



Universidad de Chile

Facultad de Derecho

Departamento de Derecho Procesal

PROBLEMÁTICAS DE LA IRRUPCIÓN DEL *DEEPPFAKE* Y LA INTELIGENCIA ARTIFICIAL EN EL PROCESO CHILENO

Memoria para optar al grado de Licenciado en Ciencias Jurídicas y Sociales

Felipe Rivera Capra

Profesor guía: Jesús Ezurmendia Álvarez

Santiago, Chile

2024

*A mi madre, padre y abuelos,
Por su amor y apoyo incondicional*

ÍNDICE

RESUMEN	4
INTRODUCCIÓN	5
CAPÍTULO I: ¿QUÉ ES LA INTELIGENCIA ARTIFICIAL?	7
1.1 Definición	7
1.2 Evolución	7
1.3 ¿Qué es el <i>deepfake</i> ?.....	10
CAPÍTULO II: DEEPFAKE COMO MATERIAL PROBATORIO: LA PRUEBA	
AUDIOVISUAL EN EL PROCESO	12
2.1 La prueba audiovisual en el proceso civil	13
2.2 La prueba audiovisual en el proceso penal.....	18
2.3 Otras disposiciones relevantes.....	19
CAPÍTULO III: DESAFÍOS DE LA APARICIÓN DEL DEEPFAKE Y LA IA EN EL	
PROCESO	21
3.1 Masificación y democratización de herramientas de IA	21
3.2 Problemáticas relevantes causadas por el <i>deepfake</i>	22
3.3 Sesgo y discriminación por los algoritmos	29
CAPÍTULO IV: REGULACIÓN COMPARADA Y PROPUESTAS PARA	
REGLAMENTAR LA INTELIGENCIA ARTIFICIAL Y EL DEEPFAKE	31
4.1 La experiencia de la Unión Europea.....	31
4.1.1 Clasificación de IA según su riesgo	34
4.1.2 Obligaciones de los desarrolladores/proveedores de sistemas de IA de alto riesgo	35
4.1.3 IA de propósito general/GPAI e IA Generativa	36
4.2 Otras experiencias relevantes	39
4.3 Propuestas para mitigar a la IA y el <i>deepfake</i>	44
CONCLUSIÓN	46
BIBLIOGRAFÍA	48

RESUMEN

El presente trabajo trata sobre cómo la irrupción de la inteligencia artificial, y en particular la tecnología denominada *deepfake*, traerá nuevos desafíos para la sociedad y para el sistema procesal chileno. En este sentido, la estructura que seguirá será la siguiente: en el Capítulo I, se definirá y expondrá la evolución de la inteligencia artificial, además de conceptualizar qué es el *deepfake*. Posteriormente, en el Capítulo II, se analizará la prueba audiovisual en relación a cómo el *deepfake* y el *voice cloning* son valorados en virtud de los principales sistemas de valoración de la prueba que convergen en nuestro ordenamiento. En el Capítulo III contextualizaremos la masificación y democratización de las herramientas de inteligencia artificial que vivimos en los últimos años, y se evidenciarán las eventuales problemáticas de su repentina aparición en la sociedad, poniendo especial énfasis en los problemas probatorios que traerá aparejada el uso del *deepfake* y la problemática del uso de algoritmos como política pública. Finalmente, en el Capítulo IV, se revisará la legislación comparada respecto a la regulación de la inteligencia artificial de distintos estados, principalmente de la Unión Europea, así como los intentos de regulación nacional. Además, se plantearán propuestas para mitigar los efectos negativos de esta tecnología.

Palabras clave: Inteligencia Artificial, *Deepfake*, *Voice Cloning*, Prueba Audiovisual, Unión Europea, Proceso

INTRODUCCIÓN

En la era digital en la que vivimos, el avance tecnológico es esencial para el desarrollo moderno de la sociedad. En este sentido, la última gran revolución tecnológica ha llegado hace ya varios años, y es la irrupción de la Inteligencia Artificial¹ (en adelante IA). La creación de Chat GPT hizo accesible esta tecnología a millones de usuarios de internet, que vieron la posibilidad de ocupar de manera fácil y rápida un chat capaz de responder con cierto margen de precisión cualquier pregunta que se le haga, tomando como base de datos la internet.

La aparición de esta tecnología causó que apareciera una variedad programas que utilizaban la IA como base para la creación de sistemas con diversos fines. Poco a poco, la IA fue abarcando cada vez más espacios, tanto por parte de empresas dedicadas a la tecnología, que buscaban crear sistemas más complejos, como por personas comunes y corrientes, que lo utilizaban con fines de entretenimiento, en donde el material era especialmente difundido en internet. En este contexto aparecen los *deepfake*. Los *deepfake* son una tecnología potenciada por IA, capaz de poner la cara de personas en videos, o de imitar la voz o forma de escribir de una persona, todo mediante un sistema capaz de aprender de fonemas, expresiones faciales o muletillas específicas de una persona².

Esta tecnología se ha utilizado para distintas funciones, principalmente de ocio y de desinformación vía redes sociales³. Sin embargo, también puede ser utilizado para fines fraudulentos, con la posibilidad de cometer ciertos delitos con esta tecnología⁴. Los estados de muchos países se han puesto en alerta tras la aparición de la IA y su masificación, ya que su creación ha tenido diversas implicancias en lo jurídico de forma transversal, desde problemáticas de propiedad industrial, derecho laboral, derecho procesal y derecho penal.

Dada su capacidad de abarcar diversas esferas de nuestro día a día, la irrupción del *deepfake* ha arremetido de manera abrupta en el mundo judicial, generando una serie de nuevos desafíos a sortear. En el ámbito procesal también ha traído desafíos de gran envergadura, especialmente relacionados con las pruebas que son aportadas al proceso. Debido a la facilidad de esta tecnología de generar o modificar material audiovisual, su aparición cambia el paradigma de los magistrados, por ejemplo, al momento de valorar una prueba que históricamente ha sido infalible debido a su alta dificultad de alterar por parte de las personas, en donde era necesario un experto en edición para hacer pruebas contundentes que puedan pasar desapercibidas en el proceso. Este es sólo un ejemplo de cómo la irrupción de la IA puede complicar a los jueces. La aparición del *deepfake* causará a largo plazo que los jueces dejen de confiar

¹ Cheng, «La Revolución de la IA».

² Smith y Mansted, «What's a deep fake?»

³ Helmus, «Artificial Intelligence, Deepfakes, and Disinformation».

⁴ Nurkin, Rodriguez, y Carter, «A Framework for Understanding Applied Ai».

en el material audiovisual como lo hacían normalmente, perdiendo la calidad de esta prueba, que es considerada como “infalible”, a una prueba que puede ser fácilmente modificada y alterada, siendo valorada por los jueces de los sistemas que utilizan la sana crítica de una forma más mesurada, a diferencia de cómo estas pruebas son valoradas en la actualidad. Otra problemática que analizaremos es la aparición de los algoritmos en la vida diaria de las personas, en donde profundizaremos en el sesgo y la discriminación en la que pueden incurrir debido a la forma de “entrenamiento” de estos algoritmos⁵.

A lo largo de esta tesis analizaremos la historia de la IA y su evolución, revisaremos los pasos que llevaron a transformar a esta tecnología en lo que es actualmente. Posteriormente, se hará un énfasis particular en el *deepfake* y en las problemáticas que esta tecnología trajo a la vida moderna, como también se analizarán problemas que eventualmente podría traer al mundo jurídico, particularmente en materia procesal y la posibilidad de constitución de ciertos delitos. Asimismo, como adelantamos hace unas palabras, se revisará la problemática de los algoritmos en la sociedad, otro de los principales problemas de la IA en la utilización de esta en políticas públicas⁶.

Después, describiremos las características de la prueba audiovisual en los sistemas procesales chilenos, principalmente en el proceso civil, caracterizado por su antigüedad y dificultad para responder al fenómeno de la IA, y el proceso penal, un proceso que se encuentra reformado y que responde de manera radicalmente diferente al proceso civil al momento relacionarse con la prueba audiovisual.

Posteriormente, contextualizaremos la masificación y la democratización del *deepfake*, tecnología que hoy en día es alcanzable para cualquier persona con un dispositivo electrónico inteligente capaz de conectarse a internet, pudiendo generar modificaciones o falsificaciones de alta calidad⁷.

Por último, analizaremos las respuestas por parte de los estados al fenómeno de la IA, haciendo especial hincapié en la regulación de la Unión Europea (en adelante UE) que es caracterizada como la piedra angular para la regulación por parte de diversos estados del mundo⁸. Analizaremos los esfuerzos que actualmente hace nuestro país para regular este fenómeno, y cerraremos con propuestas para mitigar el uso indebido de la IA y del *deepfake* en distintas aristas, dando énfasis en el proceso judicial.

⁵ Julia Angwin et al., «Machine Bias».

⁶ Ibid.

⁷ Citron y Chesney, «Deep Fakes».

⁸ Haataja y Bryson, «The European Parliament’s AI Regulation».

CAPÍTULO I: ¿QUÉ ES LA INTELIGENCIA ARTIFICIAL?

1.1 Definición

Desde tiempos antiguos, el ser humano ha buscado replicar el intelecto humano, y realizar tareas que requieran inteligencia. Es así, como en esta incesable búsqueda de replicar los procesos cognitivos, nace la IA, que busca imitar las funciones del cerebro a través de esta tecnología que está en constante evolución.

La IA tiene varias formas de ser definida, dependiendo del enfoque que se le dé en el área de la informática. La concepción más famosa, y que resulta más cercana a cómo son en la realidad, es la que realiza JOHN MCCARTHY y sus colegas investigadores en 1955, que caracteriza a la IA como “hacer que una máquina se comporte de manera inteligente, como si un ser humano se comportara de ese modo”. Por otra parte, llega a una comprensión similar TURING, FLORINI y COWLS, que definen a la IA como “un recurso creciente de agencia interactiva, autónomo y, a menudo, autodidacta (refiriéndose a la capacidad de aprendizaje automático), que puede realizar con éxito tareas que, de otro modo, requerirían la inteligencia e intervención humana⁹.”

1.2 Evolución

Los inicios de la IA se remontan a 1842, cuando la matemática y pionera en informática, ADA LOVELACE, programa el primer algoritmo que sería procesado por una máquina. LOVELACE especulaba que la máquina en cuestión, y a futuro, podría perfectamente actuar sobre otras cosas, además de solamente números, postulando incluso que se podrían elaborar piezas de alta complejidad o extensión, como lo puede ser una pieza musical¹⁰. Décadas más tarde, se marca el siguiente hito fundacional de la IA: en una conferencia celebrada en Darmouth en 1956 organizada por JOHN MCCARTHY, MARVIN MINSKY, CLAUDE SHANNON y NATHANIEL ROCHESTER se origina por primera vez el término de “inteligencia artificial”, estableciéndose así el nombre y creando el campo de estudio, formalizando esta rama científica. Esta reunión planta la semilla en los asistentes: el pensamiento como una forma de computación que no es exclusiva de los seres humanos o seres biológicos, sino que la inteligencia humana se puede simular o replicar de manera artificial en máquinas digitales¹¹.

⁹ Kavanagh, «Artificial Intelligence».

¹⁰ Fuego y Francis, «Lovelace & Babbage and the creation of the 1843 'notes'».

¹¹ Abeliuk y Gutiérrez, «Historia y evolución de la inteligencia artificial».

El mismo año en que se realizó la conferencia, dos asistentes de esta, ALAN NEWELL y HERBERT SIMON, publican el primer programa computacional de IA¹² llamado “*Logic Theory Machine*”. Este programa es el primero en poder descubrir demostraciones de teoremas en lógica simbólica, pudiendo combinar simples operaciones, logrando así construir expresiones cada vez más complejas¹³. El desafío era evidente: encontrar las combinaciones de operaciones que demuestren un teorema dado por el programador, entre la exponencial cantidad de combinaciones formadas por estas operaciones primitivas¹⁴. Es en este punto donde se le da el enfoque heurístico¹⁵ a la IA, haciendo que estos procesos encuentren de manera inteligente una solución, recortando las combinaciones a las más idóneas para resolver la mayoría de los casos. La introducción de la heurística en los procesos de la IA, según los mismos autores, es lo que denominan la “magia central” en toda resolución de problemas humanos. Este aporte le valió a HERBERT SIMON el Premio Nobel en Economía por modelar el comportamiento humano como una “racionalidad limitada”, en donde la toma de decisiones es principalmente heurística¹⁶.

Aquí nace una distinción relevante en el campo de la IA: el modelo simbólico y el modelo conexionista. El primero se basa en el razonamiento lógico y la búsqueda heurística como pilares fundamentales de su funcionamiento, buscando así solución a los problemas sin necesidad de tener un cuerpo físico ni estar materializado en un entorno real¹⁷. Lo anterior no significa que sea una exigencia prohibitiva, aunque en la práctica solía concebirse de esta manera¹⁸. Son representaciones abstractas del mundo real, modeladas mediante lenguaje informático basado en la matemática. Fue así como nacieron las IA, que eran capaces de demostrar teoremas o, con un ejemplo más popular en la época, la capacidad de jugar ajedrez, ya que no era necesario mover físicamente las piezas en un tablero real, sino que a través de coordenadas del tablero eran capaces de hacer las movidas respectivas¹⁹. Así, se llega a uno de los logros más importantes y populares en el desarrollo de la IA: el computador integrado con IA simbólica, Deep Blue, desarrollado por IBM, logra vencer al campeón mundial de ajedrez de 1997, GARRY KASPAROV²⁰.

De forma paralela al desarrollo de la IA simbólica, se desarrolla la IA conexionista, que es una escuela de pensamiento distinta, basada en modelar la biología del cerebro compuesto por redes neuronales biológicas, partiendo de la hipótesis en que la inteligencia nace a partir de la actividad

¹² Newell y Simon, «The logic theory machine—A complex information processing system».

¹³ Op. Cit. Abeliuk y Gutiérrez

¹⁴ Ibid.

¹⁵ En el campo de la inteligencia artificial, “heurístico” está relacionado con la tarea de resolver problemas inteligentemente utilizando la información disponible.

¹⁶ Wheeler, «Bounded rationality».

¹⁷ Op. Cit. Abeliuk y Gutiérrez

¹⁸ López de Mántaras, «Algunas reflexiones sobre el presente y futuro de la Inteligencia Artificial».

¹⁹ Ibid.

²⁰ Hansen Hsu, «AI and Play, Part 1».

distribuida de una gran cantidad de unidades, interconectadas entre sí, que procesan una elevada masa de información de manera paralela²¹. La forma de conectar estas redes neuronales artificiales toma la inspiración en el modelo de reforzamiento de las sinapsis entre neuronas del cuerpo humano del profesor DONALD HEBB²², pudiendo entrenar estas redes neuronales para que aprendieran funciones relacionadas con entradas y salidas mediante el ajuste de los pesos utilizados para ponderar las conexiones entre las neuronas²³. Esta forma de modelar la IA se empieza a plantear como un modelo mucho mejor que el planteado por la IA simbólica, siendo idóneos para cognición, aprendizaje y la memoria²⁴. A causa de un libro publicado en 1969 por pioneros de la IA simbólica, MARVIN MINSKY y SEYMOUR PAPERT, la investigación de la IA conexionista se ve detenida, principalmente por la falta de financiamiento de la investigación y por una sobre expectativa que no fue cumplida²⁵.

Es a comienzos de 1980 cuando GEOFFREY HINTON, quien fue en 2018 el ganador del Premio Turing²⁶, junto a otros informáticos del área, que redescubren un algoritmo central detrás de la búsqueda heurística propia de la IA simbólica, que logra encontrar parámetros del mismo modelo para minimizar sus errores en la red neuronal del modelo de la IA conexionista, pudiendo lograr que una máquina con red neuronal de múltiples capas aprenda a partir de datos. Es el popular método de la retropropagación²⁷.

Ya en 2010, gracias a los avances tecnológicos de la informática, hay disponible *hardware* altamente especializado, que han acelerado de manera drástica el entrenamiento y rendimiento de las redes neuronales, además de ver reducido su consumo de energía. Esto, sumado al aumento importante de datos a disposición -gracias a internet y a los bancos de datos de manera online, y servicios de costo reducido para etiquetar datos- le han dado el impulso necesario a la IA para llegar a un nivel de desarrollo altamente complejo²⁸.

La directa consecuencia de estos avances se traduce en las innovaciones tecnológicas que tenemos hoy en día: aplicaciones de reconocimiento facial, detección de patrones complejos -utilizados en la medicina, por ejemplo, para detectar cáncer-, traducción de idiomas, asistentes virtuales, etc.²⁹ Un ejemplo del nivel al que llega la IA en comparación con su inicio es el sistema AlphaGo de DeepMind,

²¹ Op. Cit. Abeliuk y Gutiérrez

²² Hebb, *The organization of behavior*.

²³ Ibid.

²⁴ Op. Cit. López de Mántaras.

²⁵ Minsky y Papert, «An introduction to computational geometry».

²⁶ El premio Turing es un premio otorgado anualmente por la asociación para la maquinaria computacional (ACM) relativo a las ciencias de la computación.

²⁷ Rumelhart, Hinton, y Williams, «Learning representations by back-propagating errors».

²⁸ Op. Cit. Abeliuk y Gutiérrez.

²⁹ Ibid.

que en 2017 venció al campeón mundial de Go³⁰, un complejo juego de origen chino en donde hay, en promedio, 300 movimientos posibles en cada turno, versus los 30 que hay en el ajedrez. El árbol de búsqueda de este juego tiene una cantidad de ramificaciones considerablemente mayor.

En noviembre del año 2022, OpenIA lanzó al internet Chat GPT, una herramienta de IA que cuenta con la particularidad de ser un interfaz conversacional, capaz de responder preguntas a los usuarios, así como realizar acciones que impliquen la generación de lenguaje natural, y variedad de tareas más complejas³¹.

Para poder sacarle el máximo provecho a cualquier *chatbot* de la actualidad, como lo es Chat GPT, es necesario manejar de manera hábil el uso de *prompts*. Los *prompts* son las instrucciones o indicaciones específicas que se le proporcionan al *chatbot* en cuestión, útiles para que la respuesta o generación de contenido que nos proporcione sea preciso y adecuado a la finalidad que estamos buscando. Mientras más precisos seamos al momento de utilizar el *prompt*, mejor resultado nos dará la IA³².

Hay que recalcar que el proceso de creación o entrenamiento no es 100% perfecto, es posible que la IA cometa errores, o interprete de forma errónea la instrucción dada. Una de las novedades más útiles de la tecnología GPT es la capacidad de retener en los bancos de memoria todo el chat producido con el usuario. En este sentido, si el *chatbot* comete un error, lo podemos encauzar, especificándole lo que queremos de respuesta, a la vez que le explicitamos por cuál tipo de información no nos queremos adentrar, o qué tipo de orden se debe abstener de seguir, a la luz de seguir adecuadamente el *prompt* dado.

1.3 ¿Qué es el *deepfake*?

La eficiente capacidad de aprender y de retener en su memoria de la IA ha dado pie a que muchos usuarios de internet creen aplicaciones con la capacidad de imitar la voz o rasgos faciales de las personas³³, pudiendo elaborar tanto audios como videos con finalidades fraudulentas. Lo anterior es conocido como “*deepfake*”.

El *deepfake* es una falsificación digital creada con *deep learning* (un subcampo de la IA que le da dirección al *machine learning*, que es la capacidad de la IA de aprender de forma automatizada³⁴), que tiene la capacidad de crear o manipular cualquier tipo de material audiovisual, pudiendo generar o

³⁰ Silver et al., «Mastering the game of Go with deep neural networks and tree search».

³¹ Javier Márquez, «GPT-4: así es la IA más avanzada de OpenAI, cómo funciona y todas las novedades».

³² «Prompts de Chat GPT para eCommerce».

³³ Op. Cit. Nurkin, Rodríguez y Carter

³⁴ Shinde y Shah, «A Review of Machine Learning and Deep Learning Applications».

alterar imágenes, videos, audios y textos³⁵. El nacimiento del *deepfake* trae la evolución de la forma en la que la desinformación en las redes sociales se dispersa por internet³⁶. Es una herramienta de edición que le permite al usuario, de forma completamente automática, reemplazar el rostro de cualquier persona por el rostro de otra, o falsear los fonemas de cualquiera, para ocupar su tono de voz y poder decir lo que el creador del *deepfake* quiera³⁷. Esto se logra mediante el “entrenamiento” de la IA, nutriendo al sistema de la mayor cantidad de imágenes de la víctima, de la mejor calidad y de todos los ángulos posibles, para perfeccionar la calidad de esta³⁸. Asimismo, las voces generadas por IA analizan la voz de la persona afectada, y al igual que con el *deepfake* visual, mientras más se “alimente” a la IA de material de la persona hablando, no solo puede imitar fonema por fonema la voz humana, sino que aprender de las muletillas, respiraciones y acento de la persona afectada³⁹.

³⁵ Op. Cit. Smith y Manstead

³⁶ Op. Cit. Helmus

³⁷ Ibid.

³⁸ Arcila y Pabon, «Repercusiones éticas sobre el uso indebido del deepfake en el ámbito de las TIC mediante un análisis cualitativo documental».

³⁹ Op. Cit. Nurkin, Rodríguez y Carter

CAPÍTULO II: *DEEPPFAKE* COMO MATERIAL PROBATORIO: LA PRUEBA AUDIOVISUAL EN EL PROCESO

Para adentrarnos en las problemáticas que el *deepfake* puede causar al momento de que este material audiovisual sea añadido al proceso, es necesario esclarecer, en primer lugar, a qué nos referimos con la prueba, qué medio de prueba es el que corresponde con este tipo de contenido audiovisual que se añade al proceso, en qué tipo de prueba, según la doctrina, se podrían encasillar los materiales audiovisuales creados con medios tecnológicos y la forma en que los magistrados las valoran.

En primer lugar, cuando nos referimos al concepto de prueba, entendida para este contexto como prueba judicial (debido a lo polisémico y multifuncional de la palabra⁴⁰) nos podemos aproximar a esta desde diferentes puntos de vista: desde la prueba como actividad, como resultado y como medio de prueba⁴¹. En primer lugar, se entiende como actividad, ya que esta se va desarrollando dentro del proceso, en donde las partes aportan antecedentes que sustenten sus alegaciones y el juez determina la cuestión de hecho que se debate. La activa intervención de los distintos sujetos procesales en el procedimiento probatorio es entendida como la parte más dinámica de los aspectos de la prueba⁴². Por otra parte, la prueba también se puede entender como un resultado, en el sentido de que es la conclusión a la que llega el juez sobre el hecho que debe ser probado (el *factum probandum*) basándose en las evidencias aportadas. El juez determina qué afirmaciones de las partes respecto a los hechos que suscitaron el juicio son dadas por verificadas⁴³. Por último, también podemos entender a la prueba como un medio, en el sentido de que necesita de ciertos elementos que le sirvan de base para que el juez de por acreditadas los enunciados sobre hechos de la causa⁴⁴. Así, se puede entender a la prueba, especialmente en los sistemas de *civil law* como el nuestro, como los antecedentes respecto a los que el tribunal se basa para determinar la materia fáctica de una disputa judicial⁴⁵. Al respecto, TARUFFO se refiere a los medios de prueba como “cualquier elemento que pueda ser usado para establecer la verdad acerca de los hechos de la causa”⁴⁶.

⁴⁰ Taruffo, *La prueba de los hechos*.

⁴¹ Meneses Pacheco, «FUENTES DE PRUEBA Y MEDIOS DE PRUEBA EN EL PROCESO CIVIL».

⁴² Ibid.

⁴³ Abellán, *Los hechos en el Derecho*.

⁴⁴ Op. Cit. Meneses.

⁴⁵ Ibid.

⁴⁶ Michele, *La prueba*.

Por otra parte, CARNELUTTI tiende a distinguir entre las fuentes de prueba y los medios de prueba. Las fuentes de prueba son los conceptos preexistentes al proceso (un documento añadido al juicio, un testigo, el conocimiento de un perito). Por otra parte, los medios de prueba son entendidos como conceptos que existen en y para el proceso (interrogatorio de partes, reconocimiento judicial)⁴⁷. En particular, en el proceso civil chileno, en donde podemos desprender del art. 341 del Código de Procedimiento Civil que nuestro sistema probatorio es uno legal tasado, o sea, un sistema que tiene reglas de valoración y decisión que perentoriamente señalan a los jueces lo que debe ser considerado como prueba, como por ejemplo, cuando una resulta plena o semiplena, y asimismo, contamos con un listado taxativo de medios y fuentes de prueba (siguiendo la clasificación de CARNELUTTI)⁴⁸. Esta taxatividad de los medios y fuentes de prueba del sistema civil chileno se estaría explicando por lo dispar que resultarían, a la luz de esta distinción, medios y fuentes de prueba, siendo los medios de prueba muy limitados en comparación a las fuentes de prueba que se pueden aportar al proceso⁴⁹.

Entonces, con base en todas las distinciones que hemos repasado respecto a prueba, y en particular, a medios de prueba, ¿En dónde podemos encasillar a la prueba generada por *deepfake*, ya sea un video, entendido como su formato más común, o el *voice cloning*, siendo mucho más problemático por su dificultad de detectar, dependiendo de los dos grandes sistemas procesales que tenemos?

2.1 La prueba audiovisual en el proceso civil

Es necesario volver a hacer mención del art. 341 del Código de Procedimiento Civil. Este artículo se encuentra en el Título XI del Libro II, que regula los medios de prueba en particular. Como dijimos anteriormente, hace un listado taxativo de los medios de prueba que se pueden utilizar en el proceso civil. Estos son: (i) instrumentos; (ii) testigos; (iii) confesión de parte; (iv) inspección personal del tribunal; y (v) informe de peritos y presunciones.

Para nuestro análisis, cobra especial relevancia un medio de prueba particular: el instrumento. El instrumento por antonomasia en el proceso es el documento, y este puede ser incluso entendido como sinónimo de instrumento. El sentido más vulgar de la palabra “documento” refiere a una hoja de papel que contiene determinada información de forma escrita. Sin embargo, en el contexto jurídico, el concepto de documento toma diversos significados dados por la doctrina, que buscan dotar de una significancia que sea más acorde al mundo legal. Para esto, es inevitable hacer mención del origen etimológico de la palabra “documento”.

⁴⁷ Carnelutti, *La prueba civil*.

⁴⁸ Padilla Parot, «ESE DOLOR DE CABEZA LLAMADO PRUEBA LEGAL TASADA».

⁴⁹ Op. Cit. Rodríguez

La palabra “documento” viene del griego *dék*, que se desprende del verbo del latín *docere*, o instruir, de donde proviene el vocablo *documentum*, que podríamos dotar del siguiente significado: “lo que se enseña, con lo que alguien se instruye”. De la definición precedente, podemos definir preliminarmente a la palabra “documento” como algo que muestra, que indica algo⁵⁰.

A pesar de la etimología de la palabra, la doctrina es diversa respecto a la noción de documento, en donde tiene diversos significados en el contexto jurídico y legal. Encontramos 3 concepciones relevantes de este vocablo. Dentro de los autores clásicos, CARNELUTTI entiende al documento como “una cosa representativa, o sea, capaz de representar un hecho”⁵¹. Diversos autores explican que lo que el autor se refiere es a que el documento es una imagen de la realidad, que se presenta al intelecto a través de los sentidos. Así, un “documento” sería una cosa que sirve para representar otra⁵². Por otra parte, desde el punto de vista de la cultura general, podemos dotar de otro significado: entendiendo a la cultura como lo realizado por el hombre, el documento, en su aspecto más amplio, sería “toda entidad material perceptible por los sentidos que revele intervención humana”⁵³. Dentro del ámbito más jurídico, en una de sus vertientes más amplias, es posible entender la noción de documento como “aquellos objetos que tengan una función probatoria con la sola limitación (según GUASP) de que dichos objetos sean, por su índole, susceptibles de ser llevados ante la presencia judicial”⁵⁴. Otras doctrinas plantean, a la luz del derecho instrumental de forma más genérica, al documento como un escrito de soporte en papel, y lo entienden de forma estricta, como “todo recipiente en el cual se vierten por escrito manifestaciones o declaraciones. Estrictamente entendidos, los documentos son escritos o escrituras”⁵⁵.

De las concepciones ya expuestas, las podemos resumir en dos teorías que buscan explicar la realidad jurídica del documento jurídico. Estas son: la teoría del escrito, en donde un documento siempre es un escrito, y la teoría de la representación, consistente en que un documento no se limita solamente a un escrito, sino que se entiende como todo documento representativo o que pueda informar sobre un hecho u otro objeto⁵⁶. Esta última concepción admite dentro de su haber la concepción de un documento de carácter electrónico⁵⁷. Es importante destacar que ni la doctrina ni la legislación como tal han buscado una noción unívoca de lo que significa documento, sino que más bien coexisten estas acepciones que evidencian distintas realidades respecto a la palabra misma⁵⁸.

⁵⁰ Carrascosa, Pozo, y Rodríguez de Castro, *La contratación informática: el nuevo horizonte contractual. Los contratos electrónicos e informáticos*.

⁵¹ Op. Cit. Carnelutti

⁵² Pinochet Olave, «EL DOCUMENTO ELECTRÓNICO Y LA PRUEBA LITERAL».

⁵³ Ibid.

⁵⁴ Ramón Camino, «El Documento Electrónico su admisibilidad en el Derecho español».

⁵⁵ Díez-Picazo, *Fundamentos del derecho civil patrimonial: Introducción; Teoría del contrato; Las relaciones obligatorias*.

⁵⁶ Carrascosa, Bauza, y González, «El Derecho de la Prueba y la Informática. Problemas y Perspectivas».

⁵⁷ Op. Cit. Pinochet.

⁵⁸ Ibid.

Dentro de la doctrina nacional, destacados exponentes del Derecho y las Nuevas Tecnologías, como FERNÁNDEZ, adoptan la noción de CARNELUTTI, por ser mucho más completa, y asimismo, admite la incorporación de manifestaciones documentarias más amplias, que representan un paso evolutivo inevitable y en constante evolución respecto a la noción de documento⁵⁹.

Así, llegamos la noción de documento electrónico. La principal característica de este instrumento es que tiene cabida en un soporte electromagnético, o sea, que tiene respaldo en una plataforma diferente a la forma óptica y auditiva, y lo tiene en una plataforma “informática”⁶⁰, entendida por antonomasia como existente en una computadora. Entonces, un carrete de fotos no estaría revestido con las características necesarias para ser considerada como un documento electrónico, aunque compartan determinadas de esas características, y, en cambio, un fichero electrónico que recogiese las fotos si sería considerada como un documento electrónico⁶¹. Sin embargo, parte de la doctrina ha entendido que, a pesar de que la creación misma del documento tiene que ser sobre la base de una plataforma informática, que asocian a su símil técnico “electromagnético”, no limita el documento electrónico a una base informática, ya que la electromagnética es un fenómeno físico que tiene cabida en otros soportes aparte de la informática. Respecto a esto, podemos destacar la existencia de variedad de sistemas de almacenamiento de soporte óptico, que se alejan de la electrónica o el electromagnetismo, en donde podemos encontrar variedad de documentos electrónicos⁶².

Según PINOCHET, podemos distinguir cuatro características esenciales del documento electrónico, o más bien dicho, documento informático. Estas características son las siguientes: (i) el documento debe estar escrito en lenguaje binario; (ii) el documento debe estar o poder ser almacenado en soporte informático, magnético, óptico o cualquier otra clase de soporte que pudiera ser desarrollado para tales fines; (iii) el documento debe poder ser, mediante la aplicación del correspondiente programa informático, transformado a alguna clase de lenguaje comprensible por el ser humano, y; (iv) la posibilidad de que tal información sea tratada informáticamente, lo que es llamado el “tratamiento automatizado de la información”, o sea, que pueda ser susceptible de ser copiada, modificada y transmitida por medios informáticos⁶³.

Respecto a la legislación nacional, podemos encontrar definido el concepto de documento electrónico en la ley 19.799, que regula los documentos ya mencionados, las firmas electrónicas y los servicios de certificación de estas. El artículo 2, letra d) de esta ley, entiende al documento electrónico

⁵⁹ Fernández, «El documento electrónico en el derecho civil chileno».

⁶⁰ Fernández y Grande, «La prueba y el documento electrónico».

⁶¹ Ibid.

⁶² Redondo, «La prueba de los documentos electrónicos en los distintos órdenes jurisdiccionales».

⁶³ Op. Cit. Pinochet.

como “toda representación de un hecho, imagen o idea que sea creada, enviada, comunicada o recibida por medios electrónicos y almacenada de un modo idóneo para permitir su uso posterior”. Como podemos observar, la definición legal hace referencia principalmente al seguimiento de la teoría de CARNELUTTI por parte del legislador, al señalar al documento electrónico como “toda representación de un hecho”, como también lo hace al referirse a una “imagen o idea”⁶⁴. Además de esta definición que nos da este artículo en particular, la ley prevé otra definición relevante, esta vez en la letra a) del mismo artículo 2, referida al concepto de “electrónico”: “a) Electrónico: característica de la tecnología que tiene capacidades eléctricas, digitales, magnéticas, inalámbricas, ópticas, electromagnéticas u otras similares”. Esta precisión enuncia ciertas características que, según la ley, debe contener todo lo considerado como “electrónico”, y por ejemplo, hace referencia a “inalámbrico” solo en lo referente a la transmisión de un documento. Además, se amplía la concepción de electrónico a otros soportes, como lo “óptico” o lo “electromagnético”, en relación con el lugar de almacenamiento de un documento⁶⁵, lo que se diferencia de lo ya visto planteado por CERVELLÓ Y FERNÁNDEZ⁶⁶.

Como hemos visto, y en relación con el análisis del concepto de documento que se enmarca en lo que el Código de Procedimiento Civil entiende por instrumento, podemos establecer que un video o una pista de audio que se aporte a un juicio en el término probatorio se haría con base en la amplia noción de documento, en donde encontraríamos el documento electrónico.

Es menester hacer una distinción que prevé la ley respecto a la naturaleza de los instrumentos en el proceso civil. El Código de Procedimiento Civil, en el Párrafo 2 del Título XI del Libro II, regula los instrumentos, y hace la distinción entre instrumentos públicos e instrumentos privados. Esta distinción se encuentra consagrada de manera explícita en el inciso 2 del art. 1.698 del Código Civil, en donde se plantea que “las pruebas consisten en instrumentos públicos o privados (...)”.

Según el Código Civil, el instrumento público o auténtico es el “autorizado con las solemnidades legales por el competente funcionario”. Si este es otorgado ante notario e incorporado en un protocolo o registro público, adquiere el nombre de instrumento público. Su principal particularidad, que es beneficiosa para la persona que la haga valer, es su presunción de validez que le da el artículo 1700 del Código Civil, que le da al instrumento público “plena fe en cuanto al hecho de haberse otorgado”, que no hay que confundir con la veracidad de las declaraciones que contenga.

Por otra parte, el Código Civil distingue al instrumento privado en el artículo 1701 y posteriores. Este es entendido como todo escrito que le da constancia a un hecho, y que ha sido otorgado sin

⁶⁴ Op. Cit. Fernández.

⁶⁵ Ibid.

⁶⁶ Op. Cit. Cervelló y Fernández

formalidad o solemnidad alguna⁶⁷. Lo esencial que tiene el instrumento privado es que deja testimonio de que cierto hecho sucedió. Estos instrumentos no necesitan una formalidad particular, no necesitan firma de las partes, que indiquen lugar y hora, ni siquiera es necesario que emitan de las partes. Asimismo, es irrelevante el medio utilizado para su confección⁶⁸. Estos instrumentos, a diferencia de lo que veíamos con los instrumentos públicos, no están revestidos de algún sello de autenticidad. Para solucionar estos inconvenientes, el legislador lo toma por auténtico solamente si de manera previa ha sido reconocido, asignándole fecha cierta en los casos que taxativamente enuncia en el artículo 346 y sus numerales del Código de Procedimiento Civil.

Para nuestro análisis particular, y basándose en que estamos analizando un documento creado a través de un sistema informático y la automatización de la IA, nos interesa en particular el numeral 4 del artículo 346 del Código de Procedimiento Civil, que regula la declaración de autenticidad del instrumento privado mediante una resolución judicial. Este procedimiento es el reconocimiento del instrumento por la vía judicial. En virtud al numeral 3 del mismo artículo, tras 6 días de presentado el instrumento a la parte contraria, y “no se alegue su falsedad o falta de integridad”, el tribunal podrá apercibir a la parte respectiva a tener por reconocido tácitamente dicho instrumento. El numeral 4 regula lo sucedido al alegar dicho instrumento dentro de este plazo de 6 días por falsedad o falta de integridad, por lo que el tribunal resolverá dicho incidente surgido. El tribunal dictará traslado a la parte del escrito de impugnación del instrumento, y con lo que esta exponga (o con una eventual rebeldía) lo recibirá a prueba, y con el mérito del rendimiento de estas, se fallará el incidente, que declarará si el instrumento privado es auténtico o no⁶⁹.

En virtud al artículo 1702 del Código Civil, cuando el instrumento privado es “reconocido o es mandado a tener por reconocido”, este tendrá el mismo valor probatorio que tiene el instrumento público respecto de quienes se ha mandado a tener por reconocido, para el caso particular.

Por otra parte, nuestro Código de Procedimiento Civil tiene cierta regulación respecto a los documentos electrónicos. El artículo 348 bis del Código de Procedimiento Civil plantea que tras presentar un documento electrónico, el tribunal citará al sexto día a las partes a una audiencia de percepción documental, en donde se podrá realizar una prueba complementaria de autenticidad, en conformidad a las reglas generales, a costa de la parte que la impugna. El resultado de esta prueba complementaria de autenticidad, que deberá ser llevada por peritos de acuerdo a las reglas generales (regulado en los artículos 417 a 423 del Código de Procedimiento Civil), será suficiente para tener por reconocido o por objeto dicho instrumento.

⁶⁷ Casarino, «Manual de derecho procesal».

⁶⁸ Ibid.

⁶⁹ Ibid.

En síntesis, podemos establecer que la naturaleza de la prueba de video o audio generada a través de sistemas de IA revisten el carácter de prueba instrumental, y más específicamente, de documento electrónico con carácter de instrumento privado, que para ser correctamente autenticado (o eventualmente ser declarado falso) debe ser reconocido por la vía judicial, en donde el material audiovisual será o no declarado como auténtico según los resultados del recibimiento a prueba del mismo, esto sin contar la prueba complementaria de autenticidad que se le hará al momento de presentarse, en virtud al artículo 348 bis del Código de Procedimiento Civil.

2.2 La prueba audiovisual en el proceso penal

El proceso penal en Chile tiene una lógica distinta a la del proceso civil. El procedimiento penal fue reformado en todo el país el año 2005, modernizando la legislación, a diferencia del procedimiento civil, que no ha sido reformado de forma sustancial desde la publicación del Código de Procedimiento Civil en 1905. Para nuestro análisis de la prueba audiovisual en el proceso penal, la regulación es escueta, debido a la naturaleza de esta, de la que se hará referencia en los párrafos siguientes.

Respecto a qué pruebas pueden ser introducidas en el proceso penal, el artículo 295 del Código Procesal Penal establece la libertad de prueba en este procedimiento, lo que contrasta de manera relevante con el proceso civil, que nos da una lista taxativa de los medios y fuentes de pruebas que pueden ser añadidos en un juicio⁷⁰. Esto hace el análisis de los tipos de pruebas que pueden ser añadidos al proceso penal mucho menos problemático versus la situación en el proceso civil, ya que no es necesario recurrir a distinciones técnicas y doctrinarias para encasillar a la prueba audiovisual en uno u otro medio o fuente de prueba, como sucedía con la prueba instrumental y documental del proceso civil⁷¹. La norma tiene por objeto dar flexibilidad a las pruebas que pueden ser añadidas, y busca que el acervo probatorio sea amplio, pudiendo corroborar de mejor forma las alegaciones de las partes⁷².

Asimismo, en el sistema procesal penal, a diferencia del sistema procesal civil, estamos regidos por un sistema de valoración de la prueba de sana crítica, consagrado en el artículo 297 del Código Procesal Penal. La sana crítica, a diferencia del sistema de prueba legal tasado, ocupa las reglas del correcto entendimiento humano, que serán contingentes y variables según la experiencia en un determinado tiempo y lugar, pero siempre regidos por principios lógicos en los que debe descansar una sentencia⁷³. Es un sistema que busca dar libertad al juez, añadiendo su criterio racional, pero siempre subordinado a la lógica, experiencia o máximas, que debe exteriorizar al momento de fallar.

⁷⁰ Op. Cit. Padilla

⁷¹ Op. Cit. Fernández

⁷² Ezurmendia Álvarez et al., «Incertidumbre probatoria por exceso».

⁷³ Eduardo Couture, «Estudios de Derecho Procesal Civil (Buenos Aires, Ediciones Depalma)» (1979).

En nuestra legislación procesal penal, el artículo que nos permite la introducción de prueba audiovisual a un juicio es el artículo 323 del Código Procesal Penal. Este artículo regula los medios de prueba “no regulados expresamente”, y entre ellos menciona como medios de prueba admisibles a “películas cinematográficas, fotografías, fonografías, videgrabaciones y otros sistemas de reproducción de la imagen o del sonido, versiones taquigráficas y, en general, cualquier medio apto para producir fe”. Como se puede observar, la norma se aleja de la distinción doctrinaria clásica que mantenemos en el proceso civil de instrumentos y documentos, y permite como prueba a “cualquier medio apto para producir fe”.

Otra disposición relevante respecto a las pruebas audiovisuales en el proceso penal es la del artículo 333 del Código Procesal Penal, que plantea que los documentos serán exhibidos con indicación de su origen. Respecto a las pruebas audiovisuales, la ley plantea que se reproducirán en la audiencia por cualquier medio idóneo para la percepción por los asistentes.

2.3 Otras disposiciones relevantes

Para finalizar el análisis de cómo los distintos sistemas procesales en Chile regulan la prueba audiovisual, es relevante mencionar que nuestro ordenamiento tiene contemplado en el Código Penal determinadas sanciones al cometer falsificación de un instrumento privado. El artículo 197 del Código Penal establece penas de presidio menor en cualquiera de sus grados y/o multas de once a quince unidades tributarias mensuales si, con perjuicio de un tercero, se comete alguna de las falsedades que se encuentran enunciadas en el artículo 193 que el mismo código nos entrega con ocasión de delitos de funcionarios públicos que abusen de su oficio. En este artículo, podemos encontrar las causales de falsedad: (i) contrahaciendo o fingiendo letra, firma o rúbrica; (ii) suponiendo en un acto la intervención de personas que no la han tenido; (iii) atribuyendo a los que han intervenido en él declaraciones o manifestaciones diferentes de las que hubieren hecho; (iv) faltando a la verdad en la narración de hechos sustanciales; (v) alterando las fechas verdaderas; (vi) haciendo en documento verdadero cualquiera alteración o intercalación que varíe su sentido; (vii) dando copia en forma fehaciente de un documento supuesto, o manifestando en ella cosa contraria o diferente de la que contenga el verdadero original; y (viii) ocultando en perjuicio del Estado o de un particular cualquier documento oficial. Para el análisis particular de la falsificación de instrumento público, materializada en este caso como un *voice cloning* o un *deepfake*, cobra relevancia la causal número 2, “suponiendo en un acto la intervención de personas que no la han tenido”. Esta causal es posible extrapolarla a los casos de la modificación o creación con esta tecnología de IA.

Esta norma va de la mano con los artículos 207 y 296 bis letra b) del Código Penal. El primero sanciona a quien, a sabiendas, presenta al tribunal medios de prueba falsos o adulterados, e impone las penas de “presidio menor en su grado mínimo a medio y multa de seis a veinte unidades tributarias mensuales, si se tratare de proceso civil o por falta, y con presidio menor en su grado medio a máximo y multa de veinte a treinta unidades tributarias mensuales, si se tratare de proceso penal por crimen o simple delito”. Además, este artículo establece penas distintas según el sujeto que aporte la prueba: “Los abogados que incurrieren en la conducta descrita sufrirán, además, la pena de suspensión de profesión titular durante el tiempo de la condena. Tratándose de un fiscal del Ministerio Público, la pena será de presidio menor en su grado máximo a presidio mayor en su grado mínimo”. También, establece una agravante según si la falsificación o adulteración se hace en perjuicio del imputado o acusado: “En todo caso, si la conducta se realizare contra el imputado o acusado en proceso por crimen o simple delito, la pena se impondrá en el grado máximo”.

Otra norma que va de la mano con las ya mencionadas es la del artículo 293 bis, en sus letras b) y d). El artículo se refiere a ciertos supuestos en un contexto de proceso por asociación delictiva criminal, que pena con presidio menor en su grado máximo a quién, en este contexto delictual: “b) Amenace o constriña a otro a que omita prestar declaración o testimonio, a que produzca o presente antecedentes o pruebas falsas, o a que omita producir o presentar antecedentes o pruebas relevantes” y “d) Ofrezca o entregue a otro un beneficio económico o de otra naturaleza con el objeto de que produzca o presente antecedentes o pruebas falsas u omita producir o presentar antecedentes o pruebas relevantes”. Estos artículos, a la luz del análisis de las pruebas audiovisuales y su posibilidad ser alteradas o generadas artificialmente, son aplicables a la prueba generada o modificada por IA, ya sea un *deepfake* o un *voice cloning*.

CAPÍTULO III: DESAFÍOS DE LA APARICIÓN DEL *DEEPPFAKE* Y LA IA EN EL PROCESO

3.1 Masificación y democratización de herramientas de IA

Como adelantábamos en la introducción de esta tesis, el *deepfake* y el *voice cloning* son utilizados por sus creadores para diversos fines, tanto de ocio como fraudulentos, incluso siendo algunos constitutivos de delitos. Asimismo, comentábamos la facilidad de la creación de este material, ya que no era necesaria la intervención de un experto en informática o en edición audiovisual para la creación de este tipo de pruebas. Al ser una tecnología en constante desarrollo, cada día aparece un nuevo modelo de IA capaz de crear este material de manera mucho más simple y de mejor calidad.

Un problema de esta -y de todas- las tecnologías nuevas, es su fácil y rápida difusión y democratización⁷⁴. Una de las medidas que logra retrasar la llegada de la tecnología a lo largo del mundo es lo escaso de los materiales para su construcción, lo que funciona como un cuello de botella. Sin embargo, esta medida solo funciona cuando los materiales son tangibles, como lo fue en su momento el uranio y plutonio para el desarrollo de armas nucleares⁷⁵. Cuando nos referimos a un recurso que es intangible, como la información, no solo basta las medidas contra el robo, divulgación errónea o espionaje, sino que también es necesario contar con la capacidad y voluntad de mantener confidencial la información⁷⁶.

Por esto, no fue sorpresa la rápida difusión que tuvieron estas herramientas, como en su momento lo adelantó CITRON y CHESNEY⁷⁷. Al igual que otras tecnologías, su difusión se dio tanto en el ámbito comercial como por mercados negros virtuales, como pueden ser los casos de herramientas de manipulación gráfica como *Photoshop* o ciertos *malware* que se difundieron vía *Deep web*⁷⁸. De hecho, el primer *boom* que tuvo el *deepfake* fue cuando un usuario de Reddit publicó una herramienta que

⁷⁴ Op. Cit. Citron y Chesney

⁷⁵ Allison, *Avoiding nuclear anarchy*.

⁷⁶ Op. Cit. Citron y Chesney, 1762

⁷⁷ Ibid.

⁷⁸ Ibid.

insertaba rostros de famoso en videos pornográficos⁷⁹. Posteriormente, apareció la famosa “*Fake App*”, una aplicación que creaba videos que intercambiaba cualquier rostro de manera fotorrealista a través de *deep learning*, que se hizo viral en tiempo récord⁸⁰, aplicación que actualmente no se encuentra disponible, pero que dio pie a multitud de aplicaciones que funcionan de manera similar y de manera más eficiente. Ante lo viral que fue esta aplicación, un periodista del New York Times creó un *deepfake* con multitud de imágenes del actor Chris Pratt y de él mismo, y posteriormente, lo replicó con el presentador norteamericano Jimmy Kimmel⁸¹.

Tras la masificación y democratización de estas herramientas, la proliferación de su uso es exponencial. Solamente en 2023, según herramientas de detección utilizadas por la empresa DeepMedia, se difundieron globalmente 500.000 videos generados con *deepfake* y audios generados con *voice cloning*⁸². Actualmente, la cantidad de *deepfakes* y *voice cloning* que circulan en la web se están duplicando cada 6 meses⁸³. Esto nos lleva a esperar que para el año 2025, 8 millones de copias de material falsificado con IA esté compartido vía redes sociales, siguiendo el patrón de duplicar la cantidad de material cada 6 meses, si es que no aumenta eventualmente⁸⁴. En esta línea, un sitio web con la capacidad de prácticamente desnudar mujeres recibió 38 millones de visitas en los primeros 8 meses del 2021⁸⁵.

Como veremos en los capítulos finales, la regulación por parte de los estados se ha iniciado hace pocos años, haciendo especial énfasis en regular los sistemas de IA de alto riesgo, dentro de los que no se encuentra el *deepfake* y el *voice cloning*, a los que se les regula de manera no tan estricta⁸⁶. Sin embargo, y a pesar de ser una tecnología de IA no considerada como de alto riesgo para los derechos de las personas respecto al tipo de bien jurídico que podría eventualmente transgredir⁸⁷, es por lejos la herramienta de IA más masificada⁸⁸.

3.2 Problemáticas relevantes causadas por el *deepfake*

La rápida proliferación de material audiovisual generado o modificado con *deepfake* causó que internet fuera invadida de un millar de videos con contenido falsificado⁸⁹. Esto fue especialmente

⁷⁹ Emma Grey Ellis, «People Can Put Your Face on Porn—and the Law Can’t Help You».

⁸⁰ Op. Cit. Citron y Chesney

⁸¹ Kevin Roose, «Here Come the Fake Videos, Too».

⁸² Ulmer y Tong, «Deepfaking It».

⁸³ Jacobson, «Deepfakes and Their Impact on Society».

⁸⁴ Ibid.

⁸⁵ Ministerio de Justicia del Reino Unido y Dominic Raab, «New Laws to Better Protect Victims from Abuse of Intimate Images».

⁸⁶ Parlamento Europeo, Reglamento de Inteligencia Artificial.

⁸⁷ Ibid.

⁸⁸ Op. Cit. Jacobson

⁸⁹ Ibid.

problemático cuando este material audiovisual se utilizó para desinformar en temas que eran relevantes para la población. Un ejemplo de lo anterior son videos del presidente de Ucrania, Volodimir Zelenski, comunicando información falsa, en donde en el año 2022 , tras el hackeo del sitio web de *Ukraine's 24 TV*, fue publicado un video del presidente llamando a sus tropas a bajar las armas⁹⁰.

Desde las primeras falsificaciones, la tecnología *deepfake* se ha hecho cada vez más precisa y de mejor calidad, exigiendo cada vez menos recursos. Por ejemplo, el área de investigación asiática de Microsoft presentó en abril de 2024 un sistema de IA llamado VASA-1, con la capacidad de crear avatares parlantes simplemente entregándole al sistema una única foto, con la que puede crear a un orador que sincroniza la boca del avatar con la pista de audio, que es entregada por el creador de este nuevo *deepfake*⁹¹. Además, esta IA permite cambiar distintos parámetros del avatar, cambiando la posición de la cara, la ubicación hacia donde van dirigidos los ojos al momento de hablar, las emociones que emite (sorpresa, enojo, felicidad)⁹², etc. VASA-1 aún no ha sido lanzado al mercado o puesto a disposición libremente a los usuarios, ya que Microsoft quiere tener mecanismos que aseguren a la sociedad la detección de usos fraudulentos de esta tecnología, así como tener algún método que les dé certeza que esta tecnología será utilizada de manera responsable y de acuerdo a la normativa que eventualmente le sea aplicable. Esta tecnología se crea para fomentar la creación y utilización de avatares para diversos proyectos, especialmente educativos. Sin embargo, llama la atención que los desarrolladores empiecen a ponerle especial énfasis a los posibles usos de las tecnologías que desarrollan, como es el caso de Microsoft y la creación de VASA-1⁹³.

Estos videos o cualquier tipo de material audiovisual son peligrosos para el proceso jurídico, ya que dan la posibilidad de introducir pruebas falsificadas al proceso o incluso ser constitutivo de delitos de diversa índole. Elementos como una confesión, grabadas con buena iluminación y con buen audio, sumado a una pequeña producción por parte de los creadores, o un audio comprometedor creado por la IA, puede ser un medio para contaminar el proceso judicial, capaz de generar problemáticas y desconfianza al momento de aportar material audiovisual en la etapa probatoria correspondiente.

Por una parte, han aparecido delitos asociados a estafas en donde era utilizada la tecnología *deepfake*, en particular el *deepfake* que imita la voz humana, que es mucho más difícil de identificar, conocido como *voice cloning*⁹⁴. Por ejemplo, en 2019, el gerente general de una empresa energética del

⁹⁰ Op. Cit. Limón.

⁹¹ «VASA-1».

⁹² Ibid.

⁹³ Ibid.

⁹⁴ Op. Cit. Helmus.

Reino Unido recibió una llamada de alguien que sonaba como su jefe directo, pidiendo que transfiriera €220.000 a la cuenta bancaria de un proveedor húngaro⁹⁵.

Por otra parte, el aporte de pruebas digitales falsificadas no es una novedad en los procesos judiciales, la ley reconoce esa posibilidad y la sanciona⁹⁶. Actualmente, esto se ha materializado en la facilidad de crear un chat falso de WhatsApp, simplemente accediendo a una página web, como lo es la viralizada *Prankshit*⁹⁷, la cual tiene todas las sofisticaciones necesarias para engañar a cualquier lector. Sin un debido peritaje informático que copie la base de datos de WhatsApp o la aparición de inconsistencias en el mismo dispositivo donde supuestamente ocurrió la conversación, es casi indistinguible una verdadera conversación por este medio con una falsificada⁹⁸.

Debido a la naturaleza misma de las pruebas digitales, la manipulación de estas es posible, suprimiendo o añadiendo partes a un documento o dando pie a conversaciones que nunca han tenido lugar⁹⁹. Esto se dificulta cuando en nuestros ordenamientos jurídicos cuando contamos con ciertos procedimientos que tienen un especial cuidado con los derechos fundamentales de las partes involucradas al momento de aportar pruebas al proceso, como lo es en materia de familia, en donde la evidencia aportada está constantemente al límite de la vulneración de derechos como la privacidad, la honra y la inviolabilidad del hogar, causando que mucho de este material probatorio sea excluido excesivamente¹⁰⁰.

El auge de la IA de estos últimos años proliferó de manera exponencial el uso del *deepfake* y el *voice cloning*, principalmente para fines recreativos¹⁰¹, como también para cometer ciertos delitos¹⁰². Sumado a lo anterior, esta tecnología también tiene el potencial de cambiar la lógica del desarrollo de la etapa probatoria de tres formas: en primer lugar, como hemos insinuado anteriormente, introduciendo prueba digital creada o modificada por esta tecnología, causando una eventual desconfianza de los jueces al momento de valorar este tipo de pruebas. En segundo lugar, el uso del *deepfake* y el *voice cloning* puede ser utilizado para desacreditar autoridades de todo tipo. En el clima actual en el que vivimos, con temas de especial sensibilidad, un audio o video de un juez en el que se refiera a un juicio particular de una u otra manera, ataca directamente a la legitimidad y a la confianza en el proceso¹⁰³. El juez puede

⁹⁵ Catherine Stupp, «*Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case - WSJ*», accedido 29 de marzo de 2024, <https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402>.

⁹⁶ Mayer Lux et al., «El documento como objeto material de las falsedades documentales y del sabotaje informático en el Derecho penal chileno».

⁹⁷ «Prankshit».

⁹⁸ «Conversaciones de Whatsapp falsas en procesos judiciales – Indalics Peritos Informáticos», accedido 31 de marzo de 2024, <https://indalics.com/blog/conversaciones-de-whatsapp-falsas>.

⁹⁹ Marta Rodríguez Acosta, «La prueba digital en el proceso penal», 2018, <https://riull.ull.es/xmlui/handle/915/7290>.

¹⁰⁰ Jesús Ezurmendia Álvarez, «PROBLEMAS RELATIVOS A LA PRUEBA EN CADA ETAPA DE LA ACTIVIDAD PROBATORIA EN EL PROCEDIMIENTO DE FAMILIA», Revista chilena de derecho 47, n.º 1 (abril de 2020): 101-18, <https://doi.org/10.4067/S0718-34372020000100101>.

¹⁰¹ Op. Cit. Anikin

¹⁰² Op. Cit. Stupp

¹⁰³ Op. Cit. Citron y Chesney

incurrir en una causal de implicancia, en específico la N.º10 del artículo 196 del Código Orgánico de Tribunales, fundado por el audio o video en cuestión. Asimismo, material audiovisual con connotaciones xenófobas u homofóbicas, a las que se les puede dar visibilidad pública, pueden dar vuelta por completo el resultado de un juicio, ya que la presión ejercida por la ciudadanía puede afectar directamente al resultado de un juicio¹⁰⁴, si es que no se logra comprobar la completa veracidad del audio. En tercer lugar, la posibilidad de que aumente la utilización de *deepfake* y *voice cloning* en el proceso no solo generará desconfianza en el juez que valorará la prueba audiovisual, sino que también afectará a las partes de un juicio¹⁰⁵. Las partes pueden cuestionar el valor probatorio de cualquier prueba fotográfica, de video o de audio, alegando el uso de *deepfake* y pidiendo peritajes cada vez más especializados que descarten la manipulación de esta tecnología en las pruebas digitales aportadas al proceso. En definitiva, el *deepfake* eventualmente cambiará el paradigma de cómo los magistrados y las partes confían en la prueba audiovisual.

Siguiendo la idea de cómo la irrupción del *deepfake* puede afectar un proceso judicial, y a diferencia de la realidad previa a la aparición del *deepfake* y el *voice cloning*, actualmente es necesario desconfiar de cualquier tipo de material audiovisual que sea añadido a un proceso. Dependiendo del sistema en el que esta prueba sea añadida, nos encontraremos con distintos tipos de dificultades y desafíos. Siguiendo esta línea, el sistema probatorio que rige nuestros procesos civiles se enfrentará a una serie de desafíos propios del sistema de prueba legal tasada. Hay que considerar que en este tipo de sistemas los hechos toman un papel secundario en el proceso, y cobra relevancia seguir un determinado ritual probatorio formal y simbólico, ya que estos serán considerados como la demostración efectiva de los hechos¹⁰⁶. Así, el uso de la racionalidad que podría ocupar el juez queda totalmente de lado, como también su criterio propio respecto de cómo valorar la prueba en cuestión, y más relevante, considerar las particularidades del caso en cuestión. Tal como plantea Taruffo, excluir los criterios racionales de valoración de la prueba es algo contra-racional¹⁰⁷, y ante la aparición de estas pruebas falsificadas con IA, no contar con estos criterios es un peso más al procedimiento civil.

Para los casos en que nos enfrentamos a pruebas audiovisuales falsificadas con IA, la regulación procesal civil es insuficiente. La casi nula participación del criterio racional del juez respecto a la valoración de la prueba es crucial para que se perpetúen estas peligrosas prácticas. A fin de cuentas, el sistema probatorio legal tasado se puede considerar un *checklist* de formalidades y procedimientos para darle el valor de prueba plena o semiplena a este tipo de material probatorio, que nos da la ley de manera

¹⁰⁴ Francisco J. Leturia I., «LA PUBLICIDAD PROCESAL Y EL DERECHO A LA INFORMACIÓN FRENTE A ASUNTOS JUDICIALES. ANÁLISIS GENERAL REALIZADO DESDE LA DOCTRINA Y JURISPRUDENCIA ESPAÑOLA», Revista chilena de derecho 45, n.º 3 (diciembre de 2018): 647-73, <https://doi.org/10.4067/S0718-34372018000300647>.

¹⁰⁵ Op. Cit. Citron y Chesney

¹⁰⁶ Op. Cit. Gascón

¹⁰⁷ Op. Cit. “La prueba”, Taruffo

previa. Así, queda como un “último bastión” frente al *deepfake* los informes periciales al momento de analizar el material audiovisual, ya sea de forma complementaria o al momento de que se reconozca el instrumento privado por la vía judicial. Si una prueba generada por *deepfake* o *voice cloning* logra sortear los obstáculos periciales, al concederse como plena prueba, daña la legitimidad del proceso, ya que podrían darse absurdos como un audio con contenido que favorece todas las pretensiones de la parte que añade la prueba, y sumado a un proceso en donde el material probatorio sea escaso, no tendría otra opción el juez que aceptarlo como prueba válida en el juicio en cuestión, no dejando espacio al uso de su criterio racional que ayudaría a disminuir la posibilidad de que el *deepfake* o el *voice cloning* afecte al proceso.

En este sentido, se hace una crítica a cómo el sistema de prueba legal no se puede hacer cargo correctamente de este nuevo tipo de pruebas, a diferencia de lo que sucedería si se añade a un procedimiento reformado que utilice sana crítica, como el procedimiento laboral o de familia, en donde el criterio racional del juez toma un papel relevante al momento de fallar y de valorar la prueba en el juicio. En Chile, proyectos como la reforma al proceso civil chileno, que busca reformar el Código de Procedimiento Civil, despachado por la Cámara de Diputados el 7 de mayo del 2014, presenta la coexistencia de reglas de la prueba legal, en particular para la prueba instrumental, y la valoración libre que se le entrega al juez conforme a las reglas de la sana crítica (artículos 302 y 285 del Proyecto del Código de Procedimiento Civil)¹⁰⁸. Autores como HUNTER AMPUERO han planteado la posibilidad de compatibilizar las normas del sistema de prueba legal en sistemas de valoración libre conforme a la sana crítica, en donde podrían operar como normas de presunciones legales, o sea, inferencias probatorias normativas, y en último término, como reglas que determinen la carga de la prueba¹⁰⁹.

Por otra parte, en el procedimiento penal, la problemática es menos compleja respecto al tema de la admisión probatoria, pero, en cambio, la sana crítica tiene el principal nudo que se crea por la irrupción del *deepfake*, *voice cloning* e IA en general. Al momento de valorar una prueba, en el proceso civil, el juez no usa su criterio racional, sino que solo guía el proceso mientras se discute el reconocimiento por la vía judicial del video o audio añadido al proceso. En cambio, en el proceso penal, el juez es clave al momento de percibir estas pruebas. Como planteamos anteriormente, el juez funda su decisión en la lógica y su experiencia personal¹¹⁰, por lo que sus conocimientos respecto a cómo funciona la IA y al contexto actual de la posibilidad de modificación o creación del material audiovisual mediante la IA son cruciales para valorar. Así, no es lo mismo la percepción de un juez que no tenga conocimiento acerca de la existencia del *deepfake* o *voice cloning* en comparación con un juez que esté informado de

¹⁰⁸ Hunter Ampuero, «Reglas de prueba legal y libre valoración de la prueba».

¹⁰⁹ Ibid.

¹¹⁰ González Castillo, «LA FUNDAMENTACIÓN DE LAS SENTENCIAS Y LA SANA CRÍTICA».

la irrupción de esta tecnología en la forma de alterar un material audiovisual, y de forma más relevante, el alcance que esta tiene actualmente, siendo sumamente fácil de utilizar.

La jurisprudencia respecto a la prueba generada por *deepfake* es nula en nuestro país, a fecha de cierre de esta tesis. Por otra parte, fallos respecto a la prueba generada por IA, haciendo referencia a esta de forma amplia, es escueta. El fallo más relevante en nuestro país (y el único que ha escalado a un tribunal de segunda instancia hasta la fecha de cierre de esta tesis) respecto a la prueba generada por IA es de la undécima sala de la Corte de Apelaciones de Santiago, con fecha 11 de marzo del 2024, respecto a un recurso de nulidad de una condena en sede penal¹¹¹. En esta sentencia podemos encontrar que la prueba digital es cuestionada por la posibilidad de que haya sido generada por IA. Este caso es similar, aunque no igual, a lo que pasaría con una prueba audiovisual. En estos chats de WhatsApp, el tribunal de primera instancia reconoce la falta de fiabilidad, trazabilidad y adecuada custodia legal de la prueba en cuestión. Ante esto, la corte establece: “En relación a la prueba reseñada precedentemente (chats de WhatsApp), el tribunal no la estimó apta para establecer la supuesta venta lícita del vehículo objeto del delito en estudio. Ello en primer término por cuanto el tribunal no conoció la forma de obtención de aquellos mensajes, su fiabilidad, trazabilidad y adecuada custodia legal. En efecto, dadas las nuevas plataformas que utilizan inteligencia artificial y los innumerables métodos de creación de archivos, documentos, fotografías, etc, resultaba de especial relevancia contar con antecedentes que dieran cuenta que efectivamente esa evidencia fue obtenida desde el celular del acusado, en la fecha que allí se indica, y quienes eran las personas titulares de los números de teléfonos celulares o cuentas de Facebook, para así conocer datos del presunto interlocutor y oferente de la motocicleta”¹¹².

Este considerando particular del fallo de la Corte de Santiago, que confirma la sentencia del tribunal de primera instancia, explicita que, en este caso, los chats de WhatsApp no cuentan con la suficiente fiabilidad para demostrar los hechos alegados. Lo relevante de este fallo para nuestro análisis es que la Corte pone de manifiesto las plataformas de IA y su uso para la creación de documentos y contenido audiovisual, por lo que, al momento de la valoración de esta prueba, no les da el valor probatorio suficiente para establecer, en el caso particular, que la misma sea apta para dar por probada la venta ilícita del vehículo que era objeto del delito en cuestión. Cabe destacar que la sentencia mencionada se encuentra firme y ejecutoriada, al no existir recursos pendientes, por lo que estamos privados de conocer la opinión jurídica de la Corte Suprema.

A pesar de que este fallo no versa de manera específica sobre material audiovisual, se puede notar la poca confianza que la Corte de Santiago le da a los chats de WhatsApp. Esta es una tendencia

¹¹¹ Bases Jurisprudenciales del Poder Judicial, Rol 488-2024, Bases Jurisprudenciales del Poder Judicial.

¹¹² Ibid.

que va en aumento, en particular en el procedimiento de familia, en donde están familiarizados con la aportación de este tipo de pruebas en el proceso¹¹³. Por otra parte, en las bases jurisprudenciales del país, es común hoy en día encontrar fallos de todas las cortes del país en donde la prueba audiovisual no es alegada por los abogados o fiscales como generada por sistemas de IA. Simplemente, al cumplir los criterios de admisibilidad que la materia procesal particular exige, la prueba es tomada como válida y no es común que se alegue su falsedad.

En este sentido, el contenido del material audiovisual añadido al proceso es muy relevante en relación con el proceso en el que se lleve a cabo el aporte de este material probatorio falsificado con IA. Si es en un juicio civil, las barreras de entrada de este tipo de pruebas son la admisión de esta como instrumento privado reconocido, por lo que el contenido de este es irrelevante en un sentido valorativo de la prueba, ya que, como planteábamos, la ley suple la valoración del juez, debido al sistema de prueba legal tasada consagrado en el Código de Procedimiento Civil. En cambio, en un proceso penal, contamos con el criterio racional del juez como un pilar relevante al momento de la admisión probatoria, por lo que el contenido de la prueba no es indiferente para el juzgador¹¹⁴. El juez, basándose en su experiencia y la lógica, puede restar valor probatorio a una prueba audiovisual introducida al proceso, debido a que, como pasó en el caso ya expuesto, el tribunal desconoce la trazabilidad, fiabilidad y custodia legal de la prueba. Esto causará que se cambie el paradigma de fiabilidad que tenían las pruebas audiovisuales, en donde, por ejemplo, un audio o video comprometedor, al que el juez brindaba un alto valor probatorio por ser considerado “infalible”, o como la “prueba más importante del juicio”, pase a ser simplemente una prueba que, según las características que recoja, puede ser más o menos probable su veracidad, teniendo en consideración por parte del juez la posibilidad de que sea un *deepfake* o un *voice cloning*.

Actualmente, solo queda esperar a que, tras el paso de los meses, los tribunales se sigan refiriendo a cómo las pruebas audiovisuales pueden ser eventualmente modificadas o creadas por la IA, y según se ha expuesto respecto a la masificación y democratización de esta tecnología, la tendencia a la duda de este tipo de pruebas irá en aumento. Para que se desarrolle este fenómeno, es crucial la instrucción a los participantes de los procesos judiciales respecto a que la modificación o creación de prueba generada por IA es ya una realidad cercana a todas las personas. Mientras más se alegue por las partes o por fiscalía, o se capacite a los magistrados en los peligros que genera los *deepfake* y el *voice cloning*, podremos ver reflejado en muchos más fallos que las pruebas audiovisuales no se introducirán sin más en el proceso, ya que el juez dudará de base en su fiabilidad, y se impugnará gran parte de la prueba audiovisual hasta avanzar en un estándar común para este tipo de material probatorio. Debido al avance exponencial de

¹¹³ Agudelo, Valencia, y Espinosa, «Autenticidad y debido proceso en los mensajes de Whatsapp».

¹¹⁴ Op. Cit. Couture

esta tecnología¹¹⁵, es cosa de tiempo para que un *deepfake* o un *voice cloning* aparezca como material probatorio aportado en un juicio, y probablemente en este momento, se concretará una pérdida de la “infalibilidad” del material audiovisual, e incluso es posible que se tomen medidas desde el poder político para la regulación del uso malicioso de esta adulteración creada con IA.

3.3 Sesgo y discriminación por los algoritmos

La IA y su capacidad de buscar patrones, contrastar datos y tomar decisiones son elementos que encontramos en gran parte de la tecnología actual, desde nuestros teléfonos hasta sistemas altamente complicados. La existencia de cámaras de vigilancia con reconocimiento biométrico o sistemas capaces de detectar el tipo de arma usada en algún delito solamente replicando el sonido del disparo son una realidad hace años. Esta capacidad de identificar patrones también puede ser utilizada, en vez de rasgos faciales o sonidos, en conductas o factores de riesgos sobre la comisión de un eventual delito.

Esta tecnología ya fue creada e incluso implementada. En Estados Unidos, se creó la herramienta llamada COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), encargada de “evaluar el riesgo” de un determinado acusado sobre si se le deben aplicar ciertas medidas cautelares en el proceso penal. Esta decisión es tomada por el juez en cuestión, y en este caso la IA funciona como una guía o precedente para que el magistrado dirima respecto al caso en cuestión. La forma de entrenamiento de este sistema es a través de datos históricos de los mismos acusados¹¹⁶. El sistema de IA se encarga de buscar correlaciones entre los factores dados, destacando entre ellos la edad, los antecedentes que contengan en el sistema de justicia penal, como un arresto previo, u otros factores relevantes. Posteriormente, estos datos se correlacionan y se intenta predecir la probabilidad de que el acusado en cuestión vuelva a delinquir en el periodo de espera del juicio¹¹⁷.

El problema surge cuando la forma de recopilación de los datos de los que se nutre la IA está basada en un historial de arrestos desproporcionados e incluso erróneos¹¹⁸ contra las minorías raciales, historial atribuido principalmente a la policía americana, contando incluso con manipulación de datos en su haber¹¹⁹. Supuestamente, esta práctica sería mucho más justa al sustituir el sesgo e intuición de los jueces, para recomendar una evaluación más objetiva. Sin embargo, sin una base de datos sólida, no es del todo concebible conseguir este objetivo.

¹¹⁵ Op. Cit. Jacobson

¹¹⁶ Karen Hao, «Caso práctico».

¹¹⁷ Ibid.

¹¹⁸ El margen de error del sistema se da porque los nuevos arrestos de los reincidentes generalmente ocurren por infracciones técnicas, en donde, por ejemplo, los reincidentes no comparecen en el tribunal. Entonces, no necesariamente la reincidencia se asocia a la actividad delictiva como tal.

¹¹⁹ Ibid.

A pesar de que se declare que COMPAS no tiene en cuenta la raza del acusado para el cálculo de su “puntuación de riesgo”, en 2016 una investigación de ProPublica reveló que, aun no utilizando este criterio, la herramienta sí se encontraba sesgada contra las personas afroamericanas¹²⁰. Según la metodología de ProPublica, se analizó si los acusados que fueron encarcelados antes del juicio volvieron o no a ser arrestados en los dos años posteriores a su liberación¹²¹. Posteriormente, se utilizó para calcular si los acusados hubieran sido o no detenidos nuevamente antes del juicio de no haber sido encarcelados. El resultado fue claro: entre los acusados que no volvieron a ser encarcelados, los afrodescendientes tenían el doble de probabilidades que los blancos de acabar en el grupo de alto riesgo de COMPAS¹²².

En el caso hipotético de que esta tecnología empiece a funcionar en nuestro país, es altamente probable que el fenómeno dado en Estados Unidos se dé también en Chile. Actualmente, el sesgo y discriminación hacia los extranjeros es evidente¹²³. La sensación de una oleada de delincuencia desatada exclusivamente por los migrantes sudamericanos es generalizada, tanto por la opinión pública como por parte de las autoridades, aun cuando este sesgo no se condice al 100% con los datos respecto a delitos cometidos por extranjeros¹²⁴.

El caso de COMPAS nos deja mucho por reflexionar y un gran desafío por superar: hay que regular los sesgos en los sistemas automatizados de toma de decisiones. El profesor experto en IA y derecho, Andrew Selbst, es partidario de la regulación legal de la responsabilidad de los algoritmos¹²⁵. De hecho, en 2019 se propuso en Estados Unidos un proyecto de ley respecto a la responsabilidad algorítmica. Este proyecto obligaría a las grandes empresas tecnológicas a controlar sus sistemas de aprendizaje automático, para así detectar a tiempo indicios de sesgo y discriminación, y así tomar las medidas correctivas al identificar estos problemas. Se propone el sistema de “evaluación de impacto¹²⁶”, que plantea evaluar los sistemas automatizados en función de su impacto en vez de centrarse en los datos algorítmicos, que llevarían a un debate innecesario de complejidad técnica¹²⁷.

¹²⁰ Op. Cit. Angwin

¹²¹ Ibid.

¹²² Ibid.

¹²³ Gissi et al., «Entre el miedo y la xenofobia».

¹²⁴ Ajzenman, Patricio Domínguez, y Raimundo Undurraga, «Inmigración, delincuencia y las erróneas percepciones sobre el delito».

¹²⁵ Op. Cit. Karen Hao.

¹²⁶ Kristian Lum, «Cómo fijar la responsabilidad algorítmica sin saber qué es un algoritmo».

¹²⁷ Karen Hao, «EE. UU. responsabilizará a las tecnológicas de las “fake news” y el sesgo algorítmico».

CAPÍTULO IV: REGULACIÓN COMPARADA Y PROPUESTAS PARA REGLAMENTAR LA INTELIGENCIA ARTIFICIAL Y EL *DEEPPFAKE*

4.1 La experiencia de la Unión Europea

La UE es probablemente la comunidad política que más interés y desarrollo tiene en la actualidad (y a futuro) respecto de la regulación de la IA. Fueron pioneros al momento de vislumbrar la necesidad de encauzar la prioridad del mundo digital hacia la regulación de la IA, y buscan ser líderes y referentes en la materia para otros estados¹²⁸. Nos dedicaremos a analizar las directrices y los enfoques que la UE ha dado a la regulación de la IA, poniendo especial énfasis en la primera ley integral de la IA, que busca ser un nuevo estándar normativo para el resto de legislaciones.

El acercamiento de la UE a la IA se remonta al año 2017, cuando el Parlamento Europeo emite una resolución con recomendaciones respecto a normas de Derecho Civil sobre la robótica. Así, la Comisión Europea en 2018 se compromete a proponer un enfoque desde el mundo europeo a la IA. Es aquí donde la UE pone como objetivo tener una política industrial global europea en materia de IA¹²⁹. Tras finalmente deliberar sobre la necesidad de regulación de la IA, se crea el Libro Blanco como un primer intento de marcar las directrices que guiarán a la UE en la regulación de la IA.

El Libro Blanco baraja muchas opciones de políticas públicas que buscan el desarrollo seguro de la IA en Europa. Para la UE, es relevante darle seguridad a la comunidad respecto a la implementación

¹²⁸ Parlamento Europeo, «Inteligencia artificial».

¹²⁹ Gerard Rincón Andreu, «Libro Blanco de la Comisión Europea sobre Inteligencia Artificial. Un enfoque europeo hacia la excelencia y la confianza», *Ius et Praxis* 27, n.º 1 (marzo de 2021): 264-70, <https://doi.org/10.4067/S0718-00122021000100264>.

progresiva de la IA, esta vez no solo como consumidores, sino también como productores y creadores. Esto se refleja en un considerable aumento de los fondos destinados a la investigación de la IA, que han aumentado en un 70% en los 3 años anteriores a la realización del Libro Blanco¹³⁰. Asimismo, el Libro Blanco pone énfasis en que el fomento de la IA debe ir de la mano con la fiabilidad que debe tener para el ser humano, respecto a que tenga las debidas medidas de seguridad para evitar resultados que sean dañinos (sin que necesariamente estén previamente programados) por culpa de la autonomía que caracteriza a esta tecnología, sin una supervisión humana de por medio¹³¹.

Sobre el tema de la responsabilidad de la IA, podemos encontrar implicancias relevantes en el Informe de la Comisión anexo al Libro Blanco. Este identificó problemáticas clave sobre la responsabilidad a perseguir respecto a la IA. Marca como objetivo principal tener un sistema de responsabilidad eficiente, traducido en la reparación integral de daños causados por esta tecnología, con altos estándares de protección para generar confianza en los consumidores, y que estos no tengan recelo al momento de interactuar con estas nuevas tecnologías digitales¹³².

El 18 de junio de 2020, el Parlamento Europeo aprobó la creación de la Comisión Especial sobre Inteligencia Artificial en la Era Digital (en adelante, AIDA, por sus siglas en inglés). Esta comisión tuvo por objetivo estudiar los retos y el impacto de la IA y encontrar puntos de encuentro con la UE a fin de organizar objetivos. Tuvo una duración de 18 meses, y tuvo como base informes anteriores, como el Libro Blanco o el Informe de la Comisión anexo al Libro Blanco¹³³. Los resultados de la comisión AIDA evidencian el atraso que la UE tiene respecto a la investigación de la IA, así como en desarrollo e inversión. Se propuso un camino a largo plazo para el correcto desarrollo de esta tecnología, y así alcanzar niveles prolijos de integración de la IA a lo largo de toda la comunidad¹³⁴.

Es necesario recordar que el resultado de la comisión AIDA, y la posterior transformación a una normativa aplicable, tiene que pasar por la aprobación de los órganos legislativos de la UE: el Consejo de la UE (o Consejo de Ministros) y el Parlamento Europeo. El primero es integrado por representantes de las naciones europeas que constituyen la UE, o sea, 27 ministros que ostentan de manera rotativa el cargo. El Parlamento Europeo se elige de manera democrática por todos los ciudadanos con derecho a voto que integran la UE (aproximadamente 375 millones de votantes). El último paso son las llamadas

¹³⁰ Ibid.

¹³¹ Comisión Europea, «White Paper on Artificial Intelligence».

¹³² INFORME DE LA COMISIÓN AL PARLAMENTO EUROPEO, AL CONSEJO Y AL COMITÉ ECONÓMICO Y SOCIAL EUROPEO Informe sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica.

¹³³ Parlamento Europeo, «AIDA Committee Activity Report».

¹³⁴ Op. Cit. “Inteligencia artificial: la hoja de ruta del Parlamento para la UE”, Parlamento Europeo

“negociaciones trilaterales” que llevan al Parlamento Europeo, el Consejo de la UE y a la Comisión a una deliberación final que lleva a la eventual aprobación de una normativa¹³⁵.

El principal resultado de la exhaustiva investigación que duró varios años por parte de la UE es la creación de la llamada “Primera ley de IA”¹³⁶. Esta “ley de IA” en realidad reviste la naturaleza jurídica de un reglamento, en virtud al artículo 288 del Tratado de Funcionamiento de la UE. Los reglamentos en la UE se caracterizan por tener un alcance general, ser obligatorios en todos los elementos del mismo y ser directamente aplicables a los estados miembros de la Unión¹³⁷, siempre y cuando no hayan establecido una cláusula de exclusión voluntaria, que es una manera de que un Estado miembro que no quiera participar de cierta política particular de la UE, pueda restarse sin la consecuencia de que haya un bloqueo a esta política en relación con los otros Estados participantes¹³⁸.

La elección de este instrumento jurídico se justifica por la necesidad de la UE de evitar la fragmentación jurídica, en la búsqueda de una uniformidad en la aplicación del reglamento. Esta homogeneidad en las normas, como lo pueden ser ciertas prohibiciones, las definiciones relevantes y limitaciones a los proveedores de IA (que exploraremos más adelante) son relevantes para simplificar la existencia de un mercado exclusivo en donde se trancen sistemas de IA seguros, fiables y dentro de la legalidad de la UE¹³⁹.

La elaboración del reglamento sobre la regulación de la IA impactó de manera relevante los intentos de regulación a nivel mundial. Ha permeado legislaciones de todas las latitudes, desde Brasil hasta China, y ha impactado de manera relevante el cómo se desarrolla la discusión regulatoria en Estados Unidos¹⁴⁰.

Finalmente, el miércoles 13 de marzo del 2024 fue aprobada por la Eurocámara la llamada coloquialmente “Ley de Inteligencia Artificial”. El reglamento pone énfasis principalmente en determinados temas de relevancia para la UE: la separación de los distintos tipos de IA según su riesgo, regulación respectiva a los usuarios de la IA, entendiéndolos como usuarios no finales, distinguiendo entre estos y los desarrolladores, la concentración de las obligaciones regulatorias que los proveedores/desarrolladores de la clasificación más riesgosa de IA deben cumplir, la creación de la Oficina de IA, y por último, la regulación de la IA General/de Propósito General (en donde encontramos, por ejemplo, Chat GPT, programas de *voice cloning*, o cualquier otra aplicación de IA generativa).

¹³⁵ Op. Cit. Haataja y Bryson

¹³⁶ Parlamento Europeo, «Ley de IA de la UE».

¹³⁷ Moritz y Gibello, «El Reglamento Europeo (UE) 2016/679».

¹³⁸ «Cláusula de exclusión voluntaria».

¹³⁹ Linera, «La propuesta de “Ley de Inteligencia Artificial” europea».

¹⁴⁰ Op. Cit. Haataja y Bryson.

4.1.1 Clasificación de IA según su riesgo

El Capítulo II y III del reglamento regula latamente los sistemas de IA que están estrictamente prohibidos, así como los de alto riesgo. Respecto a las prácticas que están expresamente prohibidas con la utilización de la IA, podemos encontrar, entre otras, la prohibición del uso de técnicas subliminales que manipulen o engañen a una persona, con el objetivo de distorsionar su forma de tomar decisiones, los sistemas que busquen aprovecharse de características de grupos particularmente vulnerables, como discapacitados, menores de edad o gente de escasos recursos, con el propósito de alterar su conducta y causar así un daño significativo, la utilización de sistemas que, a través de datos biométricos, categoricen a personas físicas por sus características personales para estimar su pensamiento político, religión, orientación sexual u otros, la utilización automatizada de datos biométricos utilizados a distancia y en tiempo real con fines policiales, a menos que sea en el contexto de una investigación de un delito de gravedad o de alta connotación social, como abusos sexuales, delitos asociados al terrorismo, entre otros, y la prohibición de sistemas de IA que realizan evaluaciones de riesgo de individuos para predecir el riesgo de que esta persona cometa con más o menos probabilidad un delito, basándose en su aspecto físico y su personalidad. En particular, esta última prohibición puede ser relacionada a la fallida experiencia que se tuvo con el sistema COMPAS¹⁴¹, del que ya se hizo mención¹⁴².

Respecto a la clasificación de la IA según riesgo y la regulación, en particular, de la IA altamente riesgosa, el reglamento establece altas exigencias a los desarrolladores de estas, acaparando gran parte de las obligaciones contenidas en la normativa. En primer lugar, son IA de alto riesgo las que sean subsumibles al anexo III, que enlista los sistemas que son susceptibles de esta clasificación. Este anexo también cuenta con las excepciones particulares a cada caso, que exonera a cierto sistema a no ser introducida dentro de esta clasificación. Dentro del anexo III podemos encontrar, de manera inapelable e inexcusable, todos los sistemas que tratan con elementos personales de individuos. En otras palabras, cualquier IA que trate de manera automatizada la información personal, ya sea de usuarios o de personas que no sean consumidores de este, que haga un diagnóstico de las diversas aristas de una persona, como su estado financiero, su productividad en el trabajo o cualquier opción particular que se busque ser monitoreada¹⁴³.

El resto de sistemas que también fueron considerados como de alto riesgo por el reglamento son los que traten sobre: parámetros biométricos, siempre que el país en cuestión permita su uso, los que supervisen infraestructura crítica de manera digital, como autopistas o suministros de servicios básicos,

¹⁴¹ Op. Cit. Hao

¹⁴² Op. Cit. Reglamento de Inteligencia Artificial

¹⁴³ Ibid.

trabajo y administración de personal, respecto a la utilización de la IA para la contratación o selección, el análisis de posibles aspirantes, toma de decisiones respecto a promoción de trabajadores, etc., aprendizaje tanto básico como profesional, respecto a sistemas que determinen la asignación de personas a ciertas escuelas, que califiquen evaluaciones de un estudiante o los evalúen para otros fines, policiales y de autoridad, relativos a IA utilizada por autoridades, ya sean de autoridad policial o de la judicatura, para evaluar riesgo de victimismo de una persona sobre delitos penales, que busquen funcionar como polígrafos o, lo que es particularmente relevante para lo expuesto en esta tesis, evaluar la credibilidad de ciertas pruebas presentadas en el contexto de investigaciones criminales o en el juicio. Respecto al último punto mencionado, el reglamento lo acota solamente a las infracciones penales, por lo que deja de lado o no considera de alto riesgo sistemas de IA utilizados en contexto de un juicio civil o de otra naturaleza¹⁴⁴.

Por otra parte, son también IA riesgosa las destinadas a finalidades de seguridad o que se trate de uno cubierto por el anexo II, que tiene directa relación con el artículo 5 N.º1 letra d) inciso 3, que establece sistemas de IA con la capacidad de localizar o identificar a persona sospechosa de haber infringido un delito penal establecido en el mencionado anexo. Dentro de este anexo, podemos encontrar una lista de ciertos delitos graves y de alta connotación social, como terrorismo, asesinato, violación, etc. Cabe destacar que, a pesar de que el artículo 5 del reglamento trata sobre los sistemas de IA que están estrictamente prohibidos, este inciso establece una excepción cuando su uso sea estrictamente necesario para estos fines. Las IA que se enmarquen en esta categoría deberán ser sometidas a una evaluación especial desarrollada por un tercero¹⁴⁵.

Todos los proveedores que crean que su IA no es de alto riesgo, o sea, que no se puede subsumir bajo el anexo III, deberán igualmente acreditar esta realidad previo a la inserción del sistema al mercado o a la puesta en marcha de este.

4.1.2 Obligaciones de los desarrolladores/proveedores de sistemas de IA de alto riesgo

Entre los artículos 8 y 27 del reglamento de IA, encontramos los requisitos y obligaciones que deben cumplir los proveedores o desarrolladores de IA de alto riesgo¹⁴⁶. Estos deben seguir ciertos pasos relevantes para garantizar la correcta seguridad y eficiencia de los sistemas que desarrollen. De ellos, se desprenden las siguientes obligaciones relevantes: (i) se debe establecer un sistema integral de gestión de riesgos, que debe abarcar desde la creación hasta el cese del funcionamiento de sistema de IA de alto

¹⁴⁴ Ibid.

¹⁴⁵ Ibid.

¹⁴⁶ Ibid.

riesgo. Esto implicaría la identificación y evaluación de los riesgos que puedan surgir en todas las etapas, desde lo más inicial como lo puede ser el diseño hasta su eventual puesta en funcionamiento y todos sus mantenimientos posteriores; (ii) la administración y gobernanza de los datos son fundamentales. Los proveedores o desarrolladores deberán asegurarse de que los datos que utilizan para el entrenamiento, validación y la prueba de estos sistemas sean suficientemente representativos y sin ningún error. Se justifica esta obligación para hacer al sistema lo más preciso y fiable; (iii) el requerimiento de documentación de carácter técnico, lo suficientemente detallado, para demostrar conformidad con los estándares de las regulaciones a aplicar. Estos informes le dará a las autoridades respectivas la información necesaria para evaluar en conformidad a su sistema y a su funcionamiento; (iv) En la etapa de diseño del sistema de IA se debe incluir un registro automatizado de eventos de relevancia que sean capaces de identificar posibles riesgos para el país, como asimismo cualquier tipo de modificación sustancial a lo largo del ciclo de vida del sistema de IA; (v) los proveedores y desarrolladores deben proporcionar instrucciones de manera clara y detallada para el correcto uso del sistema de IA a los que sean responsables de su implementación cuando el sistema vaya a entrar en funcionamiento. Con esto se busca garantizar que la IA se utilice de la manera correcta que establecen los creadores; (vi) el sistema de IA debe permitir que sea supervisado por un humano, lo que se traduce en que el desarrollador u otra persona calificada pueda intervenir el sistema cuando se estime pertinente; (vii) El sistema de IA debe cumplir con ciertos estándares al momento de su diseño. Debe desarrollarse con altos estándares respecto a ciberseguridad, así como a la precisión y solidez del sistema, garantizando un óptimo funcionamiento y la protección de la privacidad y de los datos de los usuarios; y (viii) la creación de un sistema que se dedique a gestionar la calidad para asegurar un cumplimiento continuo de la regulación y los estándares del reglamento, así como eventuales mejoras que se les puedan aplicar al sistema¹⁴⁷.

4.1.3 IA de propósito general/GPAI e IA Generativa

Se entiende por IA de propósito general (también conocida como IA de modelo GPAI) aquel sistema de IA que tiene la capacidad de cumplir diversas funciones de manera autónoma, y no están limitadas a una sola función, además de tener capacidades comparables o superiores a las de un ser humano. Se caracterizan por poder adaptarse a entornos desconocidos y tomar decisiones basándose en recursos limitados y a su funcionamiento en ámbitos en los que generalmente encontramos humanos, por la necesidad de contextualización de estos¹⁴⁸. El nombre de “IA de propósito general” nació como

¹⁴⁷ Ibid.

¹⁴⁸ Gutierrez et al., «A Proposal for a Definition of General Purpose Artificial Intelligence Systems».

una forma de definir a este tipo de IA al momento de la redacción del reglamento de IA, ya que para la literatura académica había un espacio gris respecto a la denominación y no había consenso¹⁴⁹.

Dentro de esos sistemas de IA que cumplen una gran variedad de funciones con bastante eficiencia, encontramos la aplicación de la IA generativa. Esta tecnología se caracteriza por poder producir información nueva e inédita, en función a los datos con los que el sistema se ha entrenado. Así, tienen la capacidad de desarrollar material fresco, buscando imitar a la mente humana y a su creatividad, y se desmarca de sistemas convencionales que se basan en previsiones numéricas y reglas internas, lo que limita las posibilidades de uso de la IA¹⁵⁰.

La relevancia de la IA con modelado de GPAI es que se pueden usar en modelos de alto riesgo. Por esto, el legislador pone foco en la utilización de los modelos GPAI y su uso e integración en los sistemas de IA que sean considerados de alto riesgo, presentes en el capítulo III del reglamento, ya sea que el modelo GPAI sea usado *per se* como sistema de IA o sea integrado en estos¹⁵¹.

Las principales obligaciones a las que son sometidos los sistemas GPAI son: (i) la elaboración de documentación técnica, que abarca el proceso de *training* del modelo, las pruebas hechas en este y los resultados de las pruebas; (ii) informar correctamente a través de documentos detallados para futuros desarrolladores que deseen integrar este sistema en sus modelos de IA. Esto les ayudará a comprender lo que pueden y no pueden hacer con el modelo GPAI; (iii) el establecimiento de una política que esté en armonía con la directiva sobre derechos de autor; y (iv) realizar y publicitar un informe detallado de todo el contenido que se utilizó para el *training* del modelo GPAI, siendo así más transparentes respecto al proceso de formación de esta IA¹⁵².

Respecto a los usuarios que utilizan estos sistemas de IA, la ley los define como personas físicas o jurídicas, que utilizan el sistema de IA de manera profesional, y los diferencia de los usuarios que recibirán el producto de la utilización de la IA. Estos usuarios son denominados implantadores (*deployers*, por su nombre en inglés), y si utilizan la IA de alto riesgo, tienen que cumplir con ciertos deberes, pero de manera bastante reducida, a diferencia de los que deben cumplir los desarrolladores/proveedores de la IA. También, se puntualiza que la aplicación de estas obligaciones aplica a los implantadores que se encuentren bajo la potestad normativa de la UE como también fuera de la misma, siempre que la producción se llegue a utilizar en la UE¹⁵³.

¹⁴⁹ *Ibid.*

¹⁵⁰ García-Peñalvo y Vázquez-Ingelmo, «What Do We Mean by GenAI?»

¹⁵¹ Op. Cit. Reglamento de Inteligencia Artificial.

¹⁵² *Ibid.*

¹⁵³ *Ibid.*

Es menester hacer una distinción relevante respecto al desarrollo de sistemas de IA que utilizan el modelo GPAI. Los códigos respecto a la forma de desarrollo del sistema de IA, esto significa, su arquitectura, los parámetros, o lo esencial del modelo, puede estar reservado por parte de la empresa que lo desarrolla, o al contrario, puede ser de acceso público, para que cualquier persona pueda desarrollar sistemas similares con respecto a los códigos de un determinado sistema¹⁵⁴. El reglamento de la UE nos plantea que, si un sistema de IA es de código abierto, solamente tiene la obligación de cumplir con los dos últimos puntos (iii y iv) de los ya mencionados. Sin embargo, el reglamento presenta una excepción: que dicho modelo GPAI sea calificado de “riesgo sistémico”.

Según el reglamento de IA¹⁵⁵, un modelo GPAI considerado de “riesgo sistémico” será tal si cuenta con ciertas capacidades que la caractericen de gran impacto, evaluadas por herramientas y ciertas metodologías técnicas adecuadas, como lo son indicadores y parámetros de referencia, como asimismo, y según el anexo XIII del reglamento, ciertos criterios de relevancia, como lo son la cantidad de parámetros del modelo, calidad o tamaño del conjunto de datos, cantidad de cálculo utilizado para entrenar el modelo, medido en FLOPs (operaciones en coma flotante), etc.

Si estos parámetros se cumplen, tienen la obligación de notificar a la comisión en un tiempo máximo de 2 semanas, en donde el desarrollador tiene la oportunidad de ejercer una defensa para evidenciar de que su sistema no conlleva riesgos sistémicos, a pesar de que objetivamente cumpla los parámetros del reglamento. La relevancia de esta determinación es que los sistemas de riesgo sistémico deben cumplir con obligaciones extra, aparte de las ya mencionadas. Estas son: (i) llevar a cabo las respectivas evaluaciones de los modelos, lo que implica realizar y documentar pruebas contradictorias, que sean funcionales a reducir el riesgo sistémico; (ii) el análisis y reducción de cualquier variable que aumente la posibilidad de riesgo sistémico, incluyendo en esto las fuentes de las mismas; (iii) el registro, documentación e información de manera oportuna a la oficina de IA y a todas las autoridades nacionales pertinentes y competentes sobre cualquier incidente grave, incluyendo medidas correctivas; y (iv) tener un alto estándar respecto a seguridad cibernética, a fin de resguardar estos particulares sistemas de IA.

En este punto cobra relevancia el código de buenas prácticas, que, según la reglamento, es una forma de demostrar el cumplimiento efectivo de todos los deberes antes mencionados¹⁵⁶. Su adhesión es completamente voluntaria. La adhesión a este código, una vez se armonicen las normas europeas, dará a los desarrolladores la presunción de conformidad con estas normas. Por otra parte, los desarrolladores que sean parte del código de buenas prácticas deberán utilizar otras formas para cumplir, y el eventual visto bueno de la comisión. El código se centrará particularmente en toda la parte técnica del

¹⁵⁴ van Dis et al., «ChatGPT».

¹⁵⁵ Op. Cit. Reglamento de Inteligencia Artificial.

¹⁵⁶ Ibid.

cumplimiento de las obligaciones mencionadas, así como el sondeo de eventuales riesgos y la forma en la que se gestionarán¹⁵⁷.

Respecto a qué organismo será el encargado de la parte administrativa del reglamento, su aplicación será supervisada por la futura Oficina de IA, que estará ligada a la Comisión Europea y velará por el efectivo cumplimiento de los desarrolladores/proveedores de sistemas de IA, en particular los que utilicen la IA generativa. La Oficina de IA estará encargada de evaluar paulatinamente los modelos que se utilicen en la UE, en particular los riesgos sistémicos, en donde científicos externos y expertos en la materia harán un estudio sobre el sistema en particular¹⁵⁸.

Como hemos visto, la UE busca ser un ejemplo a nivel mundial con este gran paso dado respecto a la regulación de los distintos sistemas de IA. El desarrollo y control de esta tecnología, es y será un elemento estratégico para el funcionamiento de la sociedad en los próximos años, ya que dada la capacidad de autoaprendizaje de estos sistemas, se espera que se apliquen a una gran variedad de esferas de la vida¹⁵⁹. Actualmente, los gigantes tecnológicos como lo son Google, Amazon, Netflix, Tesla, Uber, o incluso empresas de China, como Alibaba, Baidu y Tencent han destinado inmensas cantidades de recursos para convertirse en líderes de la innovación, acelerando estos procesos de desarrollo tecnológico para acumular una cantidad relevante de poder económico, político y cultural, creando un verdadero modelo industrial-civilizatorio¹⁶⁰. Así, la necesidad de regulación se hace cada vez más necesaria, ante la formación de empresas que se inmiscuyen cada vez más en aspectos personales de la vida de las personas. Los ciudadanos de países miembros de la UE tendrán un primer acercamiento a gran escala a la protección de estas esferas por parte de su comunidad política, que busca marcar un precedente en este tipo de legislaciones¹⁶¹.

4.2 Otras experiencias relevantes

En el resto del mundo también ha habido otros intentos relevantes para regular el fenómeno de la masificación de la IA, desde distintos organismos internacionales y cooperaciones entre varios países. Sin embargo, estos intentos no tienen el mismo detalle y son más escuetos que el reglamento de IA de la UE, y en general se encargan de establecer principios generales respecto a la utilización de la IA, y a crear marcos jurídicos de menor intensidad¹⁶².

¹⁵⁷ Ibid.

¹⁵⁸ Ibid.

¹⁵⁹ Girardi, «Digitalización, política e inteligencia artificial».

¹⁶⁰ Sadin, *La silicolonización del mundo: la irresistible expansión del liberalismo digital*.

¹⁶¹ Op. Cit. Haataja y Bryson

¹⁶² Op. Cit. Linera

Uno de los relevantes organismos que buscó, al igual que la UE, la regulación de los sistemas de IA y sobre cómo se relacionan con las personas, es el Consejo de Europa. Este es un organismo internacional no dependiente de la UE fundado en 1949 que pone especial relevancia en la salvaguarda y el fomento de los derechos humanos en los 40 estados miembros que son parte de este organismo¹⁶³. Su trabajo respecto a la regulación de la IA nace por la búsqueda de cautelar los derechos fundamentales de los europeos ante los sistemas de IA, que cuentan con la capacidad de afectar a casi la totalidad de los derechos humanos¹⁶⁴.

El 2019 se celebró en Helsinki, Finlandia, una conferencia en donde el Consejo de Europa se hacía cargo sobre el “Impacto del desarrollo de la inteligencia artificial en los derechos humanos, la democracia y el Estado de derecho”. En esta conferencia se llegan a conclusiones claves que el Consejo de Europa materializaría posteriormente, en particular el reforzamiento de que la IA es una creación humana, y como tal, es fruto de una actividad previa por parte de un ser humano que automatiza cierta cantidad de datos y que diseña e implementa el uso de esta tecnología. Así, los sistemas de IA pueden estar siempre sujetos, tanto en su funcionamiento como en su aplicación, a los derechos humanos y a una exigencia de responsabilidad de estos¹⁶⁵.

Sobre la base de las conclusiones ya mencionadas y emanadas por la conferencia en Helsinki, el Consejo de Europa, a través de su Consejo de Ministros, deciden crear el Comité sobre Inteligencia Artificial (CAI), al cual se le encomienda la creación de un Convenio Marco que sea jurídicamente vinculante sobre cómo se desarrollan y diseñan los sistemas de IA en armonía con las normas del Consejo de Europa respecto a derechos humanos, democracia y Estado de derecho¹⁶⁶.

Por otra parte, aparte de la regulación referida a derechos fundamentales y los sistemas de IA, la Asamblea Parlamentaria del Consejo de Europa despachó un conjunto de principios éticos básicos a respetar al momento del desarrollo y establecimiento de la IA, con especial foco en la seguridad, la responsabilidad humana de la toma de decisiones, la equidad, la vida privada y la protección de datos. Además, reconoce la necesidad de tener una normativa transversal para la IA en materia de derechos humanos, democracia y Estado de derecho, que es de donde nace el Convenio Marco antes mencionado¹⁶⁷.

El Convenio Marco sobre Inteligencia Artificial, Derechos Humanos, Democracia y Estado de derecho finalmente se consolida el 7 de julio del 2023, en donde se explicita una doble vertiente de la

¹⁶³ González Alonso, «La bioética en el Consejo de Europa».

¹⁶⁴ «Algorithms and Human Rights - Study on the Human Rights Dimensions of Automated Data Processing Techniques and Possible Regulatory Implications».

¹⁶⁵ Fernando Martín Diz, «Capítulo XLV. Inteligencia artificial y proceso: garantías frente a eficiencia en el entorno de los derechos procesales fundamentales».

¹⁶⁶ Op. Cit. Linera

¹⁶⁷ Ibid.

IA: los sistemas de IA se pueden desarrollar, diseñar y utilizar en función de ofrecer oportunidades incomparables respecto a la protección y promoción de los derechos humanos, como para las libertades fundamentales, la democracia y el Estado de derecho. Sin embargo, la “otra vertiente” trata sobre la capacidad del desarrollo, diseño, utilización y desmantelamiento que los sistemas de IA, usados de manera libre y sin ningún tipo de control, pueden causar y aportar a socavar la dignidad humana y la autonomía individual, los derechos humanos y las libertades fundamentales, la democracia y el Estado de derecho¹⁶⁸.

Por otra parte, de manera más globalizada y saliendo un poco del viejo continente, los 36 países que conforman la Organización para la Cooperación y el Desarrollo Económicos (OCDE), sumado a Argentina, Brasil, Colombia, Costa Rica, Perú y Rumanía suscribieron el día 22 de mayo del 2019 los principios de la OCDE sobre la Inteligencia Artificial¹⁶⁹. Esta normativa está confeccionada de una manera flexible, para que perdure por mucho más tiempo y sirva como directriz a los países miembros¹⁷⁰.

La OCDE destaca la capacidad de la IA para la promoción de los objetivos de desarrollo mundial, el progreso y la prosperidad de los países del orbe. Así, se incita a los países miembros con más avance tecnológico a asumir un rol y un deber de solidaridad con las naciones que cuentan con menos avances tecnológicos. Se busca que los beneficios de la IA se democratizen, fomentando su uso y la participación, buscando la equidad a través de la colaboración internacional de conocimientos, experiencia, recursos financieros y tecnológicos, y particularmente de la cooperación internacional de todos los interesados¹⁷¹.

Los sistemas de IA, al igual que en instrumentos anteriores, se deben desarrollar de una manera que respeten íntegramente el Estado de derecho, los derechos humanos, valores democráticos, etc., y se debe implementar la debida salvaguarda, como la intervención humana en caso de ser necesario, con la intención de garantizar principios como la justicia y la equidad¹⁷².

Otro punto que la OCDE destaca es la transparencia y la explicabilidad de los sistemas de IA; los usuarios deben saber cuándo interactúan con esta tecnología, pueden cuestionar los resultados y entenderlos, a través de la divulgación responsable para entender los mecanismos de decisión automatizada utilizada por la IA para dar determinada respuesta¹⁷³. Se le da un especial énfasis al valor de la información que los usuarios manejan respecto a los sistemas de IA, ya que tener una comprensión

¹⁶⁸ Ibid.

¹⁶⁹ OCDE, «Recommendation of the Council on Artificial Intelligence».

¹⁷⁰ Morandín-Ahuerma, «Recomendación Del Consejo Sobre Inteligencia Artificial de la Ocede».

¹⁷¹ Korinek, «Integrating Ethical Values and Economic Value to Steer Progress in Artificial Intelligence».

¹⁷² Op. Cit. Linera

¹⁷³ Op. Cit. OCDE

general de cómo funcionan, y en consecuencia, el entendimiento de los resultados, especialmente si son nocivos, permitirá el efectivo cuestionamiento de ciertas respuestas de la IA¹⁷⁴.

Siguiendo con los principios, la OCDE busca la robustez, seguridad y protección a lo largo de la vida útil de los sistemas de IA. Se espera que estos sean fiables, y estén debidamente protegidos de los riesgos que puedan tener, por lo que tendrán que ser monitoreados constantemente. No deben suponer ningún peligro excesivo para la seguridad de quienes los utilizan. Se debe garantizar la trazabilidad del sistema para que sean monitoreadas en toda su vida útil¹⁷⁵. Para esto, la conexión a internet será un factor esencial para el funcionamiento de estos sistemas, y así, será relevante la ubicación de la respectiva IA, ya que en sectores al que no llegue de manera correcta una conexión estable a internet, o en casos más extremos, no llegue, pueden crearse riesgos al uso de la IA, especialmente a las tecnologías que las implementen por primera vez, sin haber pulido por completo el funcionamiento de este.

Cualquier problema suscitado por un sistema de IA, como la responsabilidad respecto a su funcionamiento, según las directrices de la OCDE, serán de exclusiva responsabilidad de los seres humanos que las gestionen, desarrollan y despliegan, de acuerdo a su grado de participación con el sistema en cuestión¹⁷⁶. No se debe olvidar que detrás del funcionamiento de los sistemas de IA, como hemos visto latamente, hay una gran variedad de actores que participan para poner en marcha esta tecnología.

La OCDE busca y promueve el desarrollo de una cooperación y facilitación público-privado respecto a investigación y desarrollo para estimular la innovación. Se busca que los intereses estatales y particulares hagan de contrapeso para formar una IA que sea considerada fiable. La formación de un ecosistema de infraestructuras digitales, el intercambio de datos y conocimiento sería cimentar el camino para un correcto desarrollo de los sistemas de IA y el intercambio de datos cada vez más ágil. Una transición equitativa es el objetivo por lograr entre los estados miembros y todos alrededor del globo, para desarrollar estándares positivos y correctos para una buena administración de la IA¹⁷⁷.

A finales del año 2023, Estados Unidos, China, la UE y 26 países más llegan a un gran acuerdo global, llamado *The Bletchley Declaration*, que busca la cooperación científica de los países para evitar, o eventualmente frenar, posibles avances de carácter “catastrófico” de la IA¹⁷⁸. Este acuerdo busca recoger los esfuerzos internacionales para abordar el impacto, tanto positivo como negativo, de los sistemas de IA, y busca ponerlos en discusión en distintos foros¹⁷⁹. Los países se comprometieron a

¹⁷⁴ Op. Cit. Morandín-Ahuerma

¹⁷⁵ Ibid.

¹⁷⁶ Op. Cit. OCDE

¹⁷⁷ Op. Cit. Linera

¹⁷⁸ Ibid.

¹⁷⁹ Ibid.

abordar protección de los Derechos Humanos, transparencia y explicabilidad de los algoritmos, la equidad, rendición de cuentas, seguridad, debida supervisión humana, ética y protección de datos, todo en un contexto de cooperación para suministrar a los países constantes insumos científicos para el correcto desarrollo de políticas públicas relacionadas con el funcionamiento de la IA en los Estados¹⁸⁰.

Por último, es necesario mencionar el caso chileno. El primer trimestre del 2024, la superintendencia de educación recibió cuatro denuncias a nivel nacional por el uso de imágenes y videos generados o modificados por *deepfake*¹⁸¹. Estos casos, que fueron considerados aislados en nuestro país, son parte de una seguidilla de casos similares que han sucedido en distintas latitudes del mundo¹⁸², que, debido al alza de esta tecnología, no pareciera tener solución sin entrar en la regulación. Motivado por esto, el 7 de mayo del 2024, se reactivó un proyecto de política nacional de IA, que recoge la experiencia comparada, en particular la ya analizada de la UE. Este proyecto de ley es bastante similar a lo planteado por el organismo europeo respecto a la regulación del riesgo, clasificación y supervisión de la IA¹⁸³. Una de las principales novedades que plantea el proyecto para nuestra legislación nacional es la creación de una agencia de protección de datos personales, institución la cual muchos países europeos cuentan, pero que, a nivel nacional, es la gran novedad respecto a la protección de datos personales¹⁸⁴. La discusión sobre la IA se reavivó en la sociedad al conocerse a finales de mayo la denuncia de estudiantes del colegio Saint George afectadas por el uso de la IA para la creación de imágenes de compañeras desnudas¹⁸⁵. Los padres de las menores presentaron un recurso de protección, que se encuentra con carácter reservado¹⁸⁶.

El proyecto que regula la IA sigue en su primer trámite constitucional en la cámara de diputados, a la fecha de cierre de esta tesis¹⁸⁷. Se espera que prontamente sea despachado por el congreso para la regulación legal de la IA en nuestro país, y para la mitigación de las consecuencias del uso nocivo de esta tecnología.

La regulación por parte de los diversos estados es mayor o menor, dependiendo del ordenamiento analizado, pero es constante el esfuerzo por parte de estos en tener una legislación, aunque tal vez incipiente, respecto al uso y consecuencias de la IA. Hasta ahora, ningún estado ha propuesto legislar contra la aportación de prueba creada o modificada por IA en los juicios. Sin embargo, la regulación actual serán los cimientos para una regulación más especializada por parte de los países.

¹⁸⁰ Ibid.

¹⁸¹ Mondaca, «Imágenes falsas».

¹⁸² Viejo, «Decenas de menores de Extremadura denuncian que circulan fotos de falsos desnudos suyos creadas por inteligencia artificial».

¹⁸³ Carey, «Ingresa a la Cámara de Diputados proyecto de ley que regula los Sistemas de Inteligencia Artificial».

¹⁸⁴ Ibid.

¹⁸⁵ Emol, «Denuncian que alumnos del Saint George usaron IA para crear imágenes desnudas de compañeras: Fiscalía investiga».

¹⁸⁶ Ibid.

¹⁸⁷ Cámara de Diputadas y Diputados de Chile, «Proyecto de Ley: Regula los sistemas de inteligencia artificial».

4.3 Propuestas para mitigar a la IA y el *deepfake*

La aparición de la IA y de los *deepfakes* a la sociedad cambió por completo el paradigma respecto a los cuidados que debemos tener con la tecnología. Desde un punto de vista científico, es necesaria la inversión de diferentes actores del mundo de la tecnología en sistemas de detección y prevención de *deepfakes*, que sean capaces de identificar y filtrar este tipo de material, en particular de las redes sociales o de portales de noticias¹⁸⁸.

Respecto a las pruebas audiovisuales añadidas al proceso, los peritos también deberán actualizarse en un sentido técnico para contar con herramientas de identificación de material generado o modificado por IA. Se deberá ser riguroso al analizar las pruebas audiovisuales que sean añadidas y someterlas exhaustivamente a estos *softwares* para identificar pruebas generadas por *deepfake* o *voice cloning*. Hay que recordar, como mencionábamos en capítulos anteriores, que el desarrollo y masificación de esta tecnología traerá aparejado que los abogados intenten impugnar las pruebas audiovisuales, acusándolas de haber sido alteradas o creadas con IA. Es en este sentido que los peritos informáticos deben tener la capacidad de darle certeza al tribunal o a las partes de que el material audiovisual es o no auténtico.

Por otra parte, educar al público respecto a los riesgos potenciales del *deepfake* es esencial¹⁸⁹, ya sea mediante medios estatales o privados, como escuelas, o incluso por parte de los empleadores. La tecnología del *deepfake* puede dañar a distintos actores de la vida cotidiana, por lo que una educación integral respecto a sus usos es relevante en la sociedad, así como la concienciación de su correcta aplicación. Hacer extensiva la información que contamos sobre la IA y su posibilidad de crear pruebas falsas, va a acercar a los magistrados actuales y futuros a contar con estos conocimientos como parte de sus “máximas de la experiencia”, cambiando, por ejemplo, la forma en la que valoran los magistrados la prueba audiovisual, al considerar la posibilidad de *deepfake* o *voice cloning* al momento de valorarla.

Debido a que la IA en los medios digitales ya es una realidad, no debe sorprender la desconfianza y el escepticismo que los usuarios de redes sociales y cualquier persona que se relacione con un medio audiovisual tenga respecto a uno de estos materiales. Así, las plataformas de redes sociales, que es el principal medio de divulgación de material de *deepfake*¹⁹⁰, deben centrarse en que su medio sea transparente respecto a la creación de contenido generado por IA, ya sea mediante *deepfake* u otra herramienta de IA¹⁹¹. En este sentido, el uso de marcas de agua u otras formas de certificación, o

¹⁸⁸ Op. Cit. Jacobson

¹⁸⁹ Ibid.

¹⁹⁰ Op. Cit. Ulmer

¹⁹¹ Op. Cit. Jacobson

herramientas que puedan verificar el uso de IA, es relevante¹⁹². De hecho, plataformas como Instagram y Facebook, administrados por el gigante Meta, ya adoptaron medidas al respecto, en donde avanzaron de una política de eliminar cualquier contenido generado por herramientas de IA a identificarlos con una etiqueta de “Generado por IA”¹⁹³. También, avanzarán con el desarrollo de las etiquetas, en donde podrán etiquetar publicaciones que sean de “alto riesgo de engaño”, que se compatibilizará con las etiquetas de IA¹⁹⁴. La búsqueda de la etiquetación de la IA para evitar posibles engaños reduce la posibilidad de que un material de esta naturaleza pueda llegar a ser aportada en juicio como un material audiovisual no alterado o modificado por la IA.

Respecto a la regulación legal específica sobre la utilización fraudulenta del *deepfake*, hay poco desarrollo por parte de los Estados, pero es relevante mencionar la regulación que hizo el Reino Unido a propósito de la protección de víctimas del abuso de imágenes íntimas¹⁹⁵. Esta serie de cambios que se realizó el 2022 a la ley inglesa criminaliza a los difusores de *deepfakes* sin consentimiento de la víctima, y que pueden potencialmente enfrentar condenas con penas de cárcel. Sin embargo, la escasa regulación a nivel de estados al respecto es insuficiente para regular este fenómeno. En este sentido, la colaboración internacional, intercambio de información y avance tecnológico es clave para una protección sólida frente al mal uso del *deepfake*¹⁹⁶. El esfuerzo directo en contra del *deepfake* se centra netamente en perjuicios a bienes jurídicos de las personas, pero posteriormente, los estados avanzarán en cautelar bienes jurídicos de diversa índole, dado el carácter avasallador de esta tecnología, pudiendo ampliarse a la generación de la prueba creada o modificada por IA, y sancionarla de manera particular, y no extrapolando artículos del Código Penal de forma amplia a la prueba falsificada.

Respecto a la posibilidad de la prohibición desde un punto de vista normativo de los sistemas de *deepfake* y el *voice cloning*, encontramos dificultades respecto a la ponderación de derechos, ya que encontramos enfrentados, por una parte, la libertad de expresión del creador del material audiovisual generado por IA, y el derecho a la protección de la vida privada, honra y protección de sus datos personales que nos asegura la Constitución actual. En este sentido, el principal obstáculo de la prohibición del *deepfake* es su permisibilidad constitucional¹⁹⁷.

Por una parte, prohibir directamente la manipulación digital no sería deseable. Esta tecnología no es intrínsecamente negativa, ya que puede ser usada en contexto de mejoramiento de imagen, o

¹⁹² Ibid.

¹⁹³ Rodrigo Mundaca, «Instagram acusará las imágenes y videos hechos con Inteligencia Artificial».

¹⁹⁴ Ibid.

¹⁹⁵ Op. Cit. Ministerio de Justicia del Reino Unido y Raab

¹⁹⁶ Op. Cit. Jacobson

¹⁹⁷ Op. Cit. Citron y Chesney

incluso, en áreas como la historia o el arte, es posible utilizarlo para diversos fines visuales¹⁹⁸. Entonces, surge la necesidad de prohibir la creación de este tipo de manipulación digital, pero excluyendo los usos beneficiosos. Esto se podría lograr, por ejemplo, atendiendo al elemento subjetivo, según la lógica del derecho penal y la responsabilidad civil extracontractual, que tenga el creador del *deepfake*, así como la prueba que se haga de un daño que sufra la víctima a raíz de la difusión del contenido¹⁹⁹. Sin embargo, en este ejemplo es necesario distinguir la penalización de la creación del *deepfake* y la difusión de este, en cuanto esta última ya ha tenido esfuerzos por penalizarse, como vimos en el caso de la reforma en Reino Unido²⁰⁰, lo que no se ha asimilado a la penalización por parte de los Estados a la creación, que es mucho más problemática.

Es entonces, en un sentido amplio, la lucha contra la desinformación acerca de cómo funciona el *deepfake* y el *voice cloning* una de las principales formas de atenuar sus resultados negativos. Esta tecnología abarca la gran mayoría de las esferas de la vida cotidiana, y que la población tenga conciencia respecto a la IA, las hace menos propensa a caer. Esto también se vería reflejado en los jueces, como mencionábamos anteriormente, que ya no verán a la prueba audiovisual como infalible, sino que dudarán, al igual que las personas que cada día dudan más de ciertas noticias sensacionalistas, pensando que tal vez puedan estar siendo víctimas de *fake news*.

CONCLUSIÓN

A lo largo de esta tesis, repasamos la historia de la IA, desde su forma más básica hasta los complejos sistemas a los que podemos acceder hoy en día, y explicamos el nacimiento del concepto del *deepfake*, su significado y cómo fue creada en armonía con la evolución de la IA.

Analizamos cómo nuestro ordenamiento jurídico regula expresamente los materiales audiovisuales al momento de aportarlos como prueba en los principales sistemas procesales del país, como lo son el proceso civil, en donde analizamos la forma específica de entender un material audiovisual como un instrumento privado que debe ser reconocido en juicio, y el proceso penal, centrándonos en la sana crítica de este último.

¹⁹⁸ Ibid.

¹⁹⁹ Ibid.

²⁰⁰ Op. Cit. Ministerio de Justicia del Reino Unido y Raab

Revisamos la masificación y democratización del *deepfake* y del resto de sistemas de IA, estableciendo el origen de su popularización y cómo paso a ser una de las tecnologías que más rápidamente se difundieron, dado su carácter viral de la mano con las redes sociales.

Revisamos las problemáticas que causaron en los últimos años, poniendo especial énfasis en los problemas con relevancia jurídico-penal de la aparición del *deepfake* tanto en Chile como en otros estados, y sus respuestas a estas mismas problemáticas, además de revisar las problemáticas con la explicabilidad de los algoritmos para ciertos sistemas.

Examinamos cómo responden nuestros sistemas procesales a la aportación de la prueba digital, y al momento de ingresar al proceso, cómo son valoradas por los jueces en la actualidad. A pesar de que los jueces del proceso civil tienen más restricciones al momento de valorar este tipo de pruebas, los jueces que valoran mediante la sana crítica tienen la posibilidad de valorar de distinta forma a la establecida en la ley, como sucede en el procedimiento civil. Así, nuestro análisis ha revelado que, a pesar de que los *deepfake* y el *voice cloning* son actualmente una realidad, la lógica del material audiovisual producido por la IA aún no se está reflejando de la manera correspondiente en los juicios. La jurisprudencia es escueta al respecto en nuestro país, y la respuesta legal es incipiente. Los jueces no están teniendo en cuenta los sistemas de IA al momento de valorar la prueba audiovisual, lo que puede ser problemático en el corto plazo, ya que se espera que con el paso del tiempo este paradigma cambie.

Que gran parte de los jueces mantengan la “fiabilidad” de la prueba audiovisual refleja la continuidad de este tipo de pruebas como sólidas y confiables. A pesar de que todo indica que esta realidad prontamente va a cambiar debido al avance exponencial del *deepfake* y el *voice cloning*, los magistrados se encuentran al debe respecto a la valoración de este tipo de pruebas. Con los números actuales de circulación de *deepfake* vía redes sociales, es posible que algún tipo de material audiovisual alterado o generado por IA se haya colado en los sistemas procesales, y si aún no lo ha hecho, prontamente lo hará.

Es necesario tomar medidas para mitigar la aportación de este tipo de pruebas en el proceso. Para esto, es esencial informar a los actores del mundo judicial y a la sociedad en su conjunto. El manejo de los profesionales del derecho en esta materia es crucial, ya que esto se traducirá en magistrados informados y preparados para valorar de forma correcta la nueva realidad de la prueba audiovisual, así como también le dará herramientas a los abogados de impugnar este tipo de pruebas en el proceso.

De manera posterior a este análisis, repasamos las respuestas de los estados, en donde se puso énfasis en lo esencial que fue la regulación de la UE como pioneros en tener una legislación para regular los sistemas de IA. Esta regulación fue en muchos casos base para el desarrollo jurídico de otros estados u organizaciones internacionales que buscaban regular el fenómeno de la IA. Analizamos las diferencias

de la regulación de estos estados en comparación con los de la UE, en donde la legislación era mucho más escueta, incluso encontrándonos con directrices muy generales.

Finalmente, propusimos formas de mitigar la introducción de la IA en el proceso y sus efectos negativos en general, en donde establecimos como clave avanzar en dos temas: el avance tecnológico para el reconocimiento de la utilización de herramientas de IA en material audiovisual (o en cualquiera que sea susceptible de alteración por parte de esta tecnología) y en la difusión de información respectiva a la utilización del *deepfake* y *voice cloning*, un correcto uso de estas y las gravosas consecuencias que puede traer un mal uso de la IA. La prohibición de esta tecnología pareciera ser un camino imposible de seguir, por lo que el camino de la concientización es una forma de mitigar su uso nocivo.

En conclusión, a pesar de que la innovación tecnológica es en suma positiva para la sociedad, es necesario avanzar de manera responsable e integral. La experiencia con los sistemas de IA ha traído beneficios y perjuicios a la sociedad, desde distintas esferas de la vida. El derecho procesal no ha estado exento de los efectos que esta tecnología tiene en el mundo. A pesar de que la IA no es esencialmente mala, la ciudadanía y los magistrados deben relacionarse con los materiales audiovisuales con cautela, teniendo en cuenta el posible uso nocivo de esta tecnología, y tener presente que la posibilidad de que dicho material probatorio sea alterado o generado por la IA es más probable que nunca.

BIBLIOGRAFÍA

1. Abeliuk, Andrés, y Claudio Gutiérrez. «Historia y evolución de la inteligencia artificial». *Revista Bits de Ciencia*, n.º 21 (2021): 14-21.
2. Abellán, Marina Gascón. *Los hechos en el Derecho: Bases argumentales de la prueba*. 1.ª ed. Marcial Pons, Ediciones Jurídicas y Sociales, 2010. <https://doi.org/10.2307/jj.2322004>.
3. Agudelo, Dany Steven Gómez, Jenny Marcela Acevedo Valencia, y Juan Esteban Aguirre Espinosa. «Autenticidad y debido proceso en los mensajes de Whatsapp: Una revisión en los casos de divorcio». *Revista Chilena de Derecho y Tecnología* 10, n.º 2 (31 de diciembre de 2021): 123-48. <https://doi.org/10.5354/0719-2584.2021.58039>.

4. Ajzenman, Nicolás, Patricio Domínguez, y Raimundo Undurraga. «Inmigración, delincuencia y las erróneas percepciones sobre el delito». *CIPER Chile* (blog), 8 de mayo de 2021.
<https://www.ciperchile.cl/2021/05/08/inmigracion-delincuencia-y-las-erroneas-percepciones-sobre-el-delito/>.
5. Allison, Graham T. *Avoiding nuclear anarchy: containing the threat of loose Russian nuclear weapons and fissile material*. 12. MIT Press, 1996.
[https://books.google.es/books?hl=es&lr=&id=uSxEdn9UNFEC&oi=fnd&pg=PP9&dq=GRAHAM+T.+ALLISON+ET+AL.,+AVOIDING+NUCLEAR+ANARCHY+\(1996\).&ots=oaTvjr0o1-&sig=Ky8rLSODMEDJv1UWx0glX20LDcI](https://books.google.es/books?hl=es&lr=&id=uSxEdn9UNFEC&oi=fnd&pg=PP9&dq=GRAHAM+T.+ALLISON+ET+AL.,+AVOIDING+NUCLEAR+ANARCHY+(1996).&ots=oaTvjr0o1-&sig=Ky8rLSODMEDJv1UWx0glX20LDcI).
6. Arcila, Maribel Guisella Maldonado, y Freddy Oswaldo Ovalles Pabon. «Repercusiones éticas sobre el uso indebido del deepfake en el ámbito de las TIC mediante un análisis cualitativo documental». *Reto* 9, n.º 1 (2021): 36-47.
7. Bases Jurisprudenciales del Poder Judicial. Rol 488-2024 (Undécima Sala, Corte de Santiago 11 de marzo de 2024).
8. Cámara de Diputadas y Diputados de Chile. «Proyecto de Ley: Regula los sistemas de inteligencia artificial». Accedido 8 de julio de 2024.
<https://www.camara.cl/legislacion/ProyectosDeLey/tramitacion.aspx?prmID=17429&prmBOL ETIN=16821-19>.
9. Carey. «Ingresa a la Cámara de Diputados proyecto de ley que regula los Sistemas de Inteligencia Artificial». *Carey Abogados* (blog), 13 de mayo de 2024.
<https://www.carey.cl/ingresa-a-la-camara-de-diputados-proyecto-de-ley-que-regula-los-sistemas-de-inteligencia-artificial/>.
10. Carnelutti, Francesco. *La prueba civil*. Ediciones Olejnik, 2023.
11. Carrascosa, Valentín, Marcelo Bauza, y Audilio González. «El Derecho de la Prueba y la Informática. Problemas y Perspectivas». *UNED Centro Regional de Extremadura, Mérida*, 1991.
12. Carrascosa, Valentín, María A. Pozo, y E. P. Rodríguez de Castro. *La contratación informática: el nuevo horizonte contractual. Los contratos electrónicos e informáticos*. Granada: Comares, 1999.
13. Casarino, Mario. «Manual de derecho procesal». *Derecho procesal civil* 4 (2007).
14. Cheng, Carlos Chen. «The AI Revolution: A Qualitative Analysis of its Implications in Conceptual Physics». PhD Thesis, Universidad de Panamá, 2023.
https://www.researchgate.net/profile/Eduardo-Chung/publication/377695708_La_Revolucion_de_la_IA_Un_Analisis_Cualitativo_de_sus_I

mplicaciones_en_la_Fisica_Conceptual_The_AI_Revolution_A_Qualitative_Analysis_of_its_Implications_in_Conceptual_Physics/links/65b2cd411bed776ae3046f8e/La-Revolucion-de-la-IA-Un-Analisis-Cualitativo-de-sus-Implicaciones-en-la-Fisica-Conceptual-The-AI-Revolution-A-Qualitative-Analysis-of-its-Implications-in-Conceptual-Physics.pdf.

15. Citron, Danielle, y Robert Chesney. «Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security». *California Law Review* 107, n.º 6 (1 de diciembre de 2019): 1753.
16. Comisión Europea. «White Paper on Artificial Intelligence: A European Approach to Excellence and Trust». Accedido 10 de abril de 2024. https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en.
17. Council of Europe Publishing. «Algorithms and Human Rights - Study on the Human Rights Dimensions of Automated Data Processing Techniques and Possible Regulatory Implications». Accedido 27 de abril de 2024. <https://edoc.coe.int/en/internet/7589-algorithms-and-human-rights-study-on-the-human-rights-dimensions-of-automated-data-processing-techniques-and-possible-regulatory-implications.html>.
18. Couture, Eduardo. «Estudios de Derecho Procesal Civil (Buenos Aires, Ediciones Depalma) t». II, 1979.
19. Díez-Picazo, Luis. *Fundamentos del derecho civil patrimonial: Introducción; Teoría del contrato; Las relaciones obligatorias*. Tecnos, 1979.
20. Dis, Eva A. M. van, Johan Bollen, Willem Zuidema, Robert van Rooij, y Claudi L. Bockting. «ChatGPT: Five Priorities for Research». *Nature* 614, n.º 7947 (febrero de 2023): 224-26. <https://doi.org/10.1038/d41586-023-00288-7>.
21. eCommerce. «Prompts de Chat GPT para eCommerce», 19 de junio de 2023. <https://blog.soy.es/prompts-de-chat-gpt-para-ecommerce/>.
22. Emma Grey Ellis. «People Can Put Your Face on Porn—and the Law Can't Help You». *Wired*, 26 de enero de 2018. <https://www.wired.com/story/face-swap-porn-legal-limbo/>.
23. Emol. «Denuncian que alumnos del Saint George usaron IA para crear imágenes desnudas de compañeras: Fiscalía investiga», 24 de mayo de 2024. <https://www.emol.com/noticias/Nacional/2024/05/24/1131697/presentan-denuncia-delitos-indole-sexual.html>.
24. EUR-Lex. «Cláusula de exclusión voluntaria». Accedido 12 de abril de 2024. <https://eur-lex.europa.eu/ES/legal-content/glossary/opting-out.html>.

25. Ezurmendia Álvarez, Jesús, María de los Ángeles Gonzalez Coulon, Jonatan Valenzuela Saldías, Jesús Ezurmendia Álvarez, María de los Ángeles Gonzalez Coulon, y Jonatan Valenzuela Saldías. «Incertidumbre probatoria por exceso: El manejo masivo de datos y la inclusión probatoria en el proceso penal chileno». *Política criminal* 17, n.º 34 (diciembre de 2022): 635-50. <https://doi.org/10.4067/S0718-33992022000200635>.
26. Fernández, Fernando. «El documento electrónico en el derecho civil chileno: Análisis de la Ley 19.799». *Ius et Praxis* 10, n.º 2 (2004): 137-67. <https://doi.org/10.4067/S0718-00122004000200005>.
27. Fernández, Ignacio, y José María Cervelló Grande. «La prueba y el documento electrónico». En *Derecho de Internet: la contratación electrónica y firma digital*, 385-406. Thomson Reuters Aranzadi, 2000.
28. Fernando Martín Diz. «Capítulo XLV. Inteligencia artificial y proceso: garantías frente a eficiencia en el entorno de los derechos procesales fundamentales». En *Justicia: ¿Garantías versus eficiencia?*, 1a. ed. Tirant lo Blanch, 2020.
29. Fuegi, John, y John Francis. «Lovelace & Babbage and the creation of the 1843'notes'». *IEEE Annals of the History of Computing* 25, n.º 4 (2003): 16-26.
30. García-Peñalvo, Francisco, y Andrea Vázquez-Ingelmo. «What Do We Mean by GenAI? A Systematic Mapping of The Evolution, Trends, and Techniques Involved in Generative AI», diciembre de 2023. <https://doi.org/10.9781/ijimai.2023.07.006>.
31. Girardi, Enzo. «Digitalización, política e inteligencia artificial», s. f.
32. Gissi, Nicolás, Hedilberto Aguilar, Nicolás Gissi, y Hedilberto Aguilar. «Entre el miedo y la xenofobia: política migratoria, vulnerabilidad social y emergencia de un Estado penal en Chile (2018-2023)». *Revista pueblos y fronteras digital* 18 (2023). <https://doi.org/10.22201/cimsur.18704115e.2023.v18.680>.
33. González Alonso, Julia. «La bioética en el Consejo de Europa». *Revista Española de Salud Pública* 71, n.º 4 (julio de 1997): 331-34.
34. González Castillo, Joel. «LA FUNDAMENTACIÓN DE LAS SENTENCIAS Y LA SANA CRÍTICA». *Revista chilena de derecho* 33, n.º 1 (abril de 2006): 93-107. <https://doi.org/10.4067/S0718-34372006000100006>.
35. Gutierrez, Carlos I., Anthony Aguirre, Risto Uuk, Claire C. Boine, y Matija Franklin. «A Proposal for a Definition of General Purpose Artificial Intelligence Systems». *Digital Society* 2, n.º 3 (12 de septiembre de 2023): 36. <https://doi.org/10.1007/s44206-023-00068-w>.

36. Haataja, Meeri, y Joanna J. Bryson. «The European Parliament's AI Regulation: Should We Call It Progress? Special Section: AI and Its Regulation (Part 1)». *Amicus Curiae* 4, n.º 3 (2023 de 2022): 707-18.
37. Hansen Hsu. «AI and Play, Part 1: How Games Have Driven Two Schools of AI Research». Computer History Museum, 23 de julio de 2020. <https://computerhistory.org/blog/ai-and-play-part-1-how-games-have-driven-two-schools-of-ai-research/>.
38. Hebb, Donald Olding. *The organization of behavior: A neuropsychological theory*. Psychology press, 2005.
<https://www.taylorfrancis.com/books/mono/10.4324/9781410612403/organization-behavior-hebb>.
39. Helmus, Todd C. «Artificial Intelligence, Deepfakes, and Disinformation: A Primer». RAND Corporation, 2022. <https://www.jstor.org/stable/resrep42027>.
40. Hunter Ampuero, Iván. «Reglas de prueba legal y libre valoración de la prueba: Cómo conviven en el Proyecto de Código Procesal Civil». *Ius et Praxis* 23, n.º 1 (septiembre de 2017): 247-72. <https://doi.org/10.4067/S0718-00122017000100008>.
41. INFORME DE LA COMISIÓN AL PARLAMENTO EUROPEO, AL CONSEJO Y AL COMITÉ ECONÓMICO Y SOCIAL EUROPEO Informe sobre las repercusiones en materia de seguridad y responsabilidad civil de la inteligencia artificial, el internet de las cosas y la robótica (2020). <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52020DC0064>.
42. Jacobson, Neill. «Deepfakes and Their Impact on Society». CPI OpenFox, 26 de febrero de 2024. <https://www.openfox.com/deepfakes-and-their-impact-on-society/>.
43. Javier Márquez. «GPT-4: así es la IA más avanzada de OpenAI, cómo funciona y todas las novedades». Xataka, 10 de noviembre de 2023. <https://www.xataka.com/nuevo/gpt-4-que-cuando-sale-como-funciona-toda-informacion>.
44. Julia Angwin, Jeff Larson, Surya Mattu, y Lauren Kirchner. «Machine Bias». ProPublica, 23 de mayo de 2016. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
45. Karen Hao. «Caso práctico: probamos por qué un algoritmo judicial justo es imposible». MIT Technology Review, 11 de noviembre de 2021. <http://www.technologyreview.es/s/13800/caso-practico-probamos-por-que-un-algoritmo-judicial-justo-es-imposible>.
46. ———. «EE. UU. responsabilizará a las tecnológicas de las “fake news” y el sesgo algorítmico». MIT Technology Review, 26 de abril de 2021.
<http://www.technologyreview.es/s/11097/ee-uu-responsabilizara-las-tecnologicas-de-las-fake-news-y-el-sesgo-algoritmico>.

47. Kavanagh, Camino. «Artificial Intelligence». *New Tech, New Threats, and New Governance Challenges*: Carnegie Endowment for International Peace, 2019.
<https://www.jstor.org/stable/resrep20978.5>.
48. Kevin Roose. «Here Come the Fake Videos, Too». *The New York Times*, 4 de marzo de 2018.
<https://www.nytimes.com/2018/03/04/technology/fake-videos-deepfakes.html>.
49. Korinek, Anton. «Integrating Ethical Values and Economic Value to Steer Progress in Artificial Intelligence». En *The Oxford Handbook of Ethics of AI*, de Anton Korinek, 474-91. editado por Markus D. Dubber, Frank Pasquale, y Sunit Das. Oxford University Press, 2020.
<https://doi.org/10.1093/oxfordhb/9780190067397.013.30>.
50. Kristian Lum. «Cómo fijar la responsabilidad algorítmica sin saber qué es un algoritmo». *MIT Technology Review*, 3 de marzo de 2021. <http://www.technologyreview.es/s/13213/como-fijar-la-responsabilidad-algoritmica-sin-saber-que-es-un-algoritmo>.
51. Linera, Miguel Ángel Presno. «La propuesta de “Ley de Inteligencia Artificial” europea». *Revista de las Cortes Generales*, 29 de diciembre de 2023, 81-133.
<https://doi.org/10.33426/rcg/2023/116/1775>.
52. López de Mántaras, Ramón. «Algunas reflexiones sobre el presente y futuro de la Inteligencia Artificial», 2015. <https://digital.csic.es/handle/10261/136978>.
53. Mayer Lux, Laura, Jaime Vera Vega, Laura Mayer Lux, y Jaime Vera Vega. «El documento como objeto material de las falsedades documentales y del sabotaje informático en el Derecho penal chileno». *Política criminal* 14, n.º 27 (junio de 2019): 419-55.
<https://doi.org/10.4067/S0718-33992019000100419>.
54. Meneses Pacheco, Claudio. «FUENTES DE PRUEBA Y MEDIOS DE PRUEBA EN EL PROCESO CIVIL». *Ius et Praxis* 14, n.º 2 (2008): 43-86.
55. Michele, Taruffo. *La prueba*. Marcial Pons, 2008.
56. Microsoft Research. «VASA-1». Accedido 21 de abril de 2024.
<https://www.microsoft.com/en-us/research/project/vasa-1/>.
57. Ministerio de Justicia del Reino Unido y Dominic Raab. «New Laws to Better Protect Victims from Abuse of Intimate Images». GOV.UK, 22 de noviembre de 2022.
<https://www.gov.uk/government/news/new-laws-to-better-protect-victims-from-abuse-of-intimate-images>.
58. Minsky, Marvin, y Seymour Papert. «An introduction to computational geometry». *Cambridge tiass., HIT* 479, n.º 480 (1969): 104.
59. Mondaca, Gabriela. «Imágenes falsas: la IA abre una nueva forma de bullying en los colegios». *La Tercera*, 25 de mayo de 2024. <https://www.latercera.com/la-tercera->

sabado/noticia/imagenes-falsas-la-ia-abre-una-nueva-forma-de-bullying-en-los-colegios/ZWBRG32IXBDXFIGVRXTBLUW7WU/.

60. Morandín-Ahuerma, Fabio. «Recomendación Del Consejo Sobre Inteligencia Artificial de la Oede: Desigualdad e Inclusión (1st edition)». En *Principios normativos para una ética de la inteligencia artificial*, 95-102. Consejo de Ciencia y Tecnología del Estado de Puebla (Concytep), 2023.
61. Moritz, Marcel, y Valentín Gibello. «El Reglamento Europeo (UE) 2016/679: análisis de un claroscuro». *Foro: Revista de Derecho*, n.º 27 (junio de 2017): 115-28.
62. Newell, Allen, y Herbert Simon. «The logic theory machine—A complex information processing system». *IRE Transactions on information theory* 2, n.º 3 (1956): 61-79.
63. Nurkin, Tate, Stephen Rodriguez, y Ashton B. Carter. «A Framework for Understanding Applied Ai». *A Candle in the Dark*: Atlantic Council, 2019.
<https://www.jstor.org/stable/resrep20946.6>.
64. OCDE. «Recommendation of the Council on Artificial Intelligence», 2019.
65. Padilla Parot, Ricardo. «ESE DOLOR DE CABEZA LLAMADO PRUEBA LEGAL TASADA: LA ANTICIPADA APLICACIÓN DEL PROYECTO DE CÓDIGO PROCESAL CIVIL, QUE LOS TRIBUNALES ORDINARIOS DE JUSTICIA SE ENCUENTRAN LLEVANDO A CABO». *Revista chilena de derecho privado*, n.º 26 (julio de 2016): 401-10.
<https://doi.org/10.4067/S0718-80722016000100016>.
66. Parlamento Europeo. «AIDA Committee Activity Report», 22 de marzo de 2022.
https://www.europarl.europa.eu/cmsdata/248838/AIDA_Committee_Activity_Report.pdf.
67. ———. «Inteligencia artificial: la hoja de ruta del Parlamento para la UE». *Temas | Parlamento Europeo*, 27 de abril de 2022.
<https://www.europarl.europa.eu/topics/es/article/20220422STO27705/inteligencia-artificial-la-hoja-de-ruta-del-parlamento-para-la-ue>.
68. ———. «Ley de IA de la UE: primera normativa sobre inteligencia artificial». *Temas | Parlamento Europeo*, 12 de junio de 2023.
<https://www.europarl.europa.eu/topics/es/article/20230601STO93804/ley-de-ia-de-la-ue-primera-normativa-sobre-inteligencia-artificial>.
69. ———. Reglamento de Inteligencia Artificial. Accedido 12 de abril de 2024.
<https://www.diarioconstitucional.cl/wp-content/uploads/2024/03/Vea-texto-aprobado..pdf>.
70. Pinochet Olave, Ruperto. «EL DOCUMENTO ELECTRÓNICO Y LA PRUEBA LITERAL». *Ius et Praxis* 8, n.º 2 (2002): 377-412. <https://doi.org/10.4067/S0718-00122002000200012>.

71. «Prankshit». Accedido 5 de junio de 2024. <https://prankshit.com/es/generador-de-chat-de-whatsapp-falso.php>.
72. Ramón Camino, José. «El Documento Electrónico su admisibilidad en el Derecho español». *La Ley* 2 (1997).
73. Redondo, Miguel Temboury. «La prueba de los documentos electrónicos en los distintos órdenes jurisdiccionales». En *Derecho de Internet: la contratación electrónica y firma digital*, 407-38. Thomson Reuters Aranzadi, 2000.
74. Rodrigo Mundaca. «Instagram acusará las imágenes y videos hechos con Inteligencia Artificial». *VTI Universidad de Chile* (blog), 19 de abril de 2024. <https://vti.uchile.cl/instagram-acusara-las-imagenes-y-videos-hechos-con-inteligencia-artificial/>.
75. Rumelhart, David E., Geoffrey E. Hinton, y Ronald J. Williams. «Learning representations by back-propagating errors». *nature* 323, n.º 6088 (1986): 533-36.
76. Sadin, Eric. *La silicolonización del mundo: la irresistible expansión del liberalismo digital*. Caja negra, 2018.
77. Shinde, Pramila P., y Seema Shah. «A Review of Machine Learning and Deep Learning Applications». En *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, 1-6, 2018. <https://doi.org/10.1109/ICCUBEA.2018.8697857>.
78. Silver, David, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, y Marc Lanctot. «Mastering the game of Go with deep neural networks and tree search». *nature* 529, n.º 7587 (2016): 484-89.
79. Smith, Hannah, y Katherine Mansted. «What’s a deep fake?» Weaponised deep fakes. Australian Strategic Policy Institute, 2020. <https://www.jstor.org/stable/resrep25129.6>.
80. Taruffo, Michele. *La prueba de los hechos*. Trotta, 2023.
81. Ulmer, Alexandra, y Anna Tong. «Deepfaking It: America’s 2024 Election Collides with AI Boom». *Reuters*, 31 de mayo de 2023, sec. United States. <https://www.reuters.com/world/us/deepfaking-it-americas-2024-election-collides-with-ai-boom-2023-05-30/>.
82. Viejo, Manuel. «Decenas de menores de Extremadura denuncian que circulan fotos de falsos desnudos suyos creadas por inteligencia artificial: “Me dio un vuelco el corazón”». *El País*, 18 de septiembre de 2023. <https://elpais.com/espana/2023-09-18/la-policia-investiga-el-desnudo->

integral-de-varias-menores-en-extremadura-con-inteligencia-artificial-me-dio-un-vuelco-el-corazon.html.

83. Wheeler, Gregory. «Bounded rationality», 2018. <https://plato.stanford.edu/Entries/bounded-rationality/>.