, cómo generar salidas tales como predicciones, contenidos, recomendaciones o decisiones que pueden influir en entornos físicos o virtuales[66]."

[65] OCDE (2021): 7.
[66] Art. 3 AIA.

Esta definición amplia y la clasificación de sistemas basada en el riesgo que establece la ley, permite que dicha definición pueda adaptarse en función de la evolución de los sistemas de IA[67].

Otro punto relevante es que el AIA define y otorga certeza respecto de los actores involucrados en el ciclo de vida de los sistemas de IA. En este sentido, define el rol de los proveedores, implantadores, representantes autorizados, importadores, distribuidores y operadores; siendo el concepto de "operador" un término genérico usado para referirse a la generalidad de los anteriormente nombrados[68]. El proveedor, para el AIA, es la "persona física o jurídica, autoridad pública, agencia u otro organismo que desarrolle un sistema de IA o un modelo de IA de propósito general o que haga desarrollar un sistema de IA o un modelo de IA de propósito general y lo comercialice o ponga en servicio el sistema de IA bajo su propio nombre o marca, ya sea a cambio de una remuneración o de forma gratuita[69]", el implantador, por su parte, también puede ser una persona física o jurídica, agencia u otro organismo, que "utilice un sistema de IA bajo su autoridad, excepto cuando el sistema de IA se utilice en el curso de una actividad personal no profesional[70]". Respecto del representante autorizado, se señala que es aquella persona física o jurídica situada en la unión que ha recibido y aceptado "un mandato escrito de un proveedor de un sistema de IA o de un modelo de IA de propósito general para, respectivamente, ejecutar y llevar a cabo en su nombre las obligaciones y procedimientos establecidos en el reglamento[71]". Los importadores "comercializan un sistema de IA que lleva el nombre o la marca comercial de una persona física o jurídica establecida en un tercer país[72]", mientras que los distribuidores, siendo parte de la cadena de suministro y siendo distintos de los proveedores o importadores, "comercializan un sistema de IA en el mercado de la unión[73]." Para concluir y, reiterando lo anteriormente dicho, los operadores pueden ser, a la vez: "proveedor, fabricante de productos, implantador, representante autorizado, importador o distribuidor[74]".

Por último, el AIA cuenta con un sistema sancionatorio basado en el riesgo asignado a un determinado sistema de IA, las cuales operan en caso de que se hayan llevado a cabo prácticas

[67] Esto, debido a que la clasificación de los sistemas de IA se delega a los anexos de la ley, otorgando un mecanismo más flexible para su modificación que respecto de la legislación general (en este caso, el AIA).
[68] Art. 3 (3) AIA.
[69] Art. 3 (4) AIA.
[70] Art. 3 (5) AIA.
[71] Art. 3 (5) AIA.
[72] Art. 3 (6) AIA.
[73] Art. 3 (7) AIA.
[74] Art. 3 (8) AIA.

prohibidas por la ley, o que se hayan incumplido alguna de las obligaciones que la regulación establece para los operadores de sistemas de IA. El monto de la multa de dichas sanciones depende, como ya se señaló, del tipo de obligación que se infringe y del tipo de sistema de IA involucrado.

(ii) Propuesta de directiva del Parlamento Europeo relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial (Directiva sobre responsabilidad en materia de IA).

La propuesta de directiva del Parlamento Europeo relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial (de aquí en adelante "propuesta de directiva", "propuesta" y "directiva"), pretende complementar el AIA al establecer normas relativas al ejercicio de acciones de responsabilidad civil extracontractual en el contexto europeo frente a daños ocasionados por sistemas de IA. Así, el proyecto señala en su exposición de motivos que, entre sus razones y fundamentos, está la insuficiencia de los sistemas de responsabilidad civil para responder a los problemas relativos a la responsabilidad civil y sistemas de IA, señalando que:

> "(…) las normas nacionales en vigor en materia de responsabilidad civil, particularmente las que se basan en la culpa, no son adecuadas para tramitar las denuncias de responsabilidad civil por daños causados por productos y servicios en los que se recurre a la IA. Con arreglo a dichas normas, las víctimas deben demostrar que ha habido una acción u omisión ilícita por parte de una persona que ha causado daño. Las características específicas de la IA, incluidas su complejidad, su autonomía y su opacidad, pueden dificultar o hacer excesivamente costoso para las víctimas determinar cuál es la persona responsable y probar que se cumplen los requisitos para una demanda de responsabilidad civil admisible[75]".

Esta propuesta contiene tres ejes principales. En primer lugar, dispone de normas comunes sobre la exhibición de pruebas relativas a sistemas de IA, con el fin de facilitar para la víctima la obtención de la información necesaria para una llevar a cabo una demanda de responsabilidad civil. Esto, mediante la posibilidad de que un órgano jurisdiccional solicite la exhibición de pruebas directamente al demandado bajo determinadas hipótesis. En segundo lugar, y en relación con lo anterior, la propuesta establece una serie de normas que alteran la carga de la prueba en favor de la

[75] Artículo 3 propuesta de directiva.

víctima. Por último, la propuesta se encarga de delinear una serie de presunciones simplemente legales o *iuris tantum* ("refutables" en el texto) en relación tanto con el incumplimiento, por un lado, como la causalidad (o nexo causal), por otro. La presunción de nexo causal, por ejemplo, opera en caso de haberse probado la culpa en el juicio de responsabilidad civil extracontractual.

Volviendo sobre el ámbito probatorio, la propuesta sigue el esquema establecido por el AIA, en el sentido de que establece la posibilidad de que un órgano jurisdiccional pueda requerir pruebas directamente al demandado y no de la víctima (demandante) para dar por acreditada la responsabilidad, lo cual depende del nivel de riesgo bajo el cual se clasifique el sistema de IA involucrado junto con la negativa del proveedor de otorgar información voluntariamente a la víctima. Refiriéndose a estos deberes de información y a las presunciones contenidas en la directiva, CASALS señala que:

> "La norma pretende alcanzar un equilibrio entre la protección de las víctimas y el fomento de la innovación por parte del empresariado, ya que no solo elimina en favor de las víctimas los obstáculos al acceso a la información necesaria para la prueba de su acción de reclamación, sino que también establece garantías para el sector de la IA[76]."

Así, y continuando con dicha idea, el artículo 3 de la propuesta señala que, en caso de tratarse de un determinado sistema de IA de alto riesgo del que se sospeche que ha causado daños, la víctima o demandante potencial podrá hacer la solicitud directamente al órgano jurisdiccional, debiendo, eso sí, apoyar su solicitud con hechos y pruebas "suficientes para sustentar la viabilidad de una demanda de indemnización por daños y perjuicios[77]". Esta posibilidad del órgano jurisdiccional de ordenar la exhibición de pruebas procede cuando "el demandante haya realizado todos los intentos proporcionados de obtener del demandado las pruebas pertinentes[78]", en este sentido, es necesario que la víctima haya solicitado previamente la información al proveedor que se presume causante del daño.

Para determinar si una orden de exhibición o conservación de pruebas será concedida, la propuesta se remite a que los órganos jurisdiccionales deberán tomar en cuenta como criterio los intereses legítimos de todas las partes involucradas, incluyendo a terceros, pero también los intereses relativos a la información confidencial o al secreto comercial. En relación con esto último, la propuesta delega

[76] CASALS (2023): 71.
[77] Artículo 3 propuesta de directiva.
[78] *Ibid.*

en los Estados miembros el velar para que los órganos jurisdiccionales "adopten las medidas específicas necesarias a fin de preservar la confidencialidad cuando dicha prueba se utilice o se mencione en procedimientos judiciales[79]." Si, de acuerdo con lo expuesto sobre el artículo 3, el operador incumple la orden de un órgano jurisdiccional para exhibir la información requerida, la propuesta señala que "se presumirá el incumplimiento por parte del demandado de un deber de diligencia pertinente". En otros términos, operaría en este caso una presunción simplemente legal de culpa respecto del operador.

La propuesta también incluye una presunción de existencia de causalidad en caso de que se pruebe la culpa. Esto tiene mucho sentido, ya que, como se tratará en capítulos subsiguientes, la prueba de la causalidad en el ámbito de los sistemas de IA podría resultar particularmente difícil. Para que opere dicha presunción, de acuerdo con el artículo 4, es necesario que se cumplan una serie de condiciones: primero, el demandante debió haber demostrado o el órgano jurisdiccional haber supuesto la culpa del demandado, o de una persona de cuyo comportamiento sea responsable el demandado, "consistente en el incumplimiento de un deber de diligencia establecido por el Derecho de la Unión o nacional destinado directamente a proteger frente a los daños que se hayan producido[80]", debe, asimismo, "considerarse razonablemente probable, basándose en las circunstancias del caso, que la culpa ha influido en los resultados producidos por el sistema de IA o en la no producción de resultados por parte del sistema de IA[81]", y, por último, el demandante debe haber demostrado "que la información de salida producida por el sistema de IA o la no producción de una información de salida por parte del sistema de IA causó los daños[82]".

En el caso de sistemas de IA de alto riesgo, por su parte, la presunción legal de nexo causal solo operará que el demandante pruebe que el proveedor o la persona sujeta a obligaciones del proveedor haya incumplido alguna de las obligaciones señaladas en el AIA, a las cuales se remite la propuesta, y que dependen de la clasificación del sistema de IA como un sistema de alto riesgo. Entre dichos incumplimientos se encuentran, a modo ilustrativo, que no se haya diseñado cumpliendo los requisitos de transparencia de la AIA, o que no se haya desarrollado el sistema de IA del modo que permita una vigilancia efectiva por personas físicas durante todo su periodo de uso. Si el demandado, en el caso de los sistemas de alto riesgo, logra probar que el demandante "puede

[79] Artículo 3 propuesta de directiva.
[80] Artículo 4 propuesta de directiva.
[81] *Ibid.*
[82] *Ibid.*

acceder razonablemente a pruebas y conocimientos especializados suficientes para demostrar el nexo causal", no operará la presunción. Si se trata de un sistema de IA que no se clasifique como de alto riesgo bajo los criterios del AIA, la presunción de causalidad solo se aplicará cuando el órgano jurisdiccional considere excesivamente difícil para el demandante demostrar dicho nexo causal.

Por último, la propuesta señala que en el caso de demandas por daños y perjuicios respecto de un demandado que haya utilizado el sistema de IA en el transcurso de una actividad personal de carácter no profesional, la presunción de causalidad solo aplicará "en caso de que el demandado haya interferido sustancialmente en las condiciones de funcionamiento de sistema de IA o cuando el demandado tuviese la obligación y estuviese en condiciones de determinar las condiciones *(sic)* de funcionamiento del sistema de IA y no lo haya hecho[83]."

(iii) La Política Nacional de Inteligencia Artificial y la circular del Sernac

En el ámbito nacional, si bien no hay una ley propiamente tal que se refiera a la IA, sí han existido acercamientos normativos, entre los cuales destacan la elaboración de una "Política Nacional de Inteligencia Artificial"[84], por parte del Ministerio de Ciencia Tecnología, Conocimiento e Innovación y la publicación de la "Circular Interpretativa sobre protección de los consumidores frente al uso de sistemas de inteligencia artificial" por parte del Sernac.

La Política Nacional de Inteligencia Artificial se publicó en octubre de 2021, e incluye "los lineamientos estratégicos que debe seguir el país en esta materia durante los próximos 10 años, con el objetivo de empoderar a las personas en el uso y desarrollo de herramientas de IA, y participar en el debate sobre sus consecuencias legales, éticas, sociales y económicas"[85] y se estructura en torno a tres ejes: (i) factores habilitantes; (ii) uso y desarrollo de Inteligencia Artificial en Chile y (iii) aspectos de ética y seguridad[86]. La política prefiere no tomar una única definición de IA, sino referirse a varios estándares internacionales[87], y destaca la importancia que se prevé tendrán este

[83] Artículo 4 propuesta de directiva.
[84] MINISTERIO DE CIENCIA, TECNOLOGÍA, CONOCIMIENTO E INNOVACIÓN (2021). En: https://minciencia.gob.cl/areas-de-trabajo/inteligencia-artificial/politica-nacional-de-inteligencia-artificial/
[85] MINISTERIO DE CIENCIA, TECNOLOGÍA, CONOCIMIENTO E INNOVACIÓN (2021): 7.
[86] *Ibid.*:7
[87] Respecto a este punto, el documento cita una definición elaborada por la Universidad de Montreal: "*el conjunto de técnicas informáticas que permiten a una máquina (por ejemplo, un ordenador, un teléfono) realizar tareas que, por lo común, requieren inteligencia tales como el razonamiento o el aprendizaje*", y otra dada por la OCDE: "*un sistema computacional que puede, para un*

tipo de sistemas en nuestro futuro[88]. De particular relevancia para nuestros fines resulta el tercer eje. "Ética, aspectos legales y regulatorios, e impactos económicos"[89]. El punto 3.1.1. de la política llama a "impulsar la construcción de certezas regulatorias sobre los sistemas de IA que permitan su desarrollo, respetando los derechos fundamentales de acuerdo con la Constitución y las leyes" señalando, entre otras cosas, que resulta importante para Chile insertarse en la discusión normativa respecto a la IA, así como respecto del desarrollo de los requisitos necesarios para "cautelar en forma ágil el desarrollo y uso responsable de la IA".

El objetivo 3.1.2 llama, por su parte, a "impulsar la transparencia algorítmica" señalando como necesario entregar información sobre cómo funcionan los algoritmos decisionales que utilizan los órganos de la administración del estado, así como los datos involucrados en la toma de decisiones, entre otros[90].

Por su parte, el 18 de enero de 2022, el Sernac aprobó la Resolución Exenta N.º 33, titulada la "Circular Interpretativa sobre protección de los consumidores frente al uso de sistemas de inteligencia artificial"[91]. La Circular incorpora la definición otorgada por la OCDE:

> "(…) un sistema de inteligencia artificial puede ser entendido como un sistema basado en una máquina que, estando diseñado para funcionar con distintos niveles de autonomía, puede, para un conjunto determinado de objetivos definidos por el ser humano, hacer predicciones, recomendaciones, o decisiones que influyen en entornos reales o virtuales"

Añadiendo que:

> "(…) puede entenderse que las tecnologías de IA presentan diversos grados de autonomía respecto de operadores externos, ya sea en relación con las tareas que ejecuta; los niveles de control o supervisión que ejerce un humano; el proceso de desarrollo y mejoramiento del sistema para llevar a cabo sus operaciones propias, y a la complejidad de los resultados que la máquina arroja[92]."

determinado conjunto de objetivos definidos por humanos, hacer predicciones y recomendaciones o tomar decisiones que influyen en entornos reales o virtuales. Los sistemas de IA están diseñados para operar con distintos niveles de autonomía".

[88] MINISTERIO DE CIENCIA, TECNOLOGÍA, CONOCIMIENTO E INNOVACIÓN (2021): 8.
[89] *Ibid.*: 49.
[90] *Ibid.*: 53.
[91] SERVICIO NACIONAL DEL CONSUMIDOR (2022): Resolución Exenta N.º 33/2022.
[92] *Ibid.*

Asimismo, la Circular señala las formas en las que este tipo de sistemas pueden presentar tanto un gran potencial en el marco de consumo, permitiendo la automatización de ciertas tareas, mayor dinamismo a ciertos procesos de toma de decisiones (lo que para el Sernac posibilitaría una reducción de los costos de contratación), la posibilidad de celebrar contratos basados en las preferencias y necesidades individuales de los consumidores, el fortalecimiento de la ciberseguridad y ejercicio de los derechos del consumidor, entre otros[93].

Por otra parte, el Sernac identifica determinados riesgos en torno al uso de IA en las relaciones de consumo, entre ellos, la manipulación a través de "patrones oscuros" o *dark patterns* un "término utilizado para referirse a diseños utilizados en sitios web y aplicaciones que pueden tener el efecto —intencional o involuntario— de oscurecer, subvertir o perjudicar la autonomía, la toma de decisiones o la elección del consumidor"[94] así como características inherentes a la IA, como la opacidad, la imprevisibilidad de sus resultados o decisiones y su comportamiento "parcialmente autónomo"[95], entre otros. En cuanto a su contenido, la circular se estructura en cinco secciones, las cuales se relacionan con diversos derechos del consumidor contenidos en la Ley N.º 19.496 Sobre Protección de los Derechos de los Consumidores, y establece diversas recomendaciones con respecto a cómo los proveedores deben ajustar su comportamiento para asegurar la protección del consumidor en el marco del uso de este tipo de tecnologías.

Si bien no revisaremos en detalle las recomendaciones dadas por el Sernac por exceder el objeto de la presente memoria, las cinco secciones se centran en: (i) Entrega de información veraz, oportuna y transparente; (ii) resguardo de la libertad de elección; (iii) seguridad en el consumo, (iv) prohibición de toda discriminación arbitraria y (v) protección de los datos de los consumidores[96].

(iv) Principios éticos de la OCDE y Unesco

Los principios éticos de la OCDE conforman el primer conjunto de directrices de políticas intergubernamentales sobre IA, frente a los cuales se pactó someterse a normas internacionales que

[93] SERVICIO NACIONAL DEL CONSUMIDOR (2022). Resolución Exenta N.º 33/2022: 6.
[94] *Ibid.*
[95] *Ibid.*
[96] *Ibid.*

velen por sistemas de IA (i) robustos, (ii) seguros, (iii) imparciales y (iv) fiables[97]. Estos principios afirman, a modo de resumen, lo siguiente[98]:

(i) La IA debe estar al servicio de la humanidad y del planeta, impulsando un crecimiento inclusivo, el desarrollo sostenible y el bienestar.

(ii) Los sistemas de IA deben diseñarse de manera que respeten el Estado de derecho, los derechos humanos, los valores democráticos y la diversidad, e incorporar salvaguardias adecuadas —por ejemplo, permitiendo la intervención humana cuando sea necesario— con miras a garantizar una sociedad justa y equitativa.

(iii) Los sistemas de IA deben estar presididos por la transparencia y una divulgación responsable a fin de garantizar que las personas sepan cuándo están interactuando con ellos y puedan oponerse a los resultados de esa interacción.

(iv) Las organizaciones y las personas que desarrollen, desplieguen o gestionen sistemas de IA deberán responder de su correcto funcionamiento en consonancia con los principios precedentes.

De particular relevancia para esta memoria resulta el principio soslayado en el numeral (iv), dado que establece una regla general de responsabilidad respecto a las personas que desarrollen, desplieguen o gestionen sistemas de IA, los cuales deben responder por su correcto funcionamiento.

Por su parte, la Unesco desarrolló la "Recomendación sobre la ética de la IA", indicando que se trata del primer instrumento mundial que busca el uso y desarrollo de una IA ética, inclusiva, justa y al servicio de la humanidad[99]. Dicho instrumento se firmó en 2021 por los 193 Estados miembros de la Conferencia General de la Unesco[100]. Las recomendaciones incluyen una lista de principios orientadores que giran a partir de los siguientes ejes: (i) proporcionalidad e inocuidad, (ii) seguridad y protección, (iii) equidad y no discriminación, (iv) sostenibilidad, (v) derecho a la intimidad y protección de datos, (iv) supervisión y decisión humanas, (v) transparencia y explicabilidad, (vi) responsabilidad y rendición de cuentas, (vii) sensibilización y educación y, por último, (viii) gobernanza y colaboración adaptativas y de múltiples partes interesadas. De particular interés para

[97] OCDE (2019). En: https://www.oecd.org/espanol/noticias/cuarentaydospaisesadoptanlosprincipiosdelaocdesobreinteligenciaartificial.htm.
[98] *Ibid.*
[99] UNESCO (2021). En: https://es.unesco.org/fieldoffice/montevideo/EticaInteligenciaArtificial
[100] *Ibid.*

los objetivos de esta memoria resultan los puntos (v), (vi) y (vii), los cuales analizaremos al acercarnos a los distintos problemas relativos a la responsabilidad y la IA.

Dicho esto, debemos tener presente que las circulares del Sernac, al ser instrumentos normativos emitidos por dicho órgano, no son de aplicación general, sino que miran de forma directa a quienes son fiscalizados por el mismo y son, como se señala en su texto, "recomendaciones". La Política Nacional de Inteligencia Artificial, si bien establece como enfoque nacional el establecer normas relativas a la responsabilidad y a la IA, no ha tenido como consecuencia fáctica la introducción de proyectos de ley a su efecto.

Sin perjuicio de esto, los principios éticos que guían la IA, propuestos tanto por la OCDE como la Unesco, pueden servirnos como un importante insumo normativo para resolver algunas de las preguntas que plantea el uso de sistemas de IA en relación con la responsabilidad civil.

Lo anterior, no obsta a que, en atención al estado actual de las normas y al contenido que hasta esta fecha tiene el proyecto de ley relativo a los sistemas de IA, forzosamente lleguemos a la conclusión de que frente a posibles casos de responsabilidad civil extracontractual en que veamos involucrados uso de sistemas de IA, y fuera del ámbito del derecho del consumidor[101], solo contamos con la normativa civil general y supletoria, así como con los principios que, aquí delineados, pueden entenderse como incorporados en la normativa chilena al ser parte de instrumentos emitidos en el marco de tratados internacionales suscritos por nuestro país.

(v) Proyecto de ley que regula los sistemas de inteligencia artificial (boletín 16821-19).

En mayo de 2024, se presentó vía mensaje en la Cámara de diputados un proyecto de ley que regula los sistemas de IA (de aquí en adelante "proyecto ley de IA" o "proyecto"), el cual actualmente se encuentra en su primer trámite constitucional. El proyecto refleja una clara influencia de los principios éticos de la OCDE y el AIA y se compone de 31 artículos. El mensaje resume la estructura del proyecto de ley de la siguiente forma: "(a) Disposiciones generales, lo que comprende el ámbito de aplicación de la propuesta, definiciones y clasificación de los sistemas de IA; (b) Sistemas de riesgo inaceptable; (c) Sistemas de alto riesgo, (d); Sistemas de IA de riesgo limitado; (e) Incidentes graves; (g) Gobernanza; (h) Medidas de apoyo a la innovación; (i) Confidencialidad,

[101] Esto, debido a que la circular del Sernac solo tiene aplicación normativa respecto de los proveedores en el sentido que les da la Ley de protección a los derechos del consumidor.

infracciones y sanciones; (j) Disposiciones finales y modificaciones a otros cuerpos legales." Antes de concluir con un análisis sobre cómo se trata la responsabilidad civil en este proyecto, nos parece pertinente repasar algunos aspectos importantes del mismo.

i. Ámbito de aplicación

En el título I, el proyecto define el objeto de la ley, su ámbito de aplicación, y los principios aplicables a los sistemas de IA. Así, el artículo 1° señala que "la presente ley tiene por objeto promover la creación, desarrollo, innovación e implementación de sistemas de inteligencia artificial al servicio del ser humano, que sean respetuosos de los principios democráticos, el Estado de derecho y los derechos fundamentales de las personas frente a los efectos nocivos que determinados usos pueda irrogar" previniendo que "la presenta ley no será aplicable a: (a) sistemas de IA desarrollados y utilizados con fines de defensa nacional (…) (b) las actividades de investigación, pruebas y desarrollo sobre sistemas de IA de forma previa a su introducción en el mercado o puesta en servicio, siempre que dichas actividades se lleven a cabo respetando los derechos fundamentales de las personas (…) (c) componentes de IA proporcionados en el marco de licencias libres y de código abierto, salvo que sean comercializados o puestos en servicio por un proveedor como parte de un sistema de alto riesgo." En sus definiciones, el proyecto define sistema de IA como: "sistema basado en máquinas que, por objetivos explícitos o implícitos infiere, a partir de la entrada que recibe, cómo generar salidas tales como predicciones, contenidos, recomendaciones o decisiones que pueden influir en entornos físicos o virtuales. Los distintos sistemas de IA pueden variar en sus niveles de autonomía y adaptabilidad tras su implementación[102]."

ii. Actores

El proyecto de ley clasifica a los diversos actores relevantes en cuanto a los sistemas de IA. Dichos actores son, de acuerdo con el artículo 3 N°s 4, 5, 6, 7, 8, 9, 10 y 11: (i) el proveedor, (ii) el implementador, (iii) el proveedor de tecnología, (iv) el representante autorizado, (v) el importador, (vi) el distribuidor, (vii) el operador, y por último, (viii) la persona afectada. Más adelante, en esta memoria, veremos que la existencia de múltiples actores en la cadena de producción de los sistemas de IA resulta un desafío particular a la hora de consolidar el juicio de responsabilidad civil. En este

[102] Artículo 3 N. 1° proyecto ley de IA.

sentido, la individualización legal de diversos actores es útil para determinar el rol específico de cada uno de los participantes en el proceso productivo y en el ciclo de vida del sistema de IA.

El proyecto define a los proveedores como "toda persona natural o jurídica u organismo del Estado que desarrolle un sistema de IA con miras a introducirlo en el mercado o ponerlo en servicio, a título gratuito u oneroso." Los implementadores, por su parte, son "toda persona natural o jurídica u organismo del Estado que utilice un sistema de IA, salvo que se trate de un uso privado del mismo, en los términos de la ley N. 17.336 sobre propiedad intelectual". Los proveedores de tecnologías son "todo proveedor involucrado con el implementador en la comercialización y suministro de software, herramientas y componentes de software, modelos y datos previamente entrenados." Los representantes autorizados se definen, por su parte, como "toda persona natural o jurídica domiciliada en Chile que haya recibido y aceptado el mandato por escrito de un proveedor de un sistema de IA para cumplir con las obligaciones establecidas en la presente ley en representación de dicho proveedor." Asimismo, los importadores constan de "toda persona natural o jurídica domiciliada en Chile que introduzca en el mercado o ponga en servicio un sistema de IA que lleve el nombre o la marca comercial de una persona natural o jurídica establecida fuera del territorio nacional." Finalmente, los distribuidores corresponden a "toda persona natural o jurídica que forme parte de la cadena de suministro, distinta del proveedor o el importador, que comercialice un sistema de IA en el mercado nacional sin influir sobre sus propiedades." El proyecto aglutina la diversidad de actores bajo el concepto de general de "operador". Así, un operador consta, de forma similar a lo dispuesto en el AIA, de todos los actores ya mencionados, entendidos como un conjunto. Por tanto, para el proyecto, un operador se compone de: "el proveedor, el implementador, el representante autorizado, el importador y el distribuidor."

La persona afectada, por su parte, se delimita como "toda persona natural o grupo de personas naturales que expuesta a un sistema de IA que sufra un perjuicio como consecuencia de dicha exposición."

iii. Principios

El artículo 4, siguiendo el modelo de las normativas ya revisadas de la UNESCO y la OCDE, delinea principios específicos aplicables a los sistemas de IA. Respecto de estos, todos los operadores deberán observar dichos principios en su ámbito de aplicación. Estos principios son:

(a) Intervención y supervisión humana: los sistemas de IA se desarrollarán y utilizarán como una herramienta al servicio del ser humano, que respete la dignidad humana y la autonomía personal, y que funcione de manera que pueda ser controlada y vigilada adecuadamente por seres humanos.

(b) Solidez y seguridad técnica: los sistemas de IA se utilizarán como una herramienta al servicio del ser humano, que respete la dignidad humana y la autonomía personal, y que funcione de manera que pueda ser controlada y vigilada adecuadamente por seres humanos.

(c) Privacidad y gobernanza de datos: los sistemas de IA se desarrollarán y se utilizarán de conformidad con las normas vigentes en materia de privacidad y protección de datos. (…) del mismo modo, se procurará que los datos que utilicen sean inter operables.

(d) Transparencia y explicabilidad: los sistemas de IA se desarrollarán y utilizarán facilitando una trazabilidad y explicabilidad adecuadas, de modo tal que los seres humanos puedan conocer de forma clara y precisa y sean conscientes de que se comunican o interactúan con un sistema de IA, en aquellos casos en los que dicho conocimiento les ayudaría a tomar decisiones sobre sus derechos, seguridad o privacidad, informando a sus destinatarios, cuando corresponda, como el sistema ha obtenido sus predicciones o resultados, así como también sobre las capacidades y limitaciones de dicho sistema de IA.

(e) Diversidad, no discriminación y equidad: los sistemas de IA se desarrollarán y utilizarán durante todo su ciclo de vida, promoviendo la igualdad de acceso, la igualdad de género y la diversidad cultural, evitando simultáneamente los efectos discriminatorios y sesgos de selección o de información que pudieran generar un efecto discriminatorio.

(f) Bienestar social y medioambiental: los sistemas de IA se desarrollarán y utilizarán de manera sostenible y respetuosa con el medioambiente y los seres humanos. Por lo anterior, los responsables de la introducción en el mercado, la puesta en servicio o la utilización de los sistemas de IA deberán revisar los efectos a largo plazo que su aplicación genera en la sociedad, la democracia y el medioambiente.

(g) Rendición de cuentas y responsabilidad: responsabilidad: los sistemas de IA deberán proporcionar un correcto funcionamiento a lo largo de su ciclo de vida por parte de quienes los diseñan, desarrollan, operan o despliegan, en relación con sus funciones propias.

(h) Protección de los derechos de los consumidores: consumidores: los sistemas de IA se desarrollarán y utilizarán de conformidad con las normas vigentes en materia de protección de los derechos de los consumidores, debiendo asegurar el trato justo, entrega de información veraz, oportuna y transparente y el resguardo a la libertad de elección y la seguridad en el consumo.

iv. Responsabilidad.

La técnica legislativa de este proyecto implica un doble sistema de responsabilidad: regula, por una parte, la responsabilidad de sus infractores en sede administrativa y, por otra, establece una norma especial de responsabilidad civil extracontractual, la cual se remite a las reglas generales, como explicaremos más adelante.

En cuanto al primer tipo de responsabilidad, puede aseverarse que nace de la infracción a las obligaciones que el proyecto establece para los proveedores que, las cuales se resuelven mediante un proceso contencioso-administrativo (artículo 19). Para dichas infracciones, el proyecto de ley establece un procedimiento administrativo sancionador (artículo 26) instruido por la autoridad de control (denominado "la Agencia"), junto con la posibilidad de llevar a cabo un procedimiento de reclamación judicial (artículo 27) contra aquellos actos administrativos que paralicen el procedimiento o una resolución final o de término emanado de la Agencia frente a la Corte de Apelaciones respectiva. Las infracciones se clasifican según su gravedad, así, son (i) gravísimas: las que involucran la puesta en servicio o utilización de un sistema de IA de riesgo inaceptable, (ii) graves: aquellas que implican el incumplimiento de las reglas (obligaciones) especiales dispuestas para sistemas de IA de alto riesgo, (c) leves: aquellas que impliquen el incumplimiento de las obligaciones de transparencia dispuestas para los sistemas de IA de riesgo limitado. Por su parte, las sanciones varían desde las 5.000 a las 20.000 UTM, según sea la infracción, y la determinación de la cuantía de la multa considera las circunstancias particulares de cada situación (tomando en cuenta, por ejemplo, la duración de la infracción, el tamaño y volumen del proveedor que comete la infracción, entre otras).

Respecto del segundo tipo de responsabilidad tratado por el proyecto, es decir, la responsabilidad civil propiamente tal, es posible afirmar que el proyecto de ley se plantea la posibilidad de que una persona que ha sufrido un daño debido a un sistema de IA pueda reclamar la indemnización de perjuicios mediante una acción de responsabilidad civil, hipótesis que la presente memoria pretende

analizar. Sin embargo, esta regla solo complementa el sistema general de responsabilidad extracontractual en Chile, al establecer una serie de medidas que, en conjunto con la indemnización de perjuicios, puede pedir la víctima. No altera, pues, la carga de la prueba, la forma en la que la víctima puede llevar a cabo la acción de responsabilidad civil, ni establece medidas relativas a sopesar las dificultades relativas a la prueba respecto de sistemas de IA, del modo que sí lo hace la propuesta de directiva de la UE. Entre las medidas adicionales a la indemnización de perjuicios, y de forma similar a lo que ocurre en materia de propiedad intelectual[103] y competencia desleal[104], el proyecto señala que puede pedirse la cesación de actos generadores de daño, la adopción de medidas necesarias para evitar que prosiga la infracción, y la publicación de la sentencia respectiva a costa del demandado[105].

Este acercamiento nos parece incompleto, ya que, tal como señalamos, no soluciona los principales problemas que se han previsto respecto de la atribución de responsabilidad civil en el marco de utilización de sistemas de IA, problemas que se tratarán con profundidad en los capítulos subsiguientes. Reiterando, el proyecto no establece, como lo hace la propuesta de directiva de la UE, normas específicas respecto de la atribución de la culpa, causalidad, o carga de la prueba, elementos del juicio de responsabilidad civil que presentan complejidades únicas al tratarse de sistemas de IA. Asimismo, el artículo no deja en claro contra quién debe recurrir el demandante, debido a que en el término operador, como ya se ha visto, se incluyen diversos actores del ciclo de vida de la IA. Tampoco clarifica de qué forma se distribuye la responsabilidad entre quienes se

---

103 El artículo 85 B de la ley 17.336 sobre propiedad intelectual señala: "*El titular de los derechos reconocidos en esta ley tendrá, sin perjuicio de las otras acciones que le correspondan, acciones para pedir:*

a) El cese de la actividad ilícita del infractor.

b) La indemnización de los daños y perjuicios patrimoniales y morales causados.

c) La publicación de un extracto de la sentencia, a costa del demandado, mediante anuncio en un diario de circulación comercial de la Región correspondiente, a elección del perjudicado."

104 El artículo 5° de la ley 20169 que regula la competencia desleal señala: "*Artículo 5º.- Contra los actos de competencia desleal pueden ejercerse, conjunta o separadamente, las siguientes acciones:*

a) Acción de cesación del acto o de prohibición del mismo si aún no se ha puesto en práctica.

b) Acción declarativa de acto de competencia desleal, si la perturbación creada por el mismo subsiste.

c) Acción de remoción de los efectos producidos por el acto, mediante la publicación de la sentencia condenatoria o de una rectificación a costa del autor del ilícito u otro medio idóneo.

d) Acción de indemnización de los perjuicios ocasionados por el acto, sujeta a las disposiciones del Título XXXV del Libro IV del Código Civil."

105 Dicho artículo señala: "*Artículo 28.- Responsabilidad civil. La persona que sufra un daño como consecuencia de la utilización de un sistema de IA, podrá demandar civilmente y de forma conjunta respecto del operador:*

(a) La cesación de los actos generadores de daño.

(b) La indemnización de los daños y perjuicios

(c) La adopción de las medidas necesarias para evitar que prosiga la infracción.

(d) *La publicación de la sentencia a costa del condenado, mediante anuncios en un diario a elección del demandante. Esta medida será aplicable cuando la sentencia así lo señale expresamente.*"

encuentran contemplados en dicha definición de "operador". De hecho, cuando la norma utiliza el término "conjuntamente", se refiere al conjunto de sanciones señaladas en la norma (v.gr. la cesación de los actos generadores de daño en conjunto con la indemnización de perjuicios, etc.), y no al conjunto de personas que pueden considerarse como operadores. A partir de la literalidad de la norma, es forzoso concluir que, respecto de aquellos que el proyecto de ley considera como operadores, la forma de responder frente a una demanda de responsabilidad civil extracontractual es simplemente conjunta y no solidaria, al no hallarnos en la hipótesis que el Código Civil establece en el art. 2317.

Profundizando lo anterior, el proyecto de ley tampoco establece facilidades en materia probatoria o, como se señaló, normas que alteren las reglas generales que rigen respecto de la culpa y el nexo causal. Estos últimos puntos serán más evidentes al tratar cada uno de los elementos del juicio de responsabilidad civil extracontractual en nuestro ordenamiento jurídico y la forma en la que la utilización de sistemas de IA puede generar desafíos previamente inexistentes en el derecho civil. Dicho esto, es posible pensar que existan cambios en el proyecto a lo largo de su tramitación, o que un futuro proyecto de ley incorpore los aspectos problemáticos que se señalarán tanto en la presente memoria como aquellos ya incluidos en la propuesta de directiva europea.

1.6. Conclusiones

Para concluir, es posible afirmar que la regulación más general y relevante en materia de sistemas de IA es el Artificial Intelligence Act (AIA), promulgado por la Unión Europea (UE) en marzo de 2024, con la intención de establecer un estándar global para la regulación de la IA. Esta regulación apunta a ser un marco normativo general de forma análoga al Reglamento General de Protección de Datos (RGPD), el cual desempeña un rol de referencia en la protección de datos personales a nivel mundial. En complemento al AIA, la propuesta de la Directiva del Parlamento Europeo sobre responsabilidad civil extracontractual en el contexto de la IA busca establecer normas específicas para cada uno de los sistemas de responsabilidad civil de las naciones de la UE. En particular, la propuesta se dirige a aquellos sistemas de responsabilidad extracontractual basados en la culpa, ya que se estima que no se adecuan para manejar los daños causados por sistemas de IA. Su objetivo, en suma, es postular un conjunto de normas sobre la carga de la prueba, la exhibición de pruebas y presunciones legales en casos de responsabilidad civil, facilitando así a las víctimas la obtención de compensaciones justas y adecuadas.

Fuera del marco de la regulación de la UE, los principios éticos de la OCDE y Unesco sobre IA proporcionan un marco normativo y ético para el desarrollo y despliegue de estos sistemas, enfatizando la necesidad de transparencia, responsabilidad y respeto a los derechos humanos. Estos principios pretenden guiar la elaboración de políticas y regulaciones que buscan asegurar que la IA se utilice de manera beneficiosa y segura para la sociedad.

En el contexto chileno, aunque no existe una ley específica sobre IA, se han dado pasos importantes con la publicación de una Política Nacional de Inteligencia Artificial y una Circular del Sernac sobre protección de los consumidores frente al uso de sistemas de IA. La política nacional establece lineamientos estratégicos para el desarrollo y uso responsable de la IA en el país, mientras que la circular del Sernac proporciona directrices para la protección de los derechos de los consumidores en este ámbito, resaltando tanto las oportunidades como los riesgos asociados al uso de tecnologías de IA. Asimismo, recientemente se presentó en nuestro país un proyecto de ley que regula los sistemas de IA, reflejando una clara influencia de las normativas internacionales que se revisarán, como lo son el AIA y los principios de la OCDE. Este proyecto aborda la responsabilidad civil en dos vertientes: administrativa y civil, y propone un procedimiento sumario para la tramitación de demandas por daños causados por sistemas de IA. Sin embargo, quedan áreas que requieren mayor claridad, especialmente en lo referente a la atribución de culpa y la carga de la prueba, como vimos en el presente capítulo.

CAPÍTULO 2: LOS DESAFÍOS QUE PRESENTA LA INTELIGENCIA ARTIFICIAL.

En el presente capítulo se expondrán algunos de los principales desafíos técnicos que traen consigo los sistemas de IA. Estos desafíos o problemáticas resultarán relevantes para analizar la aplicación de las normas de responsabilidad civil al resolver casos que impliquen el uso de este tipo de sistemas. Sobre la explicación aquí presentada es que se podrá llevar a cabo, en el Capítulo 3, el análisis pormenorizado referido a cómo estos desafíos afectan particularmente al juicio de responsabilidad civil. Exponer estos desafíos nos ayudará a aclarar por qué, en concreto, podría resultar difícil establecer el nexo causal o la culpa dentro del juicio de la responsabilidad civil extracontractual en el que se vea involucrado el uso de un sistema de IA.

Los desafíos presentados por los sistemas de IA requieren analizarse desde una óptica tanto técnica como jurídica, esto es, multidisciplinaria, por lo que la presente memoria, en lo que a este capítulo respecta, solo pretende presentar algunos fenómenos relativos a los sistemas de IA que han sido considerados[106] como particularmente problemáticos.

Es difícil señalar de forma sistemática y precisa cuáles son los desafíos de la IA que más impacto tendrán frente a la responsabilidad civil en abstracto, por lo que en atención a los instrumentos internacionales y a la literatura actualmente disponible, se delimitará el análisis a los siguientes puntos: (a) complejidad estructural y pluralidad de actores, (b) opacidad y sesgo algorítmico, (c) autonomía, y (e) imprevisibilidad[107].

---

[106] Por ejemplo, por parte de los gobiernos y organismos multinacionales señalados *supra* en el Capítulo 1, acápite 1.5 que se refieren a estos desafíos de forma expresa al motivar la legislación relativa a sistemas de IA.

[107] Es muy difícil delimitar de forma precisa *todos* los desafíos que la IA presentan que podrían tener algún impacto en la responsabilidad civil. La lista aquí presentada toma como referencia los puntos señalados por el Expert Group on Liability and New Technologies de la Comisión Europea, la cual identifica los siguientes desafíos en las tecnologías emergentes y en particular, la IA, en relación con los sistemas de responsabilidad Civil: (a) complejidad ("complexity"), (b) opacidad ("opacity"), (c) apertura ("openness"), (d) autonomía ("autonomy"), (e) previsibilidad ("predictability"), (f) dependencia en datos ("data-drivenness") y (g) vulnerabilidad ("vulnerability"). La decisión de dejar fuera los aspectos relacionados con los datos personales y a la vulnerabilidad, así como a la apertura de los sistemas, radica en que la presente memoria se dedica a analizar exclusivamente las normas generales de responsabilidad civil y no normas especiales relacionadas con la protección de datos personales, y tampoco a las normas que puedan derivar de ilícitos relacionados a delitos informáticos, sin perjuicio de que estos temas puedan (y deban) ser tratados en otros textos relativos a esos temas en el futuro. Por su parte, el Expert Group on Liability and New Technologies surge como un grupo que asiste a la UE en la misión de: "Provide the Commission with expertise on the applicability of the Product Liability Directive to traditional products, new technologies and new societal challenges (Product Liability Directive formation) and assist the Commission in developing principles that can serve as guidelines for possible adaptations of applicable laws at EU and national level relating to new technologies (New Technologies formation)."

2.1 Complejidad y pluralidad de actores:

Como ya se ha ido abordando, los sistemas de IA son un fenómeno complejo. En este acápite, se analizará la complejidad de la IA bajo tres perspectivas: (i) la interacción entre diversos componentes de un sistema de IA, (ii) la complejidad interna de los algoritmos y (iii) la complejidad en la pluralidad de actores en la elaboración y uso de sistemas de IA:

(i) Complejidad en sistemas: La IA necesita de soportes físicos que la sustenten (el *hardware*), los cuales suelen estar compuestos de diversas partes cuya interacción requiere un alto grado de sofisticación técnica[108]. A esto se le suman los componentes digitales de los sistemas de IA, complejidad que "vuelve a dicha tecnología *(la IA)* aún más compleja y la distancia mucho más allá de los arquetipos de las fuentes potencialmente dañinas en las cuales se basan las reglas de responsabilidad (civil)[109]."

(ii) Complejidad algorítmica: los algoritmos utilizados en los sistemas de IA son complejos e intricados, lo que resulta especialmente cierto respecto de los sistemas de IA de segunda generación[110]. Analizaremos con más detalle este tipo de complejidad al explicar, en el acápite 2.1. el problema de la caja negra.

(iii) Complejidad ecosistémica o diversidad de actores: la pluralidad de actores en un ecosistema digital hace que sea mucho más difícil establecer, con claridad, quién puede ser responsable respecto de determinados daños[111]. En un caso de responsabilidad relativo a un sistema de IA nos encontraremos con un entramado complejo de actores altamente profesionalizados y especializados: programadores, productores, proveedores, y en algunos casos, usuarios finales con incidencia sobre la IA. Tampoco puede descartarse la incidencia "ecosistémica" de otros factores sociales, tales como la influencia política sobre el desarrollo de determinadas aplicaciones de la IA, el sesgo algorítmico, entre otros.

La multiplicidad de complejidades ecosistémicas, en particular aquella relativa a la pluralidad de actores, es relevante para la responsabilidad civil. Como analizaremos más adelante en el Capítulo

---

[108] GRUPO INDEPENDIENTE DE EXPERTOS DE ALTO NIVEL SOBRE INTELIGENCIA ARTIFICIAL EN LA COMISIÓN EUROPEA (2019): 32.

[109] *Ibid*.: 33. Paréntesis agregado.

[110] Ver *supra*, capítulo 1, acápite 1.3 (ii).

[111] GRUPO INDEPENDIENTE DE EXPERTOS DE ALTO NIVEL SOBRE INTELIGENCIA ARTIFICIAL EN LA COMISIÓN EUROPEA (2019): 32.

3, esta multiplicidad de actores podría configurar desafíos particularmente únicos respecto a la relación de causalidad y a la atribución de culpa o dolo.

Además de esto, y adelantando un poco nuestro análisis, la profesionalización de los operadores podría presentar un elemento que cause situaciones de particular injusticia (en contra de los fines generales del sistema de responsabilidad civil) dada la carga que tiene la víctima de probar el daño en un sistema basado en la culpa[112]. Debe considerarse, además, que la pluralidad de actores está presente en todas las fases del ciclo de vida de una IA, y que la identidad jurídica de los mismos puede variar respecto a cada una de dichas fases. Es por este motivo que, como ya se expuso, algunos acercamientos normativos agrupan a todos los actores involucrados bajo un mismo término, como lo hace el proyecto de ley de IA o el AIA al hablar de "operador".

2.2. Opacidad: la opacidad algorítmica y el "black-box problem".

El "*black-box problem"* o "problema de la caja negra" se ha convertido en el verdadero centro de debates acerca de los aspectos éticos, políticos y legales relativos a la adopción del uso de sistemas de IA. En términos simples, este problema se traduce en que estos sistemas se consideran *opacos*, es decir, son difíciles de analizar en el sentido de entender qué es lo que hacen y por qué lo hacen. Esto implica que entender los resultados producidos por los sistemas de IA es mucho más fácil, o simple, que entender el procedimiento por el cual el sistema llegó a aquella solución en particular, incluso para quienes son expertos en la materia[113]. Es decir, la imposibilidad de desentrañar el algoritmo para explicar cómo un sistema de IA llegó a una conclusión determinada puede extenderse incluso a quienes programaron el algoritmo. En otras palabras: si bien los algoritmos de los sistemas de IA se entrenan con un determinado *set* o conjunto de datos para obtener una respuesta en particular, la forma en la que la IA infiere dicha respuesta llega a ser incomprensible para los seres humanos dada su complejidad. Esto es particularmente cierto respecto de quienes resultan ser sus usuarios finales, es decir, quienes no tienen acceso alguno al código fuente utilizado por el sistema de IA. A este respecto, la Comisión Europea ha señalado que:

> "(…) la opacidad es otra característica principal de algunos de los productos y sistemas basados en la IA, que puede derivar de la capacidad de mejorar la ejecución de sus tareas

[112] Ver *infra,* capítulo 3, acápite 3.2 (iii).
[113] BATHAEE (2018): 892.

> aprendiendo de la experiencia (…) pueden caracterizarse por diversos grados de opacidad, lo que puede hacer que el proceso de toma de decisiones sea difícil de determinar[114]."

Siguiendo a YAVAR, una forma de entender por qué se genera el problema de la caja negra es prestando atención a dos atributos esenciales de los algoritmos utilizados en el ML, estos son: (i) la "complejidad" y (ii) la "dimensionalidad". Así, por una parte, tenemos la complejidad estructural del algoritmo que se presenta, por ejemplo, cuando nos hallamos frente a las redes neuronales, donde un algoritmo se conforma por miles de "neuronas" artificiales que trabajan de forma difusa, existiendo muchas capas de "neuronas" interconectadas que se utilizan de forma progresiva para encontrar patrones y efectuar conexiones entre los datos (*data points)* con tal de llegar a una determinada solución[115]. Como resulta evidente, la estructura de las redes neuronales se inspira en la funcionalidad de las neuronas del cerebro humano[116]. En las redes neuronales, de forma similar a cómo entendemos que ocurre en nuestro cerebro, los *inputs* producidos por una neurona artificial se traducen en señales que se transmiten a través de una red compuesta de múltiples otras neuronas artificiales, generándose *outputs* que pueden ser interpretados como respuestas a los *inputs* originales: así, ninguna de las "neuronas" en esta red codifica una parte distintiva del proceso de toma de decisiones, sino que resulta ser un proceso compuesto de múltiples capas o *layers*[117]. Entre esta multiplicidad compleja de capas, pueden existir además capas ocultas, y cuando existen más de dos capas ocultas, se habla entonces de "*Deep learning*[118]." Es esta complejidad la que da lugar al problema de la caja negra, ya que el proceso de toma de decisiones de una red neuronal se basa en el aprendizaje que autónomamente pueden llevar a cabo los sistemas de IA, y que es, por lo tanto, *intuitivo*. Para explicar esta particular cualidad de las redes neuronales utilizadas por sistemas de IA, se ha señalado que: "su conocimiento no puede, en la mayoría de los casos, reducirse a un conjunto de instrucciones, y no es posible, en la mayoría de los casos, señalar a una neurona o grupo de neuronas para determinar lo que el sistema señaló como interesante o importante"[119].

Por otra parte, existen algunos tipos de algoritmos que descansan en relaciones geométricas que los seres humanos no son capaces de visualizar[120]. Esto se ha denominado el "problema" o "maldición"

---

[114] COMISIÓN EUROPEA (2020).
[115] BATHAEE (2018): 902.
[116] *Ibid*.: 901.
[117] *Ibid*.: 902.
[118] BOUCHER (2020): 7.
[119] BATHAEE (2018): 903. Traducción propia.
[120] *Ibid.*

de la dimensionalidad. En términos simples, el problema de la dimensionalidad se explica por el hecho de que en el uso de ML, la cantidad de datos recopilados para llevar a cabo análisis y encontrar patrones forma, por su densidad y tamaño, figuras y estructuras que no son fácilmente comprensibles para el cerebro humano[121].

Puede concluirse, a partir de estos puntos, que la opacidad en los sistemas de IA es un concepto relativo más que absoluto, dependiente de diversos factores. Respecto de la opacidad algorítmica y la responsabilidad civil, se ha señalado que:

> "Cuanto más complejas se vuelven las tecnologías digitales emergentes, menos pueden los que aprovechan sus funciones o están expuestos a ellas, comprender los procesos que pueden haber causado daños a sí mismos o a otros. A menudo, los algoritmos ya no se presentan como un código más o menos legible, sino como una caja negra que ha evolucionado mediante el autoaprendizaje y que podemos probar en cuanto a sus efectos, pero no tanto comprender. Por tanto, cada vez es más difícil para las víctimas identificar esas tecnologías como una posible fuente de daño, por no hablar de por qué lo han causado[122]."

En suma, es la complejidad y el funcionamiento particular de los sistemas de IA lo que conlleva su opacidad (o falta de transparencia) y, por tanto, a la dificultad para los usuarios (y expertos) de entender (y explicar) la forma en la que un sistema de IA llegó a una solución o conclusión en particular.

Una forma en la que se ha intentado contrarrestar el problema de la caja negra (y en cierta forma, el problema de la previsibilidad explicado en el acápite 2.3 del presente capítulo), es mediante el uso y la garantía de que la IA sea explicable o transparente (en inglés, "*Explainable AI*")[123], de tal forma que las decisiones tomadas por una IA particular sean interpretables y entendibles por un ser humano. El grupo independiente de expertos de alto nivel sobre IA de la Unión Europea incluye la explicabilidad como un principio dentro de las "directrices éticas para una IA fiable", señalando que:

---

[121] ROWEIS y SAUL (2000): 2323.
[122] GRUPO INDEPENDIENTE DE EXPERTOS DE ALTO NIVEL SOBRE INTELIGENCIA ARTIFICIAL EN LA COMISIÓN EUROPEA (2019): 33. Traducción propia.
[123] PREECE et. Al. (2018).

> "La explicabilidad es crucial para conseguir que los usuarios confíen en los sistemas de IA y para mantener dicha confianza. Esto significa que los procesos han de ser transparentes, que es preciso comunicar abiertamente las capacidades y la finalidad de los sistemas de IA y que las decisiones deben poder explicarse —en la medida de lo posible— a las partes que se vean afectadas por ellas de manera directa o indirecta. Sin esta información, no es posible impugnar adecuadamente una decisión. No habitualmente resulta posible explicar por qué un modelo ha generado un resultado o una decisión particular (ni qué combinación de factores contribuyeron a ello). Esos casos, que se denominan algoritmos de «caja negra», requieren especial atención. En tales circunstancias, puede ser necesario adoptar otras medidas relacionadas con la explicabilidad (por ejemplo, la trazabilidad, la auditabilidad y la comunicación transparente sobre las prestaciones del sistema), siempre y cuando el sistema en su conjunto respete los derechos fundamentales. El grado de necesidad de explicabilidad depende en gran medida del contexto y la gravedad de las consecuencias derivadas de un resultado erróneo o inadecuado[124]."

Como se desprende de la cita, la explicabilidad no es, desde una perspectiva técnica, una tarea fácil, incluso existiendo perspectivas críticas respecto a la necesidad epistémica de explicabilidad en el contexto de uso de un algoritmo de ML[125]. Dicho esto, el construir modelos de IA que sean perfectamente explicables e inteligibles para los desarrolladores y usuarios finales sigue siendo un problema activamente presente entre los que desarrollan estos tipos de tecnologías[126].

2.3 El sesgo algorítmico

El sesgo algorítmico ha sido un problema reconocido desde hace décadas en la IA y en los modelos de ML. Implica, en su esencia, que en las decisiones tomadas por los algoritmos se repliquen juicios discriminatorios que reflejen aquellos presentes en la sociedad, por lo que su existencia o uso puede llevar a resultados sesgados, con las respectivas consecuencias sociales relevantes que esto conlleva[127]. FRIEDMAN y NISSENBAUM señalan que: "usamos el término sesgo para referirnos a sistemas computacionales que sistemáticamente y de forma injusta discriminan en contra de

[124] GRUPO INDEPENDIENTE DE EXPERTOS DE ALTO NIVEL SOBRE INTELIGENCIA ARTIFICIAL EN LA COMISIÓN EUROPEA (2019): 16.
[125] Véase ROBBINS (2019).
[126] PREECE et al. (2018): 1.
[127] FRIEDMAN y NISSENBAUM (1996): 332.

determinados individuos o grupos de individuos a favor de otros"[128]. El sesgo algorítmico, por lo tanto, se relaciona con variables históricamente vinculadas a la discriminación, como lo son el género, la edad, la ubicación geográfica, la raza, entre otros.

> Ejemplo 3: un sistema de IA que analiza postulaciones a un trabajo elige eliminar sistemáticamente a las candidatas de género femenino, a pesar de no tener ninguna instrucción relativa a tal efecto.

El sesgo algorítmico puede ser (i) preexistente, (ii) técnico o (iii) emergente[129]. Cuando el sesgo es preexistente, este refleja prejuicios existentes en prácticas, actitudes o instituciones sociales, y puede derivar de un individuo (que contribuye sustancialmente al diseño del sistema, como su diseñador o el cliente que mandata su desarrollo), o derivar de la sociedad misma (es decir, desde organizaciones como ciertos sectores de la industria, grupos religiosos, educativos, culturales, etc.)[130] Si el sesgo algorítmico es técnico, este deriva de limitaciones o consideraciones técnicas, como las limitaciones inherentes a herramientas computacionales. Uno de los ejemplos citados por FRIEDMAN y NISSENBAUM es el de un "sistema legal experto" que aconseja a imputados a someterse o no a un *plea bargain* asumiendo que la ley puede interpretarse de forma completamente certera, no estando sujeta a posibles injerencias contextuales relativas a la persona del imputado[131]. Por último, el sesgo algorítmico es emergente cuando surge en el contexto del uso por usuarios reales (humanos), como resultado de cambios en el conocimiento social, en la población, o en los valores culturales[132], creando, como consecuencia, dificultades para un nuevo grupo de usuarios distinto de aquel presente en su primer contexto de uso.

Los posibles daños causados por el sesgo algorítmico han significado un gran desafío para la implementación de la IA, primero, por la dificultad técnica de predecir y mitigar la presencia del sesgo, y segundo, por la dificultad de regular y sancionar el sesgo en la práctica.

Si bien es posible que en aquellos casos en los que el sesgo algorítmico sea el elemento central involucrado en un caso de responsabilidad civil, este sea fácil de identificar y determinar, también resulta posible que el sesgo sea difícil de detectar o, aún más, sea imperceptible; incluso para quienes

---

128 FRIEDMAN y NISSENBAUM (1996): 332.
129 *Ibid*.: 333.
130 *Ibid*.: 334.
131 Debemos tener presente que el texto es del año 1986 y que la industria del *legal tech* ha avanzado con creces desde el ejemplo. Dicho esto, resulta de todas formas ilustrativo.
132 FRIEDMAN y NISSENBAUM (1996): 333-335.

se ven afectados por el mismo. Pensemos en el ejemplo 3: ¿cómo podrían las candidatas saber de aquel sesgo, cuando los procesos de selección de candidatos de trabajos suelen ser bastante opacos? Y, más aún, ¿podría un daño como ese configurar un caso accionable frente a un tribunal? Lo más probable es que esto no sea así, a pesar de que exista un daño efectivo relativo a la pérdida de una oportunidad laboral. Esto resulta aún más grave si consideramos que la razón que yace tras el sesgo algorítmico recae en patrones de discriminación que se encuentran presentes profundamente en la sociedad, tal como es el caso de la discriminación basada en el género. Como veremos más adelante, es posible que el sesgo algorítmico también implique desafíos bastante interesantes en la configuración de la culpa y en la distribución de la responsabilidad entre los actores involucrados, lo que se analizará en el Capítulo 3.

Siguiendo esta línea de pensamiento, GLAUBITZ señala que:

> "Los acercamientos para tratar con los daños algorítmicos han descansado principalmente en detectar el sesgo, principalmente a través de la deducción, retrospectivamente determinando si un set de datos, modelo, *output* o sistema algorítmico es sesgado. Esto incluye auditorías de datos de entrenamiento, la revelación de códigos fuente, así como auditorías sobre los sistemas y los códigos fuente[133]".

Sumado a esto, la autora es enfática en señalar que, aunque la transparencia algorítmica[134] pareciera una buena solución al sesgo algorítmico, la misma podría tener como consecuencias negativas (i) la divulgación de información confidencial, y (ii) propiciar que determinados individuos aprendieran a "ganarle al sistema[135]". Para ilustrar este último punto, señala como ejemplo:

> "(…) si las personas conocen qué variables son las que utiliza la IRS (servicio de impuestos internos estadounidense) para determinar la evasión de impuestos, entonces los evasores podrán ajustar su comportamiento con tal de que las señales originales de evasión pierdan su valor predictivo[136]."

Otros acercamientos proponen buscar y detectar comportamientos "localizados" en determinados algoritmos[137], con tal de ofrecer una explicación sobre cómo se produce el sesgo, pero aún resulta

---

133 GLAUBITZ (2021): 10.
134 Entendida como la práctica de divulgar el código fuente mediante el cual operan los sistemas de IA o algoritmos, con el fin de que sea posible la aquiescencia de los mismos por terceros.
135 GLAUBITZ (2021): 10.
136 KROLL en GLAUBITZ (2021): 11.
137 NTOUTSI et al. (2019).

un terreno especulativo, ya que, en términos técnicos[138], el descifrar y evitar el sesgo algorítmico, se relaciona fuertemente con las dificultades relativas al fenómeno de la caja negra descrito en el acápite 2.1.

A modo de conclusión, existe debate sobre la forma de abordar los problemas relacionados con el sesgo algorítmico, a pesar de su amplio reconocimiento y temprana detección en el campo, y es un problema penetrante y sin solución definitiva hasta el momento.

2.4 Autonomía

La autonomía de la IA se define como la capacidad que este tipo de sistemas tiene para llevar a cabo tareas o tomar decisiones con cada vez menos o sin ningún tipo de *input* o supervisión por parte de un ser humano. Los sistemas de IA pueden desarrollar cada vez más tareas con menos control por parte de un ser humano, incluso prescindiendo de cualquier tipo de intervención voluntaria de nuestra parte. Ciertos sistemas de IA son capaces de alterar el algoritmo inicial que se programó en ellos por parte del diseñador o equipo de diseñadores de *software* de forma autónoma, debido a las amplias capacidades de autoaprendizaje que poseen[139]. Para esto nos sirve recordar la idea del agente racional: la forma de escoger qué datos se priorizan y qué impacto tienen dichos datos sobre el resultado final es algo que puede ajustarse de forma constante por los mismos algoritmos en razón de los determinados fines para los cuales han sido elaborados[140]. La autonomía de la IA puede variar, desde aplicaciones que requieren completa supervisión humana hasta herramientas que no requieren supervisión humana alguna. Un ejemplo de este último tipo de sistema son los vehículos autónomos, tecnología que, a pesar de no hallarse adoptada por la generalidad de la población o regulada, ya existe y se encuentra en fases de desarrollo y prueba[141].

La autonomía es la característica que de forma más obvia separa a los sistemas de IA de otro tipo de tecnologías. Otras manifestaciones de esta autonomía frente a tareas complejas (aparte del ejemplo ya citado de pilotear un vehículo) pueden hallarse como ejemplo en sistemas de IA

138 Y, por tanto, probatorios.
139 GRUPO INDEPENDIENTE DE EXPERTOS DE ALTO NIVEL SOBRE INTELIGENCIA ARTIFICIAL EN LA COMISIÓN EUROPEA (2019): 33.
140 *Ibid.*
141 PEREZ RASTELLI (2012).

utilizados para monitorear un determinado sector productivo o para generar un portafolio de inversión[142], como señalamos, sin supervisión humana.

Frente al enigma de un sistema de IA respecto del cual existe una falta de completa supervisión humana, ¿quién es responsable si un algoritmo se "desvía" u opera de forma "incorrecta" de forma que no sea imputable al hecho de un ser humano en particular? Analizaremos dichos problemas en el Capítulo 3. Dicho esto, resulta necesario prevenir en este punto que la evaluación del comportamiento de un sistema de IA como correcto o incorrecto desde una perspectiva exclusivamente técnica puede no estar dado por su desempeño (es decir, puede no derivar de un error técnico *per se*), sino que por cualidades que, como seres humanos, asignamos a los resultados obtenidos por el mismo[143].

2.5. Imprevisibilidad

La previsibilidad es un concepto importante y necesario para establecer una imputación en el contexto de la responsabilidad civil, particularmente respecto de la culpa como factor de imputación. Dicho esto, una característica definitoria de los sistemas de IA ha sido su imprevisibilidad, la cual depende del ya mencionado *black-box problem* y del hecho que en muchos casos los algoritmos son capaces de hallar "soluciones" a problemas de forma imprevista para sus diseñadores originales. SCHERER ejemplifica la imprevisibilidad de la siguiente forma:

> "Una característica importante de la IA que plantea un reto a los sistemas jurídicos es el concepto de previsibilidad. Ya hemos visto numerosos casos de IA diseñada para actuar de forma que parece creativa, al menos en el sentido de que las acciones se considerarían 'creativas' o como una manifestación de pensamiento 'fuera de lo común' si las realizara un ser humano. Algunos ejemplos ampliamente conocidos de este fenómeno son los programas informáticos de ajedrez, los cuales pueden hacer jugadas que van en contra de los preceptos básicos de la estrategia ajedrecística humana[144]."

---

[142] SCHRERER (2015): 363.

[143] Respecto a este punto, nos referimos a que los resultados obtenidos por la Inteligencia Artificial que resultan moralmente erróneos desde la perspectiva humana no necesariamente responden a un error técnico o de procesamiento, sino a un enfoque, por ejemplo, demasiado versado en datos puros relativos a la productividad. Se ha dicho que esta es la razón de por qué algunos sistemas de IA discriminan y poseen un sesgo, es porque la sociedad misma refleja tal sesgo, criterio moral que hasta el día de hoy no puede implementarse en este tipo de sistemas a la hora de discernir y procesar la información.

[144] SCHRERER (2015): 363.

Más aún, se ha señalado que esta característica no solo afecta a sistemas pre-programados para ser "creativos", por ejemplo, en los casos en los que dichas soluciones se dan producto de la interacción con estímulos posteriores o sobrevinientes: "muchos sistemas están diseñados no para responder a estímulos predefinidos, pero para identificar y clasificar estímulos nuevos y vincularlos a una reacción correspondiente que no ha sido pre programada. Cuantos más datos puedan ser procesados por los sistemas y mientras más IA sofisticada posean, más difícil será prever el impacto que tengan una vez estén operando[145]."

La imprevisibilidad conlleva, en suma, a que las acciones de los sistemas de IA puedan ser difíciles o imposibles de prever, si bien hasta ahora los casos más notables no han sido muchos[146]. Sin embargo, su efecto es que los diseñadores de IA no puedan tener claridad exacta acerca de cómo los sistemas de IA puedan tener efectos no previstos por los mismos, incluso si hubo o existió otro efecto que sí fue predicho por el equipo desarrollador[147].

2.6. Conclusiones

Lo presentado en este capítulo nos permite concluir que existen determinadas características intrínsecas a los sistemas de IA que merecen especialmente nuestra atención. Si bien hemos señalado que la mayoría de estos problemas son técnicos, no por eso dejan de ser menos relevantes desde una perspectiva jurídica, y, particularmente, desde la responsabilidad civil. En efecto, y sin perjuicio de las bondades que este tipo de tecnologías han traído consigo respecto de la agilidad de procesamiento de grandes cantidades de información, no podemos dejar de ver que muchas de estas problemáticas no podrán, sino aumentar su relevancia conforme los sistemas de IA se utilizan de forma cada vez más masiva. No será lo mismo, como emana de lo expuesto, enfrentarnos a un accidente provocado por un sistema de IA completamente autónomo que enfrentarnos a un accidente donde tenga incidencia cualquier otro tipo de tecnología, precisamente por la ubicuidad de estas problemáticas en los sistemas de IA. Además de esto, resulta patente que, si bien algunos de estos desafíos recaen en aspectos exclusivamente técnicos, otros de ellos se encuentran íntimamente relacionados con el uso que, como seres humanos, damos a los sistemas de IA. Con esto último nos referimos, por ejemplo, al hecho de que al hablar de sesgo algorítmico, necesariamente hablamos de sesgos que ya existen y que permean nuestra sociedad. También es

[145] SCHRERER (2015): 363.
[146] *Ibid*.: 366.
[147] *Ibid.*

posible aquí referirnos nuevamente a los principios éticos delineados en el Capítulo 1, los cuales pueden sernos de ayuda a la hora de reflexionar sobre el uso que damos a este tipo de sistemas.

PARTE II

DERECHO

CAPÍTULO 3: LA RESPONSABILIDAD CIVIL Y EL DERECHO DE DAÑOS O RESPONSABILIDAD EXTRACONTRACTUAL EN EL ORDENAMIENTO JURÍDICO CHILENO

Como institución jurídica, la responsabilidad civil responde a la pregunta sobre la reparación por el daño o perjuicio sufrido de forma injusta a través de la imputación de una obligación de reparar pecuniariamente el daño a quien lo ha causado[148]. En este sentido, BARROS caracteriza la responsabilidad civil como aquella que establece los criterios y requisitos para que los diversos daños que enfrentamos en nuestra vida puedan ser normativamente atribuidos a un tercero que los ha provocado, de modo que se justifique otorgar a la víctima una acción reparatoria[149]. Tradicionalmente, la doctrina ha dividido la responsabilidad civil en dos grandes estatutos: la responsabilidad civil contractual, por un lado, y la responsabilidad civil extracontractual, por el otro. La responsabilidad civil contractual se caracteriza por la existencia de una obligación o vínculo contractual que une a las partes, y que, en caso de existir incumplimiento respecto del mismo (o de aplicarse otros remedios contractuales), da origen a la acción para exigir su cumplimiento forzado o la resolución, junto con la acción indemnizatoria de perjuicios[150]. Lo fundamental en este tipo de responsabilidad es, entonces, la existencia de una obligación previa contraída por las partes. FIGUEROA YÁÑEZ resume esta idea señalando que:

> "La responsabilidad civil contractual viene a ser, en consecuencia, un efecto de las obligaciones que nacen de los contratos, específicamente, un efecto de las obligaciones en su incumplimiento, efecto que conjuntamente con el cumplimiento forzado configuran los denominados remedios legales en caso de incumplimiento[151]."

La responsabilidad civil extracontractual, en cambio, halla su fundamento normativo en la existencia de una cláusula general de responsabilidad por culpa o dolo[152], prescindiendo de la existencia de

148 MAZEAUD en ALESSANDRI RODRÍGUEZ (2005): 13.
149 BARROS (2020): 18.
150 *Ibid*.: 20; artículo 1489 del Código Civil.
151 FIGUEROA YÁÑEZ (2011): 56.
152 BARROS (2020): 21.

vínculo previo alguno entre las partes. La noción de una imputación basada en una cláusula general de responsabilidad por culpa o dolo se vincula estrechamente con la existencia de un *deber general de cuidado* o *de no dañar* (también llamado principio *alterum non ladere*), dadas las consecuencias que, en nuestra sociedad, traen consigo los hechos ilícitos o dañosos[153]. Es a partir de la infracción de dicho deber general de cuidado, por tanto, que nace la obligación de reparar el daño causado a la víctima. La finalidad de la responsabilidad civil extracontractual, en suma, radica en asegurar la reparación efectiva de todos los daños que, en el interactuar social humano, ocurran en infracción a la cláusula general de responsabilidad por culpa o dolo, siempre que se cumplan los requisitos legales para llevar a cabo el juicio de responsabilidad[154].

Si bien ha existido debate respecto de la utilidad de la distinción entre los sistemas de responsabilidad, es posible señalar que el punto central que distingue a un estatuto de responsabilidad del otro es la voluntad contractual de obligarse. Siguiendo este orden de ideas, BARROS enfatiza al respecto que:

> "la fuente de responsabilidad contractual es la *convención* que opera como un mecanismo de autovinculación, mientras que en sede extracontractual lo determinante es el *derecho*, que heterónomamente pone límites y establece consecuencias patrimoniales al ejercicio de nuestra libertad cuando este da como resultado un daño a un tercero[155]."

En el presente capítulo, se revisarán de forma pormenorizada cada uno de los elementos del juicio de la responsabilidad civil extracontractual, y se propondrán diversas formas en las que las problemáticas relativas a los sistemas de IA (estudiadas a lo largo del Capítulo 2), pueden incidir sobre cada uno de ellos y generar interrogantes novedosas para la práctica jurídica.

Resulta necesario señalar que se le dará prioridad a este estatuto de responsabilidad por dos grandes razones. En primer lugar, ya que la naciente normativa internacional reconoce e identifica la existencia de posibles problemas relativos a la responsabilidad extracontractual, particularmente en los sistemas que se rigen por un modelo de atribución basado en la culpa, tal como expusimos *supra*[156]. En segundo lugar, porque en atención a las características propias de los sistemas de IA detallados a lo largo de los capítulos 1 y 2, es posible que los daños causados por el uso de los

[153] Debemos tener presente que esta concepción de "hecho ilícito" es aquella que se entiende respecto del derecho civil y no respecto del derecho penal, véase CHOTZEN en FIGUEROA (2011): 67 y ss.
[154] CORRAL TALCIANI (2003): 65.
[155] BARROS (2020) 23.
[156] Capítulo 1, acápite 1.5 (ii).

mismos no se produzcan dentro del marco de un vínculo obligacional previo entre las partes. Es más, es posible que los daños causados por sistemas de IA comprendan un espectro tan amplio que, incluso, las víctimas desconozcan el hecho de que se esté utilizando este tipo de tecnologías de un modo que les afecte[157].

Analizaremos, entonces, cómo la IA problematiza y, en ciertos ámbitos, torna difícil la aplicación de cada uno de los diversos elementos del juicio analítico de este tipo de la responsabilidad civil extracontractual.

3.1. La responsabilidad civil extracontractual en la normativa chilena

En la responsabilidad extracontractual, como veníamos señalando, la obligación de indemnizar a la víctima se basa en la obligación de reparar el daño que nace por la comisión de un delito o cuasidelito civil[158], conforme a los artículos 2314 y 2284 del Código Civil. Este tipo de responsabilidad halla su fundamento en el resguardo del ejercicio armónico de las diversas libertades humanas en el tráfico jurídico[159]. La responsabilidad extracontractual es imputable, precisamente, a quien cause daños actuando con culpa o negligencia, de conformidad con lo dispuesto en los artículos 2284, 2314 y 2329 del Código Civil. La regla general en el sistema chileno corresponde a la atribución de responsabilidad a partir de la existencia de un factor o elemento subjetivo de imputación (culpa o dolo), modelo de atribución que en doctrina se conoce como "responsabilidad por culpa" o "responsabilidad por negligencia". Este modelo de atribución se contrasta con la llamada "responsabilidad estricta", que en nuestro derecho solo recibe una aplicación limitada a casos dispuestos en la ley.

La responsabilidad estricta, a diferencia de la responsabilidad por culpa, prescinde del requisito de la culpa o dolo como factor subjetivo de imputación, ya que su fundamento se halla en la necesidad de ofrecer una justa compensación a la víctima que sufra daños causados por la ejecución de una actividad que el legislador considera particularmente riesgosa. Respecto de este último punto, puede resumirse que para la responsabilidad estricta: "(…) quien crea el riesgo debe sufrir las consecuencias si el riesgo llega a producir un daño; quien con su actividad irroga un daño, debe repararlo, haya o no dolo o culpa de su parte[160]." Así, la responsabilidad estricta se encuentra

[157] Ver *supra*, capítulo 1, acápite 1.5 (i).
[158] CHOTZEN en FIGUEROA YAÑEZ (2011): 56.
[159] BARROS (2020): 66.
[160] CORRAL TALCIANI (2011): 89.

radicada en determinados sectores de actividad donde "parece inconveniente poner de cargo de la víctima el probar la negligencia de alguno de los agentes que intervienen en la producción del daño[161]." Este tipo de responsabilidad se regula de forma expresa[162]/[163] respecto de determinadas actividades riesgosas, como lo son los accidentes causados por aeronaves, accidentes nucleares, entre otros[164]. Resumiendo, ya que solo a partir de ley expresa resulta procedente el sistema de responsabilidad estricta, es posible concluir que el régimen general de responsabilidad en el derecho chileno es la responsabilidad por culpa.

Los casos de ejemplo que se expondrán en esta memoria, por tanto, se analizarán a través de la exposición de cada uno de los elementos necesarios para llevar a cabo el juicio de responsabilidad civil por culpa, ya que, al no contar con una regulación que establezca un modelo de atribución de responsabilidad estricta respecto de la utilización de la IA[165] (sin perjuicio que en el futuro este pueda ser el acercamiento que el legislador decida dar respecto de determinados usos y aplicaciones de esta[166]), resulta procedente la aplicación del régimen general y supletorio dispuesto en el Código Civil.

3.2. Los elementos del juicio de responsabilidad civil extracontractual por culpa o negligencia.

Según la doctrina[167], los elementos o requisitos de la responsabilidad civil por culpa o negligencia son las siguientes: (i) la existencia de un hecho imputable voluntario, (ii) dolo y culpa, (iii) daño y (iv) relación de causalidad[168].

161 CORRAL TALCIANI (2011): 222.
162 *Ibid.*
163 Parte de la doctrina ha considerado que el artículo 2326 del Código Civil contiene un caso de responsabilidad estricta, al señalar: "El daño causado por un animal fiero, de que no se reporta utilidad para la guarda o servicio de un predio, será siempre imputable al que lo tenga, y si alegare que no le fue posible evitar el daño, no será oído."
164 CORRAL TALCIANI (2011): 223.
165 Dicho esto, nos referiremos brevemente a algunas propuestas al respecto en derecho comparado.
166 El establecimiento de un régimen de responsabilidad estricta u objetiva para las aplicaciones o sistemas de IA es un tema fuertemente debatido en la actualidad, con grandes proponentes y opositores. Sus detractores señalan que dicho sistema es demasiado rígido y que, en realidad, sería contrario al progreso científico, a la innovación, mientras que quienes abogan por un régimen de responsabilidad estricta apuntan a las grandes dificultades para establecer responsabilidades específicas y determinadas mediante las normas o reglas tradicionales.
167 Entre otros: BARROS, CORRAL TALCIANI, ALESSANDIR R, RAMOS PAZOS, entre otros.
168 BARROS, CORRAL TALCIANI, ALESSANDRI R., entre otros.

(i) El hecho voluntario.

Es necesario, para que exista responsabilidad civil, que exista un acto humano, el que puede consistir en una acción u omisión. El hecho voluntario puede descomponerse analíticamente en dos elementos: uno externo y uno interno, donde el elemento externo supone la conducta del sujeto que se expresa materialmente por un comportamiento positivo (acción u omisión acometida por el sujeto), mientras que el elemento interno se refiere a la dimensión subjetiva que dicho acto supone o, en otros términos, la libertad del sujeto para actuar[169]. Ambos elementos llevan a que la conducta sea atribuible al sujeto como su hecho. Así, la doctrina considera que solo mediante un hecho humano puede surgir la responsabilidad:

> "(…) se necesita que el hecho o acto sea originado en la voluntad del ser humano. Solo las personas, y como tales, con su inteligencia y voluntad pueden incurrir en responsabilidad, (…) debe exigirse que ese hecho voluntario contraste con el derecho, es decir, sea injusto o ilícito desde un punto de vista objetivo[170]."

Por supuesto, el legislador no pensó en máquinas que pudieran poseer un intelecto similar al humano. Más bien, la exigencia de que exista un hecho voluntario para que pueda atribuírsele a una determinada persona responsabilidad civil tiene su fundamento en que tal hecho no haya sido producto de un acontecimiento que se encuentre, en su totalidad, fuera de la voluntad humana, como ocurre con el caso fortuito o fuerza mayor[171], y que, por tanto, no sea imputable a la culpa o dolo de quien causa el daño. En el caso fortuito, por ejemplo, es precisamente esta ausencia de voluntad humana lo que configura dicha institución como un eximente de la responsabilidad, en particular, al excluir la culpa[172], sin perjuicio de que su incidencia pueda excluir asimismo el nexo causal, tal como ha señalado la doctrina reciente[173].

Podría pensarse que, en aquellos casos en los que nos encontremos frente a un daño producido por sistemas de IA completamente autónomos, estaríamos frente a un escenario donde sería imposible establecer un hecho voluntario específico que pueda ser atribuible a un humano. Sin embargo, lo correcto es centrar el análisis de dichos problemas de conexión entre la aparente falta de acción

169 BARROS (2020): 70.
170 CORRAL TALCIANI (2003): 223.
171 Artículo 45 CC.
172 SAN MARTÍN (2021): 8.
173 *Ibid*.: 10.

humana (o la dificultad para asignar a un humano determinado una acción voluntaria particular que sea dañosa) y el resultado dañoso respectivo, al análisis a los juicios de culpa o dolo, por una parte, y del nexo causal, por otra. Es en esta última donde encontraremos las discusiones más activas en cuanto a atribución de responsabilidad en IA, como se detallará más adelante en este capítulo.

Respecto del requisito de voluntariedad del hecho propiamente tal como elemento del juicio de responsabilidad, de acuerdo con los principios relativos a la IA dispuestos en la recomendación sobre la ética de la IA de la Unesco, siempre deberá existir una persona humana o una personalidad jurídica responsable, de conformidad con el principio y requisito de supervisión y decisión humanas[174]:

> "Los Estados miembros deberán velar porque siempre sea posible atribuir la responsabilidad ética y jurídica en cualquier etapa del ciclo de vida de los sistemas, así como en los casos de recurso relacionados con sistemas de IA, a personas físicas o entidades jurídicas existentes. La supervisión humana se refiere, por tanto, no solo a la supervisión humana individual, sino también a la supervisión pública inclusiva, según corresponda."

Por su parte, los principios de la OCDE, en su eje referido a la responsabilidad y rendición de cuentas, señalan que[175]:

> "La responsabilidad ética y la obligación de rendir cuentas de las decisiones y las acciones basadas de alguna manera en un sistema de IA siempre deberían ser atribuibles, en última instancia, a los actores de la IA conforma a la función que tengan en el ciclo de vida del sistema de IA."

En suma, existen dos principios de los cuales podemos servirnos para enfrentar varios de los desafíos que presentan los sistemas de IA en el contexto de la responsabilidad civil, tanto respecto del hecho voluntario como de los subsiguientes elementos del juicio de responsabilidad extracontractual: el principio ético de supervisión humana (que recoge la normativa de Unesco) y el principio de responsabilidad (expresado en la normativa de la OCDE). Refiriéndose a dichos principios y proponiendo una interpretación armónica de ambos, BUSTOS señala que:

[174] UNESCO (2021): 21.
[175] OCDE (2022): 8.

> "La interpretación armónica del principio de intervención humana y de responsabilidad aplicados a la responsabilidad civil implica asumir que la decisión, recomendación o predicción que emana de un sistema de IA debe ser siempre atribuible a los actores de la IA que han debido controlarla o que han decidido aceptar el resultado de la decisión o acción de esta, conforme a la función que tengan en el ciclo de vida del sistema de IA y, por lo tanto, corresponde a los actores de la IA responder por los daños que se causen a las personas y que surjan como resultado de la falta de control sobre el funcionamiento de la IA en todo su ciclo de vida o de aceptar la decisión, recomendación o predicción basada de alguna forma en un sistema de IA[176]."

Respecto de la aplicación del principio ético de intervención humana como un criterio de interpretación, aplicación e integración de las normas de responsabilidad vigentes[177], BUSTOS propone dos argumentos que justifican su utilización: el primero de estos argumentos concibe el principio ético de intervención humana como un principio general del derecho, ya que "a pesar de no estar positivizado dentro de la normativa interna nacional se encuentra dotado de aceptación y ha sido consolidado por esta, especialmente por organismos internacionales como la OCDE y UNESCO, por la doctrina nacional y extranjera, e incluso por la normativa reglamentaria interna chilena a través de la Circular interpretativa del SERNAC[178]." El segundo argumento, que se plantea como respuesta a quienes no consideren dicho principio ético como un principio general del derecho, apunta a considerarlo, de todas formas, como un criterio ético útil para la interpretación e integración de lagunas legales. Esto, en razón de que la consagración que en las normativas citadas se otorga al principio, expresa que el mismo cumple una función de valor por parte de la comunidad, pudiendo ser un elemento útil de interpretación en vista del deber de inexcusabilidad que la Constitución establece para los jueces. Esto tiene sentido, ya que en estos casos se hallarán frente a una laguna legal frente por la falta de una legislación que actualmente regule el sistema de responsabilidad civil relativo a la IA[179]. Desde esta óptica, la aplicación de principio ético de supervisión humana está estrechamente conectada con la aplicación de los principios de equidad:

[176] BUSTOS (2024): 5.
[177] *Ibid*.: 5.
[178] BUSTOS (2024): 6.
[179] *Ibid.*

> "(…) es posible reconducir la imputación de la conducta que causa daño al sujeto que debió ejercer control sobre la IA o sobre la decisión que surge como resultado de su utilización, pues, en un escenario contrario, los daños causados quedarán sin ser resarcidos (…)[180]."

De la aplicación de este principio, es posible concluir que, respecto a la voluntariedad del hecho, la atribución de dicha voluntariedad podrá remitirse, muchas veces, a la elección de utilizar, producir, o poner a disposición de otra persona un sistema de IA. Dependiendo de cuál sea el caso, dicha decisión será nuestro punto de partida para analizar la responsabilidad civil. Incluso en determinadas hipótesis en que resulte dudosa la actuación de un ser humano, es necesario aplicar el principio ético de intervención humana, debiendo dicha voluntariedad reconducirse a un ser humano o a una personalidad jurídica, según corresponda, con tal de asegurar que la víctima del daño resulte indemne.

(ii) Capacidad

Más allá del hecho voluntario propiamente tal, es necesario que este sea subjetivamente imputable, es decir, que quien lo lleva a cabo sea capaz. Esta capacidad implica que quien causa el daño "tenga un grado mínimo de aptitud de deliberación para discernir lo que es correcto y lo que es riesgoso[181]." En este sentido, el Código Civil señala en su artículo 2319 que no son responsables de los daños los incapaces, al no tener capacidad de cometer un delito o cuasidelito civil los menores de siete años ni los dementes. Esto, sin perjuicio de que pudieran resultar responsables de los daños causados por ellos las personas a cuyo cargo dichos incapaces estén, si pudiere imputárseles negligencia.

No ahondaremos mucho más en este requisito de la responsabilidad, puesto que para la materia que nos atañe no es, hasta el momento, un elemento que se vea particularmente controvertido. La razón de esto es que, como ya se ha dicho, el estado histórico en el que nos encontramos respecto de los sistemas de IA implica que de una forma u otra se utilizan como herramientas por seres humanos para determinados fines, más allá de que en ese proceso pueda existir un determinado grado de autonomía respecto de quienes las utilizan. En ese sentido, siempre podrá establecerse una cadena (que, como veremos más adelante, puede ser bastante tenue y problemática para efectos de la responsabilidad civil extracontractual), entre los operadores y usuarios de los sistemas de IA, por

[180] BUSTOS (2024): 6.
[181] BARROS (2020): 70

lo que siempre podrá reconducirse la pregunta sobre la capacidad del hecho dañoso a un ser humano en específico[182]. Sumado a esto, dicha conexión podrá establecerse en todas las etapas del ciclo de vida de los sistemas de IA.

Además de lo ya señalado, y nuevamente en atención al desarrollo actual de la tecnología, es difícil concebir a los sistemas de IA (incluso a aquellos más avanzados) como susceptibles de ser consideradas como sujetos capaces bajo las reglas del Código Civil. Los sistemas de IA no están dotadas de personalidad jurídica, por ejemplo, y aun considerando su autonomía e imprevisibilidad, no es posible aseverar, por el momento, que su aptitud de discernimiento sea igualable al de un ser humano[183], con tal de que estos sistemas puedan ser considerados como sujetos de derecho susceptibles de ser considerados capaces para la regulación del Código.

a. La personalidad jurídica de las IA como una alternativa a la atribución de responsabilidad.

Sin perjuicio de lo señalado hasta este punto, existe un intenso debate relativo a atribuir capacidad a sistemas de IA, a través de considerarlos como personas jurídicas[184]. Esta pareciera una solución simple para los problemas que generan los sistemas de IA a los diversos estatutos de responsabilidad civil extracontractual, ya que la acción indemnizatoria podría dirigirse directamente contra el sistema de IA involucrado en el daño, el cual respondería patrimonialmente como personalidad jurídica[185]. Sin embargo, este acercamiento resulta problemático, ya que el otorgarle la calidad de personalidad jurídica a los sistemas de IA implica reconocerlos como sujetos de derecho. Respecto a esto, se ha señalado que, sin perjuicio de que los sistemas de IA cuenten con diversos atributos como la capacidad de comunicarse, alcanzar objetivos específicos, así como cierto grado de autonomía, creatividad e imprevisibilidad; la respuesta a este debate dependerá, definitivamente, de si en el

[182] Veremos más adelante, en el capítulo 3, acápite 3.2. (iii) las dificultades de determinar, en específico, a qué ser humano corresponde responder.
[183] Podrá existir discrepancia respecto a esta aseveración, sin embargo, es pertinente considerar que la regulación jurídica del código civil considera la voluntad humana como una piedra angular de su sistema y que, por lo tanto, es posible advertir que, si considerásemos a la IA como iguales a los seres humanos, tendríamos que considerar que son capaces de asumir actos voluntarios tales como los entiende la ley. Como agentes racionales, las IA pueden tomar decisiones considerando la información a mano, sin embargo, son completamente neutras a las consideraciones morales (en sentido amplio) de las mismas.
[184] ARAYA (2020): 263.
[185] Véase https://policyreview.info/articles/analysis/civil-legal-personality-artificial-intelligence-future-or-utopia

futuro los sistemas de IA serán considerados o no capaces de ejercer derechos y contraer obligaciones[186]. Respecto de esta idea, ARAYA explica que:

> "Por de pronto, expertos de la Unión Europea publicaron una carta abierta, en que solicitaban a la Comisión de la Unión Europea desistir de la iniciativa de otorgar «personalidad electrónica» a los sistemas avanzados de inteligencia artificial, dado que el otorgamiento de un estatus jurídico para los sistemas de inteligencia artificial implicaría reconocer derechos fundamentales como integridad y dignidad a entes que no lo tienen (…)[187]."

Dicho esto, dicha doctrina no puede aplicarse a nuestro país debido a que, a la fecha, no existe ley alguna que regule a los sistemas de IA como personas jurídicas.

(iii) Culpa o dolo

Es en la culpa donde encontramos una discusión doctrinaria más profunda e interesante relativa a los sistemas de IA y responsabilidad civil. Para poder llevar a cabo un juicio de responsabilidad civil, es necesario poder dirigir un juicio de reproche personal a un actor humano[188], a través de la comparación del actuar del causante del daño con un determinado estándar de comportamiento esperable. En la responsabilidad civil en general, la culpa resulta ser dicho estándar general de responsabilidad, siendo un criterio genérico que comprende tanto el ilícito intencional (dolo) y el no intencional (negligencia)[189]. El ilícito no intencional, o negligencia, puede definirse como la inobservancia del cuidado debido (o de un determinado estándar de diligencia) en una conducta susceptible de causar daño a otros[190]. Por su parte, el ilícito intencional implica, necesariamente, una conducta positiva del sujeto. Así, la culpa refleja un límite genérico de las acciones socialmente permitidas, conformando un estándar de comportamiento que tiene su fundamento en las expectativas legítimas de comportamiento recíproco que pueden existir en una sociedad[191]. La Corte Suprema ha señalado que:

---

186 ARAYA (2020): 263.
187 *Ibid*.: 263.
188 UNESCO (2021) y OCDE.
189 BARROS (2020): 82.
190 *Ibid*.: 84.
191 ALESSANDRI RODRÍGUEZ (2005).: 14

> "(…) el estándar de cuidado debido dependerá, esencialmente, del deber de previsibilidad de los daños que se siguieron de la acción, es decir, se declarará responsable al agente cuando se responda afirmativamente a la pregunta de si un hombre diligente, colocado en la misma situación y con calificaciones similares a las del demandado, habría debido prever la ocurrencia del daño que se reclama, y en consecuencia, actuar de otra forma[192]."

Este estándar de culpa corresponde, precisamente, al establecido al art. 44 del Código Civil, asimilable al raciocinio de un "buen padre de familia" o, en otros términos, a lo que un hombre (o persona) razonable hubiera llevado a cabo en una determinada situación.

Frente a la necesidad de que la culpa contraste con un estándar de cuidado o un deber general de no dañar a otros (principio *alterum non ladere*), los sistemas de IA presentan problemas particularmente agudos en los sistemas de responsabilidad civil basados en la culpa. Tres grandes desafíos que mencionamos en el Capítulo 2 son relevantes aquí: (i) la imprevisibilidad, (ii) el problema de la caja negra y, por último, (iii) la autonomía.

El dolo, por su parte, actúa como condición para atribuir la responsabilidad en casos de que el ilícito sea intencional[193]. Asimismo, puede ser positivo o negativo, siendo positivo cuando consiste en la ejecución de un hecho, como herir a otro o apropiarse de lo ajeno[194]. Siguiendo a BANFI, el dolo como elemento intencional tiene una evidente repercusión en la causalidad y, por ende, de la extensión de la indemnización de perjuicios[195]. Si bien ha existido un debate acerca de la relevancia del dolo en la responsabilidad civil[196], creemos que en el caso de los sistemas de IA es particularmente relevante, ya que en los casos en los que exista dolo, esto permitirá establecer una conexión muchísimo más directa con el ilícito civil.

a. Ausencia de culpa en sistemas autónomos

Respecto a la culpa y los sistemas de IA, una primera dificultad en establecer un juicio apropiado de culpa o negligencia frente a una acción u omisión determinada, surge a partir de la característica "imprevisibilidad" los mismos. Esto, sumado a su autonomía (pensemos en el caso extremo de sistemas de IA que funcionen sin supervisión alguna o de forma completamente autónoma respecto

192 CORTE SUPREMA. ROL 12656-2019: CONSIDERANDO CUARTO.
193 BANFI (2017): 74.
194 ALESSANDRI RODRÍGUEZ (2005): 122.
195 BANFI (2017): 79.
196 ALESSANDRI RODRIGUEZ (2005): 15.

de sus operadores humanos) llevarían a que fuese, en principio, difícil prever de forma anticipada el comportamiento de los sistemas de IA. En otras palabras, el no poder prever cuál sería el comportamiento esperable de los sistemas de IA tornaría compleja la tarea de construir un estándar razonable de cuidado esperable respecto de los operadores de los mismos, particularmente respecto de las medidas que dichos operadores puedan tomar para prevenir y mitigar los posibles daños relacionados con la autonomía y la imprevisibilidad. Al no poder establecer con claridad dicho deber de cuidado, resultaría lógicamente imposible establecer con certeza cuando estaremos frente a un comportamiento humano que configure una infracción al mismo. Como sabemos, la culpa está íntimamente ligada a la previsibilidad, donde "existe culpa cuando habiendo podido prever el daño el agente no desarrolla la suficiente diligencia para evitarlo[197]". Pero, ¿cómo conformamos el estándar de diligencia debido respecto a sistemas de IA que son, por naturaleza, imprevisibles o autónomos?

En algunas situaciones, podría suceder que la víctima del daño no pueda probar que existió culpa alguna, a menos que pudiese recurrir a un régimen legal específico que facilite la conexión entre el hecho dañoso, el sistema de IA, y una persona responsable, siendo un ejemplo el régimen de productos defectuosos propio del derecho del consumidor[198]. Esto, sin perjuicio de que los principios OCDE y Unesco apunten a que siempre es necesario que un ser humano responda por los hechos de la IA[199], debido a limitaciones técnicas inherentes a la opacidad de los sistemas algorítmicos. Para ilustrar este punto, pensemos en el siguiente ejemplo:

> Ejemplo 4: un sistema de IA completamente autónomo causa daño. El sistema de IA tiene un dueño, el cual lo compró fuera de una relación de consumo a la empresa X. Inesperadamente, el sistema de IA produce un daño.

En este caso, ¿quién debe asumir la culpa, o, en otros términos, quién no ha tenido la debida diligencia?, ¿la empresa, por decidir vender el sistema de IA, o el dueño, quien supongamos la utilizaba en condiciones normales y esperables? No parece justo dejar a la víctima sin indemnización (y tampoco se cumpliría con los principios internacionales ya citados), pero la pregunta interesante

[197] LARROCAU (2007): 103.
[198] SOYER y TETTERBORN (2023): 387.
[199] Véase *supra* en este capítulo, letra (i).

aquí radica en si sería necesaria la atribución de la culpa de ambas partes si no queda claro qué deber de cuidado o la diligencia debida debió tener cada una[200].

Este problema ha llevado a algunos juristas a proponer sistemas o esquemas alternativos de negligencia[201], pero en Chile resta preguntarnos si este sería un caso en el que ambas partes asumirían equitativamente la culpa, pero, ¿bajo qué acción u omisión? Esto último no resulta del todo claro.

La imprevisibilidad de los sistemas de IA, como sabemos, está íntimamente relacionada con su autonomía. En razón de esta característica, el problema de atribución de la culpa también es extensible, en general, a otros sistemas de IA autónomos. La adaptabilidad y autonomía de dichos algoritmos (como se explicó[202]), implica que estos sean capaces de incorporar datos e información a partir de las interacciones que el sistema tenga con los usuarios finales[203], información que podrá ser utilizada por el sistema para modificar autónomamente sus respuestas. En casos en el ciclo de la vida de la IA, donde esta ya ha sido desplegada al mercado, será difícil que exista un constante monitoreo por parte de los operadores respecto de la forma en que cada sistema de IA en particular se adapta a cada uno de los usuarios finales. En otros términos, podría ocurrir que el sistema de IA se adapte y modifique autónomamente de forma posterior a su despliegue por parte del operador, dificultando el control y monitoreo que podría ejercerse por este último.

El problema que genera la imprevisibilidad es doble: por una parte, genera dificultades probatorias para la víctima, al introducir un elemento azaroso que escapa de las nociones tradicionales de diligencia debida; y por otra, dificulta al juez la tarea de determinar el estándar de cuidado abstracto con el que contrastar la conducta material del autor. No resulta fácil construir la conducta que una persona medianamente diligente debió tener frente a situaciones que son, lógicamente, imprevisibles. Al respecto de la imposibilidad de configurar siquiera un estándar de cuidado apropiado al hablar de sistemas de IA, SELBST señala:

> "El concepto de quebrantar un deber de cuidado solo tiene coherencia, sin embargo, si hay un deber de cuidado al que la persona pueda adherirse para prever el daño[204]." Como

[200] Este ejemplo se debe matizar, debido a que determinadas acciones tomadas por una u otra parte podrían causar alteraciones al algoritmo que podrían provocar daño, constituyendo un ejemplo de falta de debida diligencia respecto a los mismos, así como casos en los que el dueño de la IA decide desplegarla fuera de sus usos comunes, etcétera.
[201] Ver SOYER y TETTERBORN (2023); así como la propuesta de la AI Directive europea (artículo 17).
[202] Ver *supra* Capítulo 2, acápites 2.4 y 2.5.
[203] PAGALLO (2013): 117.
[204] SELBST (2020): 1331.

> sabemos, dada su imprevisibilidad, es posible que usar la IA sin que ocurra un error no depende del deber de cuidado que se tome por el usuario[205].

Para ilustrar este punto, el autor se refiere a los vehículos autónomos: mucha literatura se enfoca en ellos, sin embargo, este no resulta ser el único caso que merezca la atención analizar. En muchas ocasiones, no existirá una autonomía completa del sistema de IA, por lo que será necesario que el operador de dicho sistema tome (y tenga la posibilidad de tomar) determinadas decisiones con el fin de evitar la ocurrencia de accidentes. Analizaremos esta hipótesis en el acápite siguiente, pero es pertinente señalar que incluso en los casos en los que exista un grado de control o supervisión humana, existen factores que pueden dificultar la determinación de la culpa.

> Ejemplo 5: Una persona maneja un vehículo semi-autónomo. Por una falla del sistema de IA, que no logra detectar y avisar que existe una situación riesgosa al conductor, este colisiona hiriendo a personas en el paso peatonal, sufriendo igualmente lesiones el conductor.

En casos como el del ejemplo 5, aun cuando existe una supervisión por parte de un operador humano, resulta complejo señalar con exactitud que existió, efectivamente, algún tipo de negligencia culpable, o si es que acaso existió una concurrencia de culpas o varias negligencias (por ejemplo, por parte del usuario como de alguno de los operadores del sistema de IA). Siguiendo con el mismo ejemplo, podríamos suponer que existe una falta de diligencia debida de ambas partes: el descuido del piloto, por una parte, y el error del sistema de IA, por parte del operador. Si esta situación ocurre en un periodo corto de tiempo, ¿es justo señalar que el estándar de cuidado recae enteramente en el conductor, es decir, deberíamos imputar toda la culpa al conductor? Frente a las interrogantes aquí expuestas, es posible concluir que este tipo de casos posiblemente involucrarán la concurrencia de culpas en los daños causados por sistemas de IA, debido a la particular injerencia que el actuar del usuario tuvo en el hecho dañoso. La doctrina ha señalado que, en casos en que existen una multiplicidad de hechos culpables, la responsabilidad de cada uno de los involucrados que hayan incurrido en negligencia es personal, y, por tanto, la obligación indemnizatoria respecto de cada uno de los responsables deberá cubrir el total de los perjuicios sufridos por la víctima[206]. Por otra parte, de acuerdo con el ejemplo que venimos analizando, no es posible concluir que estemos frente a una situación de responsabilidad como la descrita en el artículo 2317 del Código Civil, toda vez que no

[205] SELBST (2020): 1331.
[206] BARROS (2020): 443.

hay, en principio, coparticipación o un fin compartido por quienes intervienen en la producción del daño, tal como este requisito ha sido entendido por la doctrina[207]. Un escenario distinto es el que presenta la indeterminación del nexo causal, que analizaremos en el acápite correspondiente.

Otra pregunta distinta a la ausencia de culpa o a la concurrencia de culpas, es aquella que surge respecto a la determinación del deber de cuidado en sí. En el caso de una conducta negligente por parte del conductor, concurrente con una negligencia de parte del operador de sistema de IA, ¿podría excusarse o hallarse atenuada dicha culpabilidad atendiendo a dicha "falla"[208] del vehículo semi-autónomo, si este pasa a estar *fuera del control del piloto*?, ¿tiene el conductor un deber de cuidado acentuado al pilotear un vehículo semi-autónomo en este caso?

Para aclarar estas interrogantes, resulta necesario precisar que la situación de determinación de la culpa en el ejemplo 5 es diversa de aquella que podríamos plantearnos respecto de un accidente provocado por un defecto mecánico sufrido por un vehículo común (esto es, no autónomo o semi-autónomo).

En el caso de un accidente provocado por una falla mecánica de un vehículo común, es preciso determinar, de acuerdo a los hechos del caso, si la falla mecánica es atribuible a un descuido del fabricante, a un mantenimiento inadecuado por parte de un taller mecánico, o a una actuación negligente por parte del conductor (por ejemplo, en el caso de que haya ignorado advertencias del vehículo respecto del mal funcionamiento). A pesar de que la falla mecánica pueda imputársele a una u otra persona, el accidente, en todos los escenarios, será el resultado de la falla mecánica en sí. En el caso del vehículo semi-autónomo, en cambio, el escenario es distinto, ya que el accidente puede producirse tanto por una falla en el sistema autónomo (en el ejemplo 5, la falla del sistema autónomo llevó a que el conductor no recibiera las señales adecuadas), como por una incorrecta interacción humano-sistema de IA (el conductor del ejemplo 5, a su vez, no detectó ni tomó el control en la situación riesgosa que produjo la colisión). La falla del sistema autónomo, además, puede haberse producido por una conducta negligente de parte de una multiplicidad de actores en el ciclo de vida de la IA. Por ejemplo, puede ser que no haya existido la diligencia debida por parte de uno de los operadores involucrados en su desarrollo, que haya existido un entrenamiento del algoritmo basado en datos defectuosos o sesgados, etcétera. Dadas las características particulares

[207] BARROS (2020): 444.
[208] Dicho esto, no necesariamente habrá una "falla" en los términos usuales: a pesar de que el sistema de IA tome una mala decisión, es posible que no exista, técnicamente, un error.

de los sistemas de IA como la opacidad y la imprevisibilidad, podría ser imposible determinar cuál fue la conducta negligente en el caso en particular, o a qué actor, entre todos aquellos involucrados en el desarrollo y despliegue del sistema de IA, pueda imputársele dicha negligencia.

Resumiendo, en nuestro caso del vehículo semi-autónomo dispuesto en el ejemplo 5, la complejidad relevante resulta ser la dificultad de determinar la culpa o la concurrencia de culpas respecto del operador del vehículo semi autónomo, término que, como sabemos, engloba a la pluralidad de actores que se ven involucrados en el desarrollo del sistema de IA, dada la opacidad e imprevisibilidad del mismo. Esta complejidad se adhiere, además, a la injerencia que pudiese tener la intervención del conductor en la producción del daño.

Respecto de la segunda interrogante, referida al deber de cuidado del que responde el conductor o usuario, la complejidad que conllevan casos como el presentado en el ejemplo 5 es que pueden llevarnos a pensar que quizás exista un deber de cuidado acentuado respecto de las personas que decidan conducir este tipo de vehículos. Los jueces, para determinar el deber de cuidado a través de la construcción de la conducta exigible con tal de precisar si existió culpa, utilizan criterios de sopesamiento de intereses[209], entre ellos, la probabilidad y la intensidad del daño[210]. El hecho de que en el debate que rodea a la regulación de los sistemas de IA se haya propuesto en diversas ocasiones la implementación de un sistema de responsabilidad estricta, sumado a que normativas como el AIA sancionan de forma directa el despliegue de determinados sistemas de riesgo inaceptable, induce a pensar que en la generalidad de los usos de este tipo de sistemas existe un riesgo incrementado. Esto significaría, en suma, concluir que quien decide pilotear un vehículo semi-autónomo asume inexorablemente un riesgo aumentado, solo por el hecho de llevar a cabo dicha actividad. Por supuesto, dicha forma de pensar acercaría el uso de los sistemas de IA a un régimen de responsabilidad estricta. Debido a que, al menos hasta este momento, no puede asegurarse fácticamente que pilotear un vehículo semi-autónomo sea evidentemente más riesgoso que conducir un vehículo común, dicha idea debe descartarse.

Hasta ahora, las interrogantes aquí planteadas no tienen una respuesta clara, y corresponderá a la jurisprudencia delimitar hasta qué punto cada uno de ellos contribuye al accidente.

[209] BARROS (2020): 113.
[210] *Ibid.*

b. Dificultad de establecer la debida diligencia o el deber de cuidado.

Otra posible problemática relativa al establecimiento de la culpa en el uso de sistemas de IA se vincula con la dificultad para establecer con claridad el estándar de cuidado debido aplicable a los operadores de estos. Esto se relaciona con el ciclo de vida de los sistemas de IA, en donde podría resultar difícil saber en qué etapa determinada de este se produjo el hecho que finalmente causó el daño.

> Ejemplo 6: Un sistema de drones autónomos reparte mercadería guiándose con IA. Pensemos en el caso en el que la empresa A que diseñó dicha IA tuvo la debida diligencia esperable por parte de cualquier desarrollador de *software*, probando diversas hipótesis de seguridad en el dispositivo. El usuario (empresa B), también toma la debida diligencia respecto al uso de los dispositivos, no utilizándolos en contextos fuera de lo habitual. Sin embargo, debido a la imprevisibilidad en los algoritmos, se produce un accidente.

En este ejemplo, la complejidad del algoritmo (véase el *black-box problem* o el problema de la caja negra supra[211]) hace difícil analizar "de qué forma" se produjo el "error" en el cálculo del algoritmo. En este caso, ¿cómo puede construirse el deber de cuidado necesario respecto de un algoritmo que es opaco?, ¿quién responde respecto del riesgo? Todas estas preguntas están en el centro de la problemática relativa a la atribución de culpa en casos relativos a daños causados en el contexto de uso de sistemas de IA. Por otra parte, ¿cómo podría una empresa asegurarse de tomar todo el debido estándar de cuidado, cuando resulta difuso construir ese estándar propiamente tal?, si la autonomía característica de los algoritmos del sistema de IA lo convierte en imprevisible, nos enfrentamos a problemas bastante únicos respecto del requisito de la culpa:

> Ejemplo 7: durante el ciclo de vida de una IA, los desarrolladores tuvieron la debida diligencia de probar diversos escenarios de posible daño o error en el sistema de IA. Sin embargo, como consecuencia de la interacción normalmente esperada (es decir, sin la intermediación de hackeo o alteración del producto inicial) de la IA con usuarios finales (y su aprendizaje posterior), se produce un resultado dañoso no previsto por los desarrolladores.

Como ya se expuso, la previsibilidad es un requisito esencial de la culpa, el cual permite distinguir la acción culpable del caso fortuito, ya que este último alude a circunstancias que no pudieron ser

[211] Ver *supra* Capítulo 2, acápite 2.2

objeto de deliberación por el causante del daño al momento de actuar[212]. Refiriéndose a la previsibilidad, BARROS señala que:

> "la previsibilidad no hace referencia a un fenómeno psicológico, sino a aquello que debió ser previsto, atendidas las circunstancias. Como ocurre en general con los elementos del juicio de negligencia, la previsibilidad se valora en abstracto, considerando el discernimiento de una persona diligente[213]."

En este sentido, y volviendo al ejemplo 7, ¿podría decirse que existió negligencia por parte de los desarrolladores al no considerar un hecho que, dadas las capacidades técnicas que poseen, no pudieron haber previsto?, ¿podría imputársele, en este caso, el resultado dañoso al usuario final? Nos parece que la respuesta a esta última interrogante es negativa, debido a que la mera utilización de un sistema de IA no puede subsumirse bajo alguna hipótesis de exposición imprudente al daño en los términos del art. 2330 del Código Civil. Aun así, no parece justo señalar que, debido a que el comportamiento del sistema de IA resultó imprevisible, existe entonces una hipótesis caso fortuito, llevando a que no pueda indemnizarse la víctima.

Como veníamos señalando, el problema de la imprevisibilidad se agudiza si en el desarrollo, entrenamiento o prueba de la IA interviene más de un desarrollador, o más de una empresa dedicada a tal giro, siendo difícil determinar cuál actor, en específico, llevó a cabo una conducta negligente ¿Sería posible aplicar en estos casos algún tipo de responsabilidad solidaria por el cuasidelito, en los términos del art. 2317 del Código Civil? Si el sistema de IA es opaco, ¿cómo podría determinarse en qué etapa del ciclo de la vida del sistema de IA se produjo el hecho que tuvo como resultado el daño, con tal de imputar responsabilidad a alguno de los operadores en particular?

La incapacidad fáctica de prever determinados riesgos o daños (vinculado al problema de imprevisibilidad y opacidad que presentan los sistemas de IA), entonces, suscita dilemas interesantes respecto al cuidado debido por todos los actores involucrados en el ciclo de vida de los sistemas de IA. Una posible solución se encuentra en la imposición del cumplimiento de determinadas obligaciones de transparencia para los operadores, como se establece en el AIA. Otro acercamiento útil para prevenir y solventar este tipo de complejidades es la aplicación, por parte de los operadores,

[212] BARROS (2020): 95.
[213] *Ibid.*

de los principios éticos de responsabilidad y supervisión humanas durante todo el proceso de desarrollo y despliegue de los sistemas de IA.

Aún no existe una respuesta concreta frente a dichos casos hipotéticos, por lo que será la labor tanto de los tribunales, la doctrina y del legislador establecer los criterios respectivos, quedando todas las interrogantes aquí presentadas (por ahora) sin resolver.

c. IA como herramienta y el problema del cambio del estándar de cuidado.

Habiendo señalado las problemáticas anteriores, existen otros casos en los que pareciera que resolver la imputación de culpa del autor en el marco de la utilización de un sistema de IA puede resultar más fácil u obvia.

> Ejemplo 8: Una clínica provee a su personal médico con una IA para diagnosticar enfermedades a partir de análisis de imágenes (v.gr. resonancias magnéticas, rayos X). Frente a un falso negativo, un médico procede a no tratar a un paciente, teniendo como resultado su muerte.

En este ejemplo, a primera vista, la negligencia es imputable al médico, puesto que es razonable suponer que, su deber de cuidado profesional exige la diligencia razonable respecto de una potencial confirmación diagnóstica. El sistema de IA puede ser visto, simplemente, como una herramienta. Es posible resumir este tipo de hipótesis, presentes de diversas formas en la literatura, como hipótesis en las que un profesional altamente calificado utiliza la IA como una herramienta de ayuda en el marco de su determinada profesión o actividad.

Sin embargo, es necesario advertir que el uso de la IA en el campo médico, especialmente en el diagnóstico de enfermedades, podría complicar el establecimiento de una conducta negligente. Para ilustrar este punto, SELBST presenta una hipótesis en que, dado el fenómeno del sesgo algorítmico, los datos sobre los cuales se entrenen los algoritmos de diagnóstico presenten un sesgo basado en el género, algo que para el autor sería necesario considerar en el juicio de negligencia o culpa:

> "Pensemos en una hipotética herramienta de diagnóstico más precisa en general, pero menos fiable para las mujeres que para los hombres a la hora de diagnosticar un subconjunto de enfermedades. Dadas las tendencias actuales, es realista suponer que los beneficios de la IA se distribuirán de forma desigual y que los resultados de los hombres mejorarán más que los de las mujeres. A título ilustrativo, supongamos que la IA proporciona una mejora

> mínima pero positiva en la detección de una determinada afección en las mujeres y una mejora mayor en la detección de la afección en los hombres. Por último, supongamos que el usuario no conoce la distribución porque el fabricante solo ha probado la precisión global o no facilita documentación útil. Dadas estas premisas, un médico acabará utilizando la herramienta sin conocer los desequilibrios de género de la IA y acabará diagnosticando erróneamente a una mujer, lo que provocará la muerte de la paciente. ¿Cómo cambiaría el uso de la IA, la determinación de negligencia médica?"[214]/[215]

El caso es interesante, puesto que introduce un elemento adicional a la ecuación de determinación del actuar negligente. Dado que en un contexto normal el médico, al no haber diagnosticado la enfermedad de forma diligente, respondería simplemente por su negligencia, no siendo relevante la cuestión de género a la hora de determinar la misma[216]. Este caso también nos vincula a la posible responsabilidad compartida y al problema de la multiplicidad de actores, en este caso ¿podría existir alguna acción contra los operadores del sistema de IA, quienes fallaron en su deber de cuidado de incorporar datos carentes de sesgo, por parte del médico?, ¿y contra el hospital que le proveyó la IA en particular que llevó a cabo el diagnóstico incorrecto basado en datos que contribuyeron a un diagnóstico sesgado?, ¿existiría, entonces, una responsabilidad solidaria?

Otro factor que ha sido considerado es un posible cambio general en el estándar o debida diligencia, considerando la sofisticación actual de los modelos utilizados por los sistemas de IA. Existen ciertas áreas (como en el diagnóstico por imágenes) en las que "podrá esperarse que la IA consistentemente supere a un médico promedio o medianamente razonable[217]", lo que podría tener como consecuencia que muchas decisiones del proceso de diagnóstico se deleguen a los sistemas de IA[218]. ¿Cambiaría este escenario el estándar de cuidado esperable del médico? En aquellos casos en los que se vea involucrada una IA considerada como regularmente "segura" o de bajo riesgo, ¿cómo podría un médico mitigar la responsabilidad o asegurar que actuó con la mediana diligencia? Por ahora, dado que es un escenario meramente hipotético, no nos corresponde más que exponer el posible problema de la futura determinación de la culpa o negligencia médica.

---

[214] SELBST (2020): 1358.
[215] Este ejemplo nos remite al ejemplo 5 referido al uso de un vehículo semi-autónomo, ¿afecta en la determinación de la culpa el hecho de que el vehículo no haya funcionado como debió hacerlo debido a un error en la IA?
[216] SELBST (2020): 1358.
[217] BANJA (2022).
[218] *Ibid.*

Respecto a la pluralidad de actores en la fase de operación de la IA, es importante dar cuenta de que la situación puede ser más compleja:

> Ejemplo 9: El procesamiento de las imágenes diagnósticas por parte de una IA se lleva a cabo de un tecnólogo médico, el cual transmite el resultado de la IA al médico. El médico, como consecuencia, toma una decisión incorrecta basada en tal resultado.

Existe debate sobre la atribución de la responsabilidad en situaciones como la dispuesta en el ejemplo 9, particularmente porque, aun al utilizar una herramienta, el diagnóstico final corresponde al médico[219], por lo que podría asumirse que se trata de culpa profesional[220]. Sin embargo, ¿qué ocurre, nuevamente, con el estándar de cuidado que debió aplicarse por parte de los desarrolladores?, ¿tendría el médico algún tipo de acción en contra de los mismos? Esto nuevamente nos recuerda que existen diversas partes involucradas, tales como los desarrolladores del software/el algoritmo, el productor del hardware, los dueños del producto final de IA, entre otros[221]. Por último, el uso de IA opacas puede hacer que la decisión de determinados sistemas de IA utilizados en la práctica médica sea difícil de analizar, y, por tanto, resultará complejo para un profesional justificar la decisión tomada por el sistema de IA la forma en que dicha decisión resultó relevante a la hora de llevar a cabo un diagnóstico (sin perjuicio de que hasta el momento solo se entiende la IA como una herramienta de asistencia para la salud[222], no delegándose las tareas en ella de forma autónoma).

No obstante las problemáticas planteadas, es posible recurrir a los principios delimitados por la Unesco para dirimir el problema de culpa en las situaciones ya planteadas. Estos señalan que:

> "Puede ocurrir que, en algunas ocasiones, los seres humanos deciden depender de los sistemas de IA por razones de eficiencia, pero la decisión de ceder el control en contextos limitados seguirá recayendo en los seres humanos, ya que estos pueden recurrir a los sistemas de IA en la adopción de decisiones y en ejecución de tareas, pero un sistema de IA nunca podrá reemplazar la responsabilidad final de los seres humanos y su obligación de rendir cuentas. Por regla general, las decisiones de vida o muerte no deberían cederse a los sistemas de IA[223]."

[219] SMITH (2021): 539.
[220] ALESSANDRI RODRIGUEZZ (2005): 147.
[221] BRUYNE, VAN GOOL, GIRLS (2021): 360.
[222] SMITH (2021): 536.
[223] UNESCO (2021).

En este caso, es necesario referirnos nuevamente al principio ético de supervisión humana y al principio de responsabilidad, que, como ya señalamos al referirnos al hecho voluntario[224], pueden integrarse a las reglas generales de responsabilidad civil. En el caso de la determinación de un estándar de cuidado, aun cuando las personas puedan delegar tareas importantes a los sistemas de IA al punto de generar un "cambio" respecto de lo que actualmente se considera como el estándar de diligencia (como veíamos, por ejemplo, en el caso de los médicos), esto último no puede eliminar la necesidad de que, en caso de producirse un daño, la víctima pueda accionar contra la persona correspondiente y ser indemnizada; con mayor razón si la persona que causante del daño se encuentre respecto de la víctima en una posición que le confiera un deber de diligencia específico basado en su profesión, como es el caso de los profesionales de la salud.

Sin embargo, reiteramos que estas problemáticas deberán ser resueltas en profundidad por la jurisprudencia y la doctrina, en atención a que el uso de las aplicaciones de sistemas de IA se ha vuelto cada vez más frecuente en ámbitos profesionales.

d. IA autónomas como animales o incapaces, ¿responsabilidad por el hecho ajeno?

¿Podría aplicarse un régimen de responsabilidad por el hecho ajeno a los sistemas de IA? En nuestra opinión, esto no es posible, toda vez que para que exista una responsabilidad de un dependiente (en este caso, el sistema de IA) es requisito que este sea capaz (en términos del Código Civil), es decir, se requiere que tanto la persona civilmente responsable como la que está bajo su cuidado o dependencia sean capaces de delito o cuasidelito[225].

Descartando la teoría de entender a los sistemas de IA como dependientes (según la ley actual), no podemos dejar de mencionar que la creciente autonomía de la IA ha llevado a ciertos autores[226] a proponer su asimilación a niños, animales o incapaces[227]. Esta propuesta ha surgido en torno a los sistemas de IA utilizados en robótica, ya que la autonomía física de los mismos resulta particularmente relevante. La distinción entre la robótica y la IA, en general, está dada por la "corporalidad" de sus aplicaciones: si pensamos, por ejemplo, en la IA utilizada por un tecnólogo médico o un médico, es posible que esta tome la forma de un *software* contenido en un sistema de IA particular, mientras que un robot es (o podría ser) físicamente independiente.

---

[224] Ver *supra* Capítulo 3, acápite 3.2 (i).
[225] ALESSANDRI RODRIGUEZ (2005): 229.
[226] CHOPRA, WHITE, PAGALLO, entre otros.
[227] Para profundizar sobre esta teoría en su totalidad, véase PAGALLO (2013).

Ejemplo 10: un robot creado por la empresa Y se utiliza en salas de clases de pre-kinder para ayudar a niños neurodivergentes a reconocer sus emociones.

Ejemplo 11: un robot creado por la empresa B tiene como finalidad ayudar con determinadas tareas domésticas, como la limpieza.

En estos casos, se ha vuelto de particular notoriedad la influencia que los usuarios finales podrían tener al interactuar con un robot, ya que como vimos, los sistemas de IA son altamente adaptables y dependientes de su entorno, esto es, de los estímulos sobrevinientes a su creación y despliegue.

Respecto a la posibilidad de asimilar a los sistemas de IA robóticos con los animales, si bien esta es una teoría interesante, parece inaplicable desde el punto de vista de la legislación chilena relativa a la responsabilidad extracontractual contenida en el Código Civil, por varios motivos. En primer lugar, la asimilación a los animales resulta teóricamente compleja porque los animales son seres sintientes no racionales, mientras que los sistemas de IA pueden considerarse como seres racionales o agentes inteligentes *no* sintientes, por lo forzosamente debe concluirse que dicha analogía no resulta adecuada. En segundo lugar, el extender los principios soslayados en el artículo 2326 del Código Civil[228], se torna excesivo, debido a que atendiendo el tenor literal del artículo, pareciera que extender la situación del "extravío o soltura de un animal" mediante analogía a los accidentes provocados por robots o por sistemas de IA autónomos resulta demasiado extenso.

Por último, respecto de la de asimilación de los sistemas de IA a los incapaces, es preciso señalar que en nuestro ordenamiento jurídico la capacidad (y consecuentemente, la incapacidad) son conceptos asociados a los sujetos de derecho. Cuando el Código Civil se refiere a los incapaces absolutos y relativos, lo hace inmediatamente luego de precisar la capacidad en el artículo 1445 como un requisito para que una *persona* se obligue a otra por un acto o declaración de voluntad. Acto seguido, se refiere a los incapaces absolutos y a los incapaces relativos a través de conceptos que difícilmente podrían aplicarse a los sistemas de IA: resulta poco intuitivo concebir (al menos hasta este momento) a un sistema de IA como demente, impúber, o interdicto por disipación. Atendido el tratamiento que el Código hace de esta institución, no nos parece tampoco posible

[228] "El dueño de un animal es responsable de los daños causados por el mismo animal, aun después que se haya soltado o extraviado; salvo que la soltura, extravío o daño no pueda imputarse a culpa del dueño o del dependiente encargado de la guarda o del servicio del animal."

extender por analogía la conceptualización que el código ofrece respecto de la incapacidad a los sistemas de IA.

e. ¿Responsabilidad por el hecho de las cosas?

¿Podría aplicarse el régimen de presunción de la culpa por el hecho de las cosas a los sistemas de IA? Regulado en los arts. 2323, 2324, 2326, 2327 y 2327, dicho régimen postula que "no solo se responde por el hecho personal, mediato o inmediato del agente, sino también del que proviene del hecho de una cosa[229]". Sin embargo, a diferencia de otros regímenes de responsabilidad en el derecho comparado, los casos respecto de los cuales se aplica la responsabilidad por el hecho de las cosas son taxativos[230], por lo que fuera de los casos señalados en la ley[231] no hay otros en los que la ley presuma la responsabilidad. Si se asimilaran los sistemas de IA a las hipótesis contenidas en los artículos citados en este párrafo con el fin de aplicar el régimen relativo a la responsabilidad por el hecho en las cosas, se estaría llevando a cabo una aplicación excesivamente extensa de dichos preceptos legales. Esto se ve reforzado por el hecho de que en nuestro Código no existe una regla de aplicación general respecto al régimen de responsabilidad por el hecho de las cosas, en contraposición a otros sistemas jurídicos como Francia[232].

Por oposición a nuestra codificación, como veníamos diciendo, en Francia sería posible configurar una hipótesis de responsabilidad general por el hecho de las cosas. De acuerdo con el inciso 1.º del art. 1384 del Código Francés de 1804, y por lo interpretado por la doctrina y por la jurisprudencia de dicho país, esta consiste en una regla de carácter general, aplicable a todo daño irrogado por una cosa inanimada que no esté regida por el art. 1386 del mismo Código[233]. El texto de dichos artículos se mantuvo tras la reforma que el Código Francés tuvo en 2016, cambiando su ubicación a los artículos 1242 y 1244, respectivamente.

f. Uso doloso de sistemas de IA.

Como fue asentado en *supra*[234], la culpa implica un actuar omisivo o activo por parte de un ser humano, de acuerdo con el artículo 44 del CC. El dolo, por su parte, apunta a una intención positiva de cometer un ilícito, definido por el mismo artículo como "la intención positiva de inferir injuria

---

229 ALESSANDRI RODRIGUEZ (2005): 283.
230 *Ibid.*
231 Dichos casos corresponden al daño ocasionado por la ruina de un edificio (art. 2323 y 2324), el daño ocasionado
232 Art. 1384 del Código Francés.
233ALESSANDRI RODRIGUEZ (2005): 283.
234 Capítulo 3, acápite 3.1.

a la persona o propiedad del otro". Es posible que una persona efectúe un uso doloso (es decir, con intención de cometer un delito civil) de los sistemas de IA. Algunos ejemplos de uso doloso de sistemas de IA son:

> Ejemplo 12: Una persona utiliza un programa de IA para generar imágenes y la voz simulada de un político con tal de que aparezca diciendo, en un video, cosas contrarias a su afiliación política.
>
> Ejemplo 13: Una persona utiliza un programa de IA para generar contenido pornográfico con la cara, fisionomía y voz de una persona sin que ella consienta a esto. Como consecuencia de dichas imágenes, la persona pierde su trabajo y sufre graves secuelas psicológicas.

Creemos que los casos de uso doloso de sistemas de IA son relevantes de mencionar, ya que, dadas las dificultades de atribución de responsabilidad en los otros ejemplos aquí expuestos, la existencia de un actuar doloso conllevaría que pueda construirse un vínculo más certero y evidente entre el daño provocado por un sistema de IA y el requisito del nexo causal. Además, debe tenerse en consideración el hecho de que la gravedad de la conducta dañosa (determinada por la intención de dañar o la extrema negligencia) es un factor que suele ser considerado por los jueces al momento de indemnizar el daño moral[235]. Este punto es relevante para los dos ejemplos en cuestión, donde en ambos casos se aprecia un daño más allá de lo pecuniario. En estos casos, son responsables del daño causado el autor de este, sus cómplices y el que se aprovechó del dolo, aunque no haya participado en su ejecución ni tenido conocimiento de su existencia[236].

(iv) Daño

El daño es un requisito esencial de la responsabilidad civil: "Para que el hecho o la omisión de una persona capaz de delito o cuasidelito engendre responsabilidad delictual o cuasidelictual civil, no basta su ejecución con dolo o culpa. Es indispensable que cause daño. Sin él no hay responsabilidad civil ni interés en la acción[237]." La doctrina ha definido el daño en la responsabilidad civil como "una pérdida, disminución, detrimento o menoscabo en su persona o bienes o en las ventajas o

[235] BANFI (2017): 86.
[236] ALESSANDRI RODRIGUEZ (2005): 124.
[237] *Ibid*.: 152.

beneficios patrimoniales o extrapatrimoniales de que gozaba[238].", así como "todo detrimento, perjuicio, menoscabo, dolor o molestia que sufre una persona en sus bienes, libertad, honor, crédito, afectos, creencias, etc. El daño supone la destrucción o disminución, por insignificante que sea, de las ventajas o beneficios patrimoniales o extrapatrimoniales que goza un individuo[239]." Este, como se desprende de la cita, puede ser patrimonial o extrapatrimonial.

Dentro del daño patrimonial, se distingue entre el daño emergente (pérdida actual del patrimonio) y el lucro cesante (frustración de una legítima utilidad que hubiera incrementado el patrimonio de no haber sucedido el hecho dañoso)[240]. El daño, además, requiere ser cierto[241], personal[242] y directo[243] para poder ser indemnizado. Que el daño sea cierto hace referencia a la materialidad del daño, es decir, que efectivamente exista un daño inferido o sufrido por la víctima (por oposición a un daño hipotético o eventual)[244], mas no opuesto a daños futuros. Por su parte, la exigencia de que el daño sea personal apunta a que solo quien lo ha sufrido pueda demandar su reparación. Por último, que el daño sea directo implica que el daño por el cual se responde corresponda a aquel que se produce como consecuencia necesaria e inmediata del incumplimiento[245].

Se han identificado diversos tipos de daños relativos a la IA, algunos incidentes recientes son, por ejemplo: la falla del autopiloto de un vehículo Tesla, el cual no pudo distinguir un tráiler blanco que estaba cruzando la calle del cielo, lo que llevó un choque con resultados fatales[246], el accidente producido por un vehículo de Uber que se conducía de forma autónoma ocasionando el atropello de una transeúnte, ocasionando su muerte[247], así como un robot quirúrgico en un hospital de Minnesota no funcionó de forma correcta en una cirugía de próstata, dañando al paciente, entre otras[248].

A partir de lo expuesto, es posible señalar que los daños que puedan causarse por el uso de sistemas de IA no difieren, de forma significativa, de los daños que actualmente producen otro tipo de tecnologías. Si una falla en un sistema de IA lleva a la muerte de la víctima, o a una pérdida

238 ALESSANDRI RODRIGUEZ (2005): 213.
239 *Ibid*.: 153.
240 CORRAL TALCIANI (2003) 148.
241 BARROS (2020): 244.
242 *Ibid*.: 249.
243 *Ibid*.: 257.
244 *Ibid*.: 257.
245 *Ibid.*
246 BRUYNE, VAN GOOL, GILS. (2021) 360.
247 *Ibid.*
248 *Ibid.*

significativa de dinero, u otros, la única diferencia radica en la forma en la que se llevó a cabo el daño, y no en el daño en sí.

(v) Relación de causalidad

Según la doctrina, el requisito de causalidad se refiere, en la responsabilidad por culpa, a la relación entre el hecho culpable y el daño provocado[249]. Este requisito, según BARROS, cumple una función dual: por una parte, responde como un fundamento de responsabilidad (solo se responde por los daños que se siguen como consecuencia del hecho del demandado), y por otra, como un factor que limita la misma, toda vez que no se responde de todas las consecuencias de un hecho, sino solo de aquellas que en virtud de un juicio normativo son atribuibles a tal[250]. Siguiendo al mismo autor y a la doctrina nacional, no existe una referencia expresa en el Código Civil respecto al requisito de causalidad, pero puede desprenderse de diversos preceptos legales como lo son los artículos 1473, 2314 y 2329 del Código Civil. A partir de estas normas, es posible aseverar que debe existir una relación de causalidad entre el hecho y daño. De los artículos citados, se han sustraído los requisitos para establecer dicho vínculo en el juicio de responsabilidad civil. De esta forma, para que haya responsabilidad por un daño, este tiene que ser objetivamente imputable a la acción u omisión.[251] En suma, se entiende que hay relación de causalidad cuando el hecho —u omisión— doloso o culpable es la causa directa y necesaria del daño, es decir, cuando sin él este no se habría producido[252].

La causalidad es un presupuesto particularmente complejo y bastante discutido por la doctrina, el cual posee una dimensión fáctica y otra normativa[253], en efecto, a partir de esto se ha señalado que, a la hora de llevar a cabo el juicio de responsabilidad, no solo se consideren por el tribunal los hechos puramente fácticos al analizar la causalidad, sino que también los jueces "realizan decisiones de política jurídica para decidir si el resultado, fácticamente ligado a la acción del demandado, puede serle atribuido. En general, la lejanía causal, la imprevisibilidad o el hecho de que la acción solo haya adelantado en tiempo un resultado inevitable cuentan como razones para limitar o suprimir, según el caso, la responsabilidad[254]".

[249] BARROS (2020): 399.
[250] *Ibid.*: 394.
[251] *Ibid.*: 395.
[252] ALESSANDRI RODRIGUEZ (2005): 176.
[253] PAPAYANNIS (2014): 78.
[254] *Ibid.*

a. Indeterminación del nexo causal e IA

Los problemas de causalidad respecto a sistemas de IA presentan dificultades similares a aquellos referidos a la imputabilidad subjetiva examinados *supra*[255]. Dicho esto, los problemas relativos a la causalidad difieren en naturaleza a los relativos a la culpa, ya que, aun existiendo un vínculo entre el actuar negligente del actor y el daño provocado por un sistema de IA, no por ello resulta obvio establecer el nexo causal respectivo. En otras palabras, a pesar de existir una acción u omisión culpable determinada y un daño sufrido por la (o las) víctima(s) en el marco del uso de un sistema de IA, la determinación del nexo causal podría, aun así, ser difícil. Esto, particularmente considerando la exigencia de que el nexo causal debe ser *directo y necesario* para la producción del daño, por una parte, y la opacidad y falta de transparencia (y explicabilidad) característica de los algoritmos utilizados por los sistemas de IA, por otra.

Siguiendo las ideas ya expuestas, el principal problema que puede preverse respecto a la causalidad y los sistemas de IA está relacionado con un posible caso de indeterminación causal. Esto, por ejemplo, en casos donde sea difícil dirimir el nexo causal entre la acción humana, el funcionamiento del sistema de IA, y el resultado dañoso. En este sentido, se ha previsto que la compleja relación entre humano-máquina pueda llevar a la confusión respecto de quién es realmente responsable frente a un daño en el que se vio involucrado un sistema de IA, ya que el actuar del humano podría confundirse fácilmente con la autonomía del sistema. Al respecto, se ha señalado que:

> "En los casos de IA, la intrincada relación entre el ser humano y la máquina agrava la dificultad de probar la causalidad, especialmente cuando la IA y la supervisión humana están entrelazadas. El enredo de la responsabilidad humana frente a la de la máquina puede «dificultar la determinación técnica y fáctica de quién fue responsable [el médico o el dispositivo de IA]» para el jurado y requerirá el testimonio de un experto[256]".

De esta cita se puede concluir, como venimos desarrollando, que no siempre será fácil determinar el vínculo causal existente entre el hecho dañoso y el actuar negligente, dada la compleja interrelación entre humano (operador) y máquina.

b. Pluralidad e indeterminación de agentes

[255] Capítulo 3, acápite 3.2 (iii).
[256] GRIFFIN (2021): 101. Traducción propia.

Otro posible problema se relaciona con la pluralidad de actores presentes durante el ciclo de vida de los sistemas de IA. Similarmente a lo que ocurría con la culpa, donde podría ser difícil establecer qué persona era responsable de un determinado defecto del sistema (si es que podemos encontrarla), también podría ser difícil establecer con certeza en qué momento y en qué forma se vincula dicho actuar[257] relativo al sistema de IA con el daño, lo que podría tener implicancias relevantes en aquellos casos donde ha existido la incidencia de una pluralidad de actores a lo largo de distintos momentos del desarrollo del mismo (por ejemplo, durante la manufactura y el posterior despliegue, así como el uso por quien sea dueño del mismo, el consumidor final, etcétera).

Esto podría llevarnos a un escenario de indeterminación del agente en la causalidad o de causalidad difusa, es decir, un supuesto de causalidad alternativa en el que una pluralidad de posibles agentes ha causado el daño[258], hecha más compleja por la necesidad de una prueba de alto calibre técnico y con la cual no siempre se podrá contar[259].

c. Explicabilidad y transparencia

Una posible solución a los problemas previstos en esta sección está en la explicabilidad y transparencia algorítmica, consagrada en los distintos instrumentos normativos que ya hemos señalado antes. Los principios de la OCDE señalan:

> "Los actores de la inteligencia artificial deben comprometerse con la transparencia y la divulgación responsable en lo que respecta a los sistemas de inteligencia artificial. Con este propósito, deben proporcionar información significativa, apropiada al contexto y acorde con el estado del arte:
>
> i. para fomentar una comprensión general de los sistemas de inteligencia artificial
> ii. para que los interesados sean conscientes de sus interacciones con los sistemas de inteligencia artificial, incluso en el lugar de trabajo,
> iii. para permitir que aquellos afectados por un sistema de inteligencia artificial comprendan el resultado respectivo, y,
> iv. para permitir que aquellos afectados negativamente por un sistema de inteligencia artificial puedan cuestionar su resultado en razón de información clara y fácil de

[257] En este punto, nos referimos a acciones autónomas.
[258] BÁRCENA y PEREIRA en PAPAYANNIS (2022): 299.
[259] Ver los problemas de caja negra y falta de transparencia desarrollados *supra*, capítulo 2, acápite 2.2.

> entender acerca de los factores y la lógica que sirvieron de base para la predicción, recomendación o decisión[260]."

Mientras que el texto de la Unesco destaca:

> "La transparencia y la explicabilidad de los sistemas de IA suelen ser condiciones previas fundamentales para garantizar el respeto, la protección y la promoción de los derechos humanos, las libertades fundamentales y los principios éticos. La transparencia es necesaria para que los regímenes nacionales e internacionales pertinentes en materia de responsabilidad funcionen eficazmente. La falta de transparencia también podría mermar la posibilidad de impugnar eficazmente las decisiones basadas en resultados producidos por los sistemas de IA y, por lo tanto, podría vulnerar el derecho a un juicio imparcial y a un recurso efectivo, y limita los ámbitos en los que estos sistemas pueden utilizarse legalmente[261]."

Si bien la explicabilidad de los algoritmos podría hacer inteligible la toma de decisiones automatizada y, por consecuencia, hacer posible la identificación de la etapa del ciclo de vida del sistema de IA dónde se produjo efectivamente el error o toma de decisión que llevó al daño, no podemos dejar de mencionar que la transparencia algorítmica sigue siendo un problema técnico complejo, y que la regla general es que estos sistemas sean opacos. La transparencia, como una solución para los problemas de nexo causal (y también para aquellos relativos a la culpa), no está exenta de críticas, y reiteramos que, finalmente, es un problema relativo a la capacidad técnica[262] de quienes desarrollan este tipo de sistemas.

d. La víctima y la carga de la prueba

Otro aspecto relevante en relación con el nexo causal es la dificultad técnica de probar, para la víctima, que el sistema de IA actuó de una forma que causó daño. Consideremos que para probar el nexo causal, la víctima deberá acreditar, de forma específica, la forma en que el sistema de IA causó el daño. Evidentemente, esta situación podría llevar a un escenario injusto: en primer lugar, es difícil que la víctima tenga acceso a la información técnica relativa al sistema de IA, o, en otros términos, la información contenida en el código fuente o en los datos que el sistema haya recopilado.

[260] OCDE (2022): 8, destacado propio y traducción propia.
[261] UNESCO (2021).
[262] Véase BATHAEE (2018).

Con bastante probabilidad, tal información se hallará protegida bajo el secreto comercial. En segundo lugar, si los algoritmos involucrados en el hecho dañoso carecieren de transparencia, trazabilidad, o fuesen extremadamente complejos e ininteligibles, puede asumirse que, a menos que la víctima tuviese un conocimiento altamente avanzado relativo al funcionamiento del sistema de IA en particular, será muy difícil que pueda sustentar su acción de responsabilidad civil sin contar con este tipo de información. Como resultado de esta situación, puede asumirse que, muchas veces, la víctima carecerá de los recursos técnicos y profesionales para probar la existencia o causa del error.

Los problemas relativos a la prueba de la causalidad se reconocen en algunas de las propuestas normativas que analizamos anteriormente[263]. El ejemplo más claro de esto es la propuesta de directiva de la UE, el cual ya detallamos[264], y que en su apartado explicativo señala:

> "Puede resultar difícil para los demandantes probar que existe un nexo causal entre dicho incumplimiento y la información de salida producida por el sistema de IA o la no producción de una información de salida por parte del sistema de IA que haya dado lugar a los daños en cuestión. Por lo tanto, en el artículo 4, apartado 1, se ha establecido una presunción refutable de causalidad específica en relación con este nexo causal. Esta presunción es la medida menos gravosa para dar respuesta a la necesidad de una indemnización justa para la víctima[265]."

Esto no es algo menor, dado que ya puede resultar difícil, como se ha visto a lo largo de esta memoria, probar la negligencia de un operador respecto de un daño causado por un sistema de IA, elemento que suele encontrarse estrechamente vinculado con la causalidad. Consideramos que el acercamiento propuesto por la UE en materia de carga de la prueba refleja los fines propios de la justicia correctiva de la responsabilidad civil, es decir, que la víctima pueda efectivamente accionar y ser indemnizada al sufrir un daño.

(vi) Responsabilidad estricta o por riesgo

[263] Ver *supra,* capítulo 21, acápite 1.5.
[264] Ver *supra,* capítulo 21, acápite 1.5 (ii).
[265] COMISIÓN EUROPEA (2022). En: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A52022PC0496.

Atendiendo los puntos tratados, ha surgido la propuesta de que los sistemas de IA, particularmente aquellos que implican un alto riesgo, sigan un esquema de responsabilidad estricta[266]. La responsabilidad estricta tiene, reiterando, aplicación respecto del ámbito del riesgo que la ley atribuye a quien desarrolla una determinada actividad[267]/[268]. Lo determinante, en este tipo de responsabilidad, es que se materialice un riesgo que esté bajo el control del responsable[269]. En el derecho chileno, los estatutos de responsabilidad estricta son establecidos por el legislador, no existiendo una norma que establezca una categoría general que comprenda distintos grupos de casos sujetos a un régimen de este tipo de responsabilidad[270]. En la responsabilidad estricta, el elemento más relevante para establecer su procedencia es la realización de una actividad o la tenencia de una cosa que genera el riesgo de daño[271], siendo indiferente la calificación de la conducta efectiva del autor del daño[272].

El establecimiento de un sistema como este podría llevarnos a evitar varios de los problemas que aquí hemos analizado. Sin perjuicio de esto, se ha señalado que dicha forma de regular los sistemas de IA podría resultar perjudicial respecto de la innovación[273]. Por otra parte, debido a que no existe legislación al respecto en Chile, no puede aplicarse en la actualidad este régimen de responsabilidad, por tratarse de un régimen especial y delimitado en su campo de aplicación:

> "Es en los casos excepcionales y perfectamente definidas, en las actividades caracterizadas por su peligrosidad, donde cabe introducir derogaciones a la regla general según la cual solo la culpa engendra responsabilidad civil, pero tales derogaciones siempre estarán condicionadas a que su campo de aplicación esté estrictamente delimitado[274]."

En efecto, a pesar de que algunos acercamientos normativos (como las primeras propuestas del AIA), establezcan una categorización de los sistemas de IA basada en el riesgo con tal de determinar el régimen de responsabilidad civil aplicable (donde se aplicaría, por ejemplo, la responsabilidad

---

266 UROPEAN PARLIAMENTARY RESEARCH SERVICE (2023): 13. En: https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/739342/EPRS_BRI(2023)739342_EN.pdf
267 BARROS (2020): 469.
268 Para PAPAYANNIS (2014): 32, en la responsabilidad estricta la obligación de compensar se basa en el riesgo propio de la actividad que realiza el agente o en la peligrosidad de los objetos que se sirve, así como en el beneficio que el agente obtiene de la actividad generadora de daños. Solo es requisito la existencia de un nexo de causalidad (en este caso, normativo) entre la actividad riesgosa y el daño.
269 BARROS (2020): 469.
270 *Ibid.*: 470.
271 *Ibid.*: 499
272 *Ibid.*
273 ERDÉLYI (2020): 1310.
274 CARNAVELI DE CAMACHO en LARROCAU (2007): 52.

estricta respecto de los sistemas de alto riesgo), este tipo de sistematización no ha estado exento de críticas:

> "(…) la responsabilidad objetiva solo tendría sentido si el creador de un programa de ordenador puede prever los efectos perjudiciales del programa y ajustarlo en consecuencia. A medida que los programas informáticos se vuelven más inteligentes y menos transparentes, no solo son menos predecibles sus efectos perjudiciales, sino que su proceso de toma de decisiones también puede ser impredecible. La responsabilidad objetiva, sin embargo, supone cierto control o previsibilidad, lo que permitiría al desarrollador de la IA, por ejemplo, predecir el daño potencial del que será responsable para poder obtener un seguro adecuado[275]."

Dicho esto, solo podemos concluir que el régimen al que deban sujetarse los sistemas de IA ha de ser definido por el legislador, en miras de obtener una política pública relativa al efecto.

3.3 Conclusiones

De lo expuesto en este capítulo es posible señalar que, si bien es posible aplicar el régimen general de responsabilidad extracontractual a hipótesis de daños ocasionadas por el uso de sistemas de IA, ello resulta insatisfactorio respecto de determinados elementos del juicio de responsabilidad civil: la culpa y el nexo causal. Esto último podría llevar a resultados contrarios al fin primordial de la responsabilidad civil (la reparación del daño causado) para las víctimas.

Como se explicó, esta insatisfacción deriva de la inadecuación del criterio de la culpa como modelo de imputación subjetiva frente a las complejidades propias de los sistemas de IA, y particularmente, respecto de la autonomía, imprevisibilidad y opacidad de estos últimos. Del análisis de los casos hipotéticos presentados, es posible afirmar que será difícil atribuir y encausar un juicio de responsabilidad cuando no pueda identificarse con claridad, debido a los desafíos que los sistemas de IA presentan, un hecho culposo de un agente humano. Además de esto, la exigencia de previsibilidad y control sobre el hecho dañoso se ve socavada en aquellos casos en que los sistemas de IA operan de manera autónoma, creando escenarios donde la falta de transparencia y

[275] BATHAEE (2018): 931. Traducción propia.

explicabilidad de las decisiones algorítmicas imposibilitan determinar un estándar de cuidado aplicable.

Además de lo planteado respecto a la culpa, se demostró en este capítulo que, a partir de los desafíos planteados, otro obstáculo importante para las víctimas será la prueba del nexo causal. Esto último debido a la complejidad técnica presentada por los sistemas de IA, junto con la imposibilidad de la víctima de acceder a la información necesaria para poder acreditar la forma que un determinado algoritmo llevó a cabo una decisión que haya tenido como resultado el daño. Lo anterior evidencia una barrera adicional en la obtención de una reparación justa para las víctimas, fin primordial del sistema de responsabilidad civil.

A modo de conclusión, estas dificultades destacan la necesidad de buscar soluciones que, ajustándose a derecho, aseguren la indemnización efectiva del daño sufrido por las víctimas. Como se detalló a lo largo del capítulo, una posible forma de resolver los problemas que presentan este tipo de casos está en la aplicación, por parte de los jueces, de los principios éticos dispuestos en normativas internacionales, ya sea entendiéndolos como principios generales del derecho, o como una manifestación del principio general de equidad a partir del deber de inexcusabilidad judicial.

CONCLUSIONES

La presente memoria tuvo como objetivo analizar la aplicación del régimen general de responsabilidad civil a los daños causados por sistemas de IA, con el fin de determinar si dicha aplicación resulta procedente y satisfactoria en relación con el fin de la responsabilidad civil, esto es, la reparación efectiva del daño sufrido por las víctimas. Con este propósito, se expuso y delimitó el concepto de IA y de sistemas de IA, tanto desde un punto de vista técnico como jurídico, para luego pormenorizar la forma en la que dicha tecnología presenta desafíos complejos y novedosos. Tomando este entendimiento de la IA y los sistemas de IA, se llevó a cabo un análisis pormenorizado que contrastó las complejidades de estos con cada uno de los elementos necesarios para llevar a cabo un juicio de responsabilidad civil. Para este fin, la tesis se estructuró en tres grandes capítulos:

En el Capítulo 1, se presentó una explicación introductoria sobre la historia, concepto y funcionamiento de la IA, para transparentar cómo los sistemas de IA difieren sustancialmente de otras tecnologías de la información actuales. Además, en dicho capítulo se presentaron y analizaron los principales acercamientos normativos relacionados con la IA, tanto a nivel internacional como nacional. En el ámbito internacional, se destacaron tanto el Intelligence Act (AIA) como la propuesta de directiva sobre responsabilidad civil extracontractual, instrumentos que buscan establecer un marco regulatorio general en Europa. A nivel nacional, se examinaron la Política Nacional de Inteligencia Artificial de Chile y la Circular Interpretativa del Sernac, que proporcionan directrices para el desarrollo y uso responsable de la IA. Por último, se expuso y comentó el actual proyecto de ley que regula los sistemas de IA en Chile.

En el Capítulo 2, se detallaron los principales desafíos que presenta la inteligencia artificial (IA) desde una perspectiva técnica. Esto, con el propósito de facilitar la comprensión acerca de cómo dichos desafíos complican la aplicación de los elementos del juicio de responsabilidad civil extracontractual. Para ello, se analizaron una serie de desafíos relevantes: la complejidad estructural de los sistemas de IA, la existencia de una pluralidad de actores, la opacidad y su relación con el sesgo algorítmico y la caja negra, la autonomía, y, por último, la imprevisibilidad.

Finalmente, en el Capítulo 3 se analizaron cada uno de los elementos del juicio de responsabilidad civil extracontractual por culpa o negligencia en el ordenamiento jurídico chileno, exponiéndose en concreto una serie de desafíos aplicables a cada uno de los elementos de este, a través del análisis

de casos hipotéticos. Respecto del hecho voluntario, se subrayó la necesidad de que exista un acto humano para atribuir responsabilidad civil, y se propuso para transparentar el vínculo entre el sistema de IA y un hecho humano la aplicación de los principios éticos como la supervisión humana y la responsabilidad dispuestos por la OCDE y UNESCO.

En lo que respecta a la culpa y el dolo, se abordaron desafíos particulares, tales como la ausencia de culpa en sistemas autónomos, la dificultad para establecer el deber de cuidado, y el cambio del estándar de cuidado con la IA como herramienta; considerando la influencia de la precisión de los modelos de IA en las expectativas de cuidado de los profesionales. Respecto del daño, nos limitamos a señalar que la naturaleza de este será similar al que resulta del uso de otro tipo de tecnologías, sin perjuicio de que varíe la forma de ocurrencia. Sobre la causalidad, se destacó la dificultad de determinar el nexo causal debido a la opacidad de los algoritmos y la pluralidad de agentes en el ciclo de vida de la IA. Se discutieron soluciones como la transparencia de los algoritmos y la presunción de causalidad propuestas por la UE, las cuales buscan facilitar la prueba para la víctima ante la complejidad técnica de los sistemas de IA y el acceso limitado a la información. Para finalizar el capítulo, se discutió la posibilidad de aplicar un régimen de responsabilidad estricto respecto de los sistemas de IA.

A partir de lo expuesto podemos concluir que, si bien resulta procedente la aplicación del sistema general de responsabilidad extracontractual a casos de daños causados por sistemas de IA, existirán hipótesis en que dicha aplicación tendrá resultados contrarios al fin de la responsabilidad civil, esto es, a la reparación del daño sufrido por las víctimas. En otras palabras, si bien el juicio de responsabilidad civil extracontractual resulta generalmente aplicable, no puede dejarse de lado que dicha aplicación no siempre resultará satisfactoria.

Esto último, dado que existirán casos en que las características propias de los sistemas de IA harán compleja la tarea de establecer ciertos requisitos necesarios para el juicio de responsabilidad civil, en particular, la culpa y el nexo causal. Tal como se trató en el Capítulo 3, dichos problemas en sede de culpa responden, principalmente, a la dificultad que existirá para la existencia de un hecho culpable dada la opacidad, imprevisibilidad y autonomía, características de los sistemas de IA. Además, pueden incidir en la determinación de la culpa tanto el cambio paulatino que puede tener el deber de cuidado en atención a la sofisticación y especialización de los sistemas de IA, como la

posible dificultad a la hora de determinar el deber de cuidado esperable respecto de los operadores de estos.

Respecto de la causalidad, el problema estará relacionado con la indeterminación del nexo causal y la dificultad probatoria para la víctima de este, producto de opacidad y falta de transparencia, características de los sistemas de IA. Será difícil determinar, pues, el vínculo directo y necesario entre la acción humana, el funcionamiento del sistema de IA y el daño. Además, la pluralidad de actores en el ciclo de vida de la IA agravará dicha indeterminación, al igual que la posible incidencia que podría tener en el vínculo causal la relación entre humano y máquina. Esto último, dada la adaptabilidad que tienen los sistemas de IA respecto de la interacción e información que reciben de sus usuarios.

Reiterando, dado que las complejidades aquí resumidas podrán tener como consecuencia que la víctima no pueda acceder a la reparación del daño, es que señalamos que existen ciertas hipótesis en las que la aplicación de las normas generales de responsabilidad civil resultará insatisfactoria. Esto se ve reforzado por el hecho de que los instrumentos normativos internacionales reconocen explícitamente que la aplicación del juicio de responsabilidad civil en los sistemas regidos por un modelo de imputación basado en la culpa resulta problemática.

Sin perjuicio de que en ciertas hipótesis la aplicación del régimen general será problemática, no puede afirmarse lo mismo respecto de todos los sistemas de IA como generalidad. Como se expuso también en el Capítulo 3, existirán casos en que los problemas relativos a su aplicación no estarán presentes, o, de estarlo, tendrán una solución mucho más directa, particularmente cuando exista menos autonomía del sistema de IA, y cuando el mismo sea explícitamente usado como una herramienta.

Finalmente, esta memoria adhiere a la hipótesis que propone que una forma de solucionar las problemáticas aquí señaladas descansa en la remisión a los principios éticos delineados tanto por la OCDE como la UNESCO. Esto se debe a que dichos instrumentos establecen obligaciones que, al aplicarse respecto de los operadores, resultarían ser herramientas valiosas para asegurar que pueda determinarse de forma expedita la existencia de un responsable frente a daños causados por sistemas de IA. Además de esto, es necesario que la futura normativa nacional reconozca los desafíos particulares existentes en la responsabilidad extracontractual respecto de los sistemas de IA, reflejando las propuestas normativas internacionales.

El desafío recaerá, finalmente, en los operadores jurídicos de todo tipo y en el legislador, con tal de que puedan anticipar y responder de manera proactiva a los acelerados cambios tecnológicos que los sistemas de IA traen consigo, con el fin de asegurar que, en el caso de existir un daño, la víctima pueda ser efectivamente compensada.

Asentadas estas conclusiones, esta memoria invita a sus lectores a pensar no solo respecto de las problemáticas actuales que rodean el sistema de responsabilidad civil y los sistemas de IA, sino que también a reflexionar acerca de la permanente conjunción existente en nuestra disciplina entre normas de antaño, avances normativos del presente, y, por último, tecnologías del futuro.

De cierta forma, el Derecho es una disciplina paradójica. Por una parte, las normas jurídicas que hoy rigen nuestros actos más simples y comunes del día a día fueron escritas y pensadas hace ya más de un siglo, respondiendo a una realidad completamente distinta a la de nuestra era. Por otra parte, es una disciplina que, por su naturaleza y relevancia, debe adaptarse continuamente a las nuevas realidades y cambios que el progreso del tiempo y de la humanidad trae consigo, lo que se materializa, entre otras formas, a través del cambio legislativo. En este contexto, cuando aparece un nuevo fenómeno o forma de entender el mundo, sea producto de una revolución social o científica, este debe ser analizado desde la perspectiva del Derecho, si es que tiene o puede llegar a tener relevancia jurídica, y lo cierto es que la IA, hoy más que nunca, la tiene.

BIBLIOGRAFÍA

ABELIUK Andrés y GUTIÉRREZ Claudio. 2021. Historia y evolución de la inteligencia artificial. En Revista BITS del Departamento de Ciencias de La Computación de la Universidad de Chile. Edición N. 21 (14-21).

ABELIUK René. 2010. Las obligaciones. T. I. Santiago. 5ª ed. Jurídica de Chile.

ALESSANDRI RODRIGUEZ, Arturo. 2005. De la responsabilidad extracontractual en el derecho civil chileno.

ARAYA Carlos. 2020. Desafíos legales de la inteligencia artificial en Chile. Revisa Chilena de Derecho y Tecnología. 9. (257-290).

BARROS Enrique. 2020. Tratado de responsabilidad extracontractual. Santiago. Jurídica de Chile.

BANFI DEL RÍO Cristian. 2017. Relevancia del dolo en la responsabilidad extracontractual chilena: una relectura desde el derecho inglés. Revista de Derecho. Año 24. N. 2. (69-107).

BANJA John D et al. 2022. When Artificial Intelligence Models Surpass Physician Performance: Medical Malpractice Liability in an Era of Advanced Artificial Intelligence. J Am Coll Radiol. 19.

BATHAEE Yavar. 2018. The Artificial Intelligence Black Box and the failure of Intent and Causation. En Harvard Journal of Law & Technology. Volumen 31, Number 2. 890-934.

BISHOP Chris. 1994. Neural networks and their applications. Rev Sci Instrum. 65. (1803-1832)

BOUCHER Philip. 2020. Artificial Intelligence: How does it work, why does it matter, and what can we do about it? Panel for the Future of Science and Technology, European Parliamentary Research Service.

BUCKNALL Benjamin y DORI-HACOHEN Shiri. 2022. Current and Near-Term AI as a Potential Existential Risk Factor. AIES ’22 (1-3).

BUSTOS Magdalena. 2024. The ethical principle of human intervention in civil liability. Acta Bioethica. 30. (41-49).

CASALS Miquel. 2023. Las propuestas de la Unión Europea para regular la responsabilidad civil por daños causados por sistemas de inteligencia artificial. InDret 3. (55-100).

CHINEN M. 2019. Law and Autonomous Machines. Cheltenham. Edward Elgar Publishing.

CORRAL TALCIANI Hernán. 2003. Lecciones de responsabilidad civil extracontractual. Editorial Jurídica de Chile.

DE BRUYNE J. y VANLEENHOVE C. 2021. Artificial Intelligence and the Law. Intersentia.

DIAKOPOULOUS Nicholas. 2014. Algorithmic Accountability Reporting: On the Investigation of Black Boxes. Tow Center for Digital Journalism, Columbia University.

EL NAQA Issam y MURPHY Martin J. 2015. What is Machine Learning? Machine Learning in Radiation Oncology.) (3-11).

ESTEBAN E.H. 2018. Inteligencia artificial y vehículos autónomos: el régimen de la responsabilidad civil ante los nuevos retos tecnológicos. En Revista Aranzadi de derecho y nuevas tecnologías. (48).

ERDÉLYI y ERDÉLYI (2020). The AI Liability Puzzle and a Fund-Based Work-Around. Journal of Artificial Intelligence Research. 70. (1309-1334).

ESCHENBACH Warren. 2021. Transparency and the Black Box Problem: Why We Do Not Trust AI. En Philosophy & Technology. 34. (1607-1622)

FRIEDMAN Batya y NISSENBAUM Helen. 1996. Bias in Computer Systems. ACM Transactions on Information Systems, Vol. 14. No. 3. (330-347)

FIGUEROA YAÑEZ Gonzalo. 2011. Curso de derecho civil – Tomo IV-. Editorial Jurídica de Chile.

GRIFFIN Frank. 2021. Artificial Intelligence and Liability in Health Care. En Health Matrix: The Journal of Law-Medicine. Vol. 31, Issue 1. (65-106)

GLAUBITZ, Alina. 2021. How should liability be attributed for harms caused by biases in Artificial Intelligence? Senior thesis, Yale Department of Political Science.

HALLEVY Gabriel. 2013. When Robots Kill: Artificial Intelligence under Criminal Law. Northeastern University Press.

ISAAK, J. y HANNA, M.J. 2018 User Data Privacy: Facebook, Cambridge Analytica, and Privacy Protection. Computer. 51. (56-59).

JUHÁSZ-RÉKA Agnes. 2016 Legal questions on the appearance of self-driving cars in the road traffic with special regard on the civil law liability. En European Integration Studies. Vol. 12. (10-28)

KINGSTON J.K. 2016. Artificial intelligence and legal liability. En International Conference on Innovative Techniques and Applications of Artificial Intelligence. Springer, Cham. (269-279)

KLEINBERG Jon et al. 2017. Inherent Trade-Offs in the Fair Determination of Risk Scores. En Proceedings of Innovations in Theoretical Computer Science.

KOWERT Weston. 2017. The foreseeability of human-artificial intelligence interactions. En Texas Law Review. (96).

LARROCAU Jorge. 2007. Culpa y dolo en la responsabilidad extracontractual: análisis jurisprudencial. LexisNexis.

LEHMANN Jos, BREUKER Joost y BROUWER Bob. 2004. Causation in AI and Law. En Artificial Intelligence and Law. 12. (279-315)

LIMA Gabriel et al. 2021. Human perceptions on Moral Responsibility of AI: A Case Study in AI-Assisted Bail Decision-Making. En Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. (1-17)

MCDERMOTT Drew. Artificial intelligence and consciousness. 2007. Cambridge Handbook of Consciousness. Cambridge University Press (117-150).

MERCADAL J.J.M. 2018. Vehículos autónomos y derecho de daños. La estructura clásica de la responsabilidad civil frente al avance de la inteligencia artificial. En Revista de la Facultad de Ciencias Económicas. (20).

MOOR James. 2006. The Darthmouth College Artificial Intelligence Conference: The Next Fifty Years. En AI Magazine Volume 27 N. 4.

NILSSON Nils J. 2010. The Quest for Artificial Intelligence. A history of ideas and achievements. Cambridge University Press. http://www.cambridge.org/us/0521122937

NTOUTSI Eirini et al. 2019. Bias in data-driven artificial intelligence systems- An introductory survey. WURE's Data Mining Knowledge Discovery.

ROBBINS Scott. 2019. A Misdirected Principle with a Catch: Explicability for AI. En Minds & Machines. (495-514).

RUSSEL Stuart J y NORVIG. 2014. Artificial Intelligence: a modern approach. New Jersey. Prentice Hall.

PAPAYANNIS Diego. 2014. Comprensión y justificación de la responsabilidad extracontractual. Madrid. Marcial Pons.

PAPAYANNIS Diego (coordinador). 2022. Manual de derecho de daños extracontractuales. Direccción General de Derechos Humanos de la Nación. México.

PEREZ RASTELLI Joshué. 2012. Agentes de control de vehículos autónomos enenotrnos urbanos y autovías. Memoria para optar al grado de Doctor.

PRINO EMHART Alberto. 2013. Entre reparación y distribución: la responsabilidad civil extracontractual como mecanismo de distribución de infortunios. Revista chilena de derecho privado. N. 21. https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-80722013000200004

PREECE Alun et al. 2018. Stakeholders in Explainable AI. Arxiv Preprint.

RAMOS PAZOS René. 2008. De la responsabilidad extracontractual. Jurídica Conosur.

ROSS Philip y MAYNARD Kasia. 2021. Towards a 4th industrial revolution. Intelligent Buldings International. (13). Pp. 159-161.

Roweis, S. T. (2000). Nonlinear Dimensionality Reduction by Locally Linear Embedding. Science, 290(5500), 2323–2326. doi:10.1126/science.290.5500.2323

SAUER, Frank. 2022. Lethal autonomous weapons systems. En: The Routledge Social Science Handbook of AI. Routledge, Abingdon, Oxon. (237-250).

SAN MARTÍN, Lilian. 2021. El caso fortuito en la responsabilidad civil extracontractual. En: Revista Ius et Praxis, año 27, N.º 2. (3-20).

SCHELLEKENS Maurice. 2015. Self-driving cars and the chilling effect of liability law. En Computer Law & Security review, vol. 31. (506-517).

SCHERER Matthew. 2015. Regulating artificial intelligence systems: Risks, challenges, competencies, and strategies. En Harv. JL & Tech. (29)

SELBST Andrew. 2020. Negligence and AI's Human Users. 100 B.U.L. REV. (1315).

SMITH Helen. 2021. Clinical AI: opacity, accountability, responsibility and liability. En AI & Society 36. 535-545.

SOYER Baris y TETTENBORN Andrew. 2023. Artificial intelligence and civil liability-do we need a new regime?

SULLIVAN Hannah y SCHWEIKART Scott. 2019. Are current tort liability doctrines adequate for addressing injury caused by AI? En AMA journal of ethics. 2 (21). Pp. 160-166.

UMAR Wahyudi et al. 2023. Artificial General Intelligence (AGI) and Its Implications for Contract Law. Indonesian Journal of Artificial Intelligence and Data Mining. Vol. 6.

VLADECK, David. 2014. Machines without principles: liability rules and artificial intelligence. En Washington Law Review. (89).

YU Ronald y SPINA Gabriele. 2019. What's Inside the Black Box? AI Challenges for Lawyers and Researchers. En Legal Information Management. 19. (2-13).

VRANA Johannes y SINGH Ripi. 2020. The NDE 4.0: Key Challenges, Use Cases and Adaptation. Preprint.

WANG Pei. 2019. On Defining Artificial Intelligence. Journal of Artificial Intelligence. 10. (1-37).

WANG Pei, LIU Kai y DOUGHERTRY. 2018. Conceptions of Artificial Intelligence and Singularity. Information. 9. doi:10.3390/info9040079

ZAVRSNIK Ales. 2020. Criminal Justice, artificial intelligence systems, and human rights. En ERA forum 20. 567-583. https://doi.org/10.1007/s12027-020-00602-0