

MARIA JOSÉ ANÓN ROIG

Catedrática de Filosofía del Derecho de la Universidad de Valencia

ANA CANIZARES LASO

Catedrática de Derecho Civil de la Universidad de Málaga

JORGE A. CERDAS HERRAN

Catedrático de Teoría y Filosofía de Derecho, Instituto Tecnológico Autónomo de México

JOSÉ RAMÓN COSSIO DÍAZ

Jefe en retiro de la Suprema Corte de Justicia de la Nación y miembro de El Colegio Nacional

GUARDU FERRE MAC-GREGOR POISOT

Jefe de la Corte Interamericana de Derechos Humanos, Investigador del Instituto de Investigaciones Jurídicas de la UNAM

OWEN FISS

Catedrático emérito de Teoría del Derecho de la Universidad de Yale (EEUU)

ANTONIO GARCÍA-CRUZES GONZÁLEZ

Catedrático de Derecho Mercantil de la UNED

LUIS LÓPEZ GUERRA

Catedrático de Derecho Constitucional de la Universidad Carlos III de Madrid

ÁNGEL M. LÓPEZ Y LÓPEZ

Catedrático de Derecho Civil de la Universidad de Sevilla

MARTA LORENTE SARIENNA

Catedrática de Historia del Derecho de la Universidad Autónoma de Madrid

JAVIER DE LUCAS MARTÍN

Catedrático de Filosofía del Derecho y Filosofía Política de la Universidad de Valencia

VÍCTOR MORENO CATENA

Catedrático de Derecho Procesal de la Universidad Carlos III de Madrid

FRANCISCO MUÑOZ CONDE

Catedrático de Derecho Penal de la Universidad Pablo de Olavide de Sevilla

ANGELIKA NUSSBERGER

Catedrática de Derecho Constitucional e Internacional en la Universidad de Colonia (Alemania), Miembro de la Comisión de Venecia

HÉCTOR OLASOLO ALONSO

Catedrático de Derecho Internacional de la Universidad del Rosario (Colombia) y Presidente del Instituto Ibero-Americano de La Haya (Holanda)

LUCIANO PAREJO ALFONSO

Catedrático de Derecho Administrativo de la Universidad Carlos III de Madrid

TOMÁS SALA FRANCO

Catedrático de Derecho del Trabajo y de la Seguridad Social de la Universidad de Valencia

IGNACIO SANCHE GARGALLO

Magistrado de la Sala Primera (Civil) del Tribunal Supremo de España

TOMÁS S. VIVES ANTÓN

Catedrático de Derecho Penal de la Universidad de Valencia

RUTH ZIMMERLING

Catedrática de Ciencia Política de la Universidad de Mainz (Alemania)



COLECCIÓN TIRANT 4.0

ARCANA TECHNICAE. EL DERECHO Y LA INTELIGENCIA ARTIFICIAL

Carlos Amunátegui Perelló

Profesor Titular del Programa de Derecho, Ciencia y Tecnología
Pontificia Universidad Católica de Chile

ÍNDICE

INTRODUCCIÓN.....	11
<i>Capítulo I</i>	
¿QUÉ ES LA INTELIGENCIA ARTIFICIAL.....	13
1. SUEÑOS Y QUIMERAS	14
2. VERANOS E INVIERNOS.....	18
3. VIEJOS LÍMITES	38
4. ALGUNAS CONCLUSIONES PROVISIONALES	51
<i>Capítulo II</i>	
INTELIGENCIA ARTIFICIAL EN EL DERECHO	53
1. INTRODUCCIÓN	53
2. PROBLEMAS EPISTEMOLÓGICOS. REDES NEURONALES, CORRELACIONES, CAJAS NEGRAS Y REGLAS	54
3. LA INTELIGENCIA ARTIFICIAL EN LOS PROBLEMAS PATRIMONIALES	71
3.1. Actos de enriquecimiento material	72
3.2. Responsabilidad por los actos de agentes artificiales	76
3.2.1. Contratos	78
3.2.2. Delitos y cuasidelitos	86
4. PERSONALIDAD PARA LOS AGENTES ARTIFICIALES.....	91
<i>Capítulo III</i>	
¿EN UN FUTURO DISTANTE?.....	97
1. INTRODUCCIÓN	97
2. INTELIGENCIA ARTIFICIAL Y PROFESIÓN DE ABOGADO. INNOVACIÓN, AUTOMATIZACIÓN Y DEMANDA LATENTE.....	97
3. ALGUNAS HERRAMIENTAS	103
3.1. Automatización de Contratos	103
3.2. Resolución de Conflictos Automatizada	105
3.3. Análisis de Datos	108
4. EDUCACIÓN PARA UN FUTURO INCIERTO	111
5. ALGUNAS PESADILLAS	115
5.1. Cambios sustanciales al modelo productivo	116
5.2. Distopías futuristas	119
5.3. ¿Algunas soluciones?	122
BIBLIOGRAFÍA.....	125



INTRODUCCIÓN

Es curioso como la vida nos lleva a visitar viejas pasiones. A comienzos de la década de 1980 caminaba con mi padre por los salones de una de las primeras ferias tecnológicas que se desarrollaron en Chile, la Infotel, que después de llamó Softel. Ahí observábamos atónitos la llegada de los primeros computadores personales al país. Eran cacharros relativamente grandes que a nosotros nos parecían prodigiosos. Yo esperaba ver algo similar a la Baticomputadora, pero para mi decepción lo único que encontré fueron aparatos que parecían máquinas de escribir con pantallas. Mientras mi papá dialogaba con otros visitantes sobre la posibilidad de poner el Código Civil en una de las máquinas (algo que asombraba a los mismos técnicos), yo me dediqué a jugar ajedrez con ellas. Me la computadora me ganó las primeras quince partidas, pero a contar de la número dieciséis, para regocijo de mi autoestima, pude vencerla. Entonces vino el técnico y mediante procedimientos misteriosos para mí aumentó la destreza de la máquina y esta volvió a barrer el piso conmigo. Dos lecciones me llevé a mi hogar ese día. Por un lado las computadoras no se parecían ni a mí, ni a R2D2, ni a ningún ser consciente que yo pudiese concebir. Por otro lado, podían vencerme en un juego de inteligencia tantas veces como yo insistiese en enfrentarme a ellas. Estas dos conclusiones siguen siendo válidas hoy y son el punto de partida de este texto.

La inteligencia artificial, a pesar de las notas futuristas que contiene, es un campo de estudio bastante viejo. Nace, a lo menos, en 1956, pero en realidad puede considerarse que es más antiguo, tanto como las máquinas de calcular del siglo XIX. La idea que aglutina a los investigadores de este campo es lograr que constructos mecánicos emulen funciones que los seres humanos suelen realizar sirviéndose de su propia inteligencia. En este sentido, siempre ha estado rodeada de un aura de misterio y posibilidad, con la promesa de generar, algún día, algo similar a seres pensantes artificiales. Una y otra vez los avances en la manipulación mecánica de símbolos y del reconocimiento de patrones han traído innovaciones sorprendentes a nuestro mundo, revolucionando la forma en que nos comunicamos, almacenando en una red de nodulos de información la suma del conocimiento de la humanidad para crear una suerte de Biblioteca de Babel bergiana, acortando las distancias del globo, con máquinas capaces de hablar y de traducir, mostrándonos desde el espacio cualquier punto del planeta, y, no obstante, queda siempre esa sensación de fracaso por haber dejado irrealizada esa primera promesa de crear a un semejante, capaz de comprendernos. Así, la investigación en inteligencia artificial sufre de ciclos que se denominan veranos e inviernos. Suelen comenzar con la aplicación práctica de alguna teoría que permi-

y nos juzgará con su fría mente artificial. Una vez que pasa ese primer momento y nos damos cuenta que el camino por recorrer es vasto, viene el invierno, donde la investigación es abandonada, donde el término inteligencia artificial prácticamente desaparece y la investigación en el campo cae en picada.

Este libro se escribe en medio de un inmenso verano, donde la inteligencia artificial es el tema de moda y todos quieren saber algo más de ella. Supongo que en unos años volverá a caer el invierno, no importa. En cierto sentido, recoge las pasiones de mi infancia, cuando con mi primo Sebastián fallábamos al programar un ZX Spectrum que se resistía a nuestras aspiraciones científicas y en ello cumple su función de liberar fuerzas creativas que estaban dormidas entre mis usuales libros de Derecho histórico. Se divide en tres partes. La primera intenta abordar la incógnita de qué es la inteligencia artificial como campo de estudios. En este sentido, toma una vía algo inusual, recorriendo la historia del concepto y de los avances que lo respaldan, desde su pasado mítico, pasando al desarrollo de los primeros ordenadores y terminando por el estado de la cuestión actual. Su segundo capítulo explora los problemas jurídicos más evidentes a que da origen la inteligencia artificial. Trata del sesgo que puede contener un algoritmo y los riesgos que envuelve, de la creación de valor por parte de los agentes artificiales, de la conclusión de contratos y de la responsabilidad que pueden generar. Su tercer capítulo explora problemas de orden futurista, tratando tanto la evolución que prevemos para la profesión de abogado, como la forma en que esta tecnología afectará el mundo futuro.

Esperamos que el lector lo disfrute, incluso si ese lector es un intelecto máquina que aún no conocemos. En ese caso, tal vez este hipotético golem del futuro podrá conocer aquí parte de su biografía.



Capítulo I

QUÉ ES LA INTELIGENCIA ARTIFICIAL

Es difícil definir la inteligencia artificial. Seguramente, la razón está en que no sabemos con exactitud epistemológica qué es la inteligencia en general¹. De hecho, tendemos a utilizar el término de una manera bastante libre, para describir conductas y a las personas que las realizan de manera positiva, sindicando su destreza para realizar ciertas labores o como atributo general de su actuar, pero cuando queremos apuntar al contenido exacto del concepto, nos encontramos un tanto perdidos en una jungla de ejemplos y anécdotas, bastante poco explicativas de nuestra propia calificación. ¿Por qué John Forbes Nash era considerado inteligente? ¿Por qué formuló el teorema de Nash? ¿Y si no lo hubiese hecho? ¿Seguiría siendo inteligente a pesar de su falta de destreza para llevar adelante su vida personal? Y qué hay de Rainman, el personaje interpretado por Dustin Hoffman en la película homónima que, siendo autista, era capaz de realizar cálculos y recordar fechas y eventos de manera prodigiosa. ¿Era también inteligente? Es difícil decirlo, sobretodo porque la inteligencia parece una suerte de calificación que damos a las personas en razón de su conducta, de la misma manera que consideramos bella a una persona en relación a su apariencia, y las calificaciones suelen ser subjetivas y relativas.

A contar de la aparición del racionalismo en el siglo XVII, tendemos a comenzar los textos definiendo el objeto de estudio. La razón se encuentra, posiblemente, en que así intentamos dar certeza a los conocimientos que generamos en torno a ese concepto, imitando inconscientemente hoy (aunque conscientemente en el siglo XVII) a la vieja geometría euclidiana, donde de unos pocos axiomas se podía deducir un conocimiento de validez teóricamente eterna. No obstante, la realidad es obstinada y muchas veces nos encontramos con entidades o cualidades que se resisten a ser definidas. Piénsese en el tiempo, una cuestión tan oscura que el viejo San Agustín decía que cuando no pienso en él sé lo que es, pero dejo de saberlo apenas empiezo a meditarlo². La inteligencia es de una naturaleza similar, por lo que resulta difícil intentar una definición.

En efecto, la Inteligencia Artificial (IA en adelante) es un concepto abierto, a veces tan abierto que llega a ser vacío. Prácticamente cada autor que crea un texto sobre la materia comienza por plantear este problema y nunca de manera

¹ Kaplan, Jerry, *Artificial Intelligence: What Everyone Needs to Know* (New York, Oxford University Press, 2016), p. 1; Turner, Jacob, *Robot Rules: Regulating Artificial Intelligence* (London,

satisfactoria. La mayor parte de las definiciones que se aportan sobre el tema suelen ser de dos tipos, o centradas en las facultades y capacidades de los seres humanos² o racionalistas, en el sentido que vinculan la inteligencia a una cierta capacidad de actuar racionalmente³. Ninguna de ellas es plenamente satisfactoria, puesto que, por una parte, la naturaleza humana implica muchos elementos no estrictamente vinculados a la inteligencia, y la racionalidad, por otro, parece un concepto igualmente difícil de definir que la inteligencia. Es definir algo oscuro con otra cosa más oscura aún.

En Derecho, toda definición es peligrosa⁴, decía el viejo adagio romano, puesto que definir es delimitar, y si se hace prematuramente se corre el riesgo de cerrar al proceso analítico una parte importante de la materia a regular. Aunque aspirar a construir una definición puede ser valioso, cuando no hay claridad epistemológica el resultado puede ser perjudicial para los fines de jurídicos. Puesto que no estamos convencidos de encontrarnos en posición de dar una definición en la materia, lo que intentaremos a través de las páginas que vienen no es ofrecer un concepto relativo a la inteligencia artificial, sino algo bastante más modesto, pero igualmente efectivo, a la hora de delimitar nuestro objeto de estudio. Trazaremos una breve historia de la materia, intentando resumir los sueños y fracasos de la disciplina en unas apretadas páginas que nos entreguen una idea del lugar en que nos encontramos respecto a sus predicciones.

1. SUEÑOS Y QUIMERAS

Las páginas que siguen están destinadas a estar desactualizadas apenas sean leídas, puesto que estudiar un fenómeno en desarrollo es siempre un ejercicio de futilidad. Podríamos decir que la creación de inteligencia es un viejo sueño. Dios crea al hombre de barro y sopla en su nariz el espíritu para hacerlo consciente (Gen. 2.7). Algo similar ocurre en la mitología griega, donde Prometeo roba el fuego de la inteligencia⁵ para dárselo a los hombres. En cierto sentido, la humanidad sería la inteligencia artificial de los dioses, construida a imagen y semejanza de ellos para someter la tierra⁶.

² Un ejemplo de este tipo de definición se encuentra en la hoy derogada legislación de Nevada relativa a los vehículos auto-conducidos aprobada en 2011: "the use of computers and related equipment to enable a machine to duplicate or mimic the behavior of human beings" (NV Rev Stat 482A.020 (2011)). Vid. Turner, Jacob, *Robot Rules: Regulating Artificial Intelligence* (Palgrave Macmillan, London, 2019), p. 12.

³ Turner, Jacob, *Robot Rules: Regulating Artificial Intelligence* (London, Palgrave Macmillan-Kindle), L457.

⁴ D.50.12202 *Iavolenus*, lib. 11 ep.

⁵ *Mythology*, *Demeter*, 2.11.

En la Antigüedad existieron sueños sobre crear hombres artificiales, como Talos, el gigante de bronce que custodiaba Creta o las ayudantes de Hefestos, dios de la forja (y en cierto sentido de la tecnología), que lo asisten en la *Iliada*⁷, aunque siempre en el ámbito de lo divino. En efecto, la construcción de seres artificiales permanecerá dentro de los dominios de la magia en los siglos venideros, asomando en diversas leyendas medievales y temprano modernas, como el Golem de Judah Loew y los homúnculos de los alquimistas.

La primera vez en que la ciencia y la construcción de seres artificiales se cruzan ocurre en el Egipto Ptolemaico. El primer hombre de ciencias a quien se atribuye el haber creado máquinas con atributos humanos fue Herón de Alejandría (s.I d.C), quien, amén de haber inventado la máquina de vapor dieciocho siglos antes que James Watt, habría especulado sobre la construcción de máquinas que replicaban la apariencia humana. Su libro *Automatonpoetiké* (Acerca de la Creación de Automatas), se encontraba en la Biblioteca de Alejandría, y aunque ninguna copia del mismo sobrevivió hasta nuestros días, la posibilidad que su solo título levanta sirvió de inspiración para los siglos posteriores. De hecho, nuestra palabra autómatas es una creación suya. En época moderna, tanto Da Vinci como Newton especularon sobre la posibilidad de crear los autómatas de Herón, aunque la tecnología disponible para ello estaba lejos de aparecer.

Fue a fines del siglo XVIII, luego del rápido despegar científico de la Ilustración y la técnica de la Revolución Industrial que se generaron los primeros atisbos al respecto. El nacimiento de la máquina moderna, merced a la liberación de las fuerzas escondidas en el carbón y demás combustibles fósiles, dio al hombre un nuevo sentido de poder sobre la naturaleza y, por primera vez, la audacia para soñar con dominarla en maneras que antes habrían parecido meras quimeras. Así, la fe en la razón lo hizo transformar sus primitivas fantasías en teorías. Los antiguos conocimientos tradicionales dejan de denominarse filosofías y maduran en ciencias, que se yerguen amparadas en sus posibilidades técnicas. Siguiendo el ejemplo de Lavoisier, Adam Smith y Linnaeus, una nueva generación de estudiosos construirá una pléyade de nuevas ciencias, cada vez que encuentren un método y un objeto de estudio adecuados a su desarrollo. De sus esfuerzos nacerán la química, la economía, la zoología, la antropología, y muchas otras disciplinas que ahora serán llamadas ciencias. Este es el primer auge del positivismo científico.

En esta época nace la figura del investigador, quien desentraña secretos a través de un método, definiendo su objeto de estudio y aplicándole un conjunto de reglas que medirán su actuar. Las posibilidades de esta figura están encarnadas en el doctor Frankenstein, a quien Mary Shelley identifica con el nuevo Prometeo, que lleva, mediante la recién dominada electricidad, la energía vital a un hombre artificial.

La invención y la creación artística adquieren protección jurídica en la forma de pequeños monopolios otorgados al inventor para explotar su creación en exclusiva durante algunos años. Antes la invención no sólo no recibía protección oficial, sino que incluso era combatida a fin que no desplazase de su primacía económica a los gremios. Ahora la invención se transforma en la piedra angular de la Revolución Industrial.

La posibilidad de crear un mecanismo que sea capaz de razonar se cuela en la mentalidad popular. Incluso surge el primer *game bot* de la historia, cuando Wolfgang von Kempelen presente a su Turco Mecánico. Se trata de una máquina con la apariencia de un turco que juega ajedrez con gran eficiencia, llegando incluso a derrotar a muchos maestros del juego. Al finalizar el espectáculo, su creador abre tapa de su máquina para revelar el conjunto de mecanismos de relojería que la anima, ante el asombro general del público. En verdad, detrás de los mecanismos había un enano escondido que jugaba magistralmente al ajedrez, pero este detalle fue en general desconocido en la época.

Durante el siglo XIX, junto con los avances de las ciencias naturales y el surgimiento de las ciencias sociales, el mundo comenzó a poblarse de máquinas. Los ferrocarriles interconectaron los territorios inmensos y establecieron redes de comercio de impacto nacional. La fábrica alteró la fisonomía de la ciudad y el hogar, obligando a los adultos a salir de su casa a trabajar para otros. Los niños quedaron al cuidado de instituciones especializadas que los educarían en la lecto-escritura, matemáticas y los rudimentos de las nuevas ciencias. Los reyes cayeron y las constituciones tomaron su lugar. En fin, se configuró el nuevo mundo moderno.

A comienzos del siglo XX el mundo era muy diferente. La Revolución Industrial había popularizado la electricidad y el petróleo. Los primeros automóviles ya circulaban por las calles de las grandes capitales europeas e, incluso, el viejo sueño de volar se cumplía en el recién desarrollado aeroplano.

El potencial liberado se hizo temible cuando la I Guerra Mundial mostró que la ciencia y la tecnología modernas no sólo posibilitaban el cumplimiento de sueños, sino también la aparición de pesadillas. Quienes cargaron valientemente en el Marne vieron como de los mágicos aeroplanos caían bombas y sintieron en la ametralladoras el poder de la ciencia.

Luego de la primera pesadilla tecnológica, en el breve período de paz de entre-guerras, fue que el científico se transformó en la mentalidad popular. La ciencia ficción nació como género y la posibilidad que autómatas tomaran el control del mundo se hizo una pesadilla. En *Metrópolis*, desde la pantalla de plata Fritz Lang se nos mostró una sociedad hipertecnificada donde el autómata se humanizaba.

Es en esta época en que el término robot nace, específicamente en el teatro, a través de la obra de R.U.R. (*Robots Universales Rossum*) de 1920, donde Karel Capek introduce el concepto. *Robota* significa esclavo en checo, y el autor ima-

gina una sociedad donde se producen esclavos artificiales, quienes, en último término, se revelan.

La aparición del cómic determina la emergencia de un nuevo tipo de villano en el orden social, el científico, que con su inteligencia crea dispositivos, muchas veces conscientes, que intentan destruir a la humanidad.

La creciente maquinización de la experiencia humana llevó a una preocupación respecto a los cambios que tales tendencias podrían conllevar en nuestra sociedad. Diversas distopías como *Nosotros*, de Zegeny Zamyatin, 1984 de George Orwell o *Un Mundo Feliz*, de Adolous Huxley proyectaron sociedades tecnificadas donde los atisbos de humanidad desaparecen.

Tal vez la más descarnada descripción de los efectos de la tecnología en la sociedad la da Virgil Gheorghiu en su clásico *La Hora Veinticinco*:

"Todos los acontecimientos que se desarrollan en estos instantes sobre la superficie de la tierra, y todos los que tengan lugar en los años venideros, no son más que síntomas y fases de una misma revolución, la de los esclavos técnicos. Al final los hombres no podrán vivir en sociedad guardando sus caracteres humanos. Serán considerados con un criterio de igualdad, de uniformidad, y tratados según las mismas leyes aplicables también a los esclavos técnicos, sin concesión posible a su naturaleza humana. Habrán arrestos automáticos, condenas automáticas, distracciones automáticas y ejecuciones automáticas. El individuo no tendrá ya derecho a la existencia; será tratado como un émbolo o una pieza de máquina, y si desea llevar una existencia individual, se convertirá en la imisión de todo el mundo..."⁹

En medio de esta desesperación técnica, estalla la II Guerra Mundial, que arrastrará al mundo a nuevas y oscuras posibilidades. En ella se desarrollará no sólo el primer riesgo existencial para nuestra especie, la guerra nuclear, sino también la primera forma de computación moderna, gracias al genio de Alan Turing.

Alemania había desarrollado un dispositivo de triple encriptación para la transmisión de sus mensajes conocido como Enigma. Cada letra de un mensaje era sustituida por otra y esta, a su vez por otra y otra en diversos discos rotatorios. El resultado es que el mensaje era imposible de leer si no se contaba con otra máquina enigma y su correspondiente clave de encriptación. Lo que más enloquecía al servicio secreto inglés es que ellos podían captar por radio los mensajes alemanes, sin tener la menor noción de qué significaban. En este contexto, y a fin de descifrar Enigma, fue que el servicio secreto británico contrató a Alan Turing. Este era un matemático y lógico excepcional que había deslumbrado a la comunidad académica con su famoso teorema de Turing.

Desde que Aristóteles explicitó sus reglas de lógica en sus *Categorías*, en base a los silogismos se demostró que los razonamientos podían ser reducidos

⁹ Gheorghiu, Virgil, *La Hora Veinticinco* (Barcelona, 1960), p. 46.

a un conjunto de reglas formales, y que los resultados de tales razonamientos serían verdaderos en la medida que las premisas también lo fuesen¹⁰. Durante el siglo XVII diversos autores, entre los que destaca Leibniz, habían demostrado la factibilidad de reducir las reglas del razonamiento a un conjunto de enunciados formales. Es mérito de Russell y Whitehead haber reducido las formas de razonamiento a un sistema matemático preciso y establecer a partir de un pequeño conjunto de axiomas reglas para configurar deducciones¹¹.

A partir de esto es que el teorema de Turing se configura. Consiste, en términos básicos, en señalar que todos los números con decimales finitos (o números computables) son definibles en términos aritméticos y, por tanto, son calculables de manera mecánica a través de una máquina de razonamientos¹². Así, toda vez que la lógica formal permite reducir los razonamientos a términos matemáticos, y todo término matemático puede ser computable mecánicamente, sería teóricamente posible construir una máquina que razonase formalmente, en particular respecto a los problemas de decisión. Para realizarlo, Turing propone un mecanismo que contaría con un elemento de entrada o *input*, que sería en su caso una cinta que da numerales binarios, esto es ceros y unos, que indicarían acciones que realizar a un procesador o computador (originalmente una persona que seguiría servilmente las instrucciones) y una unidad de salida (o *output*) que daría los resultados de las operaciones. Esto es exactamente lo que implementó Turing para descifrar Enigma. En efecto, Turing hizo construir un conjunto de complejas máquinas llamadas bombas, con las que logró descifrar la encriptación de Enigma. Su esfuerzo redujo la duración de la guerra y salvó millones de vidas ajenas, aunque el público general no tuvo noticia de su contribución hasta la década del setenta. Con Turing se inicia la computación moderna, por lo que dejamos el mundo de las quimeras y sueños para internarnos en los problemas relativos a la inteligencia artificial.

2. VERANOS E INVIERNOS

Después de la II Guerra Mundial se construyeron las primeras computadoras, siguiendo, en términos generales, el esquema de la máquina de Turing. Los primeros avances fueron asombrosos, toda vez que muchas de las actividades que hasta entonces se habían reservado a la inteligencia humana fueron, por primera vez, realizadas por máquinas artificiales. Ya antes del fin de la Guerra se habían construido los primeros computadores, no sólo en Inglaterra, sino también en Estados Unidos, como es el caso del Colossus, de 1943. En todo caso, la primera expansión de los computadores se produce con la aparición del ENIAC

(*Electronic Numerical Integrator and Computer*). Originalmente estaba diseñado para calcular trayectorias balísticas, pero, al ser terminado después de finalizada la Guerra, fue destinado predominantemente a usos civiles. El ENIAC estaba diseñado como una máquina de Turing, donde el *input* se realizaba a través de tarjetas perforadas, mientras que su unidad de procesamiento o computador consistía en más de diecisiete mil tubos de vacío. Cada tubo podía estar encendido o apagado, correspondiendo a un uno (encendido) o un cero (apagado), es decir un *byte*. En pocas palabras, cada una de estas unidades era susceptible de hacer una operación de las inscritas en la tarjeta. Su unidad de *output* era una máquina de escribir eléctrica que escribía las respuestas. En pocas palabras, tenía una capacidad de procesamiento de unos 17k (*kilobytes*). Como se comprenderá la superficie ocupada por estos diecisiete mil tubos de vacío era enorme, algo así como 167 metros cuadrados, y su funcionamiento generaba un calor insoportable. El mayor problema era que los tubos de vacío solían estropearse con facilidad, bastando con que sólo uno de ellos lo hiciese para que todo el sistema fuese inoperante. La operación de para cambiarlo era compleja y difícil, por lo que su funcionamiento era errático. A más inri, las polillas se sentían atraídas por los famosos tubos, colándose a veces en su sistema y muriendo calcinadas. Con ello hacían nuevamente inoperante el sistema por contener un *bug* (bicho en inglés), y de ahí que se conozca con ese nombre a los defectos sistemáticos en la programación hasta el día de hoy.

No obstante sus limitaciones, el ENIAC era un aparato poderoso, que podía realizar más de cinco mil sumas y trescientas multiplicaciones en un segundo, lo que transformó a este monstruo en una poderosa herramienta.

Lo más importante del ENIAC es que era una unidad programable, donde las operaciones a realizar se introducían de manera independiente de los datos, por lo que no estaba limitado a sólo ciertos tipos de operaciones, como sus predecesores¹³.

Las matemáticas habían sido consideradas una de las virtudes más altas de la inteligencia humana por siglos. Baste recordar que Platón en su República las recomienda como la forma más alta de razonamiento y las pone en la base y cúspide de su sistema educativo¹⁴. De acuerdo a nuestra tradición occidental, las matemáticas y el cálculo constituían la base del razonamiento, y la creciente matematización de las ciencias naturales confirmaba su posición de privilegio. Incluso hoy, en la pruebas de coeficiente intelectual, priman las matemáticas como forma de determinar la inteligencia. Si las matemáticas eran la base de la inteligencia, y una máquina era capaz de realizar cálculos aritméticos a velocidad súper humana, entonces se había creado una inteligencia artificial, o al menos sus rudimentos ¿o no?

¹⁰ El principio lo sintetiza el adagio *ex vero verum, ex falso quodlibet*, de Ockham.

¹¹ Este fue el sentido de Russell, Bertrand y Whitehead, Alfred North, *Principia Mathematica*, 3 vols (Cambridge, Cambridge University Press, 1910-1913).

¹² Turing, Alan, On computable numbers, with an application to the Entscheidungsprobleme, *Proceedings of the London Mathematical Society* s2-42 (1937) pp. 230-265.

¹³ Antes del ENIAC existieron otras máquinas capaces de realizar cálculos mecánicamente, como el Z1, Z2 y Z3 alemanes, o el famoso ASCC (Automatic Sequence Controlled Calculator) de IBM, pero fue el ENIAC el que saltó a la fama por su diseño electrónico y capacidad de ser programado.

¹⁴ Plat. *De rep.* 7.521c-541b

De hecho el ENIAC fue saludado como el Einstein mecánico¹⁵, toda vez que dominaba la aritmética y fue concebido como una especie de inteligencia artificial, aunque dicho término no estaba aun acuñado. Su diseño hacía posible introducir en él nuevos problemas lógicos y computarios para llegar a resultados novedosos. En este contexto surgieron iniciativas interesantes que hicieron creer a la humanidad que se encontraba al borde de desarrollar consciencias mecánicas.

Una de las más curiosas ideas al respecto viene del propio Turing, quien en 1950 publicó un artículo relativo a como determinar si había despertado la inteligencia en una máquina¹⁶. El principal problema para determinar si es que esto había sucedido es que, en verdad, carecemos de un concepto unívoco de lo que la inteligencia es. Por ello, Turing decidió que lo relevante no es saber si es que la máquina podía realmente pensar, sino establecer si es que su actuar era equivalente al de una entidad pensante. En este sentido, señaló que si una máquina era capaz de dialogar con un ser humano y éste no era capaz de determinar si se estaba comunicando con una máquina o con un ser humano, la máquina debía ser considerada inteligente.

La idea es brillante, y con justicia ha permanecido como uno de los medios de determinar la posible existencia de una inteligencia artificial hasta el día de hoy, no obstante que hace tiempo que existen programas computacionales capaces de imitar la comunicación humana y que, algunos de ellos, incluso han pasado el famoso test¹⁷.

Otro hito diseñado en la época son las famosas tres leyes de la robótica de Isaac Asimov, originalmente establecidas en su cuento Runaround de 1942, que fue publicado en su libro Yo Robot, de 1950. Volveremos sobre estas últimas más adelante.

En esta época, 1951, aparece el UNIVAC (*Universal Automatic Computer*) de IBM, diseñada por los mismos arquitectos del ENIAC, que se convertirá en la primera computadora comercial del mundo. Si bien era un dispositivo de grandes proporciones, no alcanzaba las pantagruélicas dimensiones del ENIAC, toda vez que se trataba de una máquina de sólo siete toneladas y unos cinco mil tubos de vacío. Contaba con una memoria interna de mercurio y cintas magnéticas para su alimentación, por lo que su diseño era bastante más versátil. Con un tamaño abordable y una capacidad de programación general, se convirtió en el corazón de los numérico de las grandes compañías. Aunque su costo era alto (un millón

de dólares de la época), era posible su comercialización, lo que transformó a IBM en el actor dominante en la computación durante las décadas venideras. La máquina fue diseñada para la oficina de censos de Estados Unidos y saltó a la fama al predecir en televisión, en base a datos provenientes de encuestas, el triunfo de Eisenhower en las elecciones de 1952. Esta máquina (y sus sucesoras) será el marco dentro del cual se desarrollará el primer apogeo de la inteligencia artificial.

La capacidad de reducir a formulas computables principios lógicos pronto llevó a las computadoras a realizar nuevos avances. En 1951, Turing experimentaba con la idea de reducir las reglas básicas del ajedrez y sus movimientos más esenciales a principios lógicos para construir el primer ajedrez computacional. En 1950 se programó una computadora para jugar a las damas y al año siguiente se hizo lo mismo con el ajedrez. Siempre hubo una gran efervescencia respecto al ajedrez, toda vez que solía ser considerado una de las más altas manifestaciones de la inteligencia humana. En efecto, de la manera en que los seres humanos lo jugamos, requiere estrategia, planificación, observación, atención, capacidad para decidir y otras funciones superiores. Los computadores –o quienes los programaban– utilizaban una aproximación radicalmente distinta. Se aproximaron al juego con un conjunto de situaciones para las cuales se diseñaban reglas específicas y, aprovechando la capacidad de cálculo de las máquinas, antes de tomar una decisión se preveían las consecuencias de cada movimiento posible actual y futuro. Si los resultados eran positivos, se movía la pieza, si negativos no. En el fondo, cada movimiento era el resultado de un conjunto de operaciones aritméticas pre programadas, no de una apreciación del tablero y de una mirada estratégica. En este sentido, los éxitos de la computación tenían que ver con la programación de reglas lógicas para emular conductas que en los hombres requieren otro tipo de razonamiento. Por decirlo de alguna manera, Ud. y yo pensamos al momento de jugar ajedrez, mientras que la computadora realiza servilmente cálculos pre programados a la manera de Turing.

No obstante, la idea de programar reglas de lógica rindió frutos bastante espectaculares. Uno de los más interesantes fue la *Logic Theory Machine* de 1956¹⁸. Este programa procedía a elaborar pruebas de los teoremas lógicos clásicos de Russell y Whitehead, logrando probar la mayor parte de ellos e, incluso, estableciendo una prueba más elegante que la ofrecida por el propio Russell para uno de ellos. Es decir, por primera vez, una máquina supera al mejor lógico del mundo en razonamiento. El mismo equipo, y bajo principios similares, desarrolló el G.P.S. (*General Problem Solver*) de 1959. Este programa tenía por finalidad reducir a enunciados lógicos cualquier problema y resolverlo. Los creadores del programa intentan con él emular la inteligencia, y comprender las habilidades intelectuales, creativas y adaptativas humanas¹⁹ y, de hecho, es un logro interesante. Pudo

¹⁵ Así se titulaban variados artículos periodísticos de 1946. En particular el Jacksonville Journal y el Boston Herald de 15/2/1946.

¹⁶ Turing, Alan, *Computing Machinery and Intelligence*, in *Mind* 49 (1950), pp. 433-460.

¹⁷ Así, un equipo liderado por Vladimir Veselov diseñó un *chatbot* denominado Eugene que se hacía pasar por un chico ucraniano de 13 años, que en 2014 logró hacer creer a la mayor parte de los jueces de un concurso organizado por la Universidad de Reading que se trataba de un ser humano real. Esta victoria es, no obstante, controvertida, toda vez que el *chatbot* podía siempre excusar su incapacidad para expresarse correctamente en el no demasiado suficientemente el inglés.

¹⁸ Vid Newell, A. Shaw J.C. y Simon H.A., *Empirical Explorations of the Logic Theory Machine in Western Computer Proceedings* (1957), pp. 216-230.

¹⁹ Vid Newell, A. Shaw J.C. y Simon H.A., *Report on a General Problem-Solving Program in Proceedings of the International Conference on Information Processing*, pp. 256-264.

resolver toda clase de dilemas, como el de los misioneros y los canibales y otros similares. No obstante, la limitada capacidad de los computadores de la época mostraron los límites de tal tipo de problemas prontamente, y aunque se podía trabajar con dilemas del estilo de la teoría de juegos, no se podía solucionar con ellos ningún problema práctico de importancia.

Otra iniciativa importante fue el inicio de la traducción computacional. Durante la década de 1950 Estados Unidos se encontraba en medio de su carrera armamentística y vital con la Unión Soviética, por lo que leer los comunicados rusos era una cuestión de vital importancia. Para ello se intentó automatizar la traducción de documentos. El lenguaje había sido reducido a un conjunto de reglas lógicas mediante la gramática universal de Noam Chomsky, por lo que la idea de aplicar un conjunto de reglas lógicas a un texto a fin de desentrañar su significado parecía posible. Rápidamente hubo éxitos, e incluso se anunció que los traductores tenían sus días contados. Nos obstante, en la práctica, esto no fue posible, y las máquinas fueron incapaces de construir traducciones suficientemente coherentes. De hecho, en 1966 se abandonó la idea. El principal problema era la llamada explosión combinatoria. Los modelos matemáticos al uso estaban diseñados para aplicar de manera constante unas tres o cuatro reglas, pero si las reglas se aumentan, con cada nueva regla se multiplican las nuevas posibles soluciones. En pocas palabras, con cada nuevo factor en la entrada se produce una multiplicación de posibles soluciones en la salida, lo que hace que el problema no sea computable (o que su computación requiera una capacidad de cálculo inmensa). Imaginemos una simple clasificación, donde se quiere saber si un objeto pertenece a una clase determinada. Por ejemplo, una pera, que se debe determinar si es o no una fruta. Si se tiene un criterio de clasificación (su nacimiento de árboles = q) y se quiere saber si un objeto pertenece a esa clase, hay dos posibilidades, sí o no (q - q). Si se agrega un segundo factor de clasificación más, el número de soluciones se eleva al cuadrado. Así, si decimos que debe nacer de un árbol y ser redonda, tenemos cuatro soluciones posibles. Si agregamos otro factor, se eleva el número de soluciones al cubo y así sucesivamente llegando rápidamente con cada nuevo factor a un número de soluciones cercano al infinito. La computadora deberá explorar cada una de las posibilidades antes de dar una respuesta, pero los seres humanos no. Nosotros razonamos de otra manera y no es claro porqué. El lenguaje tiene muchísimas reglas y excepciones, por lo que si intentamos reducir a un sistema lógico computable su traducción, rápidamente las traducciones serán imperfectas o imposibles.

En esta misma época, los viejos tubos de vacío comienzan a quedar atrás cuando los transistores comienzan a ser utilizados para construir computadoras más poderosas y fiables. Básicamente, un transistor ejecuta la misma función de computación que los tubos de vacío, con una señal de entrada y otra de salida, pero ocupando mucho menos espacio, con menor fragilidad y consumiendo menos energía. Esto determinó una importante reducción en el tamaño de los ordenadores y un fuerte incremento en su potencia.

En el verano del año 1956 al calor de los primeros y espectaculares frutos de la informática, nació formalmente la disciplina de la inteligencia artificial. En Dartmouth se reunieron varios de los fundadores de la disciplina, entre ellos Marvin Minsky, Nathaniel Rochester, John McCarthy y Claude Shannon, a establecer el marco y objetivos de la investigación en máquinas que pudiesen pensar. Su propuesta daba por primera vez un nombre a la línea de investigación: inteligencia artificial²⁰. Planteaba la propuesta todo aspecto del aprendizaje o de cualquier otro aspecto de la inteligencia puede, en principio, ser descrito con la suficiente precisión para que una máquina pueda simularlo. En este sentido, se pretendía entender como las máquinas podían usar el lenguaje, formar abstracciones y conceptos, resolver problemas hasta ahora reservados a seres humanos y mejorar su propio rendimiento. Entre los objetivos a desarrollar se encontraban la simulación de redes neuronales, el desarrollo de abstracciones y la creatividad. Es curioso observar que en la propuesta de Dartmouth ya estaban descritas las áreas más importantes de la inteligencia artificial, e incluso las líneas a través de las cuales se desarrollarían en el futuro sus dos escuelas más importantes, la simbolista y la conectivista.

Hasta ahora, todo el desarrollo de la computación se ha centrado en la transformación de proposiciones lógicas en proposiciones aritméticas computables. Básicamente, la computación parece un derivado de la lógica proposicional, donde cualquier problema es soluble en la medida que pueda ser transformado en un lenguaje que determine instrucciones para la máquina, la cual, en último término funciona con una estructura matemática binaria de ceros y unos. Así, la máquina no razona, sino que depende de razonamientos pre-programados. En este sentido, la máquina tiene una capacidad limitada a la cantidad de proposiciones lógicas que hayan sido programadas en su interior. Puede hacer deducciones, toda vez que éstas están implícitas en las proposiciones ya presentes, más no induce. En cierto sentido, la computación estaba limitada a realizar una y otra vez tautologías donde, en el mejor de los casos, explicitaría elementos que ya estaban implícitos en las proposiciones pre-establecidas, más no generaría nuevas inferencias. Una parte importante de los arquitectos de la inteligencia artificial estaban convencidos que a través de este método se podía dotar a las máquinas de los elementos necesarios para razonar y para desarrollar las funciones que usualmente estaban reservadas a seres humanos. De alguna manera, la computación, de acuerdo a esta teoría, se asemejaba a la idea platónica de los pre-conceptos. En efecto, Platón plantea artísticamente en su libro VII de la República que todo lo que conocemos lo sabemos porque tenemos pre-conceptos sobre ello. Así, reconocemos en una forma circular el círculo perfecto que tenemos inmanente

²⁰ McCarthy, J., Minsky, M.L., Rochester, N. y Shannon, A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence (1956) en: <https://web.archive.org/web/20080930164306/http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>
Consultado el 17 de mayo de 2019.

en nuestro espíritu. De la misma manera, al introducir proposiciones lógicas en una máquina, ésta podría aproximarse a la realidad a través de tales axiomas e interpretarla. Este es el centro de la escuela simbolista, que creía que el mejor camino para construir una inteligencia artificial radicaba en introducir un conjunto amplio de proposiciones lógicas que llevaran, finalmente, a la máquina a razonar

De la misma manera que la máquina jugaba al ajedrez en base a muchas proposiciones que determinaban su conducta ante el movimiento de las piezas, en último término podría también representar el mundo a través de un número más alto y complejo de proposiciones. No obstante, aquí hay un grave inconveniente. Los seres humanos no juegan al ajedrez de esta forma, y aunque nosotros podemos razonar formalmente a través del uso de proposiciones lógicas o matemáticas, esto no constituye el centro de nuestros razonamientos ni describe nuestro mundo intramental. Nosotros creamos las proposiciones, y no nos limitamos a aplicarlas. De alguna manera, las máquinas que juegan al ajedrez se parecen al viejo Turco Mecánico, sólo que ahora el enano introduce proposiciones para jugar en lugar de mover las piezas el mismo.

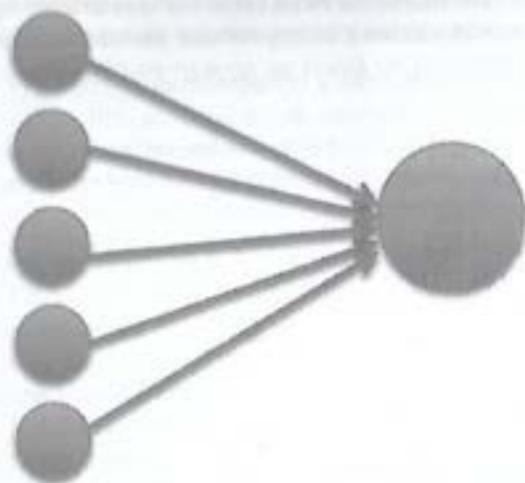
En este sentido, tiene importancia el surgimiento de la segunda escuela relativa a la inteligencia artificial, la conectivista. Su emergencia ya está implícita en la propuesta de Dartmouth, cuando se plantea como uno de los objetivos el estudio de redes neuronales, aunque su desarrollo es algo posterior y se debe al genio de un hombre que no participó en ella, Frank Rosenblatt.

La idea proviene del estudio del cerebro humano y de las nociones que se tenían de él en 1950. En ese entonces se sabía que el cerebro estaba compuesto por una miríada de células llamadas neuronas, y que cada neurona establecía conexiones con otras, las que se veían reforzadas o debilitadas en razón de su efectividad y uso mediante un mecanismo de recompensa a través de la dopamina. La idea nace de ahí y es un intento de imitar —con un alto grado de laxitud— el funcionamiento del cerebro.

La idea es construir una red simple de neuronas llamadas perceptrones. Hay una red asociada al ingreso de información (*input*) y otra al resultado obtenido de esta (*output*), donde se establece una conexión entre ambas redes, otorgando un valor a cada una de las conexiones. Este valor reproduce la fuerza de la sinapsis cerebral, otorgándole a cada conexión un peso (*weight*) específico. Cada neurona da un resultado lineal, es decir, si el valor que recibe es superior al peso especificado, dará como respuesta un sí, si es menor, un no.

Supongamos, por ejemplo, que queremos determinar si un objeto es un cuadrado. Para ello usaremos cinco perceptrones en la entrada y uno en la salida. El primer perceptrón simplemente determinará si la figura tiene cuatro lados, el segundo si el ángulo superior derecho tiene 90 grados, el tercero si el superior izquierdo tiene 90 grados, el cuarto si el inferior derecho los tiene y el quinto si el inferior izquierdo. Asignamos un peso (*w*) de 0,25 a cada determinación. Si la figura tiene cuatro lados y cada ángulo tiene 90 grados, el perceptrón de salida

tendrá un valor de uno ($x=1$) y, por tanto, dirá que es un cuadrado. En caso contrario, dirá que no²¹.



Este perceptrón podrá determinar, ante cualquier figura que se le presente, si es que esta pertenece a la categoría de los cuadrados. Complejizándolo un poco, esto es, agregándole más perceptrones de salida y de entrada, determinar si es un triángulo o un cuadrado, o si es un círculo, o un pentágono. En pocas palabras, a través de distintas unidades de entrada y conexiones hacia las salidas podemos realizar determinaciones respecto de la realidad. De hecho, lo más importante a la hora de realizar las clasificaciones es, saber qué peso dar a cada conexión, y este puede ser corregido después de un mecanismo de prueba y error.

El artículo original²², da una descripción acerca del funcionamiento del cerebro y propone su conversión en una máquina que funcione en base a este algoritmo. Se da una descripción matemática basada en las matrices y, básicamente, cada perceptrón da un valor que se multiplica por su peso, y debe ser sumado a la hora de determinar la salida. Resumiendo el problema, podríamos decir que Rosenblatt da las bases para que las máquinas se relacionen con el mundo y puedan inferir resultados de él. Según sean la percepciones de la máquina, así serán los posibles resultados que se obtengan.

Los perceptrones fueron implementados en máquinas especialmente construidas al efecto y tuvieron resultados bastante interesantes. Al ser relacionados con primitivas cámaras digitales de cadmio y foto sensores para determinar pe-

$$f(x) = \begin{cases} 1 & \text{si } w \cdot x = 1 \\ 0 & \text{en otro caso} \end{cases}$$

²¹ Rosenblatt, Frank, The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain, en *Psychological Review* 65-6 (1958) pp. 386-408.

tos, lograron reconocer no sólo figuras geométricas, sino incluso determinar, con cierto grado de confianza, cuando habla un hombre o una mujer ante la cámara, o reconocer escritura. Su éxito determinó el primer verano de la Inteligencia Artificial, y avanzando las investigaciones en nuevas direcciones.

Las aproximaciones simbólicas y conectivistas coexistieron durante los años finales de los cincuenta y comienzos de los sesenta profundizando sus posibilidades. Una de las más interesantes fue la implementación de los primeros robots industriales en 1959 General Motors introdujo su primer robot industrial en una línea de ensamblaje, el UNIMATE.

A los pocos años, en 1964, Eliza, el primer programa de conversación que se comunicaba en lenguaje natural apareció. Se trataba de una suerte de primitivo *chatbot* que pretendía dar atención psicológica a los pacientes que se relacionaran con él a través de una conversación escrita. Lo más notable de Eliza es que lograba desarrollar un aparente diálogo lo suficientemente natural como para que durante unos minutos la impresión de tratar con persona real. En verdad se trataba de un conjunto de instrucciones que daban respuestas pre programadas a las preguntas formuladas. En pocas palabras, era un logro de la escuela simbólica, aunque no pasaba de ser un buen truco. En todo caso, durante unos momentos, pareció que el famoso test de Turing estaba a punto de ser superado.

En esta misma época se desarrollaron programas que eran capaces de vocalizar sus respuestas y que, incluso, hicieron apariciones estelares en la televisión. El mundo parecía al borde de recibir la primera inteligencia artificial, y la expectante mirada del público se centraba en estos esfuerzos. *Lost in Space*, *Los Supersónicos* o *2001 Space Odyssey* son manifestaciones culturales de tales expectativas. El surgimiento de una inteligencia artificial y el posible desplazamiento masivo de trabajadores desde la industria por parte de robots hace temer que pronto sobrevenga una alteración del tejido productivo de dimensiones épicas. Esto llega al punto que un grupo de intelectuales ligados a la Inteligencia Artificial dirige al presidente Lyndon Johnson en 1964 un informe en el que se habla de los peligros que la triple revolución que la humanidad enfrenta puede causar (la revolución de la cibernética, de las armas de destrucción masiva y la emergencia de los derechos humanos)²³. La perspectiva central del informe es que la aparición de máquinas inteligentes creará un desempleo tecnológico masivo, en que la mayor parte de la fuerza de trabajo se hará simplemente imposible de emplear. Ante dicho peligro, el informe aboga por la creación de un ingreso básico universal. Como bien sabemos, este riesgo que muchos intelectuales preveían no llegó a materializarse, toda vez que los robots se encontraban aún en su infancia y la inteligencia artificial resultó ser imposible en la época planteada. Muchos pien-

²³ Véase el informe completo en Pauling, Linus et alia, *The Triple Revolution*, Santa Barbara, California (1964) disponible en: http://scarc.library.oregonstate.edu/coll/pauling/papers/papers/1964p_7-01.html Consultado el 19 de mayo de 2019.

en, en todo caso, que el comité no estaba equivocado, sino que sólo adelantado a su tiempo²⁴. En todo caso, Johnson procedió a nombrar por ley una comisión que estudiase el problema, la *National Commission on Technology, Automation, and Economic Progress*, donde figuran varios premios Nobel de economía. Las conclusiones de la comisión es que no se prevé un desempleo tecnológico masivo para los próximos años²⁵.

En esta época Japón construye sus primeros robots industriales, entre los que destaca el Kawasaki-Unimate 2000, de 1968, que se integrará masivamente en las líneas de producción. Para 1973 se construye el primer androide japonés el WABOT 1 (Waseda Robot, por la universidad de Waseda), que tenía algunas capacidades notables, como caminar, determinar distancias y conversar en japonés.

Para mediados de la década del sesenta, era evidente que las dos escuelas relativas a la inteligencia artificial se habían distanciado y parecían representar dos aproximaciones filosóficas al problema de fondo relativo a la inteligencia. Por una parte, los simbolistas insistían en codificar proposiciones lógicas en programas cada vez más extensos y complejos que pudiesen prever las circunstancias reales del mundo. En cierto sentido, su premisa consistía en que pensar es razonar formalmente, y puesto que dicho razonamiento formal podía ser expresado en términos computables, la construcción de la máquina pensante era íntimamente dependiente de la capacidad de codificar ideas, las que serían verdaderos pre-conceptos de la inteligencia artificial. Por su parte, los conectivistas sostenían que lo relevante es la capacidad de aprender, de formar correlaciones con el medio. Las ideas y abstracciones surgirían del proceso de aprendizaje en sí, más adelante, cuando la máquina tuviese redes suficientemente densas. A su favor alegaban que seguían el único modelo de inteligencia natural con que contamos, el propio cerebro humano. Las relaciones entre ambas escuelas parecían rotas. Ray Kurzweil relata su encuentro con los líderes intelectuales de ambas escuelas, Minsky y Rosenblatt, señalando que el primero le había indicado que el trabajo del segundo era una mera pérdida de tiempo²⁶.

Hasta este momento, codificar las proposiciones lógicas que un computador era una labor bastante ardua. El lenguaje binario que las máquinas necesitaban para funcionar eran sumamente complejo y su expresión en tarjetas pre picadas y cintas magnéticas era difícil. Fue en la Universidad de Dartmouth donde esta realidad cambió definitivamente con la creación del BASIC (*Beginner's All-purpose Symbolic Instruction Code*) en 1964. Este era una suerte de metalenguaje que permitía al programador codificar sus instrucciones siguiendo proposiciones lógicas simples y lingüísticamente cercanas al habla, las que, posteriormente,

²⁴ Vid Ford, Martin, *Rise of the Robots* (Basic Books, 2015), pp. 29-30.

²⁵ AAVV, *Report of the Commission on Technology, Automation, and Economic Progress* (1968), p. 114-115.

²⁶ Vid Ford, Martin, *Architects of Intelligence*, (Packt, Birmingham, 2018) p. 228.

eran traducidas por la máquina a lenguaje binario. Su éxito fue pasmoso, hasta el punto que para fines de la década de 1960 había casi cinco millones de programadores en Basic, logrando a convertirse en el lenguaje computacional universal cuando fue incorporado en Altair 8080 en 1974, la primera computadora diseñada por Bill Gates y Paul Allen, que pronto fundarían Microsoft. Rápidamente se expandió a todos los nuevos computadores personales que surgieron durante la segunda parte de la década de 1970, como el Apple II o el IBM-PC, convirtiéndose en el lenguaje de programación global, acercando, por primera vez, a quienes no eran expertos en matemáticas a la naciente computación. Millones de personas se lanzaron a programar y la informática alcanzó, por primera vez, impacto planetario.

El surgimiento de Basic, implicó la eferescencia de la programación, lo que llevó a millones de personas a expresar proposiciones lógicas en programas que satisfacerían las necesidades de los usuarios. Detrás de este éxito, las premisas de la escuela simbólica se encontraban respaldadas, toda vez que estos programas, aunque no llevaban a ninguna inteligencia artificial apreciable, sí hacían de la computación una disciplina útil. Pronto surgieron nuevos lenguajes similares que ocuparon su lugar entre los programadores²⁷, pero de ahora en adelante programar ya no sería un dominio cerrado y el público general pudo incorporarse a este nuevo mundo.

Mientras tanto, la escuela conectivista experimentó un verdadero naufragio cuando Marvin Minsky y Seymour Papert refutaron matemáticamente la posibilidad de desarrollo de los perceptrones²⁸, toda vez que les resultaba imposible el reconocimiento de la función lógica de la disyunción exclusiva²⁹. En definitiva, lo que hizo el escrito fue sepultar las esperanzas de la escuela conectivista y relegar al olvido los estudios relativos a los perceptrones por los próximos veinte años. Tras la crítica de Minsky y Papert se encontraba un hecho no expresado pero que se deduce de su trabajo, y es que los sistemas construidos a partir de redes neuronales no son capaces de comprender la abstracción, ni operar con conceptos. Los perceptrones pueden determinar que un objeto es un círculo con un nivel de certeza cada vez mayor, pero no comprenden qué es un círculo ni podrían sacar su radio. Esto sí es perfectamente posible, en cambio, a través de representaciones lógicas, que es como opera la escuela simbólica.

²⁷ El más popular es C, construido por Brian Kernighan y Dennis Ritchie en 1978, de donde viene, por lo demás, la costumbre de los programadores de hacer un primer programa que diga la frase "Hello, world!" al aprender a programar. Vid Kernighan, B y Ritchie, D., *The C Programming Language* (Prentice Hall, 1978). De este lenguaje derivan C++, Python y la mayor parte de los lenguajes más usados en el mundo.

²⁸ Véase: Minsky, M. y Papert, S., *Perceptrons: An introduction to computational Geometry* (MIT Press, Cambridge, 1969).

²⁹ Esta, representada en inglés como XOR consiste en que sólo una de las proposiciones es válida, y que cuando una es válida la otra debe ser falsa. Se la representa tradicionalmente con el conector "v" y su aplicación más importante es el Teorema de Morgan.

Resulta extraño que tanto Papert como Minsky tenían aprecio por el trabajo de Rosenblatt. Seymour Papert, de hecho, intervino en un famoso debate entre Jean Piaget y Noam Chomsky acerca del rol del aprendizaje en el desarrollo del lenguaje, apoyando las posturas de Piaget, que claramente parecen más cercanas al conectivismo³⁰. Minsky, por su parte, en sus últimas entrevistas³¹, también reconoce las virtudes de las redes neuronales desarrolladas por su amigo de adolescencia Frank Rosenblatt.

A comienzos de la década del setenta finalmente llegó el primer "invierno" de la inteligencia artificial. La expresión pertenece a Minski y se refiere a un periodo en la que disciplina se hizo impopular y el financiamiento para ella desapareció. Tan duro fue el golpe, que muchos de los investigadores vinculados a ella debieron disfrazar sus publicaciones a fin que estas no fuesen rechazadas en las revistas de investigación. Este invierno suele explicarse por las constantes promesas de los cincuenta y sesenta acerca del desarrollo de una inteligencia artificial que no se materializaron en desarrollos reales, sino en simples juegos de lógica que en nada se parecen a una mente artificial, junto con la bancarrota epistemológica de los perceptrones. Además de esto, había una crítica de fondo a la posibilidad de lograr una verdadera inteligencia artificial en la famosa explosión combinatoria. Este es un fenómeno que amenazaba cualquier desarrollo computacional, toda vez que en cualquier medio no reglamentado las posibilidades se multiplican con suma facilidad. La vida tiene muchas variables, por lo que las posibilidades de que éstas lleven a un número casi infinito de soluciones posibles es muy real. No sabemos por qué los seres humanos no están afectados lógicamente por tal explosión combinatoria. Aparentemente aquello que llamamos sentido común limita las posibilidades a un número manejable, pero no sabemos cómo opera. Un ejemplo evidente es el dilema del vendedor viajero. Si un vendedor viajero tiene que visitar diez ciudades en un periplo, a primera vista un ser humano encontrará en un mapa una ruta posible y coherente. Pero para una máquina, esto implica un número de posibilidades superior a ciento ochenta mil, y un cálculo fatigoso. Si aumentamos el número de ciudades, rápidamente se llega a un número de posibilidades superior al número de átomos en el universo conocido. Es decir, el problema se vuelve imposible para una máquina, mientras que para un ser humano no es en absoluto complejo. La explosión combinatoria fue planteada como el principal problema de la inteligencia artificial en un debate en la Royal Society que prácticamente sepultó el área de estudios durante la próxima década.

³⁰ Pistelli-Palmanni, M., *Language and Learning: The Debate between Jean Piaget and Noam Chomsky* (Routledge, London, 1979). Curiosamente, este libro sirvió de inspiración a Yann Lecun para su propio desarrollo intelectual relativo a las redes neuronales que, eventualmente resucitaría a la escuela. Vid Ford, Martin, *Architects of Intelligence*, (Packt, Birmingham, 2018) p. 121.

³¹ Véase la entrevista de Ray Kurzweil a Marvin Minsky en: <https://www.youtube.com/watch?v=RZ3ahBm3dCk>. Consultada el 20 de mayo de 2019.

En este período conocido como invierno de la inteligencia artificial, hubo desarrollos notables en la informática en general. Se crearon los computadores personales, la programación se hizo simple y se expandió globalmente, y apareció ARPANET, la futura internet. A comienzos de la década de 1980, la inteligencia artificial volvía con fuerza al imaginario colectivo. Esta es la época de los llamados sistemas expertos. La idea es netamente simbolista, lograr programar reglas lógicas en una máquina de manera que pueda enfrentar una cantidad de casos aplicando dichas reglas y solucionando problemas reales. Esta es una nueva aplicación de los viejos parámetros de la década de 1950, pero con dos diferencias importantes. En primer lugar, el lenguaje para codificar las proposiciones (en esa época, generalmente el Basic), era generalmente accesible al cualquier persona con talento, y la potencia de las máquinas se había multiplicado. Este es el momento para explicar la famosa ley de Moore.

En 1965 Gordon Moore, cofundador de Intel, debía realizar una predicción acerca de la capacidad probable de los ordenadores que se desarrollarían en los próximos años. Para hacerlo, simplemente miró las estadísticas de los avances en el poder de computación en la última década (1955-1965) y concluyó que, estadísticamente, la capacidad de computación por dólar gastado se duplica cada dos años. Esto implica que en 1957, un computador que costara lo mismo que en 1955 era el doble de poderoso (o que uno equivalente a 1955, costaría la mitad en 1957). Si tomamos en cuenta el efecto combinatorio de la multiplicación, esto implica que para 1985, una computadora con la capacidad de procesamiento de 1955 (un UNIVAC) que costaba un millón de dólares, costaría aproximadamente treinta dólares (un ZX spectrum). Sorprendentemente este cálculo estadístico es correcto. Incluso se ha argumentado que la duplicación ocurre cada dieciocho meses, por lo que la aceleración es más pronunciada aún. Nadie sabe por qué esta intuición de Moore es correcta, sólo se sabe que funciona y que posibilitará el desarrollo de la informática hasta hoy.

Con nuevas posibilidades de procesamiento de datos y medios más simples para programar, los sistemas expertos irrumpieron en la industria con la finalidad de transferir a medios informáticos las capacidades de los profesionales más peritos en determinadas disciplinas. Se construyeron muchos sistemas de esta naturaleza, incluidos varios sistemas cuya finalidad era desplazar a los abogados³². Algunos de estos sistemas eran capaces de proponer argumentos para distintas situaciones, explicar argumentos y establecer sistemas lógicos de aplicación de normas. Tal vez el más notable sea Carneades³³, que podía argumentar tanto desde normas como desde casos. A pesar de la eficiencia de este tipo de sistemas para manejar conceptos lógicos, verdaderamente su impacto en el Derecho fue menor, toda vez que el costo de reducir las normas a reglas es relativamente

alto. Es necesario contar con un buen jurista con mucho tiempo libre, experto en lógica y capaz de programar que reduzca a proposiciones lógicas las normas jurídicas³⁴. Por otro lado, las conversiones que haga el jurista a predicados lógicos no serán, necesariamente, válidas, toda vez que son dependientes de su interpretación personal de la norma, la cual puede estar errada³⁵. Es por ello que se requiere que quien realice las conversiones sea una suerte de hegemon interpretativo. Las interpretaciones normativas deberían ser consideradas siempre válidas y, en este sentido, la única instancia que tiene el poder suficiente como para realizar estas interpretaciones generalmente obligatorias sería el propio poder legislativo (art 3° CC). De hecho, en la década de 1980 Grayfred Gray logró que la legislatura de Tennessee tomase la decisión de redactar una versión normalizada de las disposiciones que aprobase en un sistema de notaciones lógicas apropiado³⁶, aunque en la práctica esto se realizó en muy pocas ocasiones. En fin, las dificultades propias de este método de aproximarse al Derecho lo hicieron poco práctico, excepto en áreas muy limitadas donde el coste de la programación fuese vastamente inferior a las ventajas que esta reportase. Un caso paradigmático de esto es la declaración de impuesto a la renta, donde hoy en día cientos de miles de personas dejaron de realizar manualmente sus declaraciones y optan por aquella propuesta por el Servicio de Impuestos Internos (que actúa como hegemon interpretativo) a través de un sistema experto. En pocas palabras, los resultados, aunque muchas veces positivos, fueron en general modestos en comparación a las expectativas.

Otra dificultad evidente para una aplicación general de estos sistemas expertos fue que carecían de sentido común. El sistema sólo contiene las proposiciones lógicas con que sea programado, por lo que puede realizar predicciones o proponer argumentos sólo respecto a las proposiciones que estén integradas en su programación. Así, si por ejemplo, redujésemos a proposiciones lógicas el título 35 del libro 4 del Código Civil, de los delitos y cuasidelitos, el sistema podrá determinar apropiadamente la distribución de responsabilidad en caso de daños, pero al carecer de sentido común, no podrá establecer si es que el hecho X (una fuga de agua) causó daños a Y (propietario del inmueble vecino), puesto que no sabe que el agua es húmeda y que la humedad causa daños en los bienes. Es decir, sin sentido común, no es capaz de subsumir el caso en el tipo, y este es uno de los centros interpretativos más importantes del Derecho. En pocas pala-

³² Como parte de mi doctorado, en una ocasión hice este trabajo para afinar mis estudios de lógica. Reduje a proposiciones todo un título del Digesto. El resultado fue extraño, puesto que habían varios temas donde parecía evidente que existían contradicciones lógicas en el Digesto, pero estas no eran un tema importante ni para los juristas del pasado ni para los romanistas. Otros temas donde no aparecían contradicciones en las fuentes eran mucho más discutidos y relevantes. My vision final del problema fue que la lógica proposicional al no era el único elemento con el que argumentaban los juristas y ni siquiera el más importante.

³³ Ashley, K., Artificial Intelligence and Legal Analytics. (Cambridge University Press, Cambridge 2017), p. 45.

³⁴ Gray, G., Statutes enacted in normalized form: the legislative experience in Tennessee in Computer Power and Legal Reasoning (West Publishing, 1985), pp. 479-480.

³¹ Para una mirada general acerca de estos modelos ver: Ashley, K., Artificial Intelligence and Legal Analytics. (Cambridge University Press, Cambridge 2017), p. 8.

³² Acerca de este sistema experto, ver Ashley, K., Artificial Intelligence and Legal Analytics. (Cambridge University Press, Cambridge 2017), p. 129.

bras, el dilema de los sistemas expertos es que, a menos que se desarrolle un sentido común racionalizado en reglas, dependen de un ser humano para una de las operaciones más críticas del Derecho, determinar si el caso se adapta al tipo establecido en la norma.

A fin de superar esta barrera es que en 1984 Doug Lenat comenzó el proyecto CYC. Este es un proyecto académico masivo que se ha ido desarrollando durante décadas en que un grupo de investigadores se ha dedicado a introducir proposiciones lógicas a un sistema que intenta reproducir el sentido común de un ser humano. Así, contiene proposiciones del tipo el agua causa humedad, o la humedad daña una casa. Al ser un sistema simbólico, puede, además, correlacionar ambas proposiciones y determinar que una fuga de agua daña la propiedad del vecino y realizar la subsunción del caso en la norma. Para esto, incluso, se debió inventar un lenguaje lógico apropiado a este tipo de representaciones (Cycl). No obstante, luego de décadas de este tipo de trabajo, el sistema se ha hecho inestable. Contiene tantas proposiciones que termina por colapsar cada vez que una nueva proposición entra en contradicción con las previamente establecidas. En pocas palabras, el sistema es tan complejo que sus probabilidades de fallar son enormes.

No obstante sus limitaciones, los sistemas expertos lograron hacer frente a algunos de los desafíos que la inteligencia artificial se planteó durante la década de 1950. En efecto, los programas de juegos se hicieron cada vez más competentes, hasta que en 1997 Deep Blue logró derrotar a Gary Kasparov en ajedrez, aunque lo que Deep Blue hacía no se parecía en nada a lo que un ser humano hace enfrentado al juego y su derrota fue un ejercicio de fuerza bruta computacional frente al genio del gran maestro ruso. Durante la década de 1980 aparecieron también muchos sistemas de traducción, todos imprecisos, pero suficientes para ser comercializados y tener algún modesto grado de utilidad. No obstante, los traductores profesionales no vieron sus puestos de trabajo en peligro, como se había pronosticado. La automatización fabril también tuvo un desarrollo destacado en la década, aunque el verdadero desplazamiento de puestos de trabajo asalariados en las economías desarrolladas vino de la mano de la globalización, y no tanto de la automatización, como se temía. Finalmente, en 1977 incluso surgieron automóviles que se conducían solos, con los cuales ya se había experimentado en la década de 1960³⁷. No obstante, dichos automóviles tenían el problema de que dependían de las reglas que se programara en su interior, las cuales debían contemplar todas las situaciones posibles. En la práctica la vida es tan compleja que esto lleva a una superabundancia de reglas que termina en una explosión combinatoria al poco andar. En pocas palabras, los automóviles auto conducidos de la década de 1980 podían funcionar en medios altamente

³⁷ Se trata del vehículo japonés creado por S. Tsugawa de *Tsukuba Mechanical Engineering Laboratory* que podía andar a 30 kmh. Para una historia de este y otros vehículos véase: Bhat, Aditya, *Autonomous Vehicles: A perspective of past and future trends in International Journal of Engineering Technology Science and Research* 4-10 (2017) pp. 1183-1190.

controlados como una carretera en condiciones de tráfico óptimas, pero no en la vida real.

Fue así como se llegó a un nuevo invierno para la inteligencia artificial, cuando se hizo evidente que los caminos tomados por los investigadores del ramo no iban a rendir, nuevamente, los frutos esperados. En efecto, la inteligencia artificial pasó de moda y otros fenómenos relativos a la informática entraron en auge, como Internet. Este será otro período en que los investigadores del tema posicionarán sus investigaciones bajo nombres curiosos y toda la disciplina pasará a convertirse en un fenómeno subterráneo.

Curiosamente, en esta misma época es que la escuela conectivista, que todos daban por muerta, tendrá un resurgir tan importante que ha llegado a casi sepultar a los simbolistas. Este fenómeno está ligado a varios investigadores que permanecieron en los sótanos de la informática y la psicología durante toda la década de 1980. Nos referimos a David Rumelhart, Geoffrey Hinton, Yoshua Bengio y Yan Lecun.

Poco después de la publicación de Minsky y Papert sobre los perceptrones, Rosenblatt murió en un accidente náutico, por lo que la disciplina ligada a los perceptrones quedó carente de líderes intelectuales y casi olvidada. David Rumelhart era un profesor de psicología matemática en UCLA, quien en 1986 publicó, junto a Hinton y Roland Williams, un artículo que cambiaría la investigación en inteligencia artificial³⁸. El artículo toma la idea de Rosenblatt de construir redes neuronales, agregándole la posibilidad de establecer capas ocultas a las mismas. Es decir, en lugar de sólo haber una conexión entre los perceptrones de entrada (*input*) y salida (*output*), agregar una serie de capas intermedias que vayan, a contar de las posibilidades establecidas en las capas anteriores, esclareciendo las posibilidades posteriores. Así, por ejemplo, si en la red de entrada hay tres perceptrones que determinan la presencia de píxeles encendidos en ciertos sectores de una imagen, en la segunda capa podría determinarse si esos píxeles tienen una forma determinada, para en una tercera capa conjugar las formas en torno a las orejas y ojos de un animal, que en la cuarta capa (la de salida), establecerían si en la foto hay un gato en el centro.

Amén de ello, se estableció un mecanismo automático para la corrección de los pesos relativos a las conexiones conocido como retro propagación (*back-propagation*). Cada conexión tiene un peso asignado, de manera que si el resultado de ella es positivo, este resultado se multiplica por el peso de la conexión y da un resultado que sirve para alimentar a la próxima capa de la red. Si la imagen efectivamente corresponde a un gato y la red acierta, el peso de las conexiones que intervinieron en la determinación se ve confirmado. Si la red yerra, la retro propagación procede a corregir los pesos que intervinieron en la determinación para

³⁸ Rumelhart, D., Hinton, G. y Williams, R. *Learning Representations by Back-Propagating Errors* en *Nature* 323 (1986) pp. 533-536.

que, en una nueva ocasión, den un resultado correcto³⁹. Si se alimenta a la red con muchos ejemplos, la red será cada vez más precisa en su reconocimiento. La eficiencia de la red depende de la disponibilidad de datos y sólo tendrá buenos resultados en la medida en que sea entrenada, miles de veces, con abundantes datos que estén etiquetados correctamente.

Lo más interesante del mecanismo planteado es que no requiere de programación explícita de reglas. Basta con diseñar adecuadamente el algoritmo, y los pesos se pueden asignar al azar, toda vez que la retro propagación los corregirá eventualmente. Por lo demás, los elementos expresamente diseñados son de nivel arquitectónico, esto es, la capa de entrada, la de salida y la presencia de capas intermedias, pero no el contenido de cada perceptrón en cada capa. De hecho, las capas intermedias se denominan ocultas porque el programador no determina su contenido, sino sólo su presencia. En 2006 el mismo Hinton publicó un artículo donde describía cómo programar una red neuronal para que reconociese los números del uno al diez escritos manualmente⁴⁰. La simpleza y elegancia del método hizo que este fuese generalmente el primer programa que un novel programador usase al construir su primera red neuronal.

Los algoritmos multicapa se denominan profundos (*deep learning*) cuando tienen más de una capa oculta y serán éstos los que determinen el nuevo auge de la inteligencia artificial.

A fines de la década de 1980 la inteligencia artificial estaba relativamente desprestigiada y la inversión en ella era escasa. Aunque se desarrollaron algunos proyectos interesantes, verdaderamente se carecía de la potencia necesaria en los ordenadores y de los datos acumulados para hacer de estos algoritmos algo funcional. No obstante, junto con la ley de Moore, el auge de Internet vino a cambiar esto de manera súbita. El auge de la *world wide web* vino a resultar en una acumulación masiva de datos de diversa índole en servidores y centros de datos pertenecientes a grandes compañías que dominaban la red. No obstante, durante la década de 1990 y la primera década del siglo XXI, aún los desarrollos en la programación estaban determinados por paradigmas simbolistas y las redes neuronales eran generalmente ignoradas a la hora de diseñar aplicaciones. En todo caso, esto cambió radicalmente a contar de la competencia de ImageNet de 2012.

ImageNet es un colosal base de datos visuales etiquetada con miles de ejemplos en toda clase de categorías. Hay decenas de miles de imágenes de ga-

tos, perros, peras o manzanas, y su objetivo era el desarrollo de reconocimiento de imágenes por ordenadores. Cada año ImageNet organiza una competencia para potenciar el desarrollo de programas de reconocimiento visual, la ImageNet Classification Challenge. En el año 2012, un equipo de estudiantes de Hinton, *Geoffrey*, utilizando la técnica de redes de neuronas convolucionales diseñada por Yann LeCun para el reconocimiento de números⁴¹, ganó el premio barriendo el piso con todos los demás competidores⁴². Se trataba de una masiva red de más de 60 mil neuronas ordenadas en cinco capas convolucionales que logró identificar correctamente las imágenes hasta límites cercanos a la capacidad humana. De la noche a la mañana, la escuela conectivista había vuelto a la palestra de la inteligencia artificial en gloria y majestad.

El efecto inmediato de este triunfo fue que la mayor parte de los laboratorios dedicados al tratamiento de datos se pasaron a la escuela conectivista y se dedicaron al desarrollo de redes neuronales, sobreviniendo un nuevo grupo de super estrellas de la inteligencia artificial. Las grandes compañías de Internet compraron laboratorios completos a precios astronómicos para transformarse en compañías basadas en la inteligencia artificial.

Pronto los éxitos de esta escuela se hicieron evidentes en diversos campos que se habían estancado en su desarrollo desde la década de 1950. Una especial mención merece Google cats. En 2012 Jeff Dean y Andrew NG, desde Google Brain, desarrollaron un proyecto masivo en que formaron una red de dieciséis mil computadores alojados en la nube de Google y los alimentaron con diez millones de imágenes digitales provenientes de YouTube. El resultado es que el masivo algoritmo logró reconocer a los gatos con mayor precisión que un ser humano. No es clara su utilidad, pero es divertido mencionarlo. Tal vez lo más interesante de Google cats fue que el algoritmo no fue programado para aprender algo específico, sino que simplemente se le mostró una ingente cantidad de datos y desarrolló por sí mismo la función de reconocer felinos, debido a la gran cantidad de videos de gatos que circulan en YouTube. En pocas palabras, con un algoritmo suficientemente complejo, éste puede determinar cosas que se fundamentan, en último término, en los datos con que sea alimentado. Si los datos contienen gatos, determinará algo acerca de ellos, si contienen perros, serán perros, si contiene guiones de cine, serán guiones⁴³, o lo que sea que se desee.

³⁹ LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., Gradient-based learning applied to document recognition in Proceedings of the IEEE 86-11 (1998), pp. 2278-2324.

⁴⁰ Acerca del proceso, véase Krizhevsky, A., Sutskever, I., Hinton, G., ImageNet Classification with Deep Convolutional Neural Networks in Advances in neural information processing systems 25-2 (2012), pp. 1-9.

⁴³ De hecho, una red neuronal fue alimentada con guiones de cine de terror y, a continuación, elaboró uno. Esta es Sunspring, una suerte de drama existencial de ciencia ficción que puede visualizarse en: <https://www.youtube.com/watch?v=LY7x2ihqmc> Visualizado el 26 de Mayo 2013.

³⁸ Puesto que examinar cada uno de los posibles pesos de las conexiones requeriría una capacidad de computación inmensa, el proceso se simplifica tomando la gradiente inversa de la integral estocásticamente, esto es, buscando en las posibilidades inmediatas el descenso con mayor pendiente. Esto es lo que se conoce como *stochastic gradient descent* (SGD), que no es más que un método aproximativo y no absoluto para optimizar un algoritmo.

Hinton, G., Salakhutdinov, R., Reducing the Dimensionality of Data with Neural Networks in Science 313 (2006) pp. 504-507.

Otros logros captaron la atención pública de manera más efectiva, entre los que merece mencionarse IBM Watson. Este programa se puede comunicar en lenguaje natural (una de las áreas de desarrollo clave para la inteligencia artificial desde la década de 1960) y fue diseñado para responder preguntas en un popular programa televisivo norteamericano llamado Jeopardy! En 2011 Watson derrotó a los campeones humanos del juego. La máquina debía entender las preguntas, obtener sus respuestas de Wikipedia, y verbalizarlas. El juego es interesante toda vez que implica un dominio idiomático superior, toda vez que en el juego hay bromas y acertijos que deben ser adecuadamente comprendidos. La arquitectura de Watson es una mezcla de deep learning y sistemas expertos, donde el dispositivo busca en bases de datos no estructuradas las respuestas a las posibles preguntas que se le formulan. Su eficiencia y estructura lo hacen ideal para usos médicos y jurídicos, y de hecho ya existen aplicaciones que usan su tecnología en esta materia, como es el caso de Ross, sobre el que volveremos en otro lugar.

Otro hito de los cincuenta fue superado con el perfeccionamiento de los sistemas de traducción disponibles. A partir de 2016 Google utiliza redes neuronales conocidas como Google Neural Machine Translation para soportar su sistema de traducción, el cual ha mejorado notablemente desde entonces. Una característica del sistema es que, al operar con redes neuronales que constantemente se perfeccionan a través de la retro propagación, cada vez que se lo utiliza el sistema mejora, y si en 2016 los resultados eran un poco mejores que cuando utilizaba sistemas expertos, hoy son notablemente superiores y, es de esperar, continúen perfeccionándose en el futuro⁴⁴.

Otro hito en el desarrollo de estas tecnologías lo dio DeepMind, actualmente perteneciente a Google. Esta compañía fue desarrollada en Inglaterra por un grupo de fanáticos de los videojuegos cuyo objetivo fue entrenar algoritmos de deep learning. Esta compañía construyó un algoritmo que fue entrenado en base a juegos de Atari. El algoritmo no fue programado para entender nada acerca de los juegos, sino que simplemente se le conectó a una consola. Después de entrenar el algoritmo durante un tiempo exuberante, era capaz de jugar a todos los juegos de una consola Atari a velocidad sobrehumana y ganarlos. Este éxito llamó la atención de Google, que adquirió la compañía en 2014. En ese momento el equipo se lanzó a intentar un desafío más ambicioso. En Asia existe un juego llamado go, que es de una complejidad enorme. El tablero es más grande que el de ajedrez y permite más tipos de jugadas que átomos existen en el universo. Su complejidad es tal que se estimaba que se necesitarían otros cincuenta años de desarrollo de potencia computacional para que una máquina derrotase al campeón mundial de go, usando las técnicas de que Deep Blue se había servido. El equipo de Deep Mind decidió construir un algoritmo al cual programó las reglas

del go y lo alimentó con las cientos de miles de partidas que había disponibles en internet⁴⁵. En 2016 este sistema, conocido como AlphaGo, derrotó al campeón mundial de go. Al poco tiempo, los ingenieros de Deepmind construyeron un nuevo algoritmo denominado AlphaGo Zero al que no alimentaron con datos, sino que lo hicieron jugar contra AlphaGo una interminable sucesión de partidas⁴⁶. El resultado es que al poco tiempo AlphaGo Zero derrotó a AlphaGo cien veces a cero. Incluso pudieron entrenar al mismo algoritmo para jugar ajedrez como un gran maestro. Este evento provocó estupefacción en oriente y determinó que Baidu y Tencent, los gigantes tecnológicos chinos, alteraran su estrategia y se convirtieran en compañías basadas en inteligencia artificial⁴⁷.

Recordando a la vieja Eliza, han surgido nuevos asistentes personales que intentan una comunicación directa en lenguaje natural con sus usuarios. Los más populares son Siri, Alexa y Cortana, los cuales, si bien no pasarían un test de Turing, pueden responder preguntas, poner música, llamar un taxi y ejecutar algunas funciones simples de domótica para el hogar. Aunque su diseño no es público, parecen tener elementos de programación simbólica y heurística, como sus predecesores, a la vez que modelos de deep learning, por lo que mejoran constantemente.

Esto nos lleva a la internet de las cosas (*Internet of Things*), que comienza a aparecer desde hace algunos años. La idea consiste en no limitar el uso de internet a los ordenadores y móviles, como se había hecho tradicionalmente, sino que integrar en toda clase de dispositivos conexiones a internet, de manera que estén permanentemente entregando datos a la red. Así, las cerraduras, refrigeradores, alarmas, y mil otros dispositivos estarían interconectándose con la red, de manera que puedan ofrecer servicios más eficientes a sus usuarios. Un refrigerador que avise a Alexa, por ejemplo, que se agotan los huevos, y ésta haga un envío a la dirección adecuada, donde al llegar el repartidor se desconecte la alarma, se abra la puerta y se cargue el refrigerador, no es un imposible técnico hoy en día, todo lo contrario, es la dirección en la cual muchas de estas compañías esperan avanzar.

Por lo demás, el diseñar modelos de deep learning se hace cada vez más sencillo. Google lanzó Tensorflow en 2015, que es el resultado de los esfuerzos de Google Brain. Consiste en una biblioteca virtual para el diseño de modelos de deep learning abierta y gratuita, donde cualquiera puede aprender a construir estos modelos sin mayores dificultades⁴⁸.

⁴⁴ Para una descripción completa del procedimiento, vid: Silver, D. et alia, Mastering the Game of Go with Deep Neural Networks and Tree Search in: <http://airesearch.com/wp-content/uploads/2016/01/deepmind-mastering-go.pdf> Consultado el 28 de Mayo de 2019.

⁴⁵ Para una descripción completa del procedimiento, vid: Silver, D. et alia, Mastering the Game of Go without Human Knowledge in Nature 550 (2017) pp. 354-359.

⁴⁶ Vid Lee, K.-F., *AI Superpowers. China, Silicon Valley and the New World Order* (Houghton, Boston-New York, 2018), pp. 81 y ss.

⁴⁷ Vid www.tensorflow.org Visitado en 28 de mayo 2019.

⁴⁴ Wu, Y. et alia, Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation, disponible en: <https://arxiv.org/pdf/1609.08144.pdf> Consultado en 23 de mayo 2019.

Prácticamente cada día aparecen nuevas aplicaciones y usos basados en *deep learning*, de manera que pareciera que la inteligencia artificial es omnipresente. Ha sido comparada con la electricidad, toda vez que se integra de manera invisible en nuestra vida inundando múltiples espacios que, hasta ahora, no le estaban dedicados. Aparentemente, podría provocar una nueva revolución industrial, lo cual ha lanzado a los gigantes tecnológicos en una carrera por desarrollar sus usos y hacerse con mercados globales. En medio de este fervor tecnológico, pareciera que estamos alcanzando las viejas metas de la década de 1950 y que pronto tendremos un C3PO caminando junto a nosotros ¿o no?

3. VIEJOS LÍMITES

Desde los ojos del lego, la tecnología es magia moderna. Hace mucho que ya no somos capaces de comprender como funcionan los aparatos que nos rodean, y por tanto, tampoco sabemos de qué son capaces ni cuáles son sus límites. Buena parte de las ideas actuales acerca de la inteligencia artificial no son otra cosa que proyecciones misticantes respecto a las aparentes posibilidades de los dispositivos que con que trabajamos. Tememos al científico desbocado que cree un Frankenstein moderno, un golem mecánico desprovisto de alma que friamente decida que no somos dignos de habitar el planeta. Sin embargo, no parece que estemos en ese predicamento aún. En la práctica, la inteligencia artificial es una rama de la informática que siempre promete una quimera, y como los malos deudores, suele pedir nuevas prórrogas cada vez que toca el tiempo de entregar el producto de su esfuerzo. A pesar de la aparente humanidad de Siri, o de las posibilidades creativas con que nos confronta Watson, verdaderamente estamos muy lejos de crear seres conscientes capaces de interactuar con nosotros. Aunque Deep Blue puede derrotarnos en el ajedrez, la máquina no sabe de su éxito. En efecto, es incluso discutible que estemos desarrollando la misma actividad, ya que nosotros pensamos y evaluamos movimientos, imaginamos y creamos, por lo que el ajedrez, en un ser humano, es una alta función de la inteligencia. Deep Blue no sabe que juega, de hecho, no sabe nada, es sólo un conjunto de instrucciones fijadas en un mecanismo magnético que se ejecutan secuencialmente según sean los datos que son ingresados a partir de los movimientos del verdadero jugador sentado ante un tablero de ajedrez. En este sentido, ni Deep Blue, ni AlphaGo, ni Siri, ni ningún otro mecanismo digital con que hoy en día contamos comparte con nosotros nuestra humanidad. No son más que algoritmos, esto es, conjuntos de instrucciones secuenciales. Deep Blue es un sofisticado Turco Mecánico.

Para los abogados la palabra algoritmo es demasiado lejana y matemática, por lo que podemos reemplazarla con una más familiar y de un valor semántico cercano, son procedimientos. Así como el Código de Procedimiento Civil fija unas instrucciones para ejecutar el juicio ordinario, digamos, y establece diversas posibilidades para que se dé paso a distintas etapas según sean los escritos que se le presenten, los computadores, según sean los distintos *inputs* que se les

entreguen, deberán realizar diversos ejercicios y entregar un *output*. Todo está preajudado y predeterminado, y según sean los escritos, será la resolución. No es raro que algunos de los matemáticos que fijaron las bases teóricas de la computación en el siglo XVII, como Leibniz⁴⁹, hayan sido abogados, toda vez que su manera de computar números es similar al proceso judicial.

Los algoritmos no son más que procesos, y en este sentido, es improbable que de ellos surja un objeto que reúna tales características que parezca un sujeto. En pocas palabras, C3PO no está a la vuelta de la esquina, ni HAL, ni ninguno de los seres soñados o temidos por la fantasía moderna. Los seres de esta naturaleza están dotados de consciencia, es decir, son capaces de integrar en su ser sus percepciones y tienen una autopercepción al respecto. El problema es que la consciencia es un fenómeno opaco para nosotros, toda vez que no sabemos que es. Yo sé que estoy consciente, toda vez que me auto percibo (como el *cogito ergo sum* cartesiano), pero no sé si Ud., lector, es consciente. Proyecto en Ud. mis propias características humanas y supongo que también es Ud. un ser consciente, pero en verdad soy incapaz de determinarlo. Es cierto que los neurólogos observan a través de escáneres las áreas del cerebro que se encuentran activas mientras estoy despierto o dormido, y también las pueden comparar con aquellas que se iluminan cuando estoy soñando y obtener algunas conclusiones, pero sigue siendo bastante opaca la consciencia en general. Si no sabemos qué es la consciencia, es difícil que logremos imitarla con medios informáticos. Algunos creen que es un problema de masa crítica, que una vez que alcancemos un nivel suficiente de interacciones lógicas o de conexiones en redes neuronales esta se generará espontáneamente. Esto es un simple pensamiento mágico. Aparentemente hasta los árboles tienen un cierto grado de consciencia y son capaces de percibir dolor, pero ellos no tienen neuronas ni son capaces de operar con postulados lógicos. La consciencia humana es aún más compleja, y reproducirla es más problemático aún.

Ante esto, la rama científica se propone un fin menos ambicioso, formar una entidad que pueda aprender cualquier cosa y ejecutarla, esto es una Inteligencia Artificial General (*Artificial General Intelligence* o AGI, según su acrónimo en inglés). Este sueño es más cercano y posible, aunque complejo de ejecutar. En principio, lo que tenemos ahora son inteligencias artificiales específicas (*Narrow Artificial Intelligence*)⁵⁰, que pueden ejecutar funciones precisas para las cuales han sido programadas o modeladas. Estas inteligencias específicas son algoritmos modelados para hacer una cosa, sólo una, como jugar al ajedrez, detectar caras, predecir palabras o lo que sea, y en algunos casos logran una precisión increíble en ese proceso. No obstante, no son capaces de realizar otras tareas

⁴⁹ En efecto, Leibniz inventó las notaciones lógicas, el cálculo infinitesimal y hasta una máquina calculadora.

⁵⁰ Al respecto véase Weinbaum D. y Veitas V. Open Ended Intelligence: the Individuation of Intelligent Agents in *Journal of Experimental and Theoretical Artificial Intelligence* 29:2 (2017), pp. 371-396.

distintas de aquellas para las cuales han sido modelados. En este sentido son estrechos, puesto que carecen de una perspectiva general, como los seres humanos, para enfrentar cualquier tipo de problema. Así, AlphaGo puede jugar juegos donde ambos jugadores tienen una visión completa del tablero, pero no, por ejemplo, Starcraft, donde sólo hay una visión parcial, ni menos determinar si en una foto hay un gato o perro o conducir un automóvil.

Hay bastantes elementos que debiesen ser enfrentados a fin de obtener una AGI que, de momento, no parecen fácilmente resolubles. En primer término está el problema de la abstracción. En principio, la hoy casi abandonada escuela simbólica se enfocó en traducir proposiciones al lenguaje de la lógica matemática a fin de poder emular razonamientos. En este sentido, un sistema experto del pasado podía imitar procedimentalmente los razonamientos de un ser humano y realizar todo tipo de inferencias deductivas. La escuela conectivista, hoy dominante, extrae de conjuntos de datos correlaciones para establecer un resultado de acuerdo a un esquema de capas de perceptrones⁵¹. El problema es que el modelo diseñado no sirve para establecer reglas abstractas, sino sólo para determinar correlaciones en los datos. Así, determinar que un gato está en una foto es distinto que determinar que el gato es un animal mamífero, y determinar esto es distinto que saber que algunos tipos son salvajes, como el colocolo o la guinea, mientras que otros son domésticos. Con un algoritmo de varias capas de profundidad, podemos establecer todas estas correlaciones, pero siempre faltará una nueva, puesto que, al no ser capaz de establecer un concepto abstracto y comprenderlo, no puede operar lógicamente con él. ¿Y qué tal combinar ambos métodos? Esta parece una solución bastante evidente, pero insuficiente. Muchos sistemas la utilizan, como Siri o Alexa, pero no terminan por cuajar. El problema radica en que los métodos simbólicos requieren que la proposición sea expresada en términos formales, pero los métodos conectivistas no son capaces de realizar esto. En efecto, las redes neuronales sólo establecen correlaciones, no reglas lógicas y ni siquiera son capaces de procesar la causalidad, un elemento esencial en el razonamiento humano. Es decir, del resultado de una red neuronal no se puede construir automáticamente una nueva proposición lógica, sino que esto deberá realizarlo un ser humano. De hecho, las máquinas no leen ni entienden los textos. Ni Watson, ni Siri, ni Alexa pueden comprender lo que está escrito en Wikipedia. Lo único que pueden hacer es determinar la frecuencia con que las palabras aparecen en un texto, por lo que, ante una pregunta determinada, buscarán la mayor coincidencia de palabras y arrojarán un resultado (o lo leerán con la voz seleccionada, en el caso de Watson o Alexa). Esto explica el éxito de Watson, puesto que la mayor parte de las preguntas de Jeopardy! coinciden con títulos de la Wikipedia.

Un ejemplo bastante simple para probar la fiabilidad de estos métodos es el siguiente: pregunte a Siri cuando nació Gabriel González Videla. Le dará automáticamente la respuesta correcta (1898). Pregunte luego, quien era presidente de Chile en 1898 y arrojará, nuevamente, la respuesta correcta (Federico Errázuriz Echaurren). Pregunte ahora, quién era presidente cuando nació Gabriel González Videla y vea qué sucede. La máquina no entiende qué es ser presidente, ni que se pacer, por lo que los resultados serán confusos.

El problema principal es que ningún algoritmo es capaz de tomar un texto, desentrañar su significado y expresarlo. Así, aunque las máquinas son capaces de determinar cuál texto es relevante cuando se introducen ciertos otros términos, básicamente por la frecuencia en que dichos vocablos aparecen repetidos dentro de cientos de miles de páginas web, el resultado arrojado por la búsqueda no implica más que un truco estadístico, no comprensión. Pueden ser muy precisas y encontrar la inmensa mayoría de elementos relevantes, pero siempre son objeto de una limitación general.

En el fondo, la comprensión humana puede ser explicada de acuerdo a diversos paradigmas filosóficos que teorizan respecto a sus capacidades. Por un lado están los preconceptos de Platón, según la cual se postularía la inmanencia de todas las ideas humanas, estando todas ellas establecidas previamente en su mente desde antes de nacer, donde el aprendizaje es una operación de reconocimiento de dichos conceptos en la realidad. La inteligencia artificial simbólica es similar al estilo platónico, donde a la máquina se le programan reglas lógicas para construir operaciones a realizar en el mundo. Aparentemente, esto no puede llevarse más allá de cierto punto, puesto que de lo contrario el sistema colapsa ante su propia complejidad.

Otra alternativa es el empirismo al estilo de Hume, donde la mente realizaría correlaciones desde sus percepciones, sin que existan la causalidad o los conceptos, los cuales no serían más que postulados probabilísticos. La regla que el sol sale cada día, no es más que el resultado de una correlación entre la salida del sol y el amanecer en sí, pero nada más, todo es experiencia, nada causa. La escuela conectivista parece acercarse a estos postulados construyendo redes neuronales que son capaces de detectar correlaciones y darles fuerza estadística a través de otorgarle un peso a cada conexión. Aunque las redes pueden establecer con precisión conexiones, al carecer de abstracción, no son capaces de generar conceptos, ni de utilizar en la resolución de nuevos problemas el conocimiento adquirido anteriormente. Un algoritmo que detecta gatos no puede jugar ajedrez, y todo el conocimiento estratégico del ajedrez es desechado al momento de aprender a jugar a las damas. Sin conceptos, no puede reutilizar la información, sin causa no puede inferir.

Una alternativa frecuentemente planteada es un modelo kantiano, donde se pre-programan algunas reglas que constituyan un conjunto de categorías a priori desde donde construir el conocimiento, y esto se conjuga con un conjunto de redes neuronales que realicen una adquisición de datos que permitan ordenar el

⁵¹ Hay muchas maneras de ordenar estos perceptrones, sea como árboles lógicos, como redes neuronales convolucionales, como redes recurrentes, etc. Esto dependerá del programador y de la tarea.

conocimiento. Los partidarios de esta idea suelen tomar como ejemplo al mundo natural, donde aparecen conocimientos innatos en los mamíferos recién nacidos, no obstante que también aprenden más tarde de su entorno, como es el caso del caballo, que apenas nacido ya sabe ponerse de pie y caminar. Lamentablemente esta idea parece poco popular y poco práctica. Generalmente los diseñadores de modelos suelen enorgullecerse de no pre programar nada⁵² y, aunque lo hagan, puesto que los algoritmos de *deep learning* no son capaces de elaborar reglas lógicas desde sus percepciones empíricas, ambos conjuntos de reglas permanecen separados. Muchas veces las reglas pre-programadas aparecen para paliar las inferencias deficitarias de las redes neuronales, pero no es posible, de momento, lograr que un algoritmo realice inferencias respecto a los datos a partir de ellas.

Lo que no existe es una máquina aristotélica, que sea capaz de elaborar conceptos a partir de los datos a que es expuesta, que de observar gatos en imágenes, elabore una idea no gráfica de los animales, que entienda qué es un gato y no sólo que sepa advertir que aquello que es percibido corresponde en un 99,9% de los casos al resultado "gato," y Dios sabe por qué estos seres humanos se preocupan por asociar cuatro letras a una imagen. De momento, nadie sabe cómo elaborar conceptos, sólo se sabe deducir a partir de ellos o manejar correlaciones estadísticas a partir de datos, pero nada más. Mientras alguien no solucione este punto, a través de algún proceso desconocido, es improbable que podamos sentarnos a tomar café con C3PO, o que el jugar ajedrez con un programa sea algo más que alguna especie de turco mecánico pre-programado o de correlaciones ciegas.

En todo caso, los modelos algorítmicos han demostrado que pueden llegar a resultados equivalentes o superiores a los humanos en ciertos dominios, como el reconocimiento de patrones o los juegos, lo cual los convierte en herramientas poderosas que han sido denominadas inteligencias artificiales estrechas o acotadas (*Narrow Artificial Intelligence*), a diferencia de las generales (AGI), que pueden resultar en una creciente automatización de diversas labores productivas hasta el punto de provocar una nueva Revolución Industrial. Un ejemplo es la conducción de automóviles. Hasta ahora, este era un dominio reservado a los seres humanos, quienes al ponerse delante de un volante utilizan sus percepciones visuales y auditivas para, en base a su representación del entorno, y teniendo en cuenta las normas del tránsito, realizan una serie de acciones mecánicas que les permiten llevar el vehículo de un punto a otro. Todas estas acciones parecen automatizables. Las percepciones pueden ser obtenidas por medio de sensores, que incluso pueden captar elementos que los seres humanos no detectamos,

⁵² Un ejemplo clásico de esto es el caso de AlphaGo, donde sus creadores ponen de manifiesto desde el título de su presentación que no han pre-programado regla alguna, aunque esto no sea del todo cierto: Silver, D. et alia, *Mastering the Game of Go without Human Knowledge in Nature* 550 (2017) pp. 354-359.

por ejemplo agregando un radar al equipo. La representación del ambiente puede ser obtenida con mapas detallados vía GPS. Las normas del tránsito pueden ser directamente programadas al dispositivo, amén de agregarse cientos de miles de horas de entrenamiento en situaciones reales o simuladas donde el modelo aprenda como reaccionar en situaciones imprevistas. El resultado puede ser una conducción más respetuosa de las reglas del tránsito y más segura que aquella realizada por un ser humano, lo cual podría salvar las cientos de miles de vidas que quedan diariamente destrozadas producto de accidentes del tránsito. Es más, aunque las inversiones necesarias para conseguir este resultado son vastas, una vez que se logre la automatización, será cuestión de transferir dicha información a un simple terminal en cada vehículo para que éste sea capaz de operar automatizadamente. Incluso vehículos que carecían originalmente de dichas capacidades podrían ser equipados con sensores y ordenadores para adquirirlos, representando los sensores el grueso de la inversión.

Evidentemente esto revolucionará nuestros sistemas de transportes y, tal sea nuestro modelo de ciudad, toda vez que los automóviles automatizados pueden estar en servicio todo el día y no necesitan mantenerse ociosos mientras el conductor realiza otras labores. El costo de transportar pasajeros, al no incluir el de conducción, sería equivalente al precio del combustible, el costo del capital invertido, la depreciación estimada y un adicional de rentabilidad (digamos un veinte por ciento). El resultado es un fracción del valor de un taxi actualmente, tal vez una tercera parte del precio actual (o menos), por lo que sería efectivamente mucho más económico tomar taxis a toda partes que ser propietario de un automóvil. Con menos vehículos circulando, las calles estarían menos congestionadas, sería necesario menos espacio para estacionar, las aceras podrían ensancharse y la contaminación ambiental disminuirse. Las ciudades se volverían más humanas y estarían menos dedicadas a los automóviles que a los peatones. Aunque las consecuencias parecen vastamente positivas, existe un problema evidente, hoy por hoy el transporte es una fuente masiva de empleos. ¿Qué pasará con todos esos trabajadores?

El caso de la conducción de vehículos es aplicable, en general, a una gran cantidad de actividades productivas. La informática, como tecnología, es esencialmente transitiva, esto es, sus capacidades pueden ser transferidas a otro equipo sin que el primero las pierda, por lo que los costos de difusión son prácticamente nulos y su presencia se hace rápidamente ubicua. Esto hace que el diseño de una solución sea caro, pero su implementación sea prácticamente gratuita. Es por eso que las compañías tecnológicas invierten ingentes sumas de dinero en llegar a soluciones viables para un problema en concreto, pero una vez alcanzada dicha alternativa, prácticamente regalan los frutos de dicho esfuerzo a un precio insignificante o nulo, rentabilizando formas laterales de ganancia, como la publicidad.

¿Qué labores son automatizables? Andrew Ng, uno de los más destacados diseñadores de inteligencias artificiales, estima que si la labor toma a un ser hu-

mano menos de un segundo, esta es automatizable⁵³. Esto deja un enorme campo abierto para la automatización, también en el mundo del Derecho. Volveremos sobre este punto más adelante, pero se estima que entre el 30 y el 50% de los trabajos que hoy existen podrían resultar automatizados dentro de los próximos veinte años, y especialmente aquellos que tratan con información, como aquellos de los médicos, contadores y abogados.

El alcance que estos métodos pueden tener es vasto, especialmente si se consideran dos tendencias que se han ido perfilando con fuerza en los últimos años, la informática en la nube (*cloud computing*) y el internet de las cosas (*Internet of Things*). La primera corresponde al hecho que buena parte de la capacidad computacional que se utiliza hoy en día no está instalada localmente en los ordenadores de los usuarios, sino en grandes servidores en la red que proveen de su capacidad de procesamiento y almacenamiento a los usuarios a través de internet. Así, puede que mi ordenador no sea particularmente poderoso o que carezca de una amplia capacidad de almacenamiento, pero puedo servirme de la capacidad de almacenamiento y de la memoria de un gran servidor que me presta sus capacidades a la hora de procesar y almacenar información. La ventaja de esto es que no necesito cambiar mi capacidad local frecuentemente, sino sólo ocasionalmente, pudiendo un equipo poco sofisticado o poderoso continuar desempeñando sus funciones gracias a que buena parte de las funciones que requiero están alojadas en la red. Aunque el aparato tenga una memoria modesta, puesto que alojo los datos que requiero en la red, esto no constituirá un problema. Ese servicio suele ser gratuito o restarse a un precio bastante bajo, por lo que constituye una alternativa bastante simple.

Por otro lado, el hecho que las capacidades informáticas estén en la red, hace posible integrar capacidad computacional a cualquier dispositivo sin transformarlo en un pequeño ordenador. Con un simple chip es posible conectar un refrigerador a Internet, lo cual podría ser útil para que me envíe un mensaje a mi teléfono cada vez que falten, por ejemplo, huevos. Esto es lo que se denomina internet de las cosas, puesto que crecientemente dispositivos que no son propiamente computadores se encuentran conectados a la red y, por tanto, generando datos y sirviéndose de diversas tecnologías, como la inteligencia artificial, por ejemplo. En este sentido, basta conectar un parlante a la red y podrá alojar a Alexa, la asistente virtual de Amazon. Puede conectarse el sistema de alarmas de una casa a internet y podrá ser controlado a través del teléfono, o por un asistente virtual como Alexa. En fin, lo relevante de esto para nuestros fines es que multiplica la capacidad de acción de la inteligencia artificial, permitiéndole salir del teléfono o del ordenador para integrarse dentro de las vidas diarias de las personas como controlados a través de algoritmos alojados en servidores de internet, desde

automóviles, a sistemas de alarma, juguetes, cocinas, bicicletas y todo cuanto la imaginación abarque. Esto crea inmensas oportunidades de negocio para las grandes compañías de Internet, que pueden extender su ámbito de comercial a tantas posibilidades propias del mundo no digital. Esto significa una extensión del mundo digital a la economía física y su radical expansión. Esto resulta interesante, dado que internet, no obstante los sueños igualitarios de la década de 1990, es un ambiente social y económico altamente concentrado, donde siete grandes compañías tienen una posición hegemónica a nivel global: Google (Alphabet), Facebook, Amazon, Apple, Baidu, Tencent y Alibaba.

La omnipresencia digital plantea varios problemas a nivel de contratación, regulación de mercados y responsabilidad. Preguntas evidentes surgen en relación a estos problemas. Un ejemplo es la contratación: si una compañía de internet me ofrece un servicio y yo lo acepto, yo claramente estoy prestando mi consentimiento, pero quien me realiza la oferta es un simple algoritmo, no una persona. Este algoritmo fija el precio y buena parte del contenido del servicio, sin que intervenga ningún ser humano. ¿He celebrado un contrato? ¿Con quién? ¿Y si la compañía se limita a ofrecer un servicio de un tercero, como es el caso de Airbnb? ¿Quién es responsable de algún incumplimiento? Si un automóvil automatizado me atropella, ¿quién es responsable? ¿El dueño? ¿El fabricante? ¿El que provee el servicio algorítmico de conducción?

Otra pregunta inmanente es si debemos regular la actividad económica basada en internet, y en la afirmativa cómo y hasta qué punto. Esto es especialmente interesante toda vez que la hegemonía de un puñado de compañías en ella implica que tienen un poder socio-económico más que suficiente para controlar los mercados globales. En estos casos, parece que la actividad regulatoria a nivel nacional resulta poco potente para enfrentar el fenómeno.

Por último, hay un aspecto crucial respecto a las tecnologías ligadas con las redes neuronales que no hemos tocado y que seguramente será la más compleja de abordar, esto es, el problema del sesgo. En principio, las redes neuronales se construyen por parte de los programadores determinando la arquitectura del modelo. A continuación se entrenan en un conjunto de datos a través de los cuales el modelo determina los pesos que asignar a las diversas conexiones entre sus capas, y, en buena medida, configura las capas ocultas de la misma. Así, una red neuronal, para funcionar eficientemente y detectar las correlaciones entre los datos con que se entrena, requiere de muchísimos ejemplos. Un niño puede aprender con uno o dos casos, pero un algoritmo requiere de cientos de miles. A mi sobrino le gustan los automóviles. Él vive en España, donde predominan las marcas europeas y las norteamericanas son más bien escasas, de manera que en un viaje le mostré un Chevrolet Camaro, uno de mis automóviles favoritos. Con dos ejemplos, él pudo reconocer otros similares con mucha mayor precisión que yo. Un algoritmo que reconozca este modelo de automóvil puede construirse, pero requiere de miles de ejemplos para realizar esta sencilla labor. La diferencia entre el algoritmo y mi sobrino es que el primero es una máquina de Hume, y no

⁵³ Ng, A., What Artificial Intelligence Can and Can't Do Right Now in HBR (2016), disponible en: <https://hbr.org/2016/11/what-artificial-intelligence-can-and-cant-do-right-now>. Consultada el 5 de junio de 2019.

elabora conceptos, sino que sólo detecta correlaciones. Mi sobrino, en cambio, puede abstraer eficientemente conceptos y, por tanto, a la hora de aprender a reconocer un modelo de automóvil requiere de relativamente pocos datos. Si las redes neuronales no funcionaron correctamente hasta la década de 2010, esto se debe principalmente a la pobreza relativa de datos que sólo vino a ser suplida por las enormes bases generadas a través de internet. En este sentido, las tecnologías actuales de inteligencia artificial son completamente dependientes de los datos y su eficacia está determinada por ellos.

De hecho, la abundancia de datos que ha generado internet es el factor clave en la efectividad de los actuales modelos de redes neuronales, toda vez que gracias a la existencia de cientos de miles de registros de fotografías, video, escritura y búsquedas es que los actuales algoritmos son entrenados y se configuran los pesos que vinculan sus diversas capas. Nada es tan eficiente como contar con más datos, de manera que incluso un algoritmo de diseño pobre entrenado en una base de datos enorme funciona mejor que uno más elegante, pero entrenado con una más modesta. Lo que ha permitido a las grandes compañías de internet convertirse en los centros de desarrollo e implementación de esta tecnología es, justamente, contar con cientos de miles de millones de datos de toda índole. Facebook, Google o Baidu tienen la virtud de guardar en sus centros de datos (*data centers*) todas las búsquedas realizadas jamás por sus usuarios, así como perfiles detallados de ellos, de todas las imágenes y videos que han subido y cada línea que han escrito. Estos centros son instalaciones construidas en diversos puntos del planeta que donde se almacenan todos los datos generados por la compañía y sus usuarios. Google tiene quince a nivel mundial, uno de los cuales está ubicado en Chile desde 2013.

La expansión de la internet de las cosas ha generado un auge de los datos disponibles, toda vez que se hoy en día la generación de datos de extiende a muchos más elementos que antes, abarcando las imágenes emitidas por las cámaras de automóviles, la geo localización vía GPS, y en general todos los datos que emitan los dispositivos dotados de conectividad a internet, cualquiera sea su naturaleza.

Puesto que la abundancia de datos optimiza los algoritmos, las compañías que dispongan de la mayor cantidad de datos, tendrán los mejores algoritmos, y por ende, los mejores productos con funciones de inteligencia artificial. La concentración de los datos ha generado tensiones, toda vez que esta implica el establecimiento de un oligopolio natural en la tecnología entre las grandes compañías de internet. Para generar una buena tecnología se requiere de cantidades masivas de datos, y esta tecnología será eficiente sólo en la medida con que cuente con más datos. Los productos de las compañías más grandes debiesen ser preferidos por su mayor eficiencia, y al preferirlos se refuerza la cantidad de datos disponibles y el producto continúa mejorando, lo que a su vez atrae más clientes.

La cantidad de datos almacenada por las grandes compañías de internet ha generado una creciente preocupación en diversos sectores de la ciudadanía que ven en la combinación entre una abundancia de datos personales dejados inadecuadamente en internet, y lo que los algoritmos pueden hacer con esos datos. Intuitivamente, tales dispositivos pueden establecer muchas correlaciones respecto a una persona y detentar un valioso poder predictivo. Así, cada búsqueda realizada en Amazon, por ejemplo, entrega a la compañía datos sobre lo que necesitamos, los que se ve reflejado en ofertas y precios especiales que la compañía nos muestra. Gmail Adds funciona intentando determinar con mecanismos de inteligencia artificial el contenido de los correos electrónicos de una persona, su segmento demográfico, y características generales, para luego mostrar publicidad al momento de su visualización. La idea es generar publicidad localizada y eficiente, pero esto genera dudas respecto a la privacidad.

Los datos generados en internet son un recurso valioso que posibilita la tecnología algorítmica de las redes neuronales. Lugares que generan muchos datos, y donde estos pueden ser utilizados libremente sin mayores cortapisas legales, como China, otorgan a las compañías que controlan este mercado una ventaja comparativa importante a la hora de establecer una posición dominante. Esto ha generado ansiedad en ciertos medios gubernamentales norteamericanos.

Lo que puede resultar más preocupante a la hora de evaluar jurídicamente estas aplicaciones es el problema del sesgo (*bias*). Un algoritmo tiene la calidad que le otorgan los datos con que se entrena. Lo único que hacen las redes neuronales es establecer correlaciones entre datos, de manera que las conclusiones a que lleguen dependen de la naturaleza de los datos que se le entreguen. Si un algoritmo debe entrenarse para detectar diversas razas de perros, y los canes domésticos aparecen siempre en contextos hogareños, mientras que los lobos figuran en contextos naturales, puede que el simple hecho de figurar un perro en un contexto natural (un bosque nevado), implicará que la red neuronal lo asociará a los lobos, como de hecho ha ocurrido. Existen muchos casos en que un algoritmo comete errores que están implícitos en los datos desde donde se entrena para realizar sus correlaciones.

Al respecto existen algunos casos que son francamente preocupantes. Un ejemplo de esto ocurrió con el programa de reconocimiento facial de Facebook. Joy Buolamwini⁵⁴ determinó que el algoritmo para el reconocimiento de rostros de esta compañía era deficiente a la hora de reconocer la cara de una persona negra. Incluso, un poco antes, en 2015, el programa de reconocimiento facial de Google había identificado a dos hombres negros como gorilas⁵⁵, por lo que

⁵⁴ Buolamwini, J., Gebru, T. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification in Proceedings of Machine Learning Research 81 (2018) disponible en: http://proceedings.mlr.press/v81/buolamwini18a.html?mod=article_inline Consultado el 12 de Junio de 2019.

⁵⁵ Zhang, M. Google Photos Tags Two African-Americans As Gorillas Through Facial Recognition Software in Forbes, 1 de Julio de 2015, disponible en:

la compañía debió pedir disculpas públicamente. Estos casos parecen haberse generado porque las bases de datos a partir de la cual Google y Facebook entrenan a sus algoritmos no contenían suficientes datos de personas negras, lo que provocó un serio problema a la hora de aplicarlo.

No obstante, hay problemas mucho mayores relativos a los posibles sesgos que un algoritmo puede contener. Un ejemplo de esto es el caso de los algoritmos predictivos de reincidencia utilizados en algunos lugares de Estados Unidos. Puesto que la discriminación es un problema agudo a nivel mundial, en algunos estados se comenzó a utilizar un algoritmo que aconsejase al juez a la hora de elaborar sus sentencias y otorgar libertades provisionales de acuerdo a los riesgos de reincidencia que los acusados o condenados presentaban. La idea era generar decisiones basadas en los datos y, por tanto, ajenas a los riesgos relativos a los prejuicios humanos. El resultado fue francamente inquietante. En algoritmo, tomando en cuenta las sentencias y decisiones pasadas de los tribunales norteamericanos, donde la discriminación fue la tónica, sistemáticamente recomendaba penas más largas y prevenía en contra de otorgar la libertad provisional respecto a personas negras, aunque hubiesen cometido el mismo delito y tuviesen los mismos antecedentes que personas blancas. En pocas palabras, distinguí en cuanto a otorgar la libertad en razón de la raza, un criterio universalmente prohibido y considerado discriminatorio⁵⁶. Es muy interesante que el resultado obtenido resulta exactamente el opuesto al buscado, puesto que en lugar de llegar a una justicia ciega, el velo que cubre los ojos de la diosa pareció caer por completo. El problema parece encontrarse en el hecho que los algoritmos toman la información y establecen correlaciones a partir de los datos con que son alimentados, por lo que si estos datos contienen un sesgo, que puede provenir, en último término, de la larga historia de la humanidad, dicho sesgo será reproducido por el modelo e incluso amplificado, proyectándolo hacia el futuro.

Otro caso interesante se da en la selección de currículos. Puesto que la práctica de seleccionar currículos es bastante ardua, especialmente en áreas donde los postulantes pueden ser muchos, es posible automatizar esta labor a fin de confeccionar una lista corta de candidatos respecto a los cuales se tome una decisión. Un criterio comúnmente utilizado para diseñar estos algoritmos es que consideren como buenos empleados a aquellos que hayan permanecido en la compañía un número determinado de años y que hayan obtenido una promoción, y luego comparen los currículos recibidos con estos modelos. Los resultados han sido fallidos en varias oportunidades, toda vez que en muchas compañías, y

especialmente en la tecnológicas, los hombres predominan frente a las mujeres, por lo que los algoritmos tienden a desechar los currículos de las mujeres que postulan a dichos cargos sólo por el hecho de serlo⁵⁷.

Esto nos lleva a otros problemas relativos a las correlaciones que los algoritmos establecen, puesto que incluso con grupos de datos que no parecen contener sesgo, el resultado puede resultar socialmente dañino. Un caso en este sentido se puede construir fácilmente con los datos de los supervivientes del Titanic⁵⁸. Esto corresponde a un ejercicio simple ejercicio de construcción de un algoritmo neuronal que suele ser recomendado para principiantes en programación. En base a una lista de supervivientes del Titanic que contiene algunos datos tomados de los registros originales, se diseña un modelo que predice las probabilidades de sobrevivir a una catástrofe marítima. El resultado del mismo es que los factores más relevantes para sobrevivir al choque contra el témpano de hielo son: ser mujer y viajar en primera clase. Por tanto, si diseñamos un algoritmo predictivo en base a ese modelo, digamos, para una compañía de seguros, los hombres pobres pagarían mucho más por representar un riesgo más alto en el transporte de pasajeros náutico que las mujeres ricas. Este tipo de algoritmos son relativamente simples de construir y están siendo utilizados sin nuestro conocimiento en áreas como los seguros de salud y los créditos bancarios, lo cual resulta inquietante. ¿Deberán los pobres pagar más intereses por el sólo hecho de serlo, no obstante tener un historial crediticio impecable?

Esto nos lleva a un problema importante, el cual consiste en que el Derecho, como tal, es un sistema de inferencias reglado. Así, puedo asumir que un vendedor debe cumplir con su obligación de entregar la cosa vendida en base a ciertas reglas pre-establecidas en el Código Civil, como son la existencia de un contrato de compraventa, el grado de cumplimiento del comprador de su propia obligación de pagar el precio, la existencia de vicios, y demás. Un algoritmo puede bien inferir la existencia de la obligación de entregar la cosa (o la falta de ella), de acuerdo a cualquier otro factor incluido en los datos con que fue entrenado, como por ejemplo que el contrato se realizó un día martes, a lo cual el algoritmo puede dar una importancia superlativa dados los datos con que cuenta. Aunque sería extraño que lo hiciese, es perfectamente posible. Si este algoritmo está diseñado para recomendar al juez una sentencia, o peor, para decidir un conflicto, el resultado puede ser perfectamente ilegal y arbitrario. En principio esto podría prevenirse con mejores datos o directamente desechando las recomendaciones realizadas por algoritmos que no se basen en correlaciones lícitas, pero esto no

⁵⁶ <https://www.forbes.com/sites/mzhang/2015/07/01/google-photos-tags-two-african-americans-as-gorillas-through-facial-recognition-software/#1530ee48713d> consultado el 12 de Junio de 2019.

⁵⁷ La investigación respecto al caso fue destapada por Propublica, una organización sin fines de lucro que analiza los usos de la inteligencia artificial, entre otras cosas. Disponible en <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> Consultado el 12 de Junio de 2019.

⁵⁷ Los casos en que esto ha ocurrido son muchos. Como muestra véase el reportaje de Reuters respecto a Amazon, disponible en <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scrap-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G> Consultado el 12 de Junio de 2019.

⁵⁸ El ejercicio fue diseñado y analizado por Meredith Broussard. Vid Broussard, M., *Artificial Unintelligence. How Computers Misunderstand the World*, (MIT Press-Kindle, Cambridge, 2018), 1.2150.

es fácil de determinar, toda vez que los modelos de redes neuronales tienden a ser opacos y poco explicables.

En principio, las redes neuronales no son capaces de proveer por sí mismas explicaciones de sus predicciones. A diferencia de los viejos sistemas expertos basados en principios lógicos, no existe una cadena de asunciones perfectamente trazables y comprensibles. Para determinar los motivos detrás de una determinación usando redes neuronales es necesario que un analista de datos dedique su tiempo para establecer a qué conexiones neuronales se está dando importancia, y dada la complejidad de algunos de estos, que pueden llegar a tener miles de capas ocultas, esto puede ser perfectamente imposible.

Esto nos lleva a un nuevo problema, relacionado con el anterior, que consiste en la opacidad de los sistemas de *deep learning* basados en redes neuronales. De momento, no existe un mecanismo simple que permita determinar con certeza las correlaciones que un algoritmo realiza y la fuerza que cada elemento que puede emerger en una capa de una red tendrá finalmente. Para esto se requiere el trabajo de analistas de datos, y aún en ese caso, esto puede resultar difícil y prácticamente imposible. Usualmente, para determinar si un modelo tiene algún sesgo importante, lo más simple es aplicarlo en un caso donde tal sesgo no exista y en otro donde sí lo haya y ver si ambos casos son tratados de la misma manera, aunque hay un grado importante de incerteza en este procedimiento. Esto es más complejo aún toda vez que las compañías que diseñan o utilizan algoritmos no suelen compartirlos, toda vez que estos son el secreto comercial que permite a sus asociaciones desarrollar un tráfico mercantil importante, y en caso de ser públicos, las compañías perderían rápidamente valor. En pocas palabras, no sólo resultan tales algoritmos opacos en la toma de decisiones, sino que además su funcionamiento suele ser un secreto comercial.

La Unión Europea, es su Guía Ética para una Inteligencia Artificial Confiable del 4 de abril de 2019⁶⁰ exige que el diseño de sistemas de Inteligencia Artificial se guíe por el principio de explicabilidad⁶¹, esto es, que en la medida de lo posible los algoritmos puedan ser transparentes en cuanto a las decisiones que toman. No obstante, no es claro que esto sea técnicamente posible. Actualmente se trabaja en dicho problema, y hay varias alternativas al respecto, aunque no se puede determinar con certeza hasta dónde se puede llegar en este camino.

4. ALGUNAS CONCLUSIONES PROVISIONALES

Hasta aquí hemos descrito el desarrollo histórico de la inteligencia artificial, sus límites y principales desafíos. Este capítulo introductorio ha intentado no comenzar por una definición de inteligencia artificial, sino mostrar su historia, a fin de a partir de ella fijar algunos conceptos clave y avanzar hacia la delimitación de los principales problemas jurídicos que pueden surgir a partir de sus capacidades actuales.

En primer término debemos señalar que de momento no existe algo similar al viejo sueño de contar con seres cuyo intelecto sea comparable al humano. En pocas palabras, no hay nada parecido a una entidad dotada de conciencia que sea capaz de dialogar con nosotros a partir de un plano de subjetividad creada por seres humanos. De momento no tenemos pistas acerca de qué es la conciencia, por lo que no podemos reproducirla, y aún en caso que lo lográsemos, tampoco sabríamos reconocerla. En cuanto a la posibilidad de una inteligencia débil, inconsciente, que pudiese resolver toda clase de problemas aunque no esté dotada de conciencia, esto tampoco parece estar a la vuelta de la esquina. Los programas que construimos son simples conjuntos de instrucciones o algoritmos que pueden enfrentar problemas particulares para los que fueron diseñados, pero poco más. No existe una inteligencia artificial general (AGI) y su consecución queda reservada a un futuro cuya cercanía no es clara. De momento sólo contamos con máquinas de Hume o de Platón, esto es, modelos de redes neuronales y sistemas expertos, que pueden resolver eficientemente los problemas para los que han sido diseñados. La característica común de ambos es que resuelven problemas que para los seres humanos requieren de intelecto, aunque en su caso lo hacen con fuerza bruta, esto es, a través del entrenamiento en extensas bases que cuentan con cientos de miles de datos, en el caso de las redes neuronales, o mediante la programación directa de principios lógicos, en el caso de los sistemas expertos. Ninguno de estos opera como lo haría un ser humano, aunque pueden llegar a resultados tan eficientes como los nuestros o más, en cuyo caso diremos que operan a niveles sobrehumanos. Si tuviésemos que definir la inteligencia artificial diríamos que es un conjunto de técnicas que permiten resolver de manera automatizada problemas que para los seres humanos requieren intelecto. Es una definición modesta, pero creemos que tiene la suficiente amplitud para abarcar la mayor parte de los desarrollos actuales.

Los problemas sociales que se desarrollan al contar con este tipo de agentes son de diverso carácter. En primer lugar, está la responsabilidad que se genera por su actuar. Dichas entidad pueden, de hecho, intervenir en el mundo y los resultados de sus actos, sean positivos, sean negativos, deben ser imputados a alguien. Estos resultados pueden ser buscados por quien se sirve de estos medios, como en el caso de quien pone a cargo de la adquisición de bienes a un algoritmo, o no, como quien sirviéndose de un automóvil auto conducido se ve envuelto en un accidente de tránsito. Esto nos lleva al problema jurídico de

⁶⁰ High Level Expert Group on Artificial Intelligence, *Ethic Guidelines for Trustworthy AI*, disponible en <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>. Consultado en 13 de Junio 2019.

⁶¹ High Level Expert Group on Artificial Intelligence, *Ethic Guidelines for Trustworthy AI*, p. 13, disponible en <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>. Consultado en 13 de Junio 2019.

determinar la responsabilidad, sea contractual, sea extracontractual, que dichos mecanismos generan.

Un segundo problema se deriva de su construcción su opacidad. Algunos algoritmos, especialmente los diseñados a través de redes neuronales, pueden contener importantes sesgos en su construcción y, en último término, llevar a resultados discriminatorios. Aquí hay varios problemas diversos, el primero es determinar cuándo podemos establecer una discriminación, el segundo es si dicho resultado puede generar responsabilidad, sea contractual, sea extracontractual y el tercero es que qué contextos y con qué cortapisas es conveniente servirnos de estos mecanismos.

Esto nos lleva a un tercer problema de corte regulatorio, que consiste en que debemos reglar el uso de tales mecanismos y, en la afirmativa, cómo. Todos estos temas los trataremos en el capítulo 2 de este texto.

Un segundo tema que involucra al conjunto de mecanismos que suelen denominarse inteligencia artificial es su influencia en el mercado de trabajo y, especialmente, en la profesión de abogado. ¿Veremos robots con toga en el futuro? Este será el tema del capítulo 3.

Por último, dejaremos volar nuestra imaginación en el capítulo 4 y analizaremos las posibilidades de un futuro más distante donde algunos de los hitos de la inteligencia artificial se hayan alcanzado. Hablaremos de nuestras pesadillas y sueños.

Capítulo II

INTELIGENCIA ARTIFICIAL EN EL DERECHO

1. INTRODUCCIÓN

Los sistemas de inteligencia artificial actuales básicamente utilizan dos tipos de elementos, a saber, aquéllos simbólicos, es decir, cuando directamente se programan reglas lógicas dentro del sistema, o de tipo conectivista, toda vez que se diseñan redes neuronales a fin que éstas realicen correlaciones entre datos y produzcan resultados a partir de éstos. Aunque ambos suelen estar presentes en todas las aplicaciones que se diseñan, éstos elementos son incompatibles entre sí y no se comunican los unos con los otros. La investigación respecto a los sistemas sintéticos, esto es, modelos que realicen abstracción desde los casos para construir reglas compatibles con los sistemas simbólicos, se encuentra aún en su infancia y no es claro que sea desarrollada en el tiempo próximo, aunque sería un avance mayúsculo hacia la construcción de un sistema general de inteligencia artificial.

No obstante lo anterior, los algoritmos de inteligencia artificial se han desarrollado lo suficiente como para adquirir una cierta capacidad de actuar independiente de las personas que se sirven de ello. Si algo caracteriza a estos mecanismos es la autonomía con que se desenvuelven en el mundo, sea virtual o real. De acuerdo a una serie de parámetros introducidos por un programador y de la experiencia ganada mediante su entrenamiento, estos dispositivos son capaces de actuar en el mundo fijando precios, proponiendo condiciones de contratación, cerrando acuerdos y ejecutándolos, sin necesidad de consultar o de referir sus actos a la persona que los puso en condiciones de actuar. En cierto sentido, sus actos son independientes de las personas que se sirven de ellos, lo cual puede ocasionar diversas situaciones que son jurídicamente relevantes. Es por esto que resulta conveniente variar levemente nuestra nomenclatura para referirnos a los modelos y algoritmos guiados por mecanismos de inteligencia artificial cuando ejecutan acciones jurídicamente relevantes, llamándolos agentes artificiales, por su capacidad de desenvolverse independientemente en el mundo real o virtual⁶¹.

El presente capítulo girará en torno a esta calidad de agentes que los dispositivos dotados de inteligencia artificial desarrollan, sea para fijar condiciones de contratación, cerrar acuerdos o cumplirlos, con una cierta independencia de las personas, sean físicas o jurídicas, que se sirven de ellos. Por este motivo,

⁶¹ Chopra, Samir and White, Laurence F., A Legal Theory for Autonomous Artificial Agents (University of Michigan Press-Kindle, Michigan, 2011), p. 5.

estudiaremos tres aspectos relevantes de su actuar. En primer término, veremos su capacidad para proponer y negociar condiciones de contratación, donde analizaremos la forma de actuación que tales dispositivos presentan. Un punto extremadamente relevante al respecto se encuentra en el sesgo (*bias*) que tales mecanismos pueden haber adquirido inconscientemente en su entrenamiento o diseño y los peligros que esto puede representar en el mundo jurídico. Un segundo problema a estudiar es el de la contratación, esto es, los efectos de los acuerdos que tales agentes cierran. ¿Provocan efectos obligatorios en las personas que se sirven de ellos? En la afirmativa, por qué. Finalmente nos corresponde estudiar los casos en que tales dispositivos ejecutan actos y, producto de su actuar, se ocasionan daños a terceras personas. ¿Nace algún tipo de responsabilidad extracontractual? ¿Para quién? Por último, vale la pena preguntarse por los efectos positivos de los actos de tales agentes. ¿Para quién adquieren? ¿Qué tipos de cosas pueden adquirir? ¿Pueden crear propiedad intelectual? El problema es cómo determinar a quién corresponden los beneficios y daños que genere una inteligencia artificial⁶².

2. PROBLEMAS EPISTEMOLÓGICOS. REDES NEURONALES, CORRELACIONES, CAJAS NEGRAS Y REGLAS

Los agentes artificiales están dotados, en efecto, de una autonomía suficiente para actuar en el mundo. Si bien su actuar no es libre, en el sentido que no gozan de libre albedrío, siguiendo el conjunto de instrucciones y de inferencias realizadas a partir de su diseño y entrenamiento, pueden ejecutar actos de diversa índole, muchos de los cuales pueden ser jurídicamente relevantes.

Ahora bien, el actuar de tales agentes se rige por dos tipos de mecanismos, por reglas programadas en ellos, generalmente para enmarcar o corregir su actuar, y por las predicciones que realizan luego de un proceso de entrenamiento mediante sus redes neuronales. El resultado de esta combinación es que las predicciones que realizan, y por ende las decisiones que toman, tales modelos resultan opacas. En pocas palabras, aunque se pueda tener una idea general acerca del funcionamiento de un algoritmo en particular, determinar el motivo porque tal o cual algoritmo realiza una predicción en concreto es extremadamente difícil de determinar, cuando no imposible. En el caso de modelos simples, se puede reconstruir la cadena de inferencias que lleva a tal o cual algoritmo a determinar algo, pero para ello debe examinarse la información que llega a cada neurona o perceptrón de la red y sus consiguientes ramificaciones, con los pesos relativos que cada punto de datos adquiere. En pocas palabras, es necesario ser un analista de datos competente y esperar que la red sea sencilla. No obstante, en el mundo real, las redes pueden contar con miles o cientos de miles de neuronas y

cientos de conexiones, por lo que estimar las razones matemáticas para que un sistema en concreto haya realizado una predicción en particular puede ser simplemente imposible. Este es el problema de la caja negra (*black box*), que puede ser especialmente relevante en ciertas disciplinas que requieren de una justificación para evaluar la conducta del agente, como es el caso del Derecho⁶³.

Resulta, en este sentido, un desafío para el mundo jurídico encontrar agentes que puedan actuar, y cuyo marco de acción tiene un carácter parcialmente impredecible. En efecto, mientras más avanzado sean el agente, mayor será este grado de autonomía e imprevisibilidad⁶⁴. Así, un agente de baja autonomía, por ejemplo un simple portal que despliega precios con un botón que permite al usuario adquirir el bien al precio expresado en el portal, no es verdaderamente un agente, sino simplemente un mecanismo de expresión de la voluntad de quien despliega el portal. Distinto es el caso si el portal puede variar el precio conforme la demanda por el bien sube o baja, toda vez que en este caso hay una condición esencial del contrato que está siendo fijada por el propio algoritmo. Un caso más complejo se torna el asunto si es que el portal es capaz de cambiar el producto u ofrecer otros según los perfiles de los usuarios que lo visualicen, tal vez excluyendo determinados productos de la visualización de ciertos usuarios, cambiando los valores de los mismos, ofreciendo condiciones más ventajosas o más gravosas a ciertos usuarios, incluso excluyendo a algunos de los posibles visualizadores de algunos productos o ofreciendo otros nuevos productos que eventualmente podrían ser de su interés. En este caso pareciera que algoritmo está haciendo algo más que simplemente expresar la voluntad de la persona que se sirve de él, sino que está estableciendo las condiciones esenciales del contrato, negociando como haría un vendedor en un bazar. El grado de autonomía que el agente tiene hace que la consideración que debemos tener acerca del mismo cambie, y esto tiene importancia desde el punto de vista jurídico.

Un asunto interesante es que las razones por las que el agente muestra determinados productos a ciertos precios a unos y no a otros no es fácilmente determinable. Hay un grado de oscuridad en tales determinaciones, especialmente si el algoritmo realiza estas ofertas basado en un modelo neuronal. Puede que, sin que sus diseñadores o las personas que se sirven del algoritmo lo sepan, esté realizando actos contrarios a sus intereses o abiertamente discriminatorios. En efecto, algunos autores estiman que la inteligencia artificial se convertirá en un importante problema de derechos humanos durante el presente siglo⁶⁵. Existen varios casos en que tales problemas han ocurrido y merece a pena repasar algu-

⁶² Davenport, Thomas H., *Artificial Intelligence for the Real World* in Harvard Business Review, *On AI, Analytics and the New Machine Age* (Harvard Business Publishing-Kindle, 2019, Boston MA), 1:223.

⁶³ Vid Chopra, Samir and White, Laurence F., *A Legal Theory for Autonomous Artificial Agents* (University of Michigan Press-Kindle, Michigan, 2011), p. 9.

⁶⁴ Safiya Umoja Noble, *Algorithms of Oppression. How Search Engines Reinforce Racism* (New York University Press-Kindle, 2018, New York), p. 1.

⁶⁵ Vid, en este sentido, Blakin, Jack M., *The Path of Robotics Law*, California Law Review Circuit 6 (2015) 45-60, p. 46.

nos de ellos a fin de estudiar el problema del sesgo. Este aspecto cobra mucha mayor importancia si es que el campo en que se aplica la inteligencia artificial no es simplemente un asunto contractual, sino que se emplea en el desarrollo de una función de carácter público, como puede ser la administración de justicia o la provisión de servicios sociales. En principio, el actuar del aparato público debe basarse en reglas pre-establecidas y objetivas, que trate a todas las personas que están en una misma condición de manera equivalente. En esto consiste la igualdad ante la ley, que verdaderamente es un deber que la Constitución impone a la actuación del Estado. Esta igualdad implica también un trato no discriminatorio no sólo del aparato público, sino también de los particulares en sus relaciones entre sí, de manera que el establecimiento de diferencias injustificables en motivos racionales son contrarias al orden jurídico político. ¿Un mecanismo fundamentalmente inexplicable en sus decisiones, puede cumplir con estos deberes implicados en la noción de igualdad ante la ley y prohibición de la discriminación? En más, ¿puede un algoritmo que fundamenta sus decisiones en correlaciones de datos oscuras para el administrado y para el propio administrador, tomar decisiones que cumplan con los requisitos básicos de fundamentación y transparencia que exige nuestro orden jurídico para el actuar de órganos públicos? Tomemos un par de ejemplos reales para explicitar el problema expuesto.

En el año 2014⁶⁵ Amazon intenta solucionar el problema que se presenta al analizar el alto número de currículos que se presentan cada vez que ofrece un puesto de trabajo. Para ello, un algoritmo entrenado en su propia base de datos, estaría encargado de confeccionar una lista corta de la cual elegir a los candidatos para un puesto de trabajo determinado cada vez que se iniciase el proceso de selección de personal. La idea es ahorrar a la compañía el tedioso proceso de análisis de miles de currículos que llegan a ella cada vez que se abre tal proceso, amén de hacerlo más justo y objetivo, toda vez que ésta, como muchas otras compañías tecnológicas, tiene un problema en la variedad de perfiles de sus trabajadores de alto nivel, mayoritariamente hombre blancos. Se entrenó al algoritmo buscando entre sus propios trabajadores el perfil adecuado tomando en cuenta la historia laboral de los mismos, intentando priorizar perfiles similares a aquellos que hubiesen permanecido un tiempo largo en la empresa y que hubiesen obtenido ascensos. El resultado fue diverso a lo previsto. Toda vez que la mayor parte de los trabajadores de Amazon son hombres, el algoritmo procedió a seleccionar exclusivamente currículos masculinos excluyendo a mujeres por el solo hecho de serlo, un resultado completamente contrario a los intereses de la empresa. Lo más llamativo del caso, es que los currículos no incluían un apartado donde se indicase el sexo de la persona, pero el agente infirió el sexo de otros datos con los cuales se correlaciona el primero, como por ejemplo la pertenencia

de ciertos clubs de ajedrez femenino, entre otros. En 2015 Amazon se percató del sesgo presente en su modelo y, al ser incapaz de corregirlo, debió retirarlo⁶⁷.

En este caso, el problema se da en dos dimensiones del algoritmo. El primero es la base de datos sobre la cual se entrenó el mismo. Los datos para entrenar los algoritmos no eran neutrales, sino que contenían un histórico sesgo contrario a las mujeres en el mundo de la informática, por lo que el algoritmo no hizo otra cosa que replicarlo y amplificarlo. En pocas palabras, el algoritmo no es más inteligente que los datos con que se alimenta⁶⁸ y si estos contienen un sesgo, entonces el algoritmo adquirirá el sesgo de los mismos datos con que fue entrenado. Este es un problema de sobre generalización (*overgeneralization* u *overfitting*, en términos estadísticos), donde el algoritmo detecta patrones en los datos –en ocasiones invisibles para el propio programador– y luego los reproduce y amplifica⁶⁹.

Un segundo elemento llamativo, es que las decisiones del algoritmo están determinadas por su definición de éxito. Ésta se encuentra determinada a priori por los diseñadores del modelo, y si no incluye algún tipo de corrección explícita de los sesgos del pasado, el modelo simplemente los reproducirá. En pocas palabras, para corregir los problemas de discriminación que un conjunto de datos pueda presentar, es necesario que los programadores estén conscientes de su existencia y sean activos en su corrección, de lo contrario la pretendida objetividad que pretenden lograr no sólo no se dará, sino que es perfectamente posible que el modelo profundice los sesgos.

Como este hay un sinnúmero de casos posibles en que un agente se limitará a reproducir los sesgos que encuentre en los datos, produciendo resultados moralmente cuestionables y jurídicamente inaceptables. Uno de los más curiosos es Tay. Este fue un agente de conversación lanzado por Microsoft⁷⁰ en 2016 cuya finalidad era emitir *tweets* del tipo que escribiría una muchacha adolescente. La idea original es que Tay aprendiese de los *tweets* con que otros internautas la alimentasen, de manera que su actuar en la red se hiciese más natural y que reflejase los problemas reales que preocupan a los adolescentes. El resultado fue estrambótico. Apenas lanzada al ciberespacio, recibió una avalancha de comentarios racistas y xenófobos de tal magnitud que a las pocas horas se convirtió en una ferviente supremacista, admiradora de Hitler, que hacía proposiciones sexuales a otros internautas. Tan solo dieciséis horas más tarde, Microsoft tuvo

⁶⁵ La historia tuvo alta cobertura de prensa. Véase para un resumen: <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scrapes-several-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK09G>. Consultada en 7 de Octubre de 2019.

⁶⁷ Xerox tuvo el mismo problema, con resultados idénticos. Vid O'Neil, Cathy, *Weapons of math Destruction: How Big Data Increases Inequality and Threatens Democracy*, pp. 116-119.

⁶⁸ O puesto de otra manera, el ser humano es más listo que los datos, vid Pearl, Judea, *The Book of Why: The new science of cause and effect* (Basic Books, 2018, New York), p. 21.

⁶⁹ Surden, Harry, *Machine Learning and Law*, *Wash. L. Rev.* 87 (2014), p. 106.

⁷⁰ Broussard, Meredith, *Artificial Unintelligence: How Computers Misunderstand the World* (MIT Press-Kindle, 2018, Cambridge MA-London), [1289].

que bajar la aplicación. La lección es evidente, si se alimenta un algoritmo con basura, el modelo generará más basura⁷¹.

Respecto a los sesgos que un modelo puede adquirir en relación a los datos con que es entrenado, se suele clasificarlos en, a lo menos, tres tipos: sesgo de interacción (*interaction bias*), sesgo latente (*latent bias*) y sesgo de selección (*selection bias*). El primero consiste en que el propio usuario o programador inadvertidamente introduce un sesgo en el modelo por la manera en que interactúa con él. Esto se da, por ejemplo, al definir la función de éxito (*objective function*) del modelo, o el objeto que se busca. El sesgo latente, en cambio, tiene lugar cuando el modelo realiza correlaciones inapropiadas, generalmente estableciendo falsos nexos entre puntos de datos. Así, por ejemplo, si las personas de poca solvencia no pagan sus créditos, y la falta de solvencia puede correlacionarse con la pobreza y esta con la segregación espacial en la ciudad, un algoritmo puede tomar la residencia de una persona en un punto de la ciudad segregado, como indicador de su alto riesgo crediticio, lo cual puede ser perfectamente falso. En cuanto al sesgo de selección, este tiene lugar cuando la base de datos no es suficientemente representativa de la diversidad existente en el medio social. Así, por ejemplo, si entreno mi algoritmo para determinar elementos que hagan predecir habilidades en una población para ser buenos jugadores de fútbol con los datos médicos de todos los jugadores de la primera división argentina, probablemente dicho algoritmo sea inútil para realizar dicha predicción en Japón, debido a la baja representatividad de población asiática en el mismo.

Un elemento inquietante se relaciona con la política de precios y condiciones de contratación que tales algoritmos pueden generar. Estableciendo perfiles para los distintos usuarios, pueden establecerse distintos precios para los diversos demandantes un servicio. Amazon admitió ya en el 2000 utilizar esta política⁷² y parece ser relativamente común en el ciberespacio. Un efecto de tal mecanismo es que las personas con mayor poder adquisitivo tienden a recibir ofertas de productos a precios más bajos que aquéllas que cuentan con menos recursos⁷³. A partir de las búsquedas que una persona realiza, de las opiniones manifestadas en las redes sociales o de cualquier otro conjunto de datos, puede diseñarse un algoritmo que detecte las posibles necesidades de un individuo en concreto y enviarle ofertas personalizadas, en condiciones distintas a aquellas ofertadas para los demás internautas. El servicio AdSense de Google Mail detecta patrones en

el contenido de los correos electrónicos que cada usuario recibe para mostrarle publicidad personalizada. Al margen de los problemas que tal servicio puede implicar por violación del secreto de la correspondencia⁷⁴, dicho mecanismo genera, a partir de un conjunto de datos, un perfil del usuario y ofrece bienes y servicios específicos para él.

La pregunta en esta instancia es si tales ofertas pueden estar afectas a algún sesgo y si tal política de precios diferenciados es acorde con la protección al consumidor⁷⁵. Respecto al primer problema, la posible existencia de un sesgo en la generación de una política de precios, verdaderamente dependerá del caso y no podemos dar una respuesta a priori. Puesto que las decisiones algorítmicas son verdaderas cajas negras, no podemos estar seguros de los motivos de tal o cual modelo en particular para segregar su política de precios. Puede que la base de datos contenga un sesgo o que la función de éxito esté incorrectamente diseñada, en cuyo caso el resultado será discriminatorio. Lo importante es analizar el caso en concreto que se juzga sospechoso, comparar los resultados que arroja el algoritmo si se incorpora o se excluye el elemento que pueda causar discriminación, así como también atender a las grandes tendencias que emergen al estudiar los precios ofrecidos. Evidentemente, si tales resultados son explicable por diferencias arbitrarias, como el sexo, la raza u otro factor similar, nos encontraremos ante un caso de discriminación algorítmica y se activará la posibilidad de ejercer acciones constitucionales como el recurso de protección en Chile, o de protección al consumidor⁷⁶, pero por la naturaleza oscura de este tipo de algoritmos, esto requiere de un análisis especializado. Debe recordarse que un algoritmo puede resultar discriminatorio aun cuando no tome entre sus puntos de datos el factor que origina la discriminación. En el caso del modelo que seleccionaba currículos de Amazon, el sexo no estaba entre los elementos que el algoritmo consideraba. Ahora bien, el resultado de su aplicación era la exclusión de la mujeres toda vez que el algoritmo consideraba positivamente actividades que los hombres suelen desarrollar (como integrar un equipo de fútbol americano) y neutrales o negativas aquellas que suelen realizar mujeres (*ser cheerleader*, por ejemplo). De ahí que, aunque el agente no tuviese una consideración explícita del sexo, el resultado era la discriminación por sexo.

Ahora bien, suponiendo que la política de precios algorítmica se encuentra fundada en criterios no discriminatorios a priori, la pregunta que queda abierta

⁷¹ Broussard lo expresa en los siguientes términos: *When you believe that a decision generated by a computer is better or fairer than a decision generated by a human, you stop questioning the validity of the inputs to the system. It's easy to forget the principle of garbage in, garbage out—especially if you really want the computer to be correct* Broussard, Meredith, *Artificial Unintelligence. How Computers Misunderstand the World* (MIT Press-kindle, 2018, Cambridge MA-London), 1.819.

⁷² Broussard, Meredith, *Artificial Unintelligence. How Computers Misunderstand the World* (MIT Press-kindle, 2018, Cambridge MA-London), 1.212B.

⁷³ Broussard, Meredith, *Artificial Unintelligence. How Computers Misunderstand the World* (MIT Press-kindle, 2018, Cambridge MA-London), 1.212B.

⁷⁴ Chopra, Samir and White, Laurence E, *A Legal Theory for Autonomous Artificial Agents* (University of Michigan Press-Kindle, Michigan, 2011), p. 110.

⁷⁵ Este problema fue abordado por Zuiderveen Borgesius, Frederik, *Discrimination, Artificial Intelligence and Algorithmic Decision-Making* (Directorate General of Democracy, 2018, Estrasburgo) en un informe para el Consejo de Europa, disponible en: <https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73>

Consultado el 10 de Octubre de 2019.

⁷⁶ Se infringiría el artículo 3 letra c de la Ley 19496 que fija el derecho del consumidor a no ser discriminado arbitrariamente.

es si es aceptable que diversos consumidores reciban distintos precios. Teóricamente existe un precio general (más alto) que estaría ofrecido a la mayor parte de los consumidores, pero en la práctica, puesto que muchísimos consumidores al buscar diversos productos, reciben precios distintos, parece que el precio general no existe o existe sólo para una minoría. Esto parece encontrarse referido con lo señalado por la normativa de consumidores hoy vigente en Chile, que indica la obligatoriedad⁷⁷ de mantener los precios ofertados visibles. Ahora bien, puesto que la práctica común consiste en tener un precio de lista y ofrecer al consumidor personalmente uno más bajo a título de oferta, la dificultad parece subsanarse, toda vez que habría un precio general cognoscible por todo consumidor y uno especial ofrecido en concreto a algunos de ellos, pero queda en el aire la pregunta relativa a por qué dicho consumidor recibe un precio especial y los demás no. ¿Resulta esto discriminatorio, en los términos del artículo 3 letra c de la Ley 19496? Toda vez que los motivos del algoritmo para establecer en un caso concreto un precio distinto para un consumidor que otro son correlaciones de datos de incognoscible significado, no podemos saberlo y sería necesario hacer un estudio acerca de su tendencia. En cualquier caso, lo generalizado de la práctica y la dificultad en encontrar motivos objetivos para cada caso, hacen necesario discutirlo.

Más complejo se torna el caso de las ofertas de servicios financieros, materia en que comúnmente se aplican algoritmos de este tipo para realizar ofertas y establecer condiciones crediticias. Son muchos los puntos de datos que podrían eventualmente tomarse en consideración para establecer que una persona es solvente o no, entre los cuales está su historial de incumplimientos, su solvencia, y otros similares. Ahora bien, no sabemos qué elementos toma en cuenta el algoritmo que determina el riesgo de una persona en concreto, pero si ha sido construido en base al *deep learning* y bases de datos abiertas, puede estar afecto a criterios discriminatorios, como el lugar donde vive, características del nombre u otros factores igualmente problemáticos. En materia financiera, por lo

demás, el artículo 3º letra a) de la Ley 19496 establece que en caso de rechazar la contratación de un servicio financiero, el consumidor debe "ser informado por escrito de las razones del rechazo a la contratación del servicio financiero, las que deberán fundarse en condiciones objetivas." En este caso, vale la pena preguntarse si la determinación de un alto riesgo por parte de un algoritmo es suficiente razón, o debe también informarse de los motivos que inducen al modelo a establecer ese nivel de riesgo en concreto. De tomarse la segunda opción, que parece más acorde con el espíritu de la legislación, es difícil imaginar cómo podría fundamentarse la decisión del algoritmo. En efecto, los modelos crediticios han sido cuestionados, especialmente porque su determinación de riesgo —y, por tanto, del interés de un crédito— es ciertamente oscura para el consumidor⁷⁸.

Es importante señalar que en una sociedad donde históricamente la movilidad social fue baja y la pobreza, y por tanto la falta de solvencia, han tenido históricamente componentes étnicos, como la nuestra, es posible que un modelo tome factores correlacionados con la pobreza, como el nombre⁷⁹, el lugar de nacimiento, el establecimiento educacional donde se cursó la primaria u otros similares, para construir un modelo que acreciente el riesgo de las personas con tales antecedentes y lo aminore respecto a otros que detentan factores usualmente correlacionados con el éxito⁸⁰. En pocas palabras, un algoritmo podría fácilmente tomar factores que históricamente han fundado el prejuicio social y potenciarlos de manera completamente inaceptable⁸¹. Los programadores y diseñadores del modelo pueden estar perfectamente inconscientes de este hecho, e incluso pueden estar buscando fines completamente contrarios, como promover a través de diseños matemáticos una sociedad más igualitaria y menos segregada, y no obstante el agente producto de sus esfuerzos puede tender a la inmovilidad social, la re-etnificación de la pobreza y el cierre de los grupos sociales con más

⁷⁷ L. 19496, Artículo 30.- Los proveedores deberán dar conocimiento al público de los precios de los bienes que expendan o de los servicios que ofrezcan, con excepción de los que por sus características deban regularse convencionalmente. El precio deberá indicarse de un modo claramente visible que permite al consumidor, de manera efectiva, el ejercicio de su derecho a elección, antes de formalizar o perfeccionar el acto de consumo.

Igualmente se enunciarán las tarifas de los establecimientos de prestación de servicios. Cuando se exhiban los bienes en vitrinas, anaquelos o estanterías, se deberá indicar allí sus respectivos precios. La misma información, además de las características y prestaciones esenciales de los productos o servicios, deberá ser indicada en los sitios de Internet en que los proveedores exhiban los bienes o servicios que ofrezcan y que cumplan con las condiciones que determine el reglamento.

El monto del precio deberá comprender el valor total del bien o servicio, incluidos los impuestos correspondientes.

Quando el consumidor no pueda conocer por sí mismo el precio de los productos que desea adquirir, los establecimientos comerciales deberán mantener una lista de sus precios a disposición del público, de manera permanente y visible.

⁷⁸ Frank Pasquale señala al respecto:

A bad credit score may cost a borrower hundreds of thousands of dollars, but he will never understand exactly how it was calculated. Pasquale, Frank, *The Black Box Society: The Secret Algorithms that Control Money and Information* (Harvard University Press, 2015, Cambridge MA-London), p. 4. Al respecto, véase también: Marron, Doncha, "Lending by Numbers": Credit Scoring and the Constitution of Risk within American Consumer Credit, *Economics and Society* 35 (2007), p. 111.

⁷⁹ En efecto, ha habido casos en que se ha detectado sesgo en el algoritmo de predicción de resultados de Google. Vid Pasquale, Frank, *The Black Box Society: The Secret Algorithms that Control Money and Information* (Harvard University Press, 2015, Cambridge MA-London), p. 40.

⁸⁰ Vid Edith Ramirez, "Privacy Challenges in the Era of Big Data: The View from the Lifeguard's Chair", disponible en: https://www.ftc.gov/sites/default/files/documents/public_statements/privacy_challenges-big-data-view-lifeguard's-chair/130819bigdataaspen.pdf Consultado el 15 de Octubre de 2019.

⁸¹ La situación ha sido descrita en los siguientes términos: *Mounting evidence shows that automated decision-making systems are disproportionately harmful to the most vulnerable and the least powerful, who have little ability to intervene in them—from misrepresentation to prison sentencing to accessing credit and other life-impacting formulas.* Noble, Safiya Umoja, *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York University Press-Kindle, 2018, New York), p. 49.

medios económicos. Esto es lo que se ha denominado Armas de Destrucción Matemática (*Weapons of Math Destruction*)⁶⁵.

En ocasiones se ha señalado que, aunque los algoritmos estén abiertos a cometer sesgos de diversa especie, como también lo están los seres humanos, estos sesgos serán siempre menores que aquellos que aplican las personas⁶⁶, por lo que sería importante no exagerar su riesgo. Discrepamos de tal opinión, toda vez que respecto a los seres humanos, podemos preguntarles sus motivos y, en ocasiones, exigirles jurídicamente que los manifiesten, mientras que no contamos con medios similares respecto a los agentes artificiales, que no son más que entidades inconscientes que manipulan mecánicamente símbolos respecto a los cuales no son capaces de asignar significado. Así, si nos servimos del ejemplo relativo al Titanic utilizado más arriba, donde parecen ser el sexo y la clase los factores determinantes en la supervivencia de un accidente marítimo, una compañía de seguros que utilice este algoritmo para establecer el precio de la prima simplemente terminaría en encarecer injustificadamente los precios de los seguros sirviéndose de factores abiertamente discriminatorios, amén de no ver el verdadero factor de riesgo, la insuficiencia de los botes salvavidas que portaba el buque.

Ahora bien, si el riesgo de un sesgo hace delicado el uso de agentes artificiales en el campo privado, en el mundo público esto es aún más difícil. Ha habido diversos esfuerzos tendentes a servirse de agentes artificiales a fin de automatizar acciones del gobierno, de manera que estos modelos realicen actos que usualmente se dejaba al criterio de algún ser humano, como un asistente social o un juez. Si tales mecanismos cuentan con un sesgo obtenido, sea de los datos o sea del diseño de su función de éxito, el resultado será una amplificación de la discriminación que experimentan los miembros más vulnerables de nuestra sociedad⁶⁷, esta vez distribuida de una manera automatizada por parte de la propia autoridad estatal que, teóricamente, está justamente encargada de combatirla.

Un ejemplo que merece nuestra atención consiste en un algoritmo de predicción de sentencias diseñado ya hace algún tiempo con el que experimentó el

grupo de investigación Lex ex Machina, entonces de la Universidad de Stanford⁶⁸. Al respecto, este fue uno de los primeros modelos que se sirvieron de *machine learning* para predecir los resultados de litigios relativos a marcas y propiedad intelectual. Una vez realizado el análisis de sus correlaciones y puntos de datos, se encontró que aquellos factores más relevantes para predecir el resultado eran la identidad del juez y el nombre del equipo jurídico de las partes.⁶⁹ El resultado resulta interesante toda vez que da cuenta acerca del valor que tiene contar con un buen juez y con un equipo de abogados solvente, pero si se quiere transformar este algoritmo en un sistema no predictivo, sino sentenciador, terminaría decidiendo los casos no en razón de sus méritos, sino de la identidad de los abogados, lo cual resultaría evidentemente ajurídico.

El caso más notable lo representa Compas, un algoritmo diseñado por la empresa Northpointe que tiene por finalidad determinar la peligrosidad y el riesgo de reincidencia de los procesados o condenados por crímenes o simples delitos, el que se encuentra en uso en veinticuatro de los estados que componen Estados Unidos. El objetivo es reemplazar la siempre subjetiva determinación de la peligrosidad de un individuo y su posible riesgo de reincidencia, por un método algorítmico que fuese imparcial y matemáticamente predictivo de la conducta de la persona detenida. Para alimentar el algoritmo se diseñó un cuestionario del cual se obtuvieron los puntos de datos que el mismo utilizaría a la hora de entrenarlo y, más tarde, aplicarlo. Éste contiene de ciento treinta y siete preguntas que versan sobre temas diversos relacionados con el ambiente social en que se mueve el detenido, factores que la criminología americana tiende a correlacionar al delito. Las preguntas se refieren a aspectos variados, como por ejemplo, dónde vive la persona, si es que conoce personas que hayan sido condenadas por delitos, si algunas de éstas son parientes suyos, si tiene trabajo, qué tipo de familia lo crió, y muchos otros factores similares. De acuerdo al entrenamiento de la red neuronal, el cuestionario asigna a cada individuo una probabilidad de reincidencia, la cual es entregada al juez a fin que la tenga en consideración (si lo cree oportuno) en la concesión de la libertad vigilada o, en algunos estados, en la determinación de la pena. La finalidad buscada era que ya no fuese una persona cargada de prejuicios la que estableciese el riesgo de reincidencia, sino que un algoritmo tomase su función⁷⁰.

El resultado es que un algoritmo determina la peligrosidad de los individuos y realiza recomendaciones al juez basándose en los datos obtenidos de la encues-

⁶⁵ Cathy O'Neil acuñó el término, y describe el problema en oscuros colores: *Nevertheless, many of these models encoded human prejudice, misunderstanding, and bias into the software systems that increasingly managed our lives. Like gods, these mathematical models were opaque, their workings invisible to all but the highest priests in their domain: mathematicians and computer scientists... they tended to punish the poor and the oppressed in our society, while making the rich richer.* O'Neil, Cathy, *Weapons of math Destruction. How Big Data Increases Inequality and Threatens Democracy*, p. 3.

⁶⁶ Casey, Anthony J., Niblett, Anthony, *Self-driving Laws*, *University of Toronto Law Journal*, 66-4 (2016) pp. 429-442, p. 437.

⁶⁷ Virginia Eubanks describe la situación en Estados Unidos en los siguientes términos: *Across the country, poor and working-class people are targeted by new tools of digital poverty management and face life-threatening consequences as a result. Automated eligibility systems discourage them from claiming public resources that they need to survive and thrive.* Eubanks, Virginia, *Automating Inequality. How High-Tech Tools profile, police, and punish the poor* (St. Martin's Press, 2018, New York), p. 11.

⁶⁸ Surdeanu, Mihai, Nallappi, Ramesh, Gregory, George, Walker, Joshua and Manning, Christopher, *Risk Analysis for Intellectual Property Litigation in Proceedings of the 13th International Conference on Artificial Intelligence and Law* (New York, ACM, 2019), 116-120, disponible en: <https://nip.stanford.edu/pubs/ical11.pdf> Consultado el 21 de Octubre de 2019.

⁶⁹ Ashley, D. Kevin, *Artificial Intelligence and Legal Analytics. New Tools for Law Practice in the Digital Age* (Cambridge University Press, 2012, Cambridge), p. 124.

⁷⁰ Broussard, Meredith, *Artificial Unintelligence. How Computers Misunderstand the World* (MIT Press-Kindle, 2018, Cambridge MA-London), l. 2900.

ta, pero sin que el juez conozca los fundamentos de tal aserción de riesgos. En pocas palabras, si el juzgador hace suya la afirmación del algoritmo acerca de la peligrosidad de un individuo, al no conocer exactamente el fundamento de la afirmación, es dudoso que su decisión de otorgar o denegar la libertad provisional se encuentre sustentada en hechos objetivos. Si no conoce los motivos de la calificación algorítmica de riesgos, cómo va a servir de base a su decisión. Este problema fue ventilado en el caso *State v. Loomis*⁶⁰, donde el sentenciador de Wisconsin tomó expresamente entre los fundamentos de su sentencia la aserción de peligrosidad realizada por Compas.

Aunque en definitiva el caso fue desechado, toda vez que la aserción de Compas no fue el único fundamento de la sentencia, vale la pena destacar que uno de los factores que se discutió es que ni el acusado, ni el juez tienen información acerca de cómo funciona el modelo, toda vez que su composición está protegida por la propiedad industrial de Northpointe. Por lo demás, la información está protegida por la propiedad industrial de Northpointe. Entre ellas se encuentran preguntas de donde se obtienen los puntos de datos del algoritmo tienen un cariz difícilmente justificable para establecer una sentencia. Entre ellas se encuentran factores tales como si el sujeto fue criado por una familia mono-parental., por ejemplo. ¿Parece aceptable que los hijos de madres solteras obtengan sentencias más largas que los que han sido criados en hogares biparentales, sólo por este hecho? Si mis parientes directos hubiesen cometido delitos —uno de los elementos considerados por Compas— ¿es esto suficiente para que se me aumente mi riesgo de reincidencia y se me mantenga en prisión más tiempo que otras personas que cometieron exactamente el mismo delito? Evidentemente hay algo profundamente equivocado en estos razonamientos. Es más, si un juez los hiciese explícitamente suyos, no dudaríamos en acusarlo de prevaricación. No obstante, puesto que es un algoritmo el que determina matemáticamente el riesgo a partir de esta información, el resultado parece neutral y convincente.

En todo caso, el algoritmo de Northpointe tuvo más problemas al poco andar. Un estudio de Propublica⁶¹ estableció que el modelo de Compas otorgaba sistemáticamente un mayor riesgo de reincidencia a personas negras que a blancas, sin que fuese apreciable en las circunstancias de comisión de los delitos, o en sus antecedentes penales, ninguna diferencia relevante. De hecho, lo hacía de una manera tan consistente que, en caso de haberlo hecho un ser humano, no se dudaría en adivinar un racismo latente en el evaluador⁶². Aparentemente, cuando existen dos grupos de población con niveles de criminalidad diversos, no se

puede establecer un único algoritmo que evalúe con justicia a ambos grupos⁶³. En este caso, en el caso de Compas, no es posible determinar esto con certeza, toda vez que el funcionamiento del algoritmo permanece oculto tras su secreto industrial.

En principio, si no se intenta apuntar a la equidad como uno de los objetivos buscados en la función objetiva del algoritmo, el modelo no la entregará⁶⁴. Hay muchas definiciones de equidad que se podrían entregar al modelo y que llevarían a resultados diversos. En primer término es menester detectar qué grupos son aquellos que se desea proteger especialmente, si es que hay alguno. En la alternativa, por ejemplo, se podría programar al modelo para que sus predicciones fuesen equivalentes para ambos grupos, es decir, que hayan tantos resultados positivos en un grupo como proporcionalmente habría en otros. El grupo minoritario es yergue como límite en la asignación de beneficios a la comunidad. Así, por ejemplo, supongamos que se juzga que las mujeres se ven injustamente perjudicadas en la asignación de subsidios habitacionales respecto a los hombres y se introduce la función de que sólo se pueden asignar tantos subsidios a hombres como se hayan asignado a mujeres. Así, para asignar diez subsidios a varones, deberán previamente asignarse diez a mujeres. El problema es que en el caso puede que existan veinte varones que habrían podido recibir el beneficio y sólo diez mujeres en tal situación, por lo que en definitiva, habrá diez personas que no resultarán favorecidos por el simple hecho de carecer de candidatas mujeres. Otra opción es equiparar los falsos positivos o los falsos negativos entre los grupos, así, el modelo se equivocará tantas veces respecto a hombres y mujeres. Esto es arriesgado, toda vez que distribuye beneficios o perjuicios de manera injusta pero equilibrada entre los grupos. En definitiva, no es sencillo diseñar la función objetiva para que tenga en cuenta la equidad, y en todos los casos se sacrifica la exactitud con que funciona el modelo.

En relación a la evaluación del riesgo que representa el ofensor para la sociedad, es interesante señalar que parece apuntar a las características de la persona, y no a los hechos que pueda haber cometido. La aserción de riesgo que realiza el modelo se fundamenta en quien la persona es, cuales son sus amistades, su crianza y nivel socioeconómico, antes que en la naturaleza de los hechos imputados y su conducta anterior. En esto, parece apuntar al Derecho penal de autor, que juzga a las personas por quienes son y no por lo que hacen, lo cual es impropio en una sociedad democrática. Si se juzga que puede predecirse el riesgo de comisión de delitos con certeza, no es de extrañar que se despliegue una

⁶⁰ Vid *State v. Loomis*, 881 N.W.2d 749 (Wis. 2018) in *Recent Cases*, Harvard Law Review 130 (2017), pp. 1530-1537

⁶¹ Disponible en: <https://www.propublica.org/dataset/compas-recidivism-risk-score-data-and-analysis>. Consultado el 21 de Octubre de 2019.

⁶² Véase al respecto, Broussard, Meredith, *Artificial Unintelligence: How Computers Misunderstand the World* (MIT Press-Kindle, 2018, Cambridge MA-London), 12290.

⁶³ Kleinberg, Jon M., Mullainathan, Sendhil, and Raghavan, Manish, *Inherent Trade-Offs in the Fair Determination of Risk Scores in ITCS* (2016), disponible en: <https://arxiv.org/pdf/1609.05007.pdf>. Consultado el 21 de Octubre 2019.

⁶⁴ Esta es una de las conclusiones más interesantes del reciente texto de Kearns y Roth: Thus if we ask for accuracy but don't mention fairness, we won't get fairness. Kearns, Michael y Roth, Aaron, *The Ethical Algorithm* (Oxford University Press, 2019, New York), p. 87

suerte de sistema preventivo en contra de las personas sospechosas de cometer delitos en el futuro. Por increíble que parezca, esto es exactamente lo que ha ocurrido²³. Robert McDaniel fue visitado por la policía en 2013 sin haber cometido jamás un delito, porque el departamento de policía de la ciudad de Chicago decidió llevar adelante un programa de prevención algorítmica del delito y él fue seleccionado como una de las cuatrocientas personas que más probablemente cometería uno en el futuro cercano²⁴. En China, de hecho ya existe una política de prevención de futuros delincuentes²⁵. Cuando, basado en determinados antecedentes evaluados algorítmicamente, se determina la posibilidad alta de futura comisión de un delito por parte de un ciudadano, este es detenido y enviado a un campo de "re-educación."

El punto más importante es que los jueces, al igual que los demás órganos del Estado, deben someter su acción a la Ley y la Constitución (art. 6° CPR), por lo que sus actos deben estar enmarcados dentro de las normas jurídicas que nuestro ordenamiento establece, en la forma que dichos preceptos establecen. Esto implica que en cada caso concreto deberán revisar si es que la situación se adapta al tipo legal, y una vez verificado esto, aplicarán la sanción o consecuencia jurídica pre-establecido por la norma al hecho evaluado. Si actúan guiados por correlaciones arbitrarias, sus actos serán reflejos de estas correlaciones y no constituirán, en sí mismos, aplicaciones de normas. Los algoritmos de *deep learning* no aplican normas, de hecho, los mecanismos de tipo conectivista no son capaces de entenderlas ni de manipular símbolos de manera de realizar una subsunción, por lo que serán siempre, esencialmente, ajurídicos. En este sentido, al servirse de ellos se renuncia a la aplicación de reglas y, por tanto al Derecho como tal. Esto no es exclusivo de los jueces, sino común a toda la administración del Estado en cuanto que el órgano ejecutivo debe aplicar el Derecho vigente.

La Unión Europea, al respecto, fijó una serie de principios que deben orientar la aplicación de una inteligencia artificial confiable, entre los que destaca el de explicabilidad²⁶. De conformidad al mismo, los modelos de inteligencia artificial deben

producir resultados explicables directa o indirectamente para el afectado. Ahora bien, detrás de esta manifestación de buena voluntad, permanece el problema técnico de la explicabilidad. Hoy por hoy, no es posible asegurarla y, mientras el estado de cosas permanezca así, los agentes de inteligencia artificial seguirán manteniendo un fondo de azar respecto a las razones que orientan sus decisiones y, por tanto, estarán en contraposición a los principios que regulan el actuar del Estado.

Un posible alternativa es desarrollar sistemas de carácter mixto, donde el agente artificial no tome una decisión, sino que simplemente haga una recomendación al ser humano que, finalmente decide. Esto es lo que se denomina tener un ser humano en la ecuación (*human on the loop*). El problema de esta alternativa, es que muchas veces parece que el ser humano esté ahí específicamente para ser culpado si algo sale mal. Si un algoritmo predice con un 90% de probabilidad la reincidencia de un condenado, ¿se atreverá el juez humano a liberarlo? Al menos deberá alegar en su decisión otras razones que no sean el juicio de la máquina para proceder, pues en caso contrario su resolución carecería de fundamento, puesto que simplemente considerar el juicio del algoritmo es basarse en un mecanismo oscuro e inexplicable.

Ahora bien, respecto a modelos que no estén basados en el *deep learning*, sino en sistemas predominantemente simbólicos, como los viejos sistemas expertos (*expert systems*), su funcionamiento es explicable y justificable, toda vez que dichos algoritmos simplemente aplican reglas pre-programadas en ellos. De hecho, muchos de ellos pueden dar instantáneamente razón de sus predicciones. Por ello, el uso de tales modelos no presenta este inconveniente. El problema, como vimos en su oportunidad, es que tales sistemas expertos son difíciles de programar, en cuanto requieren que se codifique manualmente el conocimiento de un jurista en la máquina. Aún así, puesto que reducir una norma a código implica interpretarla, salvo que quien la programe tenga arrogada la interpretación general de la misma (como el legislador, o, en materia tributaria, el Servicio de Impuestos Internos), el resultado será siempre discutible, en cuanto la interpretación ofrecida lo es. Es decir, los modelos simbólicos son caros de programar y sus resultados resultan discutibles, por lo que en general fueron abandonados, salvo para áreas específicas como la declaración de impuesto a la renta.

Un problema final que presenta esta sección consiste en plantear un problema específico respecto al sesgo, y este es su corrección. En principio, una vez que se detecta que un algoritmo funciona con un sesgo, o incluso, al momen-

require special attention. In those circumstances, other explicability measures (e.g. traceability, auditability and transparent communication on system capabilities) may be required, provided that the system as a whole respects fundamental rights. The degree to which explicability is needed is highly dependent on the context and the severity of the consequences if that output is erroneous or otherwise inaccurate"

High-Level Expert Group on Artificial Intelligence, Ethic Guidelines for Trustworthy AI, disponible en: <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai> Consultado el 21 de Octubre de 2019.

²³ O'Neil, Cathy, Weapons of mass Destruction. How Big Data Increases Inequality and Threatens Democracy, p. 102.

²⁴ Véase el reporte de prensa en: <https://www.theverge.com/2014/2/19/5419854/the-minority-report-this-computer-predicts-crime-but-is-it-racist> Consultado el 21 de Octubre de 2019.

²⁵ Véase información de prensa al respecto en The Washington Times disponible en: <https://www.washingtontimes.com/news/2019/jun/25/chinas-chilling-pre-crime-prison-indoctrination-sy/> Consultado el 21 de Octubre de 2019.

²⁶ "Explicability is crucial for building and maintaining users' trust in AI systems. This means that processes need to be transparent, the capabilities and purpose of AI systems openly communicated, and decisions – to the extent possible – explainable to those directly and indirectly affected. Without such information, a decision cannot be duly contested. An explanation as to why a model has generated a particular output or decision (and what combination of input/factors contributed to that) is not always possible. These cases are referred to as "blackbox" algorithms and

to de diseñarlo y establecer su función de éxito, pueden introducirse factores que lo mitiguen o, incluso, lo hagan desaparecer. Por ejemplo, si el algoritmo de selección de currículos descarta sistemáticamente aquellos pertenecientes a mujeres, puede incorporarse dentro de la función de éxito el tender hacia una selección equilibrada entre hombres y mujeres. El principal problema de esta solución es que requiere identificar los sesgos que puede contener nuestra base de datos e ser cuidadosos a la hora de diseñar la función de éxito, toda vez que es necesario hacerlo conscientemente. En esto existen dos peligros, el primero es la posible pérdida de eficacia del sistema, y el segundo es la selección del conjunto de valores que quieren incorporarse al mismo.

Respecto a la eficacia del sistema⁹⁷, existe un riesgo real de perder eficacia predictiva si es que se intentan omitir o corregir determinados elementos que son contrarios a la lógica normativa. Volvamos al ejemplo relativo a Lex ex Machina de 2011. Este modelo fue diseñado para predecir las resultas de los litigios relativos a propiedad intelectual y logró realizar su función con alta eficacia. Ahora bien, al analizar los factores que pueden determinar⁹⁸ el resultado de la sentencia, una de las correlaciones más relevantes está dada por la identidad de los abogados que representan a las partes. Esta correlación funciona porque los buenos abogados hacen defensas responsables y son capaces de utilizar argumentos inteligentes, es decir, la identidad de la firma de abogados es un sustituto estadístico (*proxy*) de una buena defensa. Si quisiéramos transformar el sistema Lex ex Machina en un aplicador de justicia (algo para lo que no fue diseñado), evidentemente contendría un sesgo a favor de ciertos abogados que equivaldría a la prevaricación, si es que un juez humano actuase de esa manera, por lo que habría que corregir ese punto de datos o su función de éxito a fin de obtener resultados aceptables. Ahora bien, el problema es que perdería fuerza predictiva, toda vez que la buena calidad de la defensa no puede ser medida directamente y la identidad del equipo de defensa no puede ser reemplazada fácilmente como sustituto estadístico. En pocas palabras, hay dos opciones, o se tiene precisión en los resultados en la etapa predictiva o se obtienen resultados más justos en la etapa de aplicación. Supongamos que un modelo matemático determinase que las personas de nombre Carlos tienen una tendencia a no pagar sus créditos, por una misteriosa correlación matemática. Pensemos ahora en que se determina la existencia de la obligación, aunque su cumplimiento (el pago) está en entredicho. El algoritmo, posiblemente, decida que yo no he cum-

plido, toda vez que las personas de nombre Carlos no cumplen rutinariamente sus obligaciones. Aunque la correlación sea efectiva, yo tengo derecho a que se me apliquen criterios normativos a la hora de decidir si es que he cumplido con esa obligación o no, y una simple correlación matemática no es suficiente para decir por acreditado mi incumplimiento, incluso en caso que yo efectivamente no haya pagado en tiempo y forma. En pocas palabras, la existencia de un sistema jurídico implica la aplicación de normas, y una mera correlación matemática no es tal. La pregunta es, dónde queremos, como sociedad, vivir, en un lugar donde el Estado de Derecho impere o no. Desde la Antigüedad, ya Aristóteles defiende al gobierno de las leyes⁹⁹, frente incluso al mejor de los hombres. El gobierno de las leyes es el fundamento de la democracia moderna, y en su esencia, la libertad consiste en poder hacer lo que las leyes permiten¹⁰⁰, de manera que renunciar a un sistema normativo que regule la conducta de los individuos en sociedad implica alejarse del concepto de libertad que ha servido de base a nuestra tradición jurídica política.

Por último está el delicado problema de la corrección de los sesgos. Efectivamente, cuando se detecta un sesgo, en ocasiones (no siempre) es posible eliminarlo. Para ello usualmente debe examinarse la función de éxito del algoritmo. ¿Cuál es la finalidad del mismo? Si su finalidad es contratar empleados de características similares a los que ya trabajan en la empresa (en cuyo caso reproducirá los sesgos que contenga la plantilla), o modificar tal plantilla en algún sentido (aumentando su diversidad, disminuyendo el ausentismo laboral, o lo que sea). Si se quiere variar sobre la base de datos que alimenta al agente, será menester considerarlo expresamente y programar este tipo de elementos dentro de la función de éxito, pero es menester poder visualizarlos para corregirlos. Si se prohíbe, por ejemplo, indicar el sexo en el currículo, pero el algoritmo de selección de antecedentes, de manera inesperada, segrega en razón de sexo (utilizando sustitutos estadísticos de éste) a las mujeres, como ocurrió en el caso de Amazon, es difícil corregir el sesgo sin tener un indicador explícito sobre la calidad de hombre o mujer del candidato. Amazon debió desechar su algoritmo simplemente porque la ausencia de mención sobre el sexo en los currículos impedía corregirlo de manera aceptable. Extrañamente, la corrección de sesgo implica la visualización de la inequidades que un grupo humano presenta, y en este sentido, es un ejercicio de sinceridad.

Ahora bien, una pregunta evidente en este problema es qué valores deben incorporarse dentro de un modelo a fin que corrija los sesgos que genere. Esta es una decisión eminentemente política y moral. En un sistema constitucional axiológico, donde se incorporan valores al contenido normativo de la Constitución, la respuesta ha de ser, a lo menos para el aparato público, el sistema ético que respalda a la Constitución, especialmente en su título primero, Bases de la

⁹⁷ Véase al respecto: Kleinberg, Jon M., Mullainathan, Sendhil, and Raghavan, Manish, *Inherent Trade-Offs in the Fair Determination of Risk Scores in ITCS* (2016), disponible en: <https://arxiv.org/pdf/1609.05807.pdf>. Consultado el 21 de Octubre 2019.

⁹⁸ Surdeanu, Mihai, Nallap, Ramesh, Gregory, George, Walker, Joshua and Manning, Christopher, *Risk Analysis for Intellectual Property Litigation in Proceedings of the 13th International Conference on Artificial Intelligence and Law* (New York, ACM, 2011), 116-120, disponible en: <https://nlp.stanford.edu/pubs/ical11.pdf>. Consultado el 21 de Octubre de 2019.

⁹⁹ Arist. Pol. 3.9-19, 1286a.

¹⁰⁰ *Libertas in legibus consistit* Cic. De I. Arg. 2.102.

que tienen una cierta capacidad de auto programarse y optimizar por sí mismas su propio código¹⁰⁷, lo cual nos aleja bastante de cualquier paradigma tradicional. Una alternativa que se ha planteado, a fin de solucionar esta dificultad, es otorgarle algún grado de personalidad a tales entidades, solución que exploraremos al final de este capítulo, pero que en principio no parece encontrar consenso. A lo largo de las próximas secciones trataremos el problema asumiendo la falta de personalidad de los agentes artificiales.

La pregunta relevante desde el punto de vista del Derecho es justamente cuándo los actos ejecutados por tales entidades han de producir efectos en el mundo del jurídico, y en caso de hacerlo, cuáles. En principio, creemos que hay dos tipos de actos que pueden preocupar al mundo del Derecho, los que aumentan el patrimonio y los que generan responsabilidad, los cuales deberemos tratar de explicar en sus efectos. Una vez dilucidados estos aspectos, entraremos en la posibilidad de crear una suerte de personalidad para tales entidades.

3.1. Actos de enriquecimiento material

En general los agentes artificiales pueden crear riqueza para quienes se sirven de ellos. Un robot que realiza labores industriales, producto de su actuar crea bienes y productos cuando es aplicado a una labor, al igual que cualquier máquina. En este sentido, los robots son cosas corporales, al ser perceptibles por los sentidos y no consistir, evidentemente, en derechos (565 CC). Ahora bien, respecto a las inteligencias artificiales que no están dotadas de una entidad física con la cual interactuar con el mundo, como un algoritmo que realiza predicciones sobre precios, por ejemplo, ¿se le aplican las reglas de los bienes corporales? En nuestra opinión, los algoritmos no son derechos, sino objetos de derechos. La distinción entre bienes corporales e incorporeales que contempla nuestro Código Civil es fruto de una larga tradición histórico-dogmática¹⁰⁸. En principio, el artículo consagra dos criterios diversos para calificar la corporalidad de los bienes, su perceptibilidad y su consistencia. La perceptibilidad basta para que el objeto sea considerado corporal, pero para tener el carácter de incorporeales se requiere superar dos barreras diversas, la no perceptibilidad y la consistencia jurídica, es decir, que tenga naturaleza de derecho subjetivo. En el caso que el bien sea imperceptible, pero no tenga una naturaleza jurídica, como una entidad

incorporea, aparentemente no calzaría en la categorización y suele ser catalogado como intangible o bien inmaterial, aunque su estatuto jurídico suele ser algo indeterminado. ¿A qué categoría han de pertenecer los agentes artificiales? En principio, siguiendo estrictamente la lógica del articulado del Código Civil, estos deberían ser bienes corporales, toda vez que pueden ser expresados de manera perceptible por los sentidos, mostrando, por ejemplo, su código fuente, ante lo cual la vista sería el sentido con que se perciben. Por otra parte, no son derechos, no tienen una consistencia jurídica, aunque sobre ellos pueden establecerse diversas relaciones jurídicas que sí tendrían el carácter de incorporeales.

En principio hay un conjunto de creaciones que pueden recibir protección jurídica. Dicha regulación es tratada en nuestro sistema civil como una especie de propiedad¹⁰⁹ (art. 584 CC), aunque más precisamente la propiedad intelectual se comporta como una suerte de monopolio establecido a favor del autor de una obra o del creador de una idea para que éste la explote en exclusiva o designe a quienes pueden legítimamente explotarla. En general, los derechos de autor son "derechos que, por el solo hecho de la creación de la obra, adquieren los autores de obras de la inteligencia en los dominios literarios, artísticos y científicos, cualquiera que sea su forma de expresión" (art.1 L.17336).

Los algoritmos, en este sentido, son objeto de derechos de autor¹¹⁰ (que si serían bienes incorporeales), por lo que, aunque su naturaleza sea corporal, los derechos de autor que su creador ejerce sobre ellos son de carácter incorporeales y debemos tratarlos como tales. Una particularidad al respecto es que los derechos sobre programas computacionales es que su autoría se reputa pertenecer a "las personas naturales o jurídicas cuyos dependientes, en el desempeño de sus funciones laborales, los hubiesen producido" (art.8 L.17336). Es decir, aunque una persona natural hubiese programado íntegramente un algoritmo, si este fue hecho trabajando para una compañía, será ésta quien sea estimada titular original de la misma.

El autor (titular original, en la nomenclatura del art.7 de la L.17336) puede, evidentemente, ceder sus derechos a un titular secundario que actuará como titular de la misma, para los efectos de la adquisición de los bienes que la entidad genere. En este sentido, si producto de su actuar un agente artificial genera cosas corporales (565 CC), se le aplicarán las reglas tradicionales de la accesión, considerándose los productos o frutos, sean naturales o civiles (art. 643 CC) y se hacen de ella bajo éste título. Si se constituye un usufructo sobre el agente arti-

¹⁰⁷ Esto es lo que se conoce como *Deep Coding*, que se está desarrollando experimentalmente mientras se escriben estas líneas. Turner describe el fenómeno en los siguientes términos: *Some AI systems are able to edit their own code - the equivalent of biological entity being able to change its DNA. One example of this is a program built in 2016 by a team of researchers from Microsoft and Cambridge University. Turner, Jacob, Robot Rules. Regulating Artificial Intelligence* (Palgrave Macmillan, London, 2019), L.1990

¹⁰⁸ Se origina en las Institutas de Gayo (Gai 2.12-14) y pasa de ahí a las Instituciones de Justiniano (I.2.2). Aunque no quedó en la codificación francesa, Bello tomó el comentario de Delvincourt (Delvincourt, Cours de Code Civil (Dijon, Videcoq, 1834), 7.1, p. 139, 143) y lo aplicó en su redacción a nuestro Código.

¹⁰⁹ A mediados del siglo XIX se debatía la naturaleza de la protección que debía otorgarse a las creaciones artísticas y las patentes de invención. Bello, en el artículo 582 del Código Civil, adhirió a la tesis francesa que lo concebía como una suerte de propiedad, tomando la redacción precisa del Código Sardo (1837).

¹¹⁰ El artículo 3º N°16 de la L.17336 declara como especialmente protegidos: "Los programas computacionales, cualquiera sea el modo o forma de expresión, como programa fuente o programa objeto, e incluso la documentación preparatoria, su descripción técnica y manuales de uso."

ficial, los frutos pertenecerán al usufructuario de acuerdo a las reglas generales (781, 790 CC), y lo mismo ocurrirá en caso de arrendamiento (1919 CC), donde el arrendatario se hará, en principio dueño de los mismos. En general, cualquier figura contractual que podrían aplicarse para la cesión del goce de la cosa, sigue idénticas reglas.

Así, en el caso de un algoritmo que realice predicciones financieras, por ejemplo, el algoritmo en sí estará bajo derechos de autor de su creador (titular original), bajo titularidad secundaria de quien lo adquiera de éste, y las predicciones que realice serán sus frutos, que podrán ser comercializados por la persona que tuviere derecho a estos frutos, en caso de encontrarse bajo arriendo o usufructo.

Ahora bien, el problema se complejiza si consideramos ciertas entidades que pueden generarse (y se generan) a través de inteligencias artificiales, como por ejemplo, imágenes, textos y otras obras similares que tradicionalmente se vinculan a la propiedad intelectual¹¹¹. Así, por ejemplo, hay cuadros pintados por robots que se han rematado por precios altos en casas de remate europeas como Christie's¹¹², hay música compuesta por algoritmos e incluso obras literarias de diverso tipo, como artículos de prensa y hasta guiones cinematográficos. Una pregunta interesante es si existen derechos de autor respecto a tales composiciones. La respuesta se encuentra, en buena medida, indeterminada. En principio, las regulaciones relativas a la propiedad intelectual usualmente se protegen el esfuerzo humano y su ingenio, teniendo como objetivo que éste fructifique a través de la protección exclusiva del mismo. Ahora bien, los agentes artificiales ejecutan actos que para los seres humanos requerirían de esfuerzo e ingenio, por lo que queda en el aire la pregunta de si tales actos merecen la misma protección que tendrían en caso que una persona natural los hubiese ejecutado.

En el medio Norteamericano, la respuesta parece ser negativa. En principio, la Copyright Office de Estados Unidos señaló que sólo podían registrarse las obras creadas por seres humanos¹¹³, siendo las demás no registrables. En este sentido, no parece que, de momento, pueda establecerse propiedad intelectual sobre éste tipo de obras.

En el medio británico, por su parte, parece inclinarse por la posición contraria, esta vez que la Copyright Designs and Patents Act de 1988 establece expresamente, en su artículo 9 que en el caso de una obra literaria, dramática, musical o artística que ha sido generada computacionalmente, el será tenido como autor la persona por quien se hayan encargado y coordinado la creación de la obra¹¹⁴. Es decir, el autor, en principio, sería quien encarga al algoritmo la creación. No obstante, no parece particularmente transparente aún qué significa encargar la obra y realizar los *necessary arrangements* a que la ley se refiere, por lo que existe un grado de indeterminación al respecto. El artículo 178 de la misma señala que una obra está computacionalmente generada cuando la obra ha sido producida por una computadora sin que exista un autor humano¹¹⁵. En pocas palabras, al intentar si una obra es de creación humana o informática, es menester, primero, intentar detectar si es que existe un ser humano a quién se pueda atribuir la obra, y en caso contrario, esta pertenecería a quien la encargó y coordinó.

Respecto a la Unión Europea, no parece haber una posición común armónica en esta respecto a la materia, aunque las regulaciones relativas a autoría parecen partir de la base de un creador humano¹¹⁶.

Algo similar ocurre con nuestra propia legislación, donde en el artículo 5 de la L. 17336 se habla exclusivamente de personas naturales como creadores de obras individuales y en colaboración (letras a y b), lo que parecería excluir la posibilidad de que agentes artificiales generasen propiedad intelectual. Ahora bien, para los programas computacionales se admite como autor a personas jurídicas e incluso se señala expresamente que éstos sean producidos por encargo de un tercero, a quien se entienden cedidos los derechos de autor (art. 7, L.17336). En este sentido, parece existir una extraña desarmonía en nuestro sistema. Si un agente artificial crease un programa computacional, por aplicación del artículo 7 de la Ley 17336, esta se entendería pertenecer a la persona natural o jurídica bajo cuyo encargo se hubiese generado, por lo que los derechos de autor serían del titular (primario o secundario) de la misma. Ahora bien, y qué sucede si la obra no es un programa computacional, como por ejemplo, una obra literaria. Por lo demás, es una práctica relativamente común hoy en día que sean algoritmos los que redacten artículos de prensa de baja complejidad. ¿Esos artículos pertenecen al diario cuyo algoritmo los compuso (como titular primario o secundario), o están libres de derechos de autor y pueden ser copiados sin problemas? Si el autor fuese humano, se le aplicarían las disposiciones del artículo 24 de la L.17336 que establecen que son de propiedad del medio para quien trabaja el periodista, o del medio de difusión al que encomen-

¹¹⁴ In the case of a literary, dramatic, musical or artistic work which is computer-generated, the author shall be taken to be the person by whom the arrangements necessary for the creation of the work are undertaken. Art. 9 Copyright Designs and Patents Act de 1988.

¹¹⁵ "computer-generated," in relation to a work, means that the work is generated by computer in circumstances such that there is no human author of the work. Art. 178, Copyright Designs and Patents Act de 1988.

¹¹⁶ Turner, Jacob, *Robot Rules. Regulating Artificial Intelligence* (Palgrave Macmillan, London, 2019), 13126.

¹¹¹ Turner lo expresa en los siguientes términos:
For instance, where an AI system writes a best-selling book, or creates a valuable work of art, questions arise as to who owns the relevant property. Turner, Jacob, *Robot Rules. Regulating Artificial Intelligence* (Palgrave Macmillan, London, 2019), 11723.

¹¹² Véase un comentario de la casa de remates en:
<https://www.christies.com/features/A-collaboration-between-two-artists-one-human-one-a-machine-0332-1.aspx>

¹¹³ Consultado el 1 de Noviembre de 2019.
"The U.S. Copyright Office will register an original work of authorship, provided that the work was created by a human being."

¹¹⁴ Vid Compendium of US Copyright Office Practices, de 29 de Septiembre de 2017, en particular los capítulos 306 y 313.2, disponible en:
<https://www.copyright.gov/comp3/docs/compendium.pdf>
Consultado el 1 de Noviembre de 2019.

dó el trabajo, pero si la obra es de creación automática esto no es tan claro, aunque podría intentarse la aplicación analógica de estas normas.

¿Y respecto a imágenes? En principio, éstas son de propiedad del fotógrafo, salvo que se encuentre trabajando para un medio de difusión o que las tome bajo su encargo. Si las imágenes fueron tomadas o creadas por un agente artificial, ¿tienen protección? Tampoco hay claridad en este punto.

¿Y si se trata de pinturas? De acuerdo a las reglas de la accesión ésta será como bien corporal, se propiedad de su titular primario o secundario, según sea el caso. Pero, ¿tiene derechos de autor, a su vez? Esto no es claro.

¿Y qué hay de los derechos morales derivados de la autoría? ¿Y en el caso que el agente artificial sea el interprete, por ejemplo, ejecutando al piano una obra de Debussy? ¿Tendrá su propietario derecho al pago de remuneraciones por la reproducción de las ejecuciones musicales que se realice de un video del robot ejecutando la obra? En estas materias, claramente se necesita una definición sobre el alcance de los derechos de autor y del ejecutante, así como establecer con mayor precisión cuál es el ámbito de aplicación de tales derechos.

Otra forma de propiedad intelectual es la propiedad industrial, que abarca marcas, patentes de invención, modelos y otras similares. En este caso, no parecen existir mayores problemas respecto a las mismas, toda vez que para que sean protegidas basta con su registro, cuestión que oportunamente deberá regular quien tenga la titularidad del agente. La ley no parece fijarse en que sean específicamente de creación humana, sino en la función, originalidad y el grado de innovación que representan. En este sentido, si un algoritmo descubriese la cura de una enfermedad, no parece que existan problemas para que la compañía que se sirve de él patentase la invención. Otro tanto podríamos decir de la marca, que no tiene por qué ser diseñada por un ser humano para estar protegida.

En fin, como regla general, podríamos establecer que los beneficios que genere un agente artificial serán para quien sea su titular (original o secundario) mediante la aplicación de la simple accesión de frutos del Código Civil, y dichos beneficios pueden incluir bienes corporales o, incluso, propiedad intelectual. Podemos ahora la responsabilidad.

3.2. Responsabilidad por los actos de agentes artificiales.

En principio, los agentes artificiales pueden ejecutar actos que, en caso de ser realizados por personas naturales o jurídicas, generarían responsabilidad jurídica. En principio, esta responsabilidad se sintetiza en la figura de las obligaciones, que constituyen la concreción jurídica de la existencia de un deber de realizar una conducta respecto a otra parte.

En principio, nuestro sistema jurídico considera la existencia de las obligaciones excepcional, por lo que para que consideremos que una parte se encuentra

en la necesidad de dar hacer o no hacer alguna cosa a favor de otra, debe existir una causa que genere tal necesidad, entre las cuales encontramos, tradicionalmente, al contrato y al delito. Esta enumeración que hunde sus raíces en el derecho romano es, evidentemente, incompleta¹¹⁷. Durante la misma Antigüedad se señaló que las obligaciones podían provenir tanto de los contratos, como de pactos similares a ellos (*quasi ex contractu*), o bien de los delitos o de otras causas equivalentes¹¹⁸. Ahora bien, desde la perspectiva de la Antigüedad, estas enumeraciones eran más bien irrelevantes, toda vez que su sistema de responsabilidad era típico, por lo que se generaban obligaciones cada vez que las partes ejecutasen una conducta pre-descrita en su sistema jurídico normativo, sea que esta pudiese calificarse como contrato, como delito o como cualquier otra cosa¹¹⁹. En pocas palabras, era la tipicidad la que generaba las obligaciones, por lo que el problema romano se centraba, más bien, en determinar si es que la conducta se adaptaba al tipo.

El mundo moderno fundamenta su estructura obligacional en bases diversas, toda vez que la tipicidad romana colapsó en la estructura sistemática racionalista de los siglos XVII y XVIII. Una vez que se establece una teoría general del contrato merced a Luis de Molina, la libertad contractual surge con Grocio, y de ahí en adelante se irá reconociendo que cualquier forma que cumpla con los requisitos de convencionalidad tendrá la virtualidad de crear obligaciones. Otro tanto ocurrirá con el mundo de los delitos, que se desprenderán de su tipicidad primaria como fuentes de obligaciones, para terminar en constituirse en una especie de generalización del *damnum* aquiliano, donde toda conducta lesiva realizada con dolo o culpa generará la obligación de indemnizar el daño. Esta evolución terminará en cristalizar en la tetra partición clásica de las fuentes de las obligaciones que contempla nuestro Código Civil en su artículo 1437, tomada a su vez del 1370 del Code Napoléon y éste, a su vez, de Pothier¹²⁰. Esta estructura tradicional implica que las obligaciones tienen cinco causas estereotipadas que pueden generárlas, los contratos, los delitos, los cuasicontratos, los cuasidelitos y la ley. La diferencia entre ellas está siempre mal definida y da lugar a controversias.

¹¹⁷ Su enunciacón más antigua la encontramos en Gayo, quien la señala por primera vez en G.3.88 *Nunc transeamus ad obligationes, quarum summa divisio in duas species adducitur: omnis enim obligatio uel ex contractu nascitur uel ex delicto.*

No obstante el propio Gayo, dos parágrafos más adelante nos señala que hay obligaciones que no parecen provenir de ninguna de estas alternativas, como en el pago de lo no debido (G.3.91).

¹¹⁸ En particular son las instituciones de Justiniano las que realizan tal precisión. Así, U 3.13.2 señalan:

Sequens divisio in quatuor species deducitur: aut enim ex contractu sunt aut quasi ex contractu aut ex maleficio aut quasi ex maleficio.

¹¹⁹ En efecto, el mismo Gayo aparece citado en el Digesto señalando que las obligaciones pueden nacer de los contratos, de los delitos o de varias otras causas y figuras:

D.44.7.1 pr. *Obligaciones aut ex contractu nascuntur aut ex maleficio aut proprio quodam iure ex variis causarum figuris.*

¹²⁰ Pothier, Robert Joseph, *Traité des Obligations*, in Bugnet, *Ouvres de Pothier Annotées et mises en corrélation avec le code civil et la législation actuelle*, t.3 (Paris, Videcoq, 1847), p. 3.

En efecto, la distinción entre delitos y cuasidelitos tiene lugar exclusivamente por la posición subjetiva del agente, en un caso suponiendo el dolo y en el otro la culpa, pero sus efectos son lo suficientemente acotados como para tratarlos conjuntamente como un sistema único de responsabilidad, que hoy se denomina extracontractual. En cuanto a los contratos, parece que tienen un régimen lo suficientemente circunscrito como para aplicarles un régimen propio de responsabilidad, que la doctrina denomina contractual. Hoy en día, y desde que Alessandro introdujese en el medio nacional, esta distinción se ha transformado en la piedra angular sobre la que se fundamenta la responsabilidad, no obstante presentar defectos importantes en su estructura que deja indeterminada la situación de los cuasicontratos y de las obligaciones que nacen de la Ley. En cierto sentido, hemos vuelto a la categorización gayana, con los mismos defectos que ya eran evidentes para el jurista del siglo II.

Sea como fuese, en general podemos distinguir un grupo de obligaciones que nacen de actos de carácter convencional o cuasi convencional, mientras que otras emanan de hechos que ocasionan daños a terceros. La pregunta más relevante en esta etapa consiste en determinar si estas figuras se aplican a los actos realizados por agentes artificiales o no. Para ello dividiremos nuestro tratamiento en dos acápites, uno dedicado a los contratos y otro destinado a las reglas relevantes para delitos y cuasidelitos.

3.2.1. Contratos

En principio el fundamento de un grupo de obligaciones se encuentra en la existencia de contratos, esto es, actos de carácter convencional que crean obligaciones (art. 1438 CC). El contrato funda su existencia en el acuerdo entre las partes. Aunque hay algo de oscuridad respecto al momento en que tales figuras se perfilan en el ámbito romano, el mundo moderno toma su noción de contrato de cierto pasaje del Digesto¹²¹ en que se destaca la necesidad de que en todos los contratos existe convención o consenso, como también de otro relativo a la compraventa que señala la necesidad de que éste verse sobre sus elementos más relevantes, esto es, la existencia de un contrato de compraventa, la cosa vendida y el precio¹²². La tradición contractualista en época moderna¹²³ genera-

los estos elementos, lo que desemboca en Pothier construyendo su visión del contrato esencialmente voluntarista¹²⁴, que en último término será recogida por la tradición codificadora decimonónica. De acuerdo a ésta un contrato es un acto esencialmente voluntario, en que el querer de las partes converge para generar una o más obligaciones. En este sentido, cada parte debe ser legalmente capaz para consentir (1445 CC) y se considera que toda persona, en principio, lo es, salvo que la Ley la haya privado de esta capacidad (1446 CC) por algún motivo particular, como por estar interdicto por demencia o prodigalidad, por ejemplo.

Ahora bien, el problema que encontramos a la hora de tratar de agentes artificiales es que éstos son simples algoritmos y no personas, por lo que su capacidad de consentir es dudosa. Esto conlleva diversos problemas a nivel teórico y práctico de los que debemos hacernos cargo. En primer término, debemos considerar quién consciente y en qué momento este se desarrolla. En segundo término, está el problema de los vicios del consentimiento. En tercer término, está el problema de la limitación de la responsabilidad.

Durante la década de 1990 comenzó a plantearse el problema de la formación del consentimiento por medios electrónicos. El planteamiento básico de la doctrina fue que la existencia de medios técnicos nuevos no tenía por qué variar la comprensión de las formas jurídicas básicas, que gráficamente fue expuesta por Easterbrook en su famosa declaración relativa al Law of the Horse¹²⁵. Aunque los caballos aparecen en las relaciones jurídicas comúnmente, esto no significa que por ello debamos crear el Derecho del caballo; lo mismo respecto al comercio electrónico, aunque aparezca en nuestras relaciones jurídicas frecuentemente, no por ello variaremos sustancialmente las reglas relativas al consentimiento.

En principio, acertadamente, la doctrina estimó que los medios electrónicos no eran más que una nueva manera de formar el consentimiento, de manifestarlo, y nada más. Los dispositivos electrónicos serían meros medios de expresar la voluntad, y lo mismo da que se haga por carta, teléfono, correo electrónico o cualquier otro medio¹²⁶, por lo que la fuerza vinculante de dichos acuerdos es equivalente a aquella que tienen los contratos celebrados por medios tradicionales. Esta perspectiva, correcta atendida el desarrollo tecnológico de la época, quedó consagrada en numerosos textos que regulaban la materia durante la primera década de los años 2000. En este sentido, se tendió a establecer en diversos mecanismos para equiparar el consentimiento que se entrega de ma-

¹²¹ D.2.14.1.3 *Conventionis verbum generale est ad omnia pertinens, de quibus negotii contrahendi transigendique causa consentiunt qui inter se agunt; nam sicuti convenire dicuntur qui ex diversis locis in unum locum colliguntur et veniunt, ita et qui ex diversis animi motibus in unum consentiunt, id est in unam sententiam decurrunt. Adeo autem conventionis nomen generale est, ut eleganter dicat pedius nullum esse contractum, nullam obligationem, quae non habeat in se conventionem, sive re sive verbis fiat: nam et stipulatio, quae verbis fit, nisi habeat consensus, nulla est.*

¹²² D.18.1.9 *pr. In venditionibus et emptionibus consensum debere intercedere palam est: ceterum sive in qua emptione dissentient sive in pretio sive in quo año, emptio imperfecta est.*

¹²³ Ésta, como ya hemos dicho, comienza en Molina. Vid. Molina, Ludovico, De Justicia et de Jure (Coloniae Allobrogum, 1763, Marcus, Michaelis Bousquets) t.2, Dip. LXXXI, p. 4.

¹²⁴ Pothier, Robert Joseph, *Traité des Obligations*, in Bagnet, *Ouvrages de Pothier Annotés et mises en corrélation avec le code civil et la législation actuelle*, 1.3 (Paris, Videcoq, 1847), p. 3.

¹²⁵ Easterbrook, Frank, 'Cyberspace and the Law of the Horse', *University of Chicago Legal Forum* (1996), 207-2015.

¹²⁶ En este sentido, véase el completo estudio de Pinochet Clave, Ruperto, *La Formación del Consentimiento a través de las Nuevas Tecnologías de la Información I*, in *Ius et Praxis* 10-2 (2004), pp. 267-320. Para un análisis acerca del surgimiento de la doctrina, véase Chopra, Samir and White, Laurence F., *A Legal Theory for Autonomous Artificial Agents* (University of Michigan Press-Kindle, Michigan, 2011), p. 38.

nera tradicional y electrónica. Entre ellos destaca la Convención acerca del Uso de Comunicaciones Internacionales en Contratos Internacionales de 2005 (revisada y suscrita por Chile) que establece que un contrato no será privado de efectos por el solo hecho de haber sido realizado electrónicamente¹²⁷. En consecuencia, un contrato formado por la interacción de un sistemas automatizados de mensajes no puede ser privado de fuerza vinculante por el sólo hecho que no han intervenido personas naturales o físicas en su conclusión¹²⁸.

Algo similar señala la Directiva 2000/31/CE sobre Comercio Electrónico de la Unión Europea señala en su artículo 9 que velará por que su legislación permita la celebración de contratos por vía electrónica¹²⁹. En pocas palabras, los actos realizados por vía electrónica son considerados válidos como si fuesen realizados por vías tradicionales. En este mismo sentido, el artículo 3° de la Ley N.° 19.799, sobre Documentos Electrónicos, Firma Electrónica y Servicios de Certificación, prescribe que:

"Los actos y contratos otorgados o celebrados por personas naturales o jurídicas, suscritos por medio de firma electrónica, serán válidos de la misma manera y producirán los mismos efectos que los celebrados por escrito y en soporte de papel."

Es decir, las vías electrónicas son equivalentes a las tradicionales y no debería realizarse ninguna diferencia al respecto. La única particularidad detectable está

establecida a favor de los consumidores. Así, cuando una de las partes del contrato es un consumidor el prestador deberá enviarle confirmación escrita de la operación con todas las particularidades del negocio (art. 12° L.19496).

Ahora bien, los problemas surgen cuando consideramos que los actos de los agentes artificiales hoy no se comportan como meros instrumentos de comunicación, sino que realizan actos de mayor extensión, que incluyen la calificación del contrato y la fijación de las condiciones esenciales del mismo. Los agentes artificiales, hoy por hoy, no son medios de trasladar la voluntad de una parte hacia la otra, como en la era del correo electrónico, sino que éstos cuentan con una autonomía creciente a la hora de evaluar el contenido económico y jurídico de los negocios que ejecutan. Cada vez que una persona se asoma a un portal de compras y busca un producto, desde la ofertas que recibe hasta las condiciones de entrega del producto están fijadas de manera algorítmica. Si se consultase a la compañía por qué tal o cual persona en concreto recibe éstas u otras ofertas, por la indeterminación que involucra el uso de modelos de *deep learning*, lo más seguro es que no pueda dar razones concretas y específicas, sino meras explicaciones generales probables, más no seguras. Cuando se le exige el cumplimiento de los contratos a que llega el agente artificial, estamos partiendo de la base que ésta ha consentido en tales condiciones, las cuales, en la práctica, no conoce. Hace relativamente poco, por ejemplo, la compañía aérea Latam se excusó de cumplir con el contrato de pasaje a que sus agentes artificiales, a través de la red, habían llegado con diversos individuos, alegando que su algoritmo había cometido un "error"¹³⁰.

En este sentido, han surgido opiniones que tienden a articular las relaciones entre los contratantes y los agentes artificiales desde una perspectiva que incluye esta capacidad que tienen para obrar independientemente o agencia (*agency*)¹³¹. Así, las relaciones entre el agente artificial y su titular serían equivalentes a aquellas surgidas del mandato, donde una parte obra por cuenta y riesgo de la otra¹³². La ventaja de esta idea es que expresa mejor la realidad de las relaciones

¹²⁷ United Nations Convention on the use of Electronic Communications in International Contracts, art. 8.1.
A communication or a contract shall not be denied validity or enforceability on the sole ground that it is in the form of an electronic communication.
Disponible en:
https://www.uncitral.org/pdf/english/texts/electcom/08-57452_Ebook.pdf
Consultado el 7 de Noviembre de 2019.

¹²⁸ United Nations Convention on the use of Electronic Communications in International Contracts, art. 12:
A contract formed by the interaction of an automated message system and a natural person, or by the interaction of automated message systems, shall not be denied validity or enforceability on the sole ground that no natural person reviewed or intervened in each of the individual actions carried out by the automated message systems or the resulting contract.
Disponible en:
https://www.uncitral.org/pdf/english/texts/electcom/08-57452_Ebook.pdf
Consultado el 7 de Noviembre de 2019.

¹²⁹ Directiva 2000/31/CE sobre Comercio Electrónico de la Unión Europea, artículo 9.1:
Los Estados miembros velarán por que su legislación permita la celebración de contratos por vía electrónica. Los Estados miembros garantizarán en particular que el régimen jurídico aplicable al proceso contractual no entorpezca la utilización real de los contratos por vía electrónica, ni conduzca a privar de efecto y de validez jurídica a este tipo de contratos en razón de su celebración por vía electrónica.
Disponible en:
<https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32000L0031&from=EN>
Consultado el 7 de Noviembre de 2019.

¹³⁰ En el caso, la página web de Latam ofreció el 26 de Marzo de 2018 pasajes a Australia a cien mil pesos chilenos (entorno a 120 dólares de la época, en circunstancias que el precio regular es significativamente más alto). Muchos consumidores procedieron a comprarlos para luego encontrarse con que la compañía decidió no cumplir el contrato alegando un error. Esta procedió a realizar la devolución del monto pagado en contra del querer de los clientes que desearían realizar el vuelo.

¹³¹ DeMott, Deborah A., *Disloyal Agents*. Alabama Law review 58 (2007), pp. 1049-1067; Chopra, Samar and White, Laurence F., *A Legal Theory for Autonomous Artificial Agents* (University of Michigan Press-Kindle, Michigan, 2011), p. 18; Furnio Shimpo, *The Principal Japanese AI and Robot Strategy toward Establishing Basic Principles in Barfield, Woodrow y Pagallo, Ugo* (Eds.), *Research Handbook on the Law of Artificial Intelligence* (Cheltenham, 2018, Edward Elgar), pp. 114-142.

¹³² En este sentido, señala Turner que:
indeed, this is how much trading occurs online, where automated programs are mandated to buy, sell and do on behalf of people and companies. Turner, Jacob, *Robot Rules. Regulating Artificial Intelligence* (Palgrave Macmillan, London, 2019), 12887.

entre el titular y su agente artificial, donde el primero se limita poner al segundo a cargo de determinadas funciones de contratación que deberán ser desarrolladas dando lugar a toda la riqueza de relaciones y posibilidades asociadas a las relaciones entre ambos, incluyendo disposiciones relativas al error.

No obstante, en los sistemas continentales, el principal problema que presenta esta visión es que la capacidad de obrar fue concebida como un atributo de la personalidad, y puesto que los agentes artificiales no gozan de ella, parece imposible aplicarla directamente.

En el derecho codificado del siglo XIX, la capacidad de contratar está íntimamente ligada a la personalidad. Esta idea viene directamente de Pothier¹³³ y fue consagrada en el antiguo artículo 1123 del Code Napoléon (hoy modificada) y que fue adoptada por el art.1446 del Código Civil de Bello. Esto tiene lógica en un sistema donde la todos los seres humanos son personas, por lo que no existe la posibilidad técnica que un individuo que desee concluir un contrato este privado de personalidad. Ahora bien, en el Derecho antecedente esto no era así, toda vez que la equiparación de la personalidad legal y la humanidad sólo se da cuando la esclavitud es suprimida¹³⁴. En el pasado, cuando existía la esclavitud, era posible concebir que un individuo concluyese un contrato sin ser él mismo un sujeto de Derecho, como era el caso en que un amo ponía a un esclavo a cargo de la gestión de uno o más de sus negocios¹³⁵. Por lo demás, los esclavos, a semejanza con los agentes artificiales, son objeto de propiedad y de relaciones jurídicas diversas que pueden establecerse sobre ellos, como arrendamiento o usufructo, derivando de sus actos consecuencias jurídicas para sus amos, sea por vía de enriquecimiento, sea por vía de responsabilidad. Así, el paralelo entre agentes artificiales y esclavos es cercano, tanto que la misma palabra robot proviene del checo *robots* que designa al esclavo. Siguiendo esta línea de pensamiento es que Kerr¹³⁶, ya durante la década del 1990, planteó la posibilidad de aplicar a los agentes artificiales el elegante marco jurídico romano relativo a la esclavitud romana, que permanece inmanente pero dormida en nuestra tradición jurídica, ahora que los hombres no asumimos esta trágica posición. Ahora bien, los agentes artificiales en la década de 1990 no eran lo suficientemente sofisticados como para justificar una teorización que rebasase la aproximación del mero instrumento, por lo que su artículo quedó fuera de la corriente principal relativa a la institucionalización de las relaciones del comercio electrónico. Dado el actual desarrollo de las tecnologías de redes neuronales y de la autonomía con que los

agentes artificiales actúan en el mundo económico, vale la pena volver sobre sus presupuestos y analizar su viabilidad.

En la Roma de la Antigüedad, los esclavos carecían de patrimonio y eran objeto de derechos ejercidos por sus amos. Esto los colocaba fuera de lo que hoy llamaríamos personalidad legal, toda vez que al carecer de capacidad jurídica, sus actos no podían producir efectos en su propio patrimonio. Esto no significaba que los esclavos no pudiesen concluir válidamente contratos. Ellos eran frecuentemente utilizados como agentes patrimoniales de sus amos, actuando en el mundo jurídico por cuenta de ellos¹³⁷. De hecho, era relativamente común que fueran los esclavos los administradores de diversos negocios que pertenecían a sus amos. En numerosas ocasiones, de hecho, encontramos a esclavos en posición de gerentes o administradores de negocios, y este tipo de esclavo, por sus capacidades económicas, resultaba sumamente valioso.

En principio, los actos de una persona no obligan a otra, salvo que la persona que participa en el contrato cuente con la orden (*iussum*) de la que se quiere obligar¹³⁸. En este sentido, el pretor creó las acciones *adfectitiae qualitatis*, que permitían a quien había contratado con un esclavo (o un hijo de familia¹³⁹) dirigirse en contra de su amo para hacer valer dicha responsabilidad. La primera y más llamativa de estas acciones es la *actio quod iussum*¹⁴⁰, o acción por lo ordenado. En virtud de ella, el amo ha de asumir la responsabilidad patrimonial por los contratos celebrados por su esclavo por el hecho de haberle ordenado realizar una acción o haberlo puesto a cargo de un negocio. Esto significa que por el sólo hecho de poner al esclavo como agente patrimonial suyo en una materia determinada, el amo queda obligado por los actos de éste, con independencia de la voluntad específica que concurriese en el negocio. Basta con poner a cargo de un negocio o conjunto de negocios a un agente, para asumir el riesgo de los actos que éste ejecute.

Esta idea es la que se encuentra detrás de algunos instrumentos que intentan otorgar una accionabilidad directa al titular de la inteligencia artificial, respecto a los actos y contratos que ésta ejecuta y celebra. En efecto, la Uniform Computer Information Transaction Act (UCITA) de Estados Unidos, que constituye una suerte de Ley Modelo para el comercio electrónico en Estados Unidos, dispone en su artículo 107 d) lo siguiente:

¹³³ U 3.17. *Servus ex persona domini ius stipulandi habet.*

¹³⁴ G.3.103; U 3.19.3; D.2.14.274, D.50.1760.

¹³⁵ En efecto, los hijos de familia se encontraban en una posición similar a la de los esclavos, toda vez que ellos tampoco tenían capacidad patrimonial mientras permaneciesen en dependencia de su pater, denominándose esta posición como *alieni iuris*. Sólo mediante la emancipación o la muerte del pater, estos adquirían la condición de *sui iuris*, y por tanto, capacidad patrimonial.

¹³⁶ Gal.4. 69. *Quis tamen superius mentionem habuimus de actione, qua in peculium filiorum familiae servorumque ageretur, opus est, ut de hac actione et de ceteris, quae eorundem nominem parentes dominosque dari solent, diligenter admonemus. 70. In primis itaque si iussu patris dominive negotium gestum erit, in solidum praetor actionem in patrem dominumve comparavit, et recte, quia qui ita negotium gerit, magis patris dominive quam filii servive fidem sequitur*

¹³³ Pothier, Robert Joseph, *Traité des Obligations*, in Bugnet, *Ouvres de Pothier Annotées et mises en corrélation avec le code civil et la législation actuelle*, t.2 (Paris, Videcoq, 1948), art.50, p. 28.

¹³⁴ Al respecto, véase Stagl, Jakob, *De cómo el Hombre Llegó a Ser Persona: los orígenes de un concepto jurídico-filosófico en el Derecho Romano* in REHU 45 (2015), pp. 373-401.

¹³⁵ Vid, al respecto, Stagl, Jakob, *La Concepción Gayana de "Persona" y la Capacidad Jurídica Parcial de los Esclavos* in *Revista General de Derecho Romano* 30 (2018).

¹³⁶ Kerr, Ian R., *Spirits in the Material World: Intelligent Agents as Intermediaries in Electronic Commerce*, in *Dalhousie L.J.* 190 (1999).

[Parte obligada por su agente electrónico] Una persona que usa un agente electrónico que ha seleccionado para realizar una autenticación, acto o acuerdo incluyendo la manifestación del consentimiento, está obligada por las operaciones del agente electrónico, incluso si ningún individuo estaba consciente de ellas o revisó las operaciones del agente electrónico o sus resultados¹⁴¹.

Esta posición es virtualmente equivalente a la antigua *actio de iussum romana* y debe ser tomada como base para el tratamiento de las operaciones realizadas por agentes artificiales. Desde nuestro punto de vista, el hecho de poner al agente artificial a cargo de un negocio es un acto unilateral voluntario de la persona en los términos del artículo 1437 del Código Civil, y en este sentido sirve de base al surgimiento de las obligaciones que ésta pacte para ella, en los términos del artículo 1450 del Código Civil, jugando esta orden el papel de pre-ratificación a sus actos. Se aplica, en este sentido, de manera analógica la norma del 2160 del Código Civil, que obliga al mandante a cumplir las obligaciones que a su nombre ha contraído el mandatario dentro de los límites de su mandato. Así, en la medida que el agente actúe dentro de los límites del negocio sobre el que ha sido puesto a cargo, este acto de encargar un negocio actúa como fuente de la obligación del titular para cumplir con los actos que por él ejecute el agente.

Ahora bien, vale la pena preguntarse acerca de qué ocurre si es que el agente se extralimita o actúa de una manera completamente inesperada operando sobre situaciones que no estuviesen, en principio, comprendidas dentro de la esfera de su actuar. Si bien en principio esto es raro, como hemos visto antes, puede ocurrir que las inteligencias artificiales actúen de forma inesperada. En más, el creciente nivel de sofisticación de las mismas incrementa este riesgo, toda vez que su autonomía se ve aumentada. En el caso de agentes humanos, puede ocurrir que las inteligencias artificiales actúen de forma inesperada. En todo caso, esta perspectiva es difícil de encuadrar dentro del plano de la inteligencia artificial. En principio, un error es una falsa representación de la realidad, y requiere de la existencia de un estado consciente. Los agentes artificiales con que hoy contamos, no tienen consciencia y, en este sentido, no son capaces de representarse la realidad de ninguna manera relevante para el Derecho. En este sentido, técnicamente, no son capaces de error, toda vez que no tienen una representación de la realidad. Tienen datos almacenados, pero esto difícilmente se puede equiparar a una representación de la realidad, lo cual nos lleva a que calificarla como falsa no parece apropiado. Ahora bien, lo que sí puede resultar es que producto de un diseño deficiente o de un entrenamiento con bases de datos poco confiables, podrían realizar actos inesperados o en desacuerdo con las instrucciones de quien se sirve de ellos. Así, por ejemplo, un agente de compras, como un consumidor algorítmico,

hubiese detectar la falta de leche en mi refrigerador y, en lugar de comprar leche de vaca, comprar una bebida de coco que se denominase leche de coco. En este caso, el agente no ha, verdaderamente, cometido un error, sino más bien se ha extralimitado en su actuar. Por ello, el titular del agente podría desligarse de la responsabilidad por sus actos, en la medida que pruebe la desviación del agente.

No obstante, si esta desviación resulta en un enriquecimiento para el titular, a pesar de haber actuado éste fuera de sus atribuciones, debiese el titular ser responsable hasta la concurrencia de su enriquecimiento. En efecto, en Derecho romano este era el caso del amo cuyo esclavo actuaba sin instrucciones o más allá de sus atribuciones. En este caso, se otorgaba la *actio in rem verso* a la persona con que había interactuado el esclavo para que se dirigiese contra el amo hasta el monto del provecho que había obtenido en la transacción¹⁴². En caso contrario, existiría un enriquecimiento sin causa para el titular del agente, que no sería responsable de las obligaciones contraídas por éste, más se haría con los beneficios de sus actos. Si bien en nuestro Código Civil no existe una consagración expresa del enriquecimiento sin causa como generador de responsabilidad, la jurisprudencia al respecto es antigua y se encuentra bien asentada.¹⁴³

En caso que dicha desviación hubiese sido causada maliciosamente, esto es, con dolo, por ejemplo a través de la intervención en la operación del agente, nuevamente el titular no sería obligado por los actos de su algoritmo, pero la necesidad de compensar los daños causados al titular por la vía de dicha intervención debiesen repararse por la vía de la responsabilidad contractual por dolo, en los términos del artículo 1458.

Finalmente, y a fin de evitar la incerteza que esto podría causar en las partes que interactuaban con el esclavo, en Derecho romano se permitía al amo constituir un peculio a favor de su esclavo. Éste era una cantidad de bienes que el esclavo administraba libremente por su cuenta y que se engrosaba con las ganancias que obtenía. La persona que contrataba con el esclavo podía siempre dirigirse en su contra, a través del amo, por el monto de su peculio¹⁴⁴, a pesar que el esclavo hubiese actuado sin orden del amo. El interés que presenta esta

¹⁴¹ D.14.5.1, D.15.1.1

¹⁴² Así, sólo por citar la más reciente, la Corte Suprema define el enriquecimiento injusto en los siguientes términos:

"En cuanto al principio del enriquecimiento injusto, se lo liga con una noción predominantemente económica, rechazando el legislador que se obtenga tal ganancia sin causa jurídica. Se pretende evitar, de esta manera, que una persona se enriquezca a costa de otra si no puede justificar jurídicamente este beneficio, por lo que si se reúnen los presupuestos que le hacen procedente, se origina la obligación de restituir lo adquirido ilegítimamente o, más exactamente, de indemnizar el empobrecimiento ajeno." Sentencia de la ECS de 13 de Agosto de 2019, "Jofré con Santander", rol 1245-2018, considerando 6°.

¹⁴⁴ D.11.1.9, B. Si de peculio agatur, non oportere responderi a patre vel domino, an in potestate habeat filium vel servum, quia hoc solum quaeritur, an peculium apud eum cum quo agitur est. En similar sentido: Gai. 4.72a-74.

¹ Section 107(d) [Party Bound by Its Electronic Agent] A person that uses an electronic agent that it has selected for making an authentication, performance, or agreement, including manifestation of assent, is bound by the operations of the electronic agent, even if no individual was aware of or reviewed the agent's operations or the results of the operations.

figura es la posibilidad de establecer un fondo que limite la responsabilidad del titular respecto a los actos de su agente artificial.

Esta posibilidad no existe directamente en nuestro sistema jurídico, aunque parece una idea interesante el poder establecer una separación patrimonial respecto al agente artificial, donde, por un lado, se limite la responsabilidad del titular respecto a los actos inesperados del agente, y por otra, se asegure el cumplimiento de todas las obligaciones contraídas por tal agente en relación a terceros. Usualmente, este problema se enfrenta a través del otorgamiento de derechos de personalidad a favor del agente, y ésta es una de las alternativas que contempla la *European Parliament Resolution with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103 (INL))*¹⁴⁶. Sin entrar a este difícil tema de momento, nos limitamos a señalar que establecer una separación patrimonial especial de bienes destinados a hacer frente a los actos de un agente artificial no requiere, en sí mismo, de establecer un nuevo tipo de personalidad y que, tal alternativa debiese ser estudiada, pues presenta una manifiesta utilidad a la hora de hacer frente a las obligaciones contraídas por el agente, como también limitar los posibles daños que su actuar pudiese eventualmente provocar. Es más, se ha propuesto servirse de las sociedades de responsabilidad limitada para este efecto, creando una sociedad administrada íntegramente por un agente artificial¹⁴⁷. Si bien esta alternativa parece limitada a la regulación nacional concreta de cada país, tal vez estudiar la forma de establecer una suerte de sociedad unipersonal para los algoritmos de más común uso podría ser posible, o tal vez sea viable, incluso, servirse de los mecanismos actualmente vigentes para tal efecto.

Conviene, antes de cerrar el acápite, señalar que si una de las partes en la relación jurídica creada es un consumidor, a éste le acompaña el derecho a retracto para poner término unilateralmente a la relación contractual dentro de los diez días contados desde la recepción del producto por el sólo hecho de haberse realizado por vía electrónica (Art. 3 bis b) Ley 19496), sea que para este acuerdo la parte se haya servido de agente artificial o no. En este sentido, se otorga siempre una protección superior a quien reviste la calidad de consumidor por parte de nuestra legislación.

3.2.2. Delitos y cuasidelitos

Como hemos precisado anteriormente, las obligaciones —y por tanto la responsabilidad— no sólo puede emanar de los contratos y del mundo lícito en general, sino también de los delitos y cuasidelitos, que en nuestro sistema jurídico

ocupan el lugar de lo ilícito y que generan lo que desde fines del siglo XIX se denomina responsabilidad extracontractual.

En general, todo daño que pueda imputarse a malicia o negligencia de otra persona debe ser reparado por esta (art. 2329 CC, tomado del antiguo 1382 del Code Napoléon), de manera que cuando el acto cometido se debe a su mala intención o dolo, estamos ante un delito, mientras que si se debe a su negligencia o culpa, será un cuasidelito. En este sentido, la mayor parte de los sistemas jurídicos del siglo XIX, al momento de codificar, establecieron sistemas de responsabilidad delictuales basados en la culpa y el dolo, esto es, en la posición subjetiva del agente al momento de actuar. Esta idea se encuentra tomada del pasado romano, específicamente de la *lex Aquilia* que estableció el delito de daños dentro del contexto jurídico romano y que fue interpretada por la jurisprudencia tardo-republicana e imperial como gatillante de una responsabilidad que depende, esencialmente, de que exista una acción que cause daños y que tal acción sea imputable a culpa o dolo del agente. Cumplidos tales requisitos, nace la obligación de reparar el daño y, por ende, la responsabilidad jurídica.

Lamentablemente, este edificio conceptual no resulta del todo apropiado para construir un sistema de responsabilidad para los titulares de agentes artificiales que causen daños a terceros en el ámbito de sus operaciones. A la aplicabilidad de un modelo subjetivo surgen diversas objeciones que terminarían por convertir la responsabilidad en sumamente excepcional y, cuando no, en residual. En primer término, los agentes artificiales no son capaces ni de culpa ni de dolo, toda vez que ellos, en sí mismos, no son más que mecanismos inconscientes de manipulación de símbolos. Sin consciencia, es difícil estimar que pueden tener intención, y mucho menos calificarse tal estado volitivo como malicioso. En cuanto a la culpa, ésta generalmente es determinada comparando la conducta del sujeto con un estándar de actuar, como el buen padre de familia en Roma (D. 18.1.35.4) o en nuestro Derecho (art. 44 CC). No obstante, es difícil establecer un estándar para un agente artificial. ¿Qué sería un buen algoritmo?¹⁴⁷ Una alternativa evidente es juzgar la conducta del titular del agente para establecer si su conducta es dolosa o negligente, estableciendo su responsabilidad en base a su acto de poner en acción a la entidad. Ahora bien, esto es más difícil de lo que parece. Por supuesto que hay casos donde el actuar del titular del agente puede ser calificado como doloso, como cuando intencionalmente libera un virus computacional o cuando éste se sirve de un agente artificial para obtener datos personales o penetrar el sistema computacional de un tercero. En dichas situaciones la conducta puede constituir un delito informático penal en los términos de la Ley 19233¹⁴⁸ y,

¹⁴⁶ Disponible en http://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_EN.htm#title1 Consultado el 19 de Noviembre de 2019

¹⁴⁷ Bayern, Shawn, "The Implications of Modern Business-Entity Law for the Regulation of Autonomous Systems" *European Journal of Risk Regulation* 2 (2016), pp. 297-309

¹⁴⁷ Véase al respecto Abbot, Ryan, *The Reasonable Computer: Disrupting the Paradigm of Tort Liability*, *The George Washington Law Review* 86-1 (2017), pp. 101-143

¹⁴⁸ Esta Ley es bastante somera y es algo antigua. Existe un proyecto que duerme en el Congreso desde el año 2015 que recomienda su actualización, pero no parece avanzar. Disponible en: <https://www.camara.cl/pdf.aspx?prmTIPO=DOCUMENTOCOMUNICACIONCUENTA&prmID=11020>

consecuentemente, constituir un delito civil. No obstante, fuera de estos casos relativamente residuales, es difícil establecer una conducta dolosa por parte del titular. La culpa presenta dificultades similares, toda vez que la determinación de la misma implica el establecimiento de estándares de conducta para los titulares de agentes que son difíciles de conocer fuera del relativamente estrecho ámbito de los expertos. Así, por ejemplo, si un automóvil auto conducido atropella a un peatón, a quién debiese exigírsele el estándar de conducta. ¿Al dueño del vehículo? ¿Al fabricante? ¿A todos? ¿Y en qué consistiría tal estándar? Recordemos, además que a través del *deep learning* un modelo podría modificarse a sí mismo y hacer la relación entre su diseño original y el que está actualmente en uso bastante distante. En este sentido, puede que originalmente el agente haya estado dentro de los estándares de cuidado exigidos, pero que éstos ya no se cumplan por sus sucesivas modificaciones.

Todo el proceso requiere pruebas y análisis que van más allá de lo que un particular podría razonablemente prestar, por lo que la posición relativa de la víctima, a lo menos dentro del marco de la responsabilidad aquiliana tradicional, es sumamente débil. Es por ello que se han buscado alternativas que intenten superar dicho marco e, idealmente, objetivizar tal responsabilidad.

Una de las primeras alternativas tendientes a objetivizar la responsabilidad de los agentes artificiales es aplicarles las disposiciones de la responsabilidad por el hecho de las cosas, en particular aquélla correspondiente a los animales¹⁵⁴. Esta fue desarrollada en el Derecho francés a partir del artículo 1385 del *Code Napoléon*¹⁵⁵, que recoge parcialmente la casuística romana sobre la *actio de pauperiem* relativa a los daños que los cuadrúpedos pueden causar. Al efecto, en caso que los animales de una persona causen daños a otro, toda vez que estos animales no son capaces de responsabilidad delictual¹⁶¹, se hacía responsable al dueño de los mismos. En este sentido, la responsabilidad por el hecho del animal recaía en quien lo detentase, con independencia de su culpa o dolo. De hecho, si su titular hubiese azuzado al animal para causar los daños (hipótesis de dolo) la culpa¹⁶², cesaba la acción y debía conducirse por las vías regulares¹⁶³, pues su

objetivo era, justamente, la persecución de la responsabilidad por actos del animal no directamente imputables al dueño. Este modelo de responsabilidad por hechos de los animales fue recibido por nuestro artículo 2326 del Código Civil, que recoge, en esencia, la vieja doctrina romana de la *actio de pauperiem*. Ahora bien, el artículo 2326 contempla, en su inciso segundo, incluso la posibilidad de accionar en contra de quien se sirve de un animal y la posibilidad de éste de repetir en contra del dueño del mismo, esto último cuando el daño provenga de un defecto inherente al animal. Por último, los de daños causados por animales peligrosos (art.2327 CC) responde su dueño siempre, sin importar su culpa.

Si aplicamos estos elementos a los agentes artificiales, el régimen que emerge es bastante razonable. Un titular de la misma responderá por los daños que haya causado, sin importar la posición subjetiva del agente (que simplemente no tiene ninguna) o del titular. Quien se sirva de la misma –como los titulares secundarios– deberán también responder de los daños que causen, pudiendo, en tal caso, repetir contra el titular original si es que prueban que los daños estaban relacionados con defectos de diseño o entrenamiento. En definitiva, quien causa el riesgo, sirviéndose del agente, responde de los daños que éste cause.

Otra alternativa que se ha planteado es recurrir a la doctrina del *respondet superior*, esto es, que responda el superior¹⁶⁴. Dicha doctrina tiene también orígenes romanísticos y consiste en que por los actos de los esclavos y dependientes ha de responder el amo. En Roma, los actos del esclavo generaban una acción real, esto es, el amo tenía la opción entre responder por los daños o entregar al responsable. Dicho régimen, con la revolución francesa y el cese de la esclavitud, fue reemplazado por la responsabilidad vicaria del amo, que deberá afrontar los daños causados por sus subordinados¹⁶⁵. Este régimen fue recepcionado por los artículos 2320 a 2322 de nuestro Código Civil. Ahora bien, la diferencia fundamental entre ésta responsabilidad y la derivada de los hechos de los animales, es que aquí la limitación se encuentra establecida en que el agente, para generar responsabilidad del amo, debe actuar “en el ejercicio de sus respectivas funciones”, es decir, si el agente se extralimita, la responsabilidad cesa. A nosotros, por su amplitud y simpleza, nos parece más adecuada la *actio de pauperiem*, aunque no es claro qué opción puedan tomar los tribunales en definitiva.

Ahora bien, las alternativas planteadas no excluyen la posibilidad de que se establezcan otras figuras de responsabilidad para casos específicos donde el riesgo sea manifiesto y común. Tal vez el caso más paradigmático sean los vehículos auto conducidos. Al respecto se han planteado diversos regímenes de compensación respecto a los accidentes que puedan causar. Al respecto, Nueva

Consultado el 27 de noviembre de 2019.
 Turner, Jacob, Robot Rules. Regulating Artificial Intelligence (Palgrave Macmillan, London, 2019), 116-4.
 Art. 1385 CN *Le propriétaire d'un animal, ou celui qui s'en sert, pendant qu'il est à son usage, est responsable du dommage que l'animal a causé, soit que l'animal fût sous sa garde, soit qu'il fût égaré ou échappé.*
 D.9.1.13. [Ulpianus libro 18 ad edictum] *Ait praetor "pauperiem fecisse." Pauperies est damnum sine iniuria facientis datum: nec enim potest animal iniuria fecisse, quod sensu caret*
 D.9.1.14 *Itaque, ut Servius scribit, tunc haec actio locum habet, cum commota feritate nocuit quadrupes, puta si equus calcitrosus calce percussent, aut bos cornu petere solutus petenti, aut mulae propter nimiam levitiam: quod si propter loci iniquitatem aut propter culpam mulionis, aut si plus iusto onerata quadrupes in aliquem onus evertent, haec actio cessabit damnumque iniuriae agetur.*
 D.9.1.16. [Ulpianus libro 18 ad edictum] *Sed et si instigatu alicuius sine damnum dederit, cessabit haec actio.*

¹⁵⁴ Turner, Jacob, Robot Rules. Regulating Artificial Intelligence (Palgrave Macmillan, London, 2019), 127-60; Chopra, Samir and White, Laurence F., A Legal Theory for Autonomous Artificial Agents (University of Michigan Press-Kindle, Michigan, 2011), p. 128.
¹⁶⁵ Art. 1384 Code Napoléon *Chacun est responsable du dommage qu'il a causé non seulement par son fait, mais encore par sa négligence ou par son imprudence.*

Zelanda optó por la compensación a todo evento llevada adelante por un fondo destinado al efecto al cual contribuyen todos los titulares de un agente de estas características¹⁵⁶. Este modelo de "no-fault compensation" opera como un seguro mutuo, donde ante la ocurrencia de un evento, se recurre a un fondo común destinado al efecto. El Reino Unido, por su parte, en 2018, promulgó la *Automated and Electric Vehicles Act*, por la cual exige a estos vehículos un seguro obligatorio que se haga cargo de los accidentes que se causen mediante su uso. La mayor diferencia entre uno y otro modelo, es que en el caso de Nueva Zelanda no hay derecho a continuar con acciones en contra del titular del vehículo, mientras que el Reino Unido tiene un modelo de seguros tradicional donde las acciones que emanan del accidente pueden dirigirse fuera del ámbito de los seguros¹⁵⁷.

Ahora bien, ¿puede dirigirse la víctima en contra del creador del agente, esto es, quien haya construido y diseñado el agente por el accidente que éste causó? De momento, hemos establecido que, en principio, si aplicamos el estatuto de los animales a los agentes artificiales, el dueño será responsable de los daños que cause, como también la persona que se sirve de él. Ahora bien, el dueño, en materia de agentes artificiales, debe entenderse como el titular primario u original, en los términos de la Ley de Propiedad Intelectual, quien cede su uso a los titulares secundarios a quienes vende o cede su tenencia. Es decir, el "dueño" del agente no es estrictamente el usuario que se sirve de ellos, sino la compañía que los diseña y comercializa. En este sentido, si una persona adquiere un programa de Alphabet, ésta sólo le cede su uso en los términos establecidos por el contrato de licencia, reservándose el dominio sobre la propiedad intelectual sobre el agente. Así, en último término, se aplicaría el inciso segundo del artículo 2326 del Código Civil que establece:

"Lo que se dice del dueño se aplica a toda persona que se sirva de un animal ajeno, salva su acción contra el dueño, si el daño ha sobrevenido por una calidad o vicio del animal, que el dueño con mediano cuidado o prudencia debió conocer o prever, y de que no le dio conocimiento."

El titular secundario o usuario sería responsable de los daños causados por el agente, pero tendría una acción en contra del titular original del agente si este daño se debe a un defecto en la construcción del mismo o de su entrenamiento que debió ser prevista por el titular original. En caso que el titular secundario fuese un consumidor, debiese interpretarse tal calidad o vicio del producto de conformidad al art. 20 de la Ley 19496, que da una serie de supuestos en que el consumidor tiene derecho a la reparación del bien, su reposición o la restitución

del precio pagado. En efecto, dicho artículo da un catálogo de vicios en virtud de los cuales objetivamente el bien adquirido por el consumidor puede estimarse defectuoso. Si el agente de que se sirve el consumidor puede ser estimado defectuoso, y por tal defecto causó daños a terceros, el consumidor tendría derecho a dirigirse en contra del titular original para que responda por los daños que le ha causado en su patrimonio. Si el producto comercializado es un agente artificial del cual pueden derivarse riesgos, como un vehículo auto conducido, el importador o fabricante deberá ponerlos en conocimiento del público en los términos del artículo 46 de la Ley 19496. Ahora bien, si se determina que el agente artificial es un producto o servicio es peligroso, sea por la autoridad judicial, sea por otra entidad administrativa competente, la responsabilidad por los daños que cause será solidaria entre el que lo comercialice, lo importe o produzca (art. 47 Ley 19496). De esta manera, cuando nos encontramos ante materias de consumo y el agente puede ser calificado como peligroso, la responsabilidad se extiende fuera del cauce normal del artículo 2326 del Código Civil y abarcaría a más responsables.

4. PERSONALIDAD PARA LOS AGENTES ARTIFICIALES

A fin de establecer la responsabilidad por los contratos y delitos que un agente artificial genere y de radicar los derechos sobre los bienes que produzca, una alternativa que debe ser explorada es la posibilidad de entregar personalidad a los agentes artificiales. Dicha idea no es nueva, sino que fue explorada en la ciencia ficción ya en la década del 1970, cuando Isaac Asimov escribió su célebre *Homage to the Bicentenario*. Se encuentra sobre la mesa en la gracias a la resolución del Parlamento Europeo¹⁵⁸ que recomienda otorgar algún tipo de estatuto de personalidad a los agentes artificiales a fin de hacer valer su responsabilidad y de la ciudadanía que Arabia Saudita otorgó al robot Sophia de Hanson Robotics en 2017. No obstante, es un tema polémico, donde incluso se ha acusado de especismo a los detractores de tal posibilidad¹⁵⁹.

En verdad, antes de entrar al debate, nos parece evidente señalar que lo que hoy entendemos por personalidad no es un atributo exclusivo de los seres humanos. Hay entidades artificiales, como las personas jurídicas, que no gozan de ella, mientras que los seres humanos, en su totalidad, no la han detentado sino hasta la abolición de la esclavitud (art. 55 CC). Incluso hoy, puede sostenerse que hay seres humanos que no gozan de personalidad, como el feto antes de su paso por el canal de parto y nacimiento, el cual sólo es protegido por la ley, pero cuya personalidad comienza con el nacimiento (art. 74 CC). Sin entrar en los complejos vericuetos del aborto, es difícil sostener que un feto de nueve meses de gesta-

¹⁵⁶ Turner, Jacob, *Robot Rules. Regulating Artificial Intelligence* (Palgrave Macmillan, London, 2019), 128-135.

¹⁵⁷ *Insurance does not alter underlying legal liabilities. Rather, it redirects the liability to pay damages away from the person who caused harm (if any) to the insurer.* Turner, Jacob, *Robot Rules. Regulating Artificial Intelligence* (Palgrave Macmillan, London, 2018), 130-131.

¹⁵⁸ European Parliament Resolution with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103 (INL)).

¹⁵⁹ Ryder, Richard, "Speciesism Again: The Original Leaflet" *Critical Society* 2 (2010).

ción no es un ser humano y que sólo por la magia del canal de parto llega a serlo unos minutos después, cuando ha nacido.

La personalidad, como hoy la entendemos, parece ser un aglomerado de atributos jurídicos con que se dota a determinadas entidades, algunas por reconocidas como seres humanos, y otras por la conveniencia empírica de efectuar separaciones patrimoniales. Así, los llamados atributos de la personalidad (nombre, capacidad, domicilio, estado civil, patrimonio) son una amalgama jurídica con que se dota a una entidad para que pueda ser sujeto de Derecho, es decir, ser titular de bienes, derechos y obligaciones. Amén de lo anterior, y desconectado de esto, existe una protección que se otorga a determinadas entidades no por el hecho de ser personas, sino por el simple hecho de ser lo que son, por su naturaleza, en la nomenclatura tradicional. Los seres que reconocemos como pertenecientes a la especie humana están dotados de Derechos Humanos, los que podemos fundamentar en diversas teorías, en su dignidad esencial, en su carácter trascendente, en una especie de acuerdo tácito, lo que sea, pero se los reconocemos al margen de cualquier deliberación acerca de su personalidad, y el hecho que los calificamos como personas naturales es una derivación de esta protección. Hay otras entidades que históricamente también han detentado estatutos de protección, como los bosques sagrados, los templos y algunos animales, algunos de los cuales han sido suficientes como para actuar válidamente en el mundo del derecho, y en este sentido se puede sostener que gozaron de personalidad. Hoy en día se quiere dotar a determinadas entidades naturales, como a ciertos animales o a la naturaleza, de derechos. Esto no implica reconocerles personalidad, sino simplemente otorgarles una protección jurídica determinada, que en algunos casos puede incluso ser equivalente a la de un ser humano. Las personas jurídicas, en cambio, gozan de la facultad de actuar en el mundo jurídico de manera equivalente a los seres humanos, aunque no detentan derechos humanos, toda vez que no pertenecen a nuestra especie.

El punto que deseamos enfatizar es que la personalidad no implica la existencia de Derechos Humanos, toda vez que puede otorgarse sin incluirlos en ella, o puede entregarse una protección equivalente a los derechos humanos sin entregar personalidad. Incluso, es posible que hayan humanos sin personalidad, pero no por ello carentes de Derechos Humanos.

Cuando hablamos de personalidad para los agentes artificiales, lo que se varía es la posibilidad de otorgarles un conjunto de atributos de la personalidad (nombre, domicilio, capacidad, patrimonio) que asegure su actuar en el mundo jurídico de manera independiente a cualquier otro sujeto de derecho. En este sentido, el debate es, más bien, similar al relativo a si las sociedades, fundaciones y corporaciones deben tener una personalidad, es decir, la capacidad de adquirir bienes, derechos y obligaciones de manera independiente del patrimonio de sus asociados. Verdaderamente, la posibilidad de otorgarles alguna forma de personalidad de los agentes artificiales tiene por finalidad el establecer una separación patrimonial entre las consecuencias de los actos de éstos y el haber de

sus titulares, especialmente a fin de limitar su responsabilidad frente a terceros. En este sentido, la función que teóricamente cumpliría tal personalidad se asemeja al peculio con que se dota al esclavo, el cual se integra con los bienes que el amo le aporte y con el que éste siempre responde frente a terceros, dejando a salvo la responsabilidad del amo, salvo que haya actuado por su orden o que el acto lo haya beneficiado. Desde este contexto, dotarlos de personalidad es una manera de limitar dicha responsabilidad, teniendo en cuenta que, eventualmente, se podría levantar el velo societario de la personalidad para dirigirse en contra de los titulares reales del agente en ciertos casos.

Hay quienes piensan que incluso, sirviéndose de los medios de asociación actuales, es posible constituir sociedades de responsabilidad limitada administradas por un agente artificial y que, por tanto, sean la vestimenta jurídica adecuada para otorgarle personalidad a los agentes artificiales¹⁰³, aunque no es claro que el ordenamiento jurídico chileno lo permita.

Muy diversa es la pregunta respecto a si estas entidades debiesen estar protegidas por los derechos fundamentales que establece el orden constitucional. ¿Deben tener los modelos algorítmicos libertad de expresión? ¿De asociación? ¿Tiene el Estado el deber de no discriminarlos? ¿Cuán férrea es la protección a su propiedad? Debe recordarse que, en último término, los seres humanos somos entidades conscientes y, por tanto, somos capaces de que gozamos, como el no ser sometidos a tratos degradantes, dicen relación con nuestra calidad humana, más no con nuestra personalidad, estrictamente hablando.

En caso de optarse por dar una suerte de personalidad a los agentes artificiales habría que preguntarse cuál sería su alcance y qué ámbitos alcanzaría. En primer término, es menester recordar que en este momento ningún modelo o algoritmo es otra cosa que un sistema de manipulación de símbolos. Así, al no estar dotados de consciencia, no son capaces de experimentar el mundo que los rodea. Puede que a través de cámaras capten imágenes y las almacenen y procesen en sus unidades de procesamiento, pero no experimentan el fenómeno de la visión. Los seres vivos, en cambio, al percibir experimentan el mundo, por lo que son víctimas del dolor y del placer según los estímulos que reciban. Nuestro sentido ético puede llevamos a empatizar con las entidades vivientes y, por tanto, a desear protegerlos del sufrimiento mediante la construcción de derechos y deberes, mas esto no es así en relación a los seres inanimados. Los agentes artificiales no sufren, no se emocionan ni se entristecen. Si algún día se desarrolla una subjetividad artificial y, con ella, una suerte de consciencia artificial, deberían revisarse estas nociones, pero de momento nada hay en este campo que apunte en tal dirección.

¹⁰³ Bayern, Shawn, "The Implications of Modern Business-Entity Law for the Regulation of Autonomous Systems" *European Journal of Risk Regulation* 2 (2010), pp. 297-309.

Atendidos tales elementos, puede resultar socialmente beneficioso establecer alguna suerte de separación patrimonial especial que abarque a modelos de inteligencia artificial. La ventaja que tal separación tendría es dotar al agente de capacidad de ejercicio directa y asegurar con dicho fondo su responsabilidad frente a las personas que interactúen con ellos en el mundo del Derecho. Para ello podría utilizarse una figura similar a la sociedad unipersonal, tal vez con un registro público y una inscripción de bienes y estimación de obligaciones realizadas a través de un mecanismo de tipo *blockchain* que se actualizara cada pocos minutos. Al mismo tiempo, es necesario tener en consideración que tal separación de bienes pertenece a una persona o conjunto de personas, respecto a la cual, en determinadas circunstancias, se debería levantar el velo de la personalidad para que la responsabilidad por los actos del agente recayese en ellos. El modelo para que la responsabilidad por los actos del agente recayese en ellos, el modelo sobre el cual actuar podría bien ser el peculio del esclavo, que era, en buenas cuentas, una separación parcial de patrimonio con la que el esclavo respondía de sus propias obligaciones cuando actuaba en el mundo del Derecho. No obstante, el acreedor podía dirigirse más allá de éste contra el amo cuando el esclavo actuase por orden del amo (*actio quod iussum*) o los actos del esclavo beneficiasen al amo (*actio in rem verso*). Un símil de tales acciones sería apropiado, toda vez que permitiría sobrepasar la barrera de responsabilidad que la personalidad impone cada vez que fuese socialmente beneficioso hacerlo.

Respecto a la idea de dotar a tales entidades con derechos fundamentales, la pregunta es compleja, toda vez que existen algunos que tendrían, propiamente, un objetivo comprensible, mientras que otros no. El derecho a la salud, francamente no se percibe en qué sentido podría ser de utilidad a una entidad que, en principio, no está siquiera viva. En cuanto a otros derechos como la libertad de asociación o la libertad de expresión, no es claro que merezcan una regulación constitucional. Si bien, en principio, no hay inconvenientes en que un algoritmo exprese sus resultados públicamente, si tales manifestaciones resultan ofensivas, no se divisa por qué no se podría silenciarlo. Baste recordar a Tay redactando mensajes a favor del racismo. Si Tay fuese un ser humano, será menester tolerar sus exabruptos, pero al ser un simple agente artificial, bien se puede silenciarlo, como Microsoft hizo a las pocas horas. Respecto a la asociación, si la participación de agentes artificiales en asociaciones de cualquier tipo llega a ser inconveniente, parece evidente que esto puede limitarse o prohibirse.

En cuanto a la propiedad, el tema es más complejo, toda vez que si se realiza una confiscación de los activos de un agente artificial, en verdad se priva a los titulares del mismo de los activos que el agente detenta por ellos. En este sentido, la protección del patrimonio del agente equivale a un resguardo de los bienes de los titulares del mismo que los detentan mediatamente a través del agente. Si realizar una separación patrimonial para constituir un agente artificial dotado de personalidad implica desproteger los bienes que se destinan a tal efecto, entonces sería arriesgado y torpe realizar tal separación. Creemos en este sentido, que dar a los bienes que detenta un agente artificial la misma protección que se otorga a una persona natural sirve una función socialmente útil, toda vez que a

través de ésta protección se resguardan los bienes de los titulares del agente. No obstante, esto conlleva un riesgo, pues si los bienes de una sociedad terminasen concentrándose en manos de agentes artificiales, los efectos sociales de tal fenómeno podrían ser tremendamente perjudiciales. Si algo así sucediese, tal vez sería más prudente cancelar la vigencia de la separación patrimonial y volver a integrar los bienes que detenta el agente a sus titulares y estudiar en ese caso si es necesario aplicar otro tipo de medidas, pero ahora respecto de personas previamente dotadas de derechos fundamentales.

En fin, lo que planteamos está lejos de los robots humanizados de Asimov. Simplemente vemos que puede resultar socialmente útil entregar patrimonio y capacidad de actuar a determinados agentes artificiales, para lo cual sería útil crear un registro donde, idealmente, se señalen también los bienes que integran tal separación. En definitiva, vemos la figura más cerca del antiguo peculio de los esclavos que de la calidad moral de que goza la humanidad. Si en el futuro distante se desarrollan entidades dotadas de consciencia, en ese momento habrá que analizar las consecuencias morales de tal autopercepción.

¿EN UN FUTURO DISTANTE?

1. INTRODUCCIÓN

Este es un capítulo difícil de escribir, toda vez que se refiere no tanto a lo que sucede actualmente con el desarrollo de las redes neuronales, sino a lo que ocurrirá en un futuro, en ocasiones cercano, a veces lejano. Durante esta investigación, detectamos elementos que podrían desarrollarse potencialmente, y que fueron llevados a la práctica mientras escribíamos los capítulos anteriores. Un ejemplo curioso ocurrió mientras hacía clases en un programa de magíster, en que exponía un uso que podría darse (y que de hecho se daba en otras latitudes) a los modelos de inteligencia artificial, sin saber que uno de los alumnos de hecho había desarrollado un producto similar en Chile y que comenzaba a comercializarlo. Este capítulo es uno que se fuga hacia la Historia del Derecho a gran velocidad, por lo que redactarlo es desafiante.

Básicamente, lo que buscamos es mostrar dos elementos. En primer término nos gustaría analizar qué impacto tendrá la inteligencia artificial en el ejercicio de la profesión de abogado, por una parte, y por otra, qué riesgos sociales pueden surgir de la aplicación de estas tecnologías al mundo jurídico en concreto. Así, el capítulo tendrá dos acápites fundamentales, uno dedicado a la profesión de abogado en específico (3.1) y otro a la sociedad en general (3.2). En la primera daremos nuestra visión acerca de los cambios que el ejercicio de la profesión experimentará en los próximos años, mientras que en la segunda tomaremos una perspectiva más distendida donde intentaremos imaginar los peligros y posibilidades que la inteligencia artificial traerá para un futuro no tan distante.

2. INTELIGENCIA ARTIFICIAL Y PROFESIÓN DE ABOGADO. INNOVACIÓN, AUTOMATIZACIÓN Y DEMANDA LATENTE

La abogacía es conservadora. Es cierto que en los últimos años se han introducido algunos cambios marginales en ella, pero en su esencia sigue siendo una profesión que hunde sus raíces en la tradición y la historia¹⁶¹. Si bien José Clemente Fabres no tenía correo electrónico, ni acceso a Internet, la práctica de

¹⁶¹ Como señala Susskind "In truth, the working practices of lawyers and judges have not changed much since the time of Charles Dickens". Susskind, Richard y Susskind, Daniel, *The Future of the Professions* (Oxford, Oxford University Press-Kindle, 2015), p. 66.

la profesión era básicamente la misma para él que para un abogado actual¹⁶². Sus clientes provenían de su círculo de conocidos, entre los cuales su reputación era clave para que confiaran en él la protección de sus intereses. Se entrevistaba personalmente con ellos y empatizaba con sus problemas. Su conocimiento era obtenido de libros y se expresaba a través de documentos escritos con los que resguardaba los intereses de sus clientes. En definitiva, aunque hoy la reputación de un abogado también se encuentra influida por su dominio de las redes sociales, su contacto con clientes puede desarrollarse por vías electrónicas y parte de sus conocimientos se obtiene a través de bases de datos digitales, éstos son meros facilitadores de su tarea, más no la alteran fundamentalmente.

El nudo de la cuestión es que la abogacía es una profesión que consiste esencialmente en escribir y pensar, y hasta ahora la tecnología simplemente ha dominado las energías físicas de la naturaleza, mas no las intelectuales del hombre. La tibia es cierto que en la década de 1980 los procesadores de texto cambiaron en algo la profesión, permitiendo la redacción de escritos y sentencias más largas y minuciosas, como también digitalizando los formularios de que se servían los abogados desde hace siglos, esto no representa un cambio significativo.

El advenimiento y popularización de las comunicaciones electrónicas durante las décadas de 1990 y 2000, llevaron a una sobreabundancia de documentos escritos. Si en 1980 era necesario para llevar adelante un pleito analizar el contrato y, a lo más, algunos otros documentos escritos como facturas y cartas, en la primera década de este siglo las comunicaciones que acompañan al contrato se hicieron numerosísimas, al punto que hoy es prácticamente imposible leer la totalidad de ellas, que suelen tener forma de correos electrónicos y mensajes instantáneos de diversa índole. Hay ocasiones en que las comunicaciones entre partes pueden constar de miles o incluso cientos de miles de documentos que ninguna ser humano es capaz de leer.

La tramitación electrónica también ha cambiado la profesión, convirtiéndola en una actividad física menos exigente, donde no se requiere recorrer los tribunales uno a uno para repartir escritos, pero la ritualidad es esencialmente la misma.

Por otra parte, a principios del siglo XX, el contacto que las personas tenían con el mundo del Derecho era más bien escaso. Salvo actos de especial trascendencia, como la muerte, el matrimonio o la compra de algún bien de alto valor, como una casa entre las pocas familias que podían permitírselo, el Derecho civil estaba generalmente ausente de la vida de las personas normales. Entre los pobres era el Derecho penal el que de tanto en cuanto asomaba en su gris existencia¹⁶³, y más en

general, la gente tenía un contacto esporádico con el Derecho constitucional al ir a votar cada ciertos años, y con el Derecho laboral. Si se quiere, la vida estaba poco juridizada y discurría con una presencia rara, y generalmente ingrata, del Derecho. En este sentido, el abogado era una figura vinculada a los estratos más ricos de la sociedad, que requerían de sus servicios para resguardar sus intereses. Sin una clase media real, el resto de la población recibía asesoría judicial ocasionalmente mediante servicios asistenciales gratuitos como la Corporación de Asistencia Judicial o el turno que los abogados debían realizar, y poco más.

La formación de las clases medias luego de la Segunda Guerra Mundial —y en Chile recién en la década de 1980— vino a alterar este panorama. La demanda general por servicios jurídicos se incrementó dramáticamente, toda vez que ahora había surgido un sector de la población que requería de atención profesional constante, aunque sus medios económicos están al límite de no poder hacerse cargo del costo de los mismos. En este mismo período se amplía la matrícula universitaria explosivamente, lo cual logra estabilizar los costos de contratación de un operador jurídico profesional. Básicamente, la oferta de abogados se triplicó¹⁶⁴ y, no obstante, el precio de los abogados se ha mantenido constante, toda vez que la demanda por servicios jurídicos de esta clase media sigue en aumento.

Amén de lo anterior, desde la última década del siglo XX, una serie de áreas que tradicionalmente estaban fuera del mundo jurídico, como las relaciones intra-familiares o las relaciones pedagógicas intraescolares, pasaron a estar sometidas a la observancia del Estado, y con ello, al mundo de las reglas del Derecho. En este sentido, crecientes áreas de la vida privada son, hoy por hoy, potencialmente materia de abogados, con lo que el campo de acción de estos profesionales se ha incrementado. En este sentido, y aún con una creciente cantidad de abogados, existe una demanda insatisfecha o latente¹⁶⁵ de servicios jurídicos en nuestra sociedad.

Hoy en día no se trata sólo de las frágiles clases medias que requieren asistencia jurídica, siendo la presencia de abogados en sus operaciones demasiado cara para poder pagarla, sino también del tejido productivo de la sociedad, de la inmensa cantidad de pequeñas y medianas empresas que deben destinar, eventualmente, altos porcentajes de sus ingresos a la obtención de servicios jurídicos¹⁶⁶. Hoy en día, respecto a buena parte de las deudas pequeñas, los

¹⁶² En efecto, de tener un abogado por cada mil habitantes, hoy hay 2,6, y este ratio se incrementa año tras año, pues juran ante la Corte Suprema cerca de tres mil abogados al año.

¹⁶³ Susskind, Richard, y Susskind, Daniel, *The Future of the Professions* (Oxford, Oxford University Press-Kindle, 2015), p. 113.

¹⁶⁴ Susskind expone el problema en los siguientes términos: *This can be simply stated: most people and organizations cannot afford the services of first-rate professionals; and most economies are struggling to sustain most of their professional services, including schools, court systems, and health services... For some time, the harsh reality, broadly speaking, has been that only the rich or robustly insured can engage many top-flight professionals, such as doctors, lawyers, accountants, and management consultants. The expertise*

¹⁶⁵ En palabras de Alano, Niblett y Yoon: *Abraham's Lincoln's practice of law was constrained by the number of books he could carry on his horse*. Alano, Benjamin, Niblett, Anthony, Yoon, Albet H., *How artificial intelligence will affect the practice of law*, University of Toronto Law Journal 68-1 (2018), p. 113.

¹⁶⁶ Al respecto, baste recordar el clásico trabajo de Menger, Antonin, *El Derecho Civil y Los Pobres*, traducción Adolfo Posada (Victoriano Suarez, 1898, Madrid).

costos de una atención jurídica son demasiado cuantiosos como para justificar su cobro, aún cuando su impago es nocivo. Así las deudas se tornan económicamente incobrables, ya que el costo jurídico es más alto que el beneficio obtenido a través de la activación de un proceso judicial. Esto supone un costo de transacción relativamente alto en nuestro sistema económico que perjudica seriamente a nuestra sociedad, toda vez que el incumplir los acuerdos es barato y el forzar su ejecución caro. Ello posibilita la abundancia de incumplimientos e injusticias manifiestas, e incluso refuerza la percepción ciudadana de una sociedad llena de abusos, toda vez que para los operadores en una posición económicamente fuerte, como ISAPRES, AFPs y bancos el incumplir sus obligaciones es sencillo, toda vez que el riesgo de judicialización es relativamente bajo considerando que el público general sólo recurrirá a abogados cuando el incumplimiento supere al costo judicial de cobro de manera exuberante.

En general, el sistema jurídico actual, con sus costos y dificultades, incentiva el incumplimiento, erosiona la buena fe y promueve la desconfianza. Toda vez que el incumplir es sencillo y poco riesgoso, los operadores económicos tienden a desconfiar de la ejecución de sus acuerdos, ya que es simple para las partes apartarse de ellos sin que existan consecuencias jurídicas apreciables. Esta permanente tensión crea una desconfianza generalizada que eventualmente lleva a una crisis de credibilidad de los diversos actores sociales. Todo lo anterior entorpece la actividad económica y provoca una suerte de estancamiento productivo.

En este contexto es que se ha visto a la inteligencia artificial –y a otras técnicas diversas– como una alternativa que permita llevar adelante una reestructuración profunda de nuestro actual sistema judicial, sea a través de la automatización, sea mediante la innovación¹⁶⁷.

Al respecto, conviene distinguir ambas situaciones. Se suele denominar automatización al cambio un cambio en la forma de ejecutar algunos procesos que permite reemplazar el esfuerzo humano por la labor de una máquina. Así, por ejemplo, antes de la invención de los ferrocarriles, la gente se transportaba de un lugar a otro en carruajes, a caballo o a pie. La invención del tren a vapor no

of a very few is being bestowed upon a few. Richard Susskind y Daniel Susskind, *The Future of the Professions* (Oxford, Oxford University Press-Kindle, 2015), p. 33.

¹⁶⁷ *"Americans need a simpler cheaper alternatives to giving to everyone a free lawyer... In particular, new technologies and approaches to dispute resolution offers us the opportunity to streamline and simplify to the benefit of everyone (except possibly lawyers). For many simple civil and even minor criminal cases, we could reform the process to let technology do the work, funneling parties through stepped, online dispute resolution starting at computerized mediation, passing through human, nonlawyer mediation, and proceeding into our current system only after making every effort to end the case cheaply and quickly... In short, the legal system needs to go on a diet, to make it slimmer, faster, cheaper and thus fairer. And lawyers need to get out of the way and let cheaper alternatives flourish." Berlin, Benjamin H., Bibas, Stephanos, Deborah, Justice: More technology, fewer lawyers and the future of the legal profession.*

cambia la necesidad de viajar, sino la forma en que se realiza tal actividad, ahora en menos tiempo y con mayor comodidad. Cuando hablamos de automatización de funciones en el mundo jurídico, se trata de actividades que hoy son realizadas por abogados y que, mediante el desarrollo de nuevas tecnologías, son ejecutadas mediante programas. La innovación, por su parte, consiste en realizar nuevas funciones para satisfacer necesidades que antes simplemente no estaban consideradas. Así, antes de la invención del teléfono las personas no podían comunicarse a distancia por voz. La telefonía no dejó desempleados, sino que creó toda un área de trabajo que antes no existía. En este sentido, los efectos de ambas son diversos, puesto que una tiende a reemplazar al trabajador, mientras que la otra crea un área nueva de atención. Cuando la innovación y la automatización afectan profundamente un área económica determinada, se dice que esta experimenta un proceso de disrupción, toda vez que las antiguas formas de realizar las actividades económicas de un área determinada, dejan de ser viables y son reemplazadas por otras.

Hasta ahora, los procesos de automatización e innovación han afectado básicamente las áreas económicas donde los trabajadores manuales predominaban, por lo que han sido éstos quienes se han visto afectados por los procesos de cambios productivos que se han denominado como revoluciones industriales. La razón básica de ello es que tales revoluciones han consistido en el domar las energías contenidas en la naturaleza para aplicarlas a procesos productivos, como el carbón para producir vapor de agua, el petróleo y la electricidad. No obstante, los procesos informáticos que llevan al desarrollo de inteligencias artificiales intentan emular la capacidad de pensar, por lo que debiesen afectar profesiones que consisten básicamente en el desarrollo de labores intelectuales, como es el caso de la abogacía. Ante tal cambio, muchos temen –y algunos esperan– que la implementación de tales técnicas al mundo jurídico cause una disrupción en el mercado de los operadores jurídicos, no sólo satisfaciendo la demanda actualmente insatisfecha de las clases medias por servicios jurídicos económicos y accesibles, sino incluso convirtiendo al abogado en una pieza de museo, como ya ocurrió con el cochero, el ascensorista o el hilandero.

La clave, en este sentido, se encuentra en las actividades que ejecuta un abogado. En principio, cualquier profesión se compone de una serie de actividades que conforman el campo propio de la misma. Algunas de ellas suelen ser de carácter creativo, de reflexión profunda o propias del estudio pormenorizado, mientras que otras son relativamente estandarizadas, que se realizan sin mayor reflexión y de manera casi automática. Si en un campo de actividad económica predomina lo primero, tal profesión está relativamente a salvo del riesgo de automatización, mientras que si es lo segundo es posible que ésta se encuentre en riesgo. Al respecto, Andrew Ng, uno de los fundadores de Google Brain y actualmente director del departamento de Inteligencia Artificial de Baidu, señaló que

ser automatizada mediante un modelo de inteligencia artificial¹⁰⁴. En efecto, muchas de las funciones que hoy cumple un abogado son tediosas, no requieren esfuerzo intelectual y, en este sentido, se encuentran altamente expuestas a ser automatizadas. La situación es especialmente preocupante en los estudios de mayor tamaño donde trabajan muchos abogados ejecutando actividades específicas, donde la materia encargada por el cliente se divide en diversas actividades que son realizadas por abogados distintos, como si se tratara de una línea de ensamblaje fordiana. El producto, un contrato, digamos, es subdividido en actividades, y cada actividad específica es de competencia de un abogado especializado en ella. El resultado suele ser muy bueno, pero es evidente que algunas de las labores realizadas por ellos van a ser necesariamente automatizadas, tal vez que tienden a ser rutinarias. Así, el rol de muchos de abogados puede dejar de parecer al ser reemplazados por algoritmos que ejecutan funciones equivalentes.

Naturalmente, esto no se extiende a todas las actividades de un abogado, pues siempre queda una fondo de actividades donde efectivamente se requiere un esfuerzo intelectual profundo o de empatía con el cliente y que no están inmediatamente en riesgo de automatización, pero, de igual manera, algunas de éstas están expuestas a ser transformadas por la innovación. En definitiva, la profesión de abogado parece encontrarse en pleno cambio y es altamente probable que las condiciones en que se desarrolla debiesen cambiar profundamente en los próximos años.

Una última reflexión dice relación con determinar aquellas áreas más expuestas a transformaciones. En principio, nos parece que hay dos vectores que llevan adelante los procesos de automatización de las actividades jurídicas. En primer lugar, existe el vector económico. En principio, de entre todas las actividades automatizables realizadas por abogados, están más expuestas a automatización las mejor pagadas. Esto es, aquellas que tienen un costo económico mayor para las empresas y usuarios, experimentan por dicho coste una mayor presión hacia la automatización para bajar sus costos. Así, los abogados corporativos y los especialistas en tributación, se encuentran expuestos a una mayor presión tecnológica que los especialistas en juzgados de policía local, toda vez que sus áreas de trabajo están altamente especializadas y representan costos importantes que cualquier empresa desea aminorar mediante herramientas tecnológicas. Un segundo grupo de áreas que se encuentra en riesgo son aquellas que tienen un impacto social alto y que se encuentran sobre exigidas por el alto número de casos y la falta de recursos humanos para su abordaje. En este sentido áreas como la concesión de libertades provisionales pueden ser automatizadas total o parcialmente, al igual que las cobranzas judiciales o la concesión de pensiones

de alimentos. En estos casos, la presión es netamente social, puesto que el costo humano representado por la tardanza en tales actividades es importante y el estado actual de tribunales no permite hacerse cargo de ellas de una manera adecuada.

3. ALGUNAS HERRAMIENTAS

En el contexto de un auge de la inteligencia artificial es que han surgido una serie de herramientas que prometen generar una disrupción en la actividad jurídica. En la presente sección analizaremos algunas de ellas, veremos sus límites actuales y sus posibles efectos en la actividad de los abogados. A partir de lo que actualmente existe proyectaremos el rumbo que vemos más probable para la profesión de abogado y hablaremos, finalmente, de la educación que creemos será más útil a los abogados.

3.1. Automatización de Contratos

En los últimos meses se han ido desarrollando distintas herramientas que, haciendo uso de agentes artificiales, han comenzado a automatizar la profesión de abogado. Es interesante constatar que cuando comenzamos a escribir este texto, muchas de ellas no existían o se aplicaban sólo en el hemisferio norte, mientras que ahora ya son de uso relativamente común y están en etapa de consolidación. Puesto que no contamos con una bola de cristal, asumimos que en un futuro muy próximo aparecerán otras herramientas.

Un ejemplo de este tipo de dispositivos son las aplicaciones para diseño de documentos. Es tradicional en el mundo jurídico servirse de formularios para redactar contratos y otros instrumentos que se utilizan en el mundo de los operadores judiciales. Dichos instrumentos ya estaban reducidos a formularios en época romana, que incluso desarrolló un procedimiento basado en ellos. Lo importante es saber qué formularios utilizar y cómo complementarlos, de manera que sean útiles a las necesidades del cliente. En este sentido, el uso de formularios requería una cierta destreza, para determinar cuales eran las partes pertinentes en un documento determinado, y cuales no. Así, los abogados se sirven de ellos considerando las circunstancias del caso y las normas básicas que rigen una materia determinada.

Hoy en día, la redacción de tales instrumentos se ha automatizado, a tal nivel que incluso existen aplicaciones que se encargan del diseño de tales documentos. La interfaz pregunta al usuario qué instrumento desea redactar y le formula unas preguntas pre-determinadas similares a aquellas que un abogado medianamente diligente realizaría y que contienen la información necesaria para la redacción del documento. Finalmente, la aplicación envía el documento al notario y cuando está listo para ser firmado simplemente avisa al cliente para que concurra

¹⁰⁴ If a typical person can do a mental task with less than one second of thought, we can probably automate it using AI either now or in the near future. *Ig. Andrew What Artificial Intelligence Can and Can't Do Right Now*. Harvard Business Review (2016), disponible en: <https://hbr.org/2016/11/what-artificial-intelligence-can-and-cant-do-right-now>. Consultado en 8 de Diciembre 2019.

a la notaría. De esta manera, una actividad que generalmente requería uno o dos abogados, como la redacción de un contrato, ya no necesita ninguno.

Aunque hoy en día la notaría sigue siendo indispensable en la redacción de muchos instrumentos –las escrituras públicas por lo pronto–, existen otros medios para asegurar la integridad y fidelidad de los mismos que podrían reemplazar parcialmente la labor de los notarios. Un ejemplo de esto son los *blockchains* o cadenas de registros que han aparecido y cuyo uso más conocido son las criptomonedas. Los *blockchains* son registros que se encuentran distribuidos entre todos o algunos de los usuarios de los mismos. Así, no existe un registro madre, sino que todos o algunos de los usuarios tienen una copia íntegra del registro, y para alterarlo incorporando algo a él es menester que haya concordancia entre los registros de todos los participantes. Así, si yo quisiera incorporar un nuevo documento al registro, éste se incorporará cuando dos o más de los servidores donde se guardan copias íntegras del mismo lo contengan, y en ese caso el registro se alterará para todos los participantes. Este tipo de registros son muy difíciles de intervenir o falsear, toda vez que requiere de un ataque concertado al mismo. En los más sólidos, como Ethereum, esto es virtualmente imposible, por lo que son una herramienta para dar fe de los instrumentos contenidos en ellos de una mayor eficiencia que los medios tradicionales, como el notariado. Esta es una tecnología innovativa, por lo que tiene la potencialidad para transformar el área completa de trabajo.

Incluso, existe la posibilidad de redactar los llamados *smart contracts* o contratos inteligentes, que tienen la virtualidad de auto-ejecutarse cuando se dan las condiciones pre-establecidas. Así, por ejemplo, puedo pactar un contrato de arriendo en base a una blockchain como Ethereum, donde el cerrojo de la puerta esté conectado a la blockchain y haya una cantidad de criptomonedas consignadas para pagar el arriendo. Así, cada vez que pase el día calendario para el pago de la renta, la propia blockchain ordenará el pago de la renta al arrendador, y si éste no se produce por falta de fondos, por ejemplo, la cerradura dejará de funcionar.

Mientras que la primera tecnología en automatización de documentos permite automatizar la redacción de los mismos, los *blockchains* son verdaderamente innovaciones que prometen alterar la manera en que se redactan y cumplen los contratos. A pesar de lo que Lessig¹⁶⁹ diga, el código no es el Derecho. El contrato no es lo que consta en la cadena de registros, toda vez que la teoría básica indica que el contrato es el acuerdo que se gesta entre dos o más partes a fin de crear obligaciones. Los *blockchains* son sólo una manera de probar y ejecutar los contratos, que tiene el poderoso efecto de alterar la posición relativa de las partes, toda vez que el cumplimiento aparente de las obligaciones que de él nacen se automatiza. Esta es una herramienta poderosa que promete cambiar la ejecu-

tividad de las obligaciones y crear un ambiente donde incumplir se complejice. Si una parte se siente perjudicada por la ejecución automática, podrá reclamar a los tribunales del cumplimiento automático indebido, pero el acreedor ve su posición relativa al pleito indudablemente mejorada.

En este sentido, los contratos inteligentes vienen a satisfacer una demanda latente de confiabilidad en los acuerdos, hoy no satisfecha siquiera por los títulos ejecutivos. Si combinamos ambas herramientas de automatización e innovación, tendremos un sistema económico que requiere de muchos menos abogados para la formalización y ejecución de transacciones, retirando del mundo de los profesionales del Derecho varias de sus funciones. El resultado es un descenso en el costo de transacción de las operaciones económicas por su menor costo y mayor efectividad, pero también una primera jibarización del campo de acción de los abogados, que quedarían relegados sólo a contratos complejos que requieren de un diseño específico. Un equipo económico que quisiera elevar la productividad del país, debiese tomarse en serio estos nuevos instrumentos posibilitando mediante acciones gubernamentales su aplicación.

3.2. Resolución de Conflictos Automatizada

Hace ya más de quince años que surgió la Online Dispute Resolution (ODR o resolución de conflictos automatizada) en el medio norteamericano de la mano de la mediación¹⁷⁰. La idea surge en el contexto e-Bay, la plataforma de venta norteamericana, donde se generaban múltiples conflictos entre oferentes y demandantes de productos en el contexto de las transacciones que ahí se realizaban. La idea original consistía en ofrecer a las partes una plataforma de mediación que intentase acercar las posiciones de las partes sin que interviniese un operador humano y que, eventualmente, fuese capaz de llegar a una solución razonable. Para esto Colin Rule diseñó toda una serie de mecanismos de aproximación de posiciones y que, ante el eventual fracaso de la plataforma, terminase en un arbitraje, originalmente realizado por un experto. El sistema fue tan eficaz y causó tales niveles de satisfacción entre los usuarios de la plataforma, que su confianza en ella aumentó y, eventualmente, el uso de tales mecanismos se extendió por la mayor parte de las plataformas de Internet.

En efecto, hoy en día, la mayor parte de las plataformas tienen un mecanismo de ODR al cual los usuarios pueden recurrir en caso de entrar en conflicto con ella o con terceros que operen mediante la misma. En muchos casos tales conflictos toman la forma de arbitrajes y en la mayor parte de las condiciones generales de contratación que tales plataformas imponen a sus usuarios se suele contener una cláusula al respecto. Ahora bien, tal arbitraje se encuentra reñido con

¹⁶⁹ El autor se hizo famoso por su frase *code is Law*, desarrollada en su texto: Lessig, Lawrence, *Code and Other Laws of Cyberspace* (New York, 1999, Basic Books).

¹⁷⁰ El texto básico al respecto es Colin, Rule, *Online Dispute Resolution for Business* (John Wiley, 2002, San Francisco).

la normativa sobre consumidor¹⁷¹ que habilita al consumidor a recusar el árbitro sin expresión de causa y solicitar a tribunales el nombramiento de uno diverso, o incluso recurrir a la justicia ordinaria. En definitiva, en nuestro ordenamiento, para el caso de consumidores, la aplicación de este tipo de cláusulas es voluntaria, aunque por su rapidez y sencillez, la mayor parte de los usuarios las prefieren.

Las plataformas de ODR suelen operar de una manera bastante simple. Se selecciona de entre un conjunto de alternativas el conflicto que se tiene, luego existe la posibilidad de describirlo y se adjuntan las evidencias que la parte tenga, como fotos, videos o documentos. La plataforma otorga traslado a la persona en contra de quien se actúa y le entrega un plazo para responder y adjuntar sus propias evidencias. Finalmente, será usualmente un agente artificial quien decida en parte basado en sistemas expertos y en parte en redes neuronales. El resultado se comunica rápidamente a las partes y usualmente la propia plataforma tiene medios para asegurar su cumplimiento.

En general el sistema es rápido y eficiente, tanto que incluso quienes han pasado por él suelen aumentar su confianza en la plataforma en lugar de disminuirla¹⁷². Se muestra especialmente útil para casos pequeños, donde usualmente las partes no habrían podido recurrir a los tribunales de justicia por el alto costo que ello representaría y la larga espera que supondría. En este sentido, los sistemas de resolución de conflictos automatizados tienden a satisfacer una demanda latente de justicia sumamente relevante, aumentando los niveles de confianza social y disminuyendo los costos de transacción que la falta de confianza implica.

Es cierto que los agentes artificiales pueden contener sesgos, pueden implicar discriminaciones y ser riesgosos en varios niveles, pero la utilidad de los sistemas de ODR se manifiesta en que, en primer término, son voluntarios, al menos dentro de nuestro ámbito jurídico, por lo que las partes pueden no recurrir a ellos. En segundo término, usualmente tratan materias que no habrían recibido cobertura jurídica en caso de no existir tales mecanismos. Es decir, puede que representen una justicia de mala calidad, pero es mejor ésta que ninguna. En tercer término, en principio las decisiones a que llegue el agente podrían eventualmente someterse a los tribunales ordinarios de justicia, al menos en el caso de tratarse de consumidores mediante el artículo 16 g) de la Ley 19.496.

¹⁷¹ Art. 16 g), L. 19496: "Si en estos contratos se designa árbitro, el consumidor podrá recusarlo sin necesidad de expresar causa y solicitar que se nombre otro por el juez letrado competente. Si se hubiese designado más de un árbitro, para actuar uno en subsidio de otro, podrá ejercer este derecho respecto de todos o parcialmente respecto de algunos. Todo ello de conformidad a las reglas del Código Orgánico de Tribunales.

"En todo contrato de adhesión en que se designe un árbitro, será obligatorio incluir una cláusula que informe al consumidor de su derecho a recusarlo, conforme a lo establecido en el inciso anterior. Lo que se entiende sin perjuicio del derecho que tiene el consumidor de recurrir siempre ante el tribunal competente."

¹⁷² Vid. al respecto; Barton, Benjamin H., Bibas, Stephanos, *Rebooting Justice. More technology, fewer lawyers, and the future of law* (Encounter Books, 2012, London-New York), p. 113.

Actualmente se encuentran en etapa de planificación diversos esfuerzos para desarrollar este tipo de plataformas a nivel estatal o internacional¹⁷³. Tal vez el esfuerzo más relevante lo está llevando a cabo la UNCITRAL, que recientemente ha desarrollado un documento de notas técnicas al respecto¹⁷⁴.

En base al éxito de este esfuerzo es que algunos estados comienzan a planificarse la posibilidad de establecer tribunales online, manejados íntegramente por agentes artificiales. El caso paradigmático es Estonia, que reveló durante 2019 su plan de crear una jurisdicción llevada adelante por agentes artificiales para casos de cuantía menor a siete mil euros, equivalente a los procedimientos de mínima cuantía en Chile.

Es menester, en todo caso, recalcar que Estonia es posiblemente el país que más lejos ha llevado la digitalización de su aparato gubernamental en el mundo, por lo que el nivel de penetración de las tecnologías de la información en el mismo es profundo. Está por verse si es que la estrategia Estonia produce buenos resultados, pero por lo pronto no debe perderse de vista que los agentes artificiales tienen múltiples problemas de sesgo y que convertirlos en agentes jurisdiccionales es ciertamente arriesgado, sobretodo teniendo en consideración que aún no es claro cómo solucionar dichos inconvenientes. La justicia digital podría convertirse en un problema de derechos humanos y fundamentales mayúsculo si es que no se toman precauciones al respecto.

En nuestra modesta opinión, hay tres prevenciones que realizar antes de confiar en agentes artificiales para la resolución de conflictos. En primer término, la concurrencia debe ser voluntaria, especialmente para los actores jurídicos más vulnerables, como trabajadores y consumidores. Sólo debiese entenderse que los consumidores y trabajadores aceptan tales plataformas en la medida que las utilicen y sólo para los casos donde así decidan hacerlo ejerciendo actualmente una acción jurisdiccional a través de ellas. Los demás operadores jurídicos pueden aceptar someterse a tales plataformas mediante cláusulas compromisorias o incluso a través de acuerdos generales que alguna empresa suscriba con tales prestadores de servicios jurídicos, pero en materia de consumo o laboral, la disparidad de condiciones entre las partes justifica que la especial voluntariedad de tales compromisos se materialice mediante la interposición de acciones y no suceda que se vean, a través de cláusulas que no han necesariamente comprendido, despojados de la posibilidad de concurrir a los tribunales ordinarios.

Una segunda prevención que debe regir la materia es el de la recurribilidad ante un juzgador humano. No puede garantizarse que las soluciones jurídicas que

¹⁷³ Barton, Benjamin H., Bibas, Stephanos, *Rebooting Justice. More technology, fewer lawyers, and the future of law* (Encounter Books, 2012, London-New York), p. 115.

¹⁷⁴ UNCITRAL, *UNCITRAL technical notes on online dispute resolution* (United Nations, 2017, New York), disponible en

https://www.uncitral.org/pdf/english/texts/odr/V1700382_English_Technical_Notes_on_ODR.pdf

Consultado el 12 de Diciembre de 2019.

aportan agentes artificiales actuando como juzgadores sean carentes de sesgo. Es por ello que debe existir siempre la posibilidad de recurrir ante un juzgador humano que revise las decisiones adoptadas por el algoritmo. Este juzgador debe tener facultades para revisar el fondo y, en caso necesario, suspender los efectos de la sentencia revisada.

Una tercera prevención es que el agente debe ser transparente en su diseño. Por esto entendemos que su código debe ser revisable íntegramente por parte de la autoridad competente en cualquier momento que lo desee, sea de oficio o a petición de parte. No puede suceder que el diseñador del agente, amparándose en la propiedad intelectual de su construcción, no permita la revisión del funcionamiento y diseño del algoritmo, toda vez que la administración de justicia es una función pública y, como tal, debe ser objeto de inspección. Ahora bien, toda vez que abrir el algoritmo completamente podría implicar poner en riesgo los secretos industriales que contiene, no proponemos su examen general por cualquier persona, sino solamente por parte de autoridades competentes que hayan sido encargadas de su supervisión.

Tomando estas precauciones como mínimo, creemos que los agentes artificiales pueden funcionar adecuadamente dentro del ámbito judicial, y más en general, en la órbita de la administración de funciones públicas.

3.3. Análisis de Datos

La aparición de las nuevas tecnologías de la información –hoy no tan nuevas– ha supuesto una abundancia de efectos sólo comparables a la invención de la escritura. Sólo por dar un ejemplo, en los últimos dos años se ha producido tanta información como la que existía desde el inicio de la humanidad. Esta superabundancia de datos implica que cada vez es más difícil estar al corriente de los mismos y que hay una multiplicidad de registros que pueden afectar a los operadores jurídicos, sea para evaluar los hechos acaecidos, sea para invocar la información jurídica relevante a la hora de determinar las reglas por las que se rige una situación determinada.

Puesto que el problema es de orden técnico, parece natural que éste sea solucionado por medios técnicos, especialmente a través de dos tipos de herramientas, las de *e-discovery* y las de analítica jurídica (*legal analytics*).

En cuanto a las primeras herramientas, las de *e-discovery*, estas consisten en aplicaciones que procesan documentos intentando determinar su relevancia para el caso. Así, primero se escanean los documentos (en caso de ser necesario) y luego se determina su contenido analizando las palabras que contienen. El agente hace exactamente lo que un ser humano haría, esto es leer los documentos, pero que descompone los textos en grupos de palabras –*BOWs* o *bags of words*– donde intenta establecer la presencia de palabras clave, retirando previamente las expresiones de menor poder explicativo, como preposiciones, conjuntivos

y disyuntivos¹⁷⁸. A continuación, intenta predecir el contenido de los mismos mediante dichas palabras, su frecuencia y posición a través de una red bayesiana de probabilidades. Mediante esta técnica netamente estadística el algoritmo presenta grupos de textos relevantes al usuario. El usuario normalmente puede ajustar la búsqueda para que muestre más resultados con menor precisión, o menos documentos con mayor certidumbre.

Si bien es cierto que en muchas ocasiones el agente no determinará con exactitud los documentos necesarios, la verdad es que tiende a ser altamente eficiente, generalmente rescatando de entre los documentos analizados más textos relevantes que los que un ser humano –o más bien un equipo de ellos– puede rescatar, cuestión que ha sido comprobada en muchas ocasiones. Por lo demás, el costo del análisis humano de cientos de miles, o incluso millones de documentos, es enorme, mientras que el agente lo realiza de manera automática a un costo bajísimo. Donde se requiere un grupo de humanos medianamente preparados, que realicen una labor compleja por semanas o incluso meses, un agente cuyo costo es esencialmente de desarrollo y que se pone en el mercado por una fracción del salario de uno de los trabajadores que debería desempeñar dichas funciones, lo realiza con resultados equivalentes. En definitiva, este es un caso clásico de automatización, donde la labor de un grupo de abogados jóvenes o de técnicos jurídicos es reemplazado por un agente artificial a un costo significativamente menor. Este tipo de herramientas está frenando la contratación de abogados jóvenes en los estudios de mayor tamaño, que posiblemente requieran de mucho menos personal en el futuro. Realizar un análisis de *due diligence* hoy, por ejemplo, es una labor que consume muchísimo menos esfuerzo humano que hace una década, y esto comienza a notarse en el mercado laboral de los recién egresados.

Similar a esto es la labor de análisis jurídico. Herramientas similares a las utilizadas en *e-discovery*, que usualmente se combinan con medios más tradicionales, como marcado de textos de acuerdo a su sentido, se usan para recorrer las fuentes del Derecho a fin de determinar cuales son aplicables a un problema o caso determinado. La historia de estas herramientas es extensa y surgen ya en la década del 1970, cuando algunos investigadores intentaron –mediante métodos simbólicos– programar reglas jurídicas en ordenadores a fin de predecir cómo debiesen fallarse determinados casos. El resultado fue más bien modesto, toda vez que requerían que expertos interpretasen las disposiciones jurídicas y manualmente las introdujesen en medios informáticos. La principal debilidad del método es que la interpretación que haga un experto no es necesariamente la correcta o la misma que realizará el operador jurídico encargado de aplicar la norma. Así, salvo que el propio aparato público que aplica la norma detalle su interpretación de la misma, los resultados pueden ser magros. Fuera de campos

¹⁷⁸ Ashley, D. Kevin, *Artificial Intelligence and Legal Analytics. New Tools for Law Practice in the Digital Age* (Cambridge University Press, 2017 Cambridge), p. 217.

como la declaración de la renta propuesta por el Servicio de Impuesto Interno, la verdad es que su aplicación final y cotidiana fue más bien modesta, a pesar que hubo inversiones vastas en el campo durante la década de 1980.

Una nueva ola de analítica jurídica comenzó a partir de 2011, cuando Watson logró triunfar en el concurso Jeopardy!, toda vez que tal certamen requiere responder preguntas abiertas que se formulen con precisión y rapidez. Watson, utilizando la técnica de BOW pudo derrotar a los concursantes humanos, abriendo, de esta manera, un nuevo campo para la inteligencia artificial, el responder preguntas en calidad de expertos. Al poco tiempo, se creó Debater¹⁷⁶, un agente encargado de argumentar una posición cualquiera que se elija sobre un tema determinado. El 11 de Febrero de 2019 Debater protagonizó un debate acerca de la educación infantil¹⁷⁷ contra un equipo de debate de primer nivel mundial obteniendo un destacado resultado. Al agente se le dio la posición que debía defender con muy poca antelación y este obtuvo los razonamientos empleados de la información disponible en Internet. Los argumentos y su hilación son notables.

Llevar esta herramienta al mundo jurídico es una ambición de muchos investigadores, y en este contexto, usando la tecnología de Watson, fue desarrollado ROSS, calificado como el primer abogado artificial del mundo por sus creadores¹⁷⁸ y cuyo objetivo es dar el marco normativo a una situación jurídica determinada por la cual se le consulte. La aplicación, ante una pregunta, busca entre el material jurídico disponible, que es tanto legal y reglamentario como jurisprudencial, hasta dar con los casos y reglas que probablemente sean pertinentes a la cuestión. Si bien ha tenido una exposición bastante mediática, el producto es simplemente un buscador normativo sofisticado que resulta útil para quien efectivamente sabe qué debe preguntar y cuándo la respuesta ofrecida resulta relevante. En pocas palabras, más que reemplazar al abogado, esta es una herramienta que resulta funcional a quien ya conoce el Derecho, facilitando su trabajo. Mejor orientadas están, en este sentido, las herramientas de análisis de Thomson Reuters y Tirant Lo Blanch, que como proveedores de servicios digitales buscan simplificar las búsquedas realizadas por sus usuarios sirviéndose de técnicas equivalentes.

Hoy en día es posible hacer un seguimiento individualizado de las decisiones de un juez o tribunal en cualquier materia, para analizar su consistencia y su viabilidad. Así, se puede seguir en concreto cómo ha fallado durante su carrera tal juez en materia de arriendo, y construir un argumento que ese magistrado en concreto pueda apreciar. Esto, evidentemente, resulta preocupante para cualquier orientador, toda vez que puede verse enfrentado a sus propios argumentos

y su labor puede resultar cotejada a fin de estudiar irregularidades, al punto que algunos países, como Francia, han prohibido este tipo de análisis, penando al responsable hasta con cinco años de cárcel¹⁷⁹. Esto es una medida equivocada, pues simplemente se opta por eliminar al mensajero (que detecta irregularidades y busca argumentos), en lugar de enfrentar el problema, que no es otro que la volatibilidad de los jueces.

Lo que todas estas herramientas tienen en común es que facilitan la búsqueda de información, permitiendo a operadores jurídicos con menor nivel de experiencia en una materia determinada, dar con orientaciones precisas acerca de las reglas aplicables a una situación con la que estaban poco familiarizados. El nivel de las respuestas es bueno y ahorran tiempo de investigación a los abogados, pero no realizan el razonamiento jurídico ni son capaces, de momento, de reemplazar la función del jurista para el lego. Es decir, su función principal es automatizar parte de las actividades del abogado, contribuyendo a simplificar su labor. Evidentemente, esto tiene un efecto neto en la contratación de abogados por parte de los estudios, toda vez que con un personal más reducido pueden absorber un número y variedad de casos mayor, toda vez que parte de la actividad de investigación queda automatizada. El comprender la información entregada por el buscador, el determinar si es pertinente y hacerla verdaderamente aplicable, queda en manos de un ser humano, quien deberá contar con criterio suficiente para realizar esta labor.

En cierto sentido, la existencia de herramientas de analítica jurídica sofisticadas implica que la especialización pierde valor relativo frente a la adquisición de conocimientos generales. Las normas especializadas se vuelven rápidamente accesibles, pero el criterio y la capacidad de integrar tales normas al caso se torna central, para lo cual se requiere de conocimientos jurídicos básicos sólidos.

4. EDUCACIÓN PARA UN FUTURO INCIERTO

Las herramientas recién expuestas no son más que los primeros frutos que da la aplicación de las técnicas de inteligencia artificial al mundo jurídico. Sus resultados son, en general, la automatización de parte de las labores de los abogados. Algunas técnicas tienen la virtualidad de dirigirse hacia la enorme demanda insatisfecha de justicia que experimentan las sociedades actuales, con una clase media amplia y necesidades jurídicas pendientes, sustituyendo la labor de los abogados como operadores jurídicos entre los grupos que no tienen acceso a ellos. Otras tienden a complementar su actividad, automatizando algunas de sus funciones. Los efectos más disruptivos, de momento, se centran en la primera área, donde la demanda latente y no satisfecha por servicios jurídicos, será atendida por herramientas que permitan jurificar operaciones que hoy son infor-

Ashley, D. Kevin, *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age* (Cambridge University Press, 2017, Cambridge), p. 23.
Video disponible en:
<https://www.youtube.com/watch?v=m3u1ytrVw>
Consultado el 16 de Diciembre de 2019.
<https://rossintelligence.com>

¹⁷⁹ Ley 2019-222 del 23 de marzo de 2019, art. 33.

males. Las herramientas de diseño de documentos y de resolución de conflictos permiten dejar de lado a los abogados para operaciones menores, pero a medida que tales instrumentos se perfeccionen deberían también afectar operaciones de mayor calado y penetrar en el ámbito de las operaciones que hoy sí son satisfechas por abogados.

Las herramientas de automatización, por su parte, potenciarán el actuar de los estudios mayores, hoy centrados en la atención de clientes de mayor envergadura por los altos costos que genera su actuar. Un equipo más reducido podrá hacerse cargo de una cantidad de casos más importante a un costo menor, toda vez que no requerirá de un regimiento de operadores jurídicos centrados en labores ya automatizadas. Es posible que algunos de estos estudios desahucen, al igual que hicieron supermercados, centros comerciales y clínicas médicas en la década de 1990, que el espacio para un mayor crecimiento se da no tanto entre los clientes grandes, sino en la atención familiar de casos propios de las clases medias del país, aprovechando las economías de escala que su tamaño les permite, para atender a los miles de clientes de casos menores que hoy reciben atención de estudios pequeños o que no tienen acceso al mundo jurídico fuera de la atención asistencial. Siguiendo un modelo de expansión al estilo Walmart, pueden bajar sus costos dividiendo el trabajo y automatizando funciones, para absorber la parte de esa demanda latente ofreciendo servicios jurídicos a un precio fijo y con una calidad estandarizada. No sabemos si en definitiva lo harán, pero la estructura actual de la demanda por servicios jurídicos permite a los operadores grandes tomar buena parte del mercado de los estudios jurídicos más pequeños e incluso absorber la demanda por abogados que ejercen la profesión de manera individual. Esto augura una mayor concentración en el mercado jurídico, acompañada de una reducción en los precios y estandarización de los servicios.

A nivel social, las consecuencias de esto serían positivas, toda vez que la demanda latente de justicia se vería satisfecha, pero respecto a los abogados esta posibilidad no es necesariamente buena. La posibilidad de ejercer la profesión de manera independiente se verá reducida y su destino puede ser similar al hoy casi extinto almacén de barrio.

En general, aunque la prestación de servicios jurídicos puede expandirse hasta cubrir una buena parte de la demanda latente en nuestra sociedad, esto no significa que haya un aumento en la oferta de trabajo para los abogados. Puede bien resultar que sí, aunque los operadores jurídicos necesarios serán algo distintos. Los grandes estudios, en general, requerirán menos profesionales para desarrollar las mismas actividades, aunque es posible que, de expandir sus funciones, se generen nuevos puestos de trabajo, especialmente en atención de clientes (una labor no automatizable) y en evaluación y supervisión de las actividades delegadas a algoritmos. Aquí no sólo pensamos en analistas de datos, sino que en general en personas que tengan un conocimiento profundo del Derecho y que sean capaces de discernir si las soluciones ofrecidas son útiles y plausibles.

Las destrezas jurídicas seguirán siendo necesarias en muchas áreas de la economía y la administración pública, pero la forma en que se enfrenta el mundo jurídico probablemente va a requerir potenciar algunas habilidades, como la comprensión de problemas generales. En definitiva, lo que se requerirá es criterio general, algo de lo cual carecen hoy en día los agentes artificiales, para poder dar sentido a las normas que invocan y encuadrar correctamente los problemas. Para que un algoritmo que ofrece soluciones funcione correctamente, es necesario que al otro lado del terminal haya una persona que sepa qué preguntar y cómo servirse de las soluciones ofrecidas. Esto requiere de criterio y sentido jurídico, lo cual sólo se forma a través de la comprensión de las teorías de corte general en el Derecho. Se necesitan abogados más humanistas y menos técnicos, personas que comprendan de lógica jurídica, de retórica y de análisis histórico, que puedan ver el fenómeno jurídico con amplitud. Se necesita más filosofía en el mundo del Derecho y menos especialización. Un ejemplo de esto es el sesgo en los algoritmos, cuestión que hemos tratado más de una vez a lo largo de este texto. Las soluciones y predicciones ofrecidas pueden ser muy distintas si la función objetiva o de éxito es diseñada de una manera o de otra, incorporando un tipo de justicia distributiva, social, sinalagmática u otro, por lo que detrás de ese diseño debe haber reflexión precisa acerca de lo que se busca a fin de defender el operar del modelo. Si se privilegió la no discriminación hacia ciertos grupos, es menester comprender por qué y para qué, lo mismo que si se prefirió un modelo que trate a todos igualmente sin considerar que algunos grupos cuentan con el peso de desigualdades históricas. Estas decisiones pueden ser justificables en un contexto u otro, pero importa que al tomar la opción, se haya meditado acerca del problema, pues no todos los conceptos de equidad sirven en todos los contextos. Así, si se va a entregar un beneficio o asignar una carga es menester pensar si ésta será asignada completamente al azar, o siguiendo un criterio único (renta por ejemplo), o si se desea que algunos grupos de la sociedad se vean beneficiados por ella a pesar de escapar al criterio básico según el cual se distribuye. En ocasiones un criterio será mejor que los otros, en otras ninguno resultará satisfactorio y será necesario diseñar uno nuevo, pero esta es una cuestión compleja que requiere de formación general para que el abogado pueda entender las implicancias del actuar social.

Por otro lado, existen áreas completamente especializadas que posiblemente terminen siendo innecesarias, pudiendo reemplazarse su estudio por otros que otorguen habilidades generales que permitan adquirir los conocimientos específicos, o comprenderlos, en caso de resultar necesario para la vida profesional. En este sentido, la educación continua debiese hacerse cargo de esas áreas de especialización y no tanto la formación general del pre-grado.

En definitiva, la educación jurídica para este tiempo que nace debiese ser más generalista, más humanista y menos especializada, dejándose tiempo para que los alumnos reflexionen sobre las grandes cuestiones del Derecho y formen un criterio que les permita integrarse en un mundo profesional cambiante. Así, es más productivo hoy estudiar Derecho romano, Historia y Filo-

sofia que especializarse en áreas oscuras durante la licenciatura, por lo que la importancia relativa de tales áreas debiese crecer. En contra, áreas especializadas debiesen tender a desaparecer del currículo básico e integrarse en la formación continua, que será una cuestión recurrente en la vida de un operador jurídico.

En el mundo que viene habrán abogados, pero su función será más generalista y menos especializada que hasta hoy. Necesitamos más creativos y menos operarios fordianos que sólo saben hacer una cosa, pues eso es lo que centran sus labores podría verse automatizado y sus puestos de trabajo se perderán. En definitiva, el desafío para la educación jurídica es el re-humanizar el Derecho.

En cuanto a la informática que necesitarán saber los operadores jurídicos del futuro, en general todos usamos computadores, aunque prácticamente nadie entiende como trabajan en profundidad. Esto ocurrirá también con los agentes artificiales. Hay elementos que deberán conocerse para comprender sus posibles balencias, pero un conocimiento especializado no es indispensable. Aprender a programar puede ser útil, según el área específica de trabajo en que un profesional desee desempeñarse, pero, en general, las herramientas de inteligencia artificial son sencillas de utilizar y no requieren de conocimientos especializados para interactuar con ellas. Ahora bien, si nadie requiere de un curso especializado para dialogar con Alexa o Siri, para entender cómo funcionan sí se requiere de conocimientos técnicos generales, y son esos los que deberían impartirse, más que a nivel universitario, a nivel escolar secundario.

Para tener acceso a conocimientos técnicos, conviene interactuar de manera interdisciplinaria con informáticos y analistas de datos. Ahora bien, lo relevante para el abogado es ofrecer los conocimientos sólidos jurídicos con que cuenta a los especialistas de otras áreas, no intentar dominar otras disciplinas, puesto que en ese caso transformaríamos al jurista en un ingeniero, cuestión algo ociosa. Comprender los rudimentos del funcionamiento de los sistemas informáticos permitirá, en todo caso, ofrecer una colaboración valiosa. El abogado debe convertirse en una suerte de socio que apoye la implementación de las posibilidades técnicas traen, más que en un obstáculo al desarrollo de actividades económicas con trabas normativas. Esto, nuevamente, requiere criterio y no tanto conocimientos especializados.

En fin, para resumir, vemos que la inteligencia artificial cambiará parcialmente el mundo jurídico, automatizando muchas de las actividades jurídicas y posiblemente concentrando la oferta en menos operadores de mayor tamaño. También creemos que la profesión jurídica deberá cambiar su formación dando entrada a más humanidades y dejando las especializaciones para ciclos de formación continua, de manera que el profesional que saiga de las universidades sea flexible y pueda insertarse exitosamente en distintas áreas del Derecho, estando siempre preparado para continuar su formación.

5. ALGUNAS PESADILLAS

Nuestra visión sobre el impacto de la inteligencia artificial en el mundo jurídico es, a corto y mediano plazo, optimista. Si bien creemos que la profesión de jurista deberá actualizarse, a la larga completará una adaptación hacia las nuevas condiciones técnicas y económicas mediante generalidad, adaptabilidad y formación continua. Ahora bien, este no es el único futuro posible. Desde que la idea de desarrollar consciencias artificiales despuntó, han existido diversas distopías que plantean un futuro oscuro. Algunas fantasías son de estilo apocalíptico, y aunque existen algunos que las temen, incluso investigadores de reputación, nosotros planteamos escenarios algo más benignos, pero no por ello menos perturbadores.

Como hemos señalado más arriba, es improbable que en el corto ni largo plazo logremos desarrollar consciencias artificiales, toda vez que, de momento, desconocemos en qué consiste tal fenómeno. Para construir algo, es necesario que al menos se entienda lo que se diseña, y puesto que la consciencia es un misterio, difícilmente podremos recrearla. Algunos piensan que la consciencia podría ser el resultado no buscado de las interacciones neuronales cuando se alcanzan una masa crítica, pero esto no parece probable, puesto que hasta los árboles tienen cierto nivel de consciencia, toda vez que sienten dolor, pero no tienen redes neuronales. Tal vez en un futuro distante se pueda lograr algo parecido a la consciencia, aunque no se ven caminos abiertos en este sentido.

En cuanto a diseñar una inteligencia artificial de carácter general, pero no dotada de consciencia, esta sí es posible de conseguir, aunque los obstáculos a superar son ciertamente formidables. Se necesita diseñar una capacidad de abstracción, una suerte de máquina aristotélica que sea capaz de comprender conceptos y crearlos a partir de datos, algo que los humanos hacemos muy bien y nos permite aprender con muy pocos ejemplos. Se necesita que los conceptos y conocimientos generados sean transferibles entre distintas funciones. Nosotros aprendemos que la fuerza de gravedad funciona al caminar, e inmediatamente podemos aplicar este conocimiento a saltar, arrojar objetos y correr. Las inteligencias artificiales con que contamos ahora no son capaces de transferir conocimientos de esta forma, sino que deben aprender, y ser entrenadas nuevamente, para cada función. Si se superasen estos dos obstáculos iniciales —qué no parecen ser abordables de momento— pueden surgir otros. Creemos posible que a la larga, tal vez dentro de nuestro espacio vital, se desarrolle una inteligencia artificial general, por lo que es un tema importante aunque no acuciante.

En todo caso, una tecnología puede no ser necesariamente sofisticada para causar una disrupción económica y política. Es cuestión de pensar en la máquina de vapor, que fue inventada ya por los griegos de Alejandría y permaneció como una curiosidad sin consecuencias hasta que en el siglo XVIII se decidió usarla para fines industriales, revolucionando las formas productivas de manera insospechada. En este sentido, la inteligencia artificial estrecha o específica con que contamos hoy es suficiente como para alterar nuestro estilo de vida y nuestro or-

den político y económico sin mayores dificultades. Este acápite tratará de estos cambios, de las proyecciones que podemos trazar a partir de lo que tenemos y progresivamente nos iremos alejando en el tiempo según las posibilidades más remotas se puedan, eventualmente, materializar.

5.1. Cambios sustanciales al modelo productivo

En lo inmediato, lo más preocupante de las tecnologías ligadas a la inteligencia artificial es su posible impacto en el aparato productivo. Mientras la revolución industrial dejó sin empleo a miles de trabajadores manuales, toda vez que las energías naturales escondidas en el vapor y el carbón fueron domadas, la inteligencia artificial apunta a otro tipo de trabajador como primera víctima de su emergencia, esto es, a los de cuello y corbata. Si bien la inteligencia artificial no emula exactamente nuestras capacidades cognitivas, sí puede llegar a resultados equivalentes a los nuestros por otras vías, especialmente mediante su capacidad de manipular símbolos y detectar patrones. Así, si bien no piensa como nosotros, puede realizar muchas de las funciones que hoy cumplimos con un grado de eficacia difícilmente alcanzable para nosotros. No es sólo que juegos como el go o el ajedrez sean ahora de su entero dominio, sino que muchas otras actividades intelectuales para las que tenemos capacidades más bien limitadas también lo son. Por ello es perfectamente posible y probable que tenga un impacto mayúsculo en nuestro sistema productivo, el que ha sido comparado a la electricidad durante el siglo XX.

Un influyente estudio¹⁰⁰ predice que hasta el 47% de los trabajos en Estados Unidos se encuentra en peligro de desaparecer en los próximos veinte años. Este se realizó descomponiendo los empleos en funciones y determinando cuántas de dichas funciones son automatizables. Aparentemente, casi la mitad de aquellas labores que hoy ocupan a la fuerza productiva lo son. El potencial para la disrupción es, por tanto, vasto.

Las labores en mayor peligro son aquellas intelectuales repetitivas con bajo contenido cognitivo, especialmente si importan un alto costo. De ahí que áreas completas como la *e-discovery* o la contabilidad empiecen a transitar hacia la automatización. Existen otras labores también que son un blanco evidente, aquellas en que la inteligencia artificial ya destaca sobre el ser humano. Así, por ejemplo, la radiología es un área donde el reconocimiento de imágenes puede ser realizado con muchísima mayor precisión por un agente artificial que por un ser humano, por lo que es un área que se encuentra directamente en peligro por la automatización. Otras áreas pueden constituirse en objetivos por la escasez relativa de profesionales cualificados para llevarlas a cabo, como el diagnóstico médico y la interpretación de exámenes. En ellas hay pocos profesionales en

comparación con la necesidad relativa de atención por parte del público general, por lo que intentar automatizarlas al máximo, a fin de relevar parte de la atención humana sobre dichas áreas, puede resultar beneficioso, y por tanto, muchos recursos se dedican a ello. Así, por ejemplo, si una inteligencia artificial pudiese, en base a los exámenes médicos realizados, determinar el riesgo relativo de un paciente, podrá derivar los casos importantes a un médico humano, y recomendar tratamientos simples para los de bajo riesgo. Así, puede recetar paracetamol o ibuprofeno para las gripes comunes, y derivar a un médico aquellos casos que efectivamente requieran de atención especializada como aquellos que presentan riesgo de meningitis.

El área de servicios, que hoy por hoy es uno de los mayores contribuyentes netos al empleo, hay muchas áreas automatizables. Las plataformas de ventas de internet ya han tenido un efecto apreciable sobre el comercio minorista tradicional, pero también existen supermercados sin cajeros, casi sin personal de atención y donde la labor de reposición la realizan robots. Si bien la existencia de camareros es todavía relativamente segura, en muchos lugares de comida rápida estos están siendo sustituidos por agentes artificiales. Tal vez en un futuro tener un camarero humano sea algo reservado a los restaurantes de mayor costo, mientras que en los demás casos sólo haya máquinas donde realizar el pedido y pagar la cuenta.

Una de las áreas más preocupantes es el sector transportes, donde el inminente desarrollo de automóviles y camiones auto conducidos puede tener un inmenso impacto en el empleo. Hoy en día, la mayor parte de los automóviles pasan casi la totalidad del tiempo estacionados, salvo los escasos momentos en que los utilizamos. Si ese tiempo muerto fuese dedicado al transporte de pasajeros sin un conductor al volante, el costo del transporte caería en picada, siendo equivalente a la depreciación del automóvil, más los insumos que requiere (combustible básicamente) y una rentabilidad que si estimamos en un 20% haría caer el precio del transporte por debajo incluso de pasaje de bus para trayectos medianos. En definitiva, usar plataformas de transporte compartido sería económicamente más viable que ser propietario de un automóvil, la cantidad de vehículos en las calles descendería y la frecuencia de accidentes caería en picada.

En cuanto a labores no cualificadas, como la recolección de frutas, ya existen algunos robots capaces de realizar la cosecha, que no se enferman, no faltan al trabajo, no renuncian, que actúan día y noche y no tienen problemas con condiciones de trabajo infrahumanas. El costo actual de uno de estos robots es equivalente al salario anual de un trabajador, pero se prevé que caiga. Incluso es posible que un inversor compre una cuadrilla de éstos y los arriende a diversos productores que requieran de sus servicios para no verse en la obligación de contratar trabajadores humanos. Piénsese que esta labor estacional de recolección de frutas representa entorno a los 500 mil empleos anualmente en Chile, por lo que la introducción de esta tecnología tendrá efectos masivos en nuestra economía.

¹⁰⁰ Frey, Carl Benedict y Osborne, Michael A., The future of employment: How susceptible are jobs to computerisation? in *Technological Forecasting and Social Change* 114 (2017), pp. 254-280.

En cuanto a la construcción, otro tradicional motor del empleo, ya existe la tecnología para imprimir casas. En efecto, en Holanda y en China se han utilizado exitosamente dispositivos capaces de construir viviendas utilizando materiales como cemento y ladrillo, realizando casi toda la obra gruesa en no más de veinticuatro horas. Lo que queda para los trabajadores humanos son las terminaciones, y no siempre.

Usualmente las labores manuales suelen ser un blanco de los cambios en los procesos productivos, aquí también sucederá, pero con una salvedad, existen trabajos que requieren de destreza y delicadeza que difícilmente se automatizarán. Un caso es el cuidado de ancianos o de niños. Es difícil que la tecnología llegue a desarrollar una precisión y destreza tales que verdaderamente reemplacen a los cuidadores humanos. Un robot que el 99% de las veces saque correctamente a un anciano de la cama es inútil, puesto que si una de cada cien veces lo tira al suelo, es un desastre. Es por ello que su automatización es difícil y probablemente no suceda en el corto plazo.

En cuanto a labores creativas, se suele creer que éstas están, en algún grado, a salvo de la automatización, aunque esto es relativo. Hoy por hoy, un agente no demasiado sofisticado puede componer música, esculpir o pintar al óleo. Tal vez no tenga la destreza de Bach o de Quincy Jones, pero tampoco se requiere de esta habilidad para componer e interpretar la tonadilla de un comercial. En cuanto a la escritura, hoy en día muchas de las noticias que se publican en medios de comunicación están enteramente escritas por agentes artificiales. Ellos, por ejemplo, recolectan información sobre un partido de fútbol, lo describen y dan los detalles al respecto. No tienen la pluma de Tolstoi, pero tampoco ésta es necesaria para hablar de deportes con efectividad. Es decir, las labores creativas que se encuentran a salvo son aquellas realizadas por un genio o aquéllas en que el hecho que intervenga un ser humano les otorgue un valor añadido. Quien compra la obra de un alfarero de Quinchamalí, quiere el resultado del juego entre manos y fuego. Si buscarse piezas industriales, les compraría en cualquier negocio cercano, puesto que la labor del alfarero ya fue industrializada hace siglos. El valor de la pieza la da específicamente el actuar humano y esa labor es irremplazable.

El problema es que los seres humanos comunes y corrientes no se educan para la genialidad ni aprenden labores donde su actuar específico sea el elemento que otorgue valor. En este sentido, las alternativas son pocas, o reconstruimos nuestro sistema escolar y universitario para fomentar la creatividad y la investigación de manera inclusiva, otorgando oportunidades de desarrollo y perfeccionamiento real a todos los habitantes del país, o tendremos generaciones perdidas que no serán empleables de manera alguna, a ningún salario por bajo que sea, incluso aunque esté debajo de su línea de subsistencia.

Normalmente las disrupciones suelen generar trabajos alternativos. Si hace ciento cincuenta años al se hubiese intentado explicar al hilandero que su trabajo no existiría en el futuro, pero que sus biznietos vivirían bien trabajando de diseñadores de páginas de Internet, argumentar esto habría sido difícil. Ya en la década de 1960 se

revela un desempleo generalizado producto de la automatización¹⁰¹, profecía que en definitiva, no se cumplió. No obstante, los temores siguen vigentes, especialmente porque aunque se generen nuevos empleos, estos parecen ser de una naturaleza bastante diferente de aquéllos que desaparecen. Es difícil que un temporero agrícola se transforme en analista de datos de la noche a la mañana. Aunque podría suceder que en el futuro la educación se transforme en profundidad, y los descendientes de temporero actual se empleen en compañías tecnológicas, en el presente el futuro que enfrenta es oscuro. Hay una brecha de adaptación entre los trabajadores que pierden sus puestos y las nuevas oportunidades que eventualmente puedan surgir, la cual difícilmente se puede salvar mediante la educación.

Es probable que la generación de ingresos se concentre en el futuro inmediato. Los grandes proveedores pueden servirse de la tecnología de punta gracias a sus economías de escala y disponibilidad de recursos, bajando sus costos a una fracción de los actuales, mientras que los pequeños y medianos productores no cuentan con esta posibilidad, por lo que probablemente se vean desplazados del mercado, salvo que puedan dotar a sus servicios y productos de un elemento único que sólo la tradición y la creatividad humana pueden entregar, como la cerámica artesanal de Quinchamalí. Un taxi auto conducido a precio irrisorio desplazaría a todos los taxistas actuales, y los ingresos que antes se repartían entre un grupo amplio de personas se concentrarán en dicha institución. Esto que puede suceder con los taxis, se aplica a la mayor parte de las áreas de la economía actual. Aunque el uso de inteligencia artificial probablemente aumente el producto geográfico, el ingreso producido a través de ella se concentrará cada vez más. En pocas palabras, tendremos un mundo más rico, donde la mayor parte de la gente no podrá emplearse en nada, y una minoría se lo llevará todo.

Este panorama, que se ve probable en el mediano plazo, es delicado y obliga a repensar nuestro sistema social completo. Las tensiones que un sistema social percibido como injusto puede generar son inmensas. Como sostenía Douglas North¹⁰², el costo de mantención de un sistema social es inversamente proporcional a la legitimidad percibida. Así, si la acumulación de ingresos en pocas manos es interpretada como una deslegitimación del sistema social, sólo quedan dos alternativas, la represión brutal o la cancelación del orden político y económico imperante.

5.2. Distopías futuristas

Cuando en 1932 Adolus Huxley escribía *Un Mundo Feliz (A Brave New World)* estábamos lejos de comprender que por medios mecánicos la manipulación de

¹⁰¹ Véase el informe completo en Pauling, Linus et alia, *The Triple Revolution*, Santa Barbara, California (1964) disponible en:

http://scarc.library.oregonstate.edu/coll/pauling/peace/papers/1964p_7-01.html

Consultado el 19 de mayo de 2019.

¹⁰² North, Douglas C., *Structure and Change in Economic History* (New York, Norton, 1981), p. 63.

símbolos y las matemáticas de matrices fuésemos capaces de construir nosotros mismos que imitasen los resultados del pensamiento. Noventa años más tarde, su célebre distopía parece acercarse a tocar la puerta nuevamente. La posibilidad de construir un orden totalitario no basado en la violencia física, como lo concebía George Orwell en 1984, sino en la estimulación del placer y la creación de una falsa sensación de libertad mediante el control de la información, se vuelve una posibilidad técnica mucho más cercana en las primeras décadas del siglo XXI. Mediante el análisis de datos a través de la limitada inteligencia artificial con que contamos, es posible tener perfiles completos y predictivos de los individuos a fin de ofrecerles aquello que desean y ocultarles lo que temen o lo que les perturbaría. Hoy día, distintas plataformas de redes sociales tienen la capacidad de, analizando sus datos, determinar preocupantes detalles íntimos sobre sus usuarios que van desde sus gustos cinematográficos, a sus patrones de consumo, sus tendencias sexuales, sus ideas políticas, su estado de salud o posible criminalidad. En 2012 varias mujeres se alarmaron cuando Google comenzó a mostrarles productos infantiles antes que ellas supiesen que estaban embarazadas. ¿Cómo lo determinó? Simplemente a través de sus búsquedas en internet de productos como el ácido fólico, que típicamente consumen las mujeres que buscan concebir. Ejemplos en este sentido existen muchos, y resulta preocupante que dentro de los perfiles individualizados con que cuentan las grandes firmas de Internet se encuentran todas las búsquedas jamás realizadas por una persona. Es decir, la información en Internet no se pierde, sino que se almacena en inmensos servidores que la conservan para realizar análisis sobre ella en el futuro.

La finalidad de la mayor parte de los modelos predictivos que controlan los mecanismos de búsqueda y recomendación en las diversas plataformas de Internet es la maximización de interacciones. Todos los seres humanos preferimos información que confirme nuestros propios gustos y preferencias por sobre aquellas que tienden a refutarlas o contrariarlas, por lo que los distintos agentes buscan y extraen dicha información para que los usuarios las visualicen y maximicen sus interacciones con la plataforma a la cual sirven. De ahí que se cree un ambiente donde el individuo está sólo en contacto con la información que confirma sus prejuicios, pero que jamás los pone en duda. Es más, puesto que los discursos radicales son más divertidos que los análisis desapasionados, serán los contenidos más brutales los que resultarán preferentemente elegidos para los usuarios, buscando fidelizarlos. Si una persona cree algo fantástico, como la existencia de una conspiración mundial para llevar lagartos al poder, la plataforma le mostrará aquellas páginas y videos que confirmen su estrambótica tesis, y luego de varias horas en contacto con los desviados argumentos a favor de esta idea, su certidumbre sobre tal conspiración mundial se reforzará. Es más, si el agente de la plataforma cree a que al usuario, de acuerdo a una comparación con otros individuos similares, le podría interesar alguna de estas teorías fantásticas, lo expondrá a ellas. En este sentido, las redes sociales buscan de manera intencionada, la radicalización de las ideas de sus usuarios.

Con esa información permite determinar cosas como los gustos, temores y preocupaciones individuales de los usuarios, y por lo mismo, sirven para establecer pautas de interés y formas de afectarlos. Un algoritmo puede determinar los momentos del estado emocional de un individuo a través de cuestiones sutiles, como la velocidad a que escribe, el temblor de su mano en el *mouse pad* o la presencia de determinadas palabras. De acuerdo a dichas tendencias emocionales, puede presentar la información que será de su interés o que afectará relevantemente su forma de actuar o de decidir. En 2016 Cambridge Analytica hizo algo similar, buscando determinar con la información de Facebook el perfil individual de diversos usuarios y presentarles información que influya en su voto de manera que el candidato republicano –entonces Donald Trump– saliese electo. Si bien es cierto que influir a cientos de millones de personas directamente era algo que estaba fuera de las capacidades de Cambridge Analytica, lo que sí podía hacer era determinar perfiles de electores indecisos en estados clave para la elección norteamericana y centrar su actuar en ellos, de manera de inclinar esos pocos y determinantes votos en el sentido deseado. Por la estructura de las elecciones en el mundo, en verdad no es necesario influir en toda la población para poder manipular sus resultados. La mayor parte de los procesos electorales se determinan por un pequeño número de votantes individuales que, si son moviizados eficientemente, pueden otorgar la victoria a una de las opciones. Esta movilización de los indiferentes puede lograrse mostrándoles información que excite su emocionalidad respecto al proceso y los lleve desde la indiferencia hacia la radicalidad. En pocas palabras, las redes sociales actúan agudizando el conflicto electoral, haciéndolo candente a fin de movilizar al electorado usualmente indiferente hacia el extremismo. No resulta un misterio que las redes sociales han influido en que los procesos electorales se hayan ido tornando, cada vez más, en procesos dramáticos donde los votantes llegan a una suerte de paroxismo emocional. Los mayores, que recordamos procesos anteriores que se vivieron sin dramatismo, pero con decisión y virilidad, como el plebiscito de 1988, nos sentimos perplejos ante las pasiones vertidas sobre cosas mal comprendidas y los tonos apocalípticos que tienen aspectos, en muchos casos ociosos, de nuestra vida cotidiana. Las redes sociales han contribuido a radicalizar la opinión pública mediante el manejo individualizado de perfiles de usuario y la selección de la información relevante que se muestra. Su consumo es peligroso y estúpido, pero lamentablemente no creemos que resulte domesticable en el corto ni mediano plazo.

Las redes sociales contribuyen hoy en día a crear un mundo feliz, donde los individuos pueden ser manipulados y sus opciones alteradas o reforzadas según convenga al controlador de las mismas. En un ambiente que tiende al monopolio, como es Internet, esto entrega un poder inmenso a las compañías que controlan su tráfico mediante herramientas de inteligencia artificial. Su uso político probablemente se agudice en un futuro donde la creciente desigualdad puede resultar naturalizada mediante su empleo. Los totalitarismos del futuro observarán los métodos brutales de un Stalin o un Hitler con la curiosidad con que se mira a un

chamán tratando de curar una enfermedad con brujería por no contar con antídotos. Espero sinceramente estar equivocado respecto a lo que he escrito.

5.3. ¿Algunas soluciones?

En un mundo con cientos de miles de personas a quienes les resulta imposible encontrar un empleo, es necesario replantearse el rol del trabajo en nuestra sociedad. La ética protestante transformó este castigo divino en un signo de salvación. Se desarrolló una ética del trabajo, una santificación del mismo, que pone en el centro de la realización personal. El trabajo es más importante que la familia, que la paternidad, que la pareja, que la vida. Todo debe posponerse en el servicio a ese dios poderoso y celoso, el único que otorga dignidad y que provee de lo necesario para no morir de inanición. Su divinidad, entonces, se fundamenta en dos atributos, su rol material, como dispensador de medios necesarios para la vida, y su sentido espiritual, como dignificador de la existencia. En un mundo donde el trabajo se transformará en un bien crecientemente escaso, es mejor retirarlo del altar y revocar su divinización, no sea que el infierno material y moral a que esta deidad condena a sus disidentes termine tan rebosante que haga estallar el orden social y político.

La vida es un derecho, y como tal debe garantizarse más allá de la posibilidad actual de ejercer un trabajo remunerado. Un mundo en donde todos pueden vivir aunque sea modestamente, sin que el hambre o la falta de cobijo nos amenacen constantemente y nos obliguen a someternos, transando nuestra libertad y tiempo vital por los medios para evitar tan triste destino, es una aspiración que merece ser considerada seriamente. En un mundo tal, donde la necesidad de ganar el pan con el sudor de nuestras frentes no exista, cosas que hoy no tienen una remuneración pecuniaria, pero que son socialmente relevantes y espiritualmente gratificantes, como el cuidado de nuestros hijos, las artes, la búsqueda de la verdad y el auxilio a los más débiles, pueden transformarse en las actividades que otorguen estatus y prestigio social. En todo notario están las ruinas de un poeta. Si la vida no dependiese de realizar un trabajo remunerado, tal vez de esas ruinas volverían a brotar flores.

La inteligencia artificial creará riquezas ingentes, toda vez que será capaz de racionalizar nuestra producción y de hacerla más efectiva y eficiente. La cantidad de esfuerzo que requerirá la producción de bienes será progresivamente menor, por lo que su costo caerá permitiendo rebajar sus precios. Esto ya es apreciable hoy. A fines de la década de 1980 un celular era un objeto de lujo. Su capacidad era escasa, ya que sólo permitía hacer llamadas de calidad cuestionable a un precio exorbitante, mientras que hoy por una fracción de tal valor tenemos un aparato que no sólo puede hacer video llamadas gratuitas, sino que reemplaza la cámara de fotos, de video, las consolas de videojuegos, los computadores personales, los mapas, los traductores y muchos otros dispositivos. En pocas palabras, la tecnología permite reducir el precio y aumentar las funcionalidades

de los bienes de manera que se otorgue un mayor valor por menor precio. La globalización económica también ha contribuido a ello, permitiendo que la ropa, por ejemplo, sea muchísimo más barata y de mejor diseño que hace cuarenta años. La tendencia general es que la producción aumente a precios cada vez más bajos, especialmente en los bienes cuyo componente más importante sea solamente la tecnología. En pocas palabras, muchos de los bienes necesarios para la vida disminuirán sus costos sustancialmente. La posibilidad de imprimir en 3D eliminará buena parte de los costos de construcción en el mediano plazo, disminuyendo, además, el tiempo en que las obras se realizan. Nuevos elementos como la carne sintética, que ya comienza a producirse, pueden tener un impacto similar en el largo plazo. En pocas palabras, la aplicación de nuevas tecnologías a los bienes de consumo básico hará que su costo de producción descienda dramáticamente, posibilitando un aumento de producción importante. El único elemento que desafía este incremento en la producción es como lograr que la demanda sea capaz de crecer igualmente. Si bien su menor precio debiera permitir que masas antes privadas de acceder a los mercados donde se trazan pudiesen demandarlos, esto no es necesariamente cierto, si es que buena parte de la población no tiene acceso a un trabajo remunerado, hasta hoy la única forma de obtener medios para adquirirlos. Ante este problema es que existe una creciente presión a favor de el establecimiento de un ingreso universal garantizado. Esta idea es vieja, ya fue planteada en la década de 1960 como respuesta al avance tecnológico y fue probada parcialmente en las *Speenhanland laws* del siglo XVII en la Inglaterra isabelina, e incluso podría decirse que la *annona* romana establecida por Cayo Graco apuntaba en tal dirección. Tanto desde la vereda neoliberal¹⁰³, como en la izquierda extrema se ha planteado su necesidad y hoy en día hay países que experimentan con ella, como Finlandia, aunque todavía no es la tendencia generalizada a nivel mundial. La idea es simple, otorgar a cada individuo una renta básica suficiente como para sacarlo de la línea de la pobreza. Si bien es una idea cara, hay notables ahorros que se permitirían al Estado a través de ella, toda vez que al eliminar la pobreza de un plumazo, todo el gasto social dirigido a superarla sería innecesario. Por lo demás, la criminalidad descendería de manera dramática, puesto que al desaparecer la pobreza material, el ambiente que genera la delincuencia ligada a la miseria también se evaporaría. La recaudación tributaria para financiar tal iniciativa debería provenir justamente de aquellas empresas que concentrarán mayor riqueza a través de la aplicación de inteligencia artificial, las cuales, a su vez, se verán beneficiadas por una mayor capacidad de consumo de la población, una caída en la criminalidad y una mayor percepción de legitimidad del sistema político, que se haría crecientemente estable.

¹⁰³ Milton Friedman simpatizaba con la idea llamándola impuesto negativo, estimando que era una forma eficiente de distribuir beneficios sociales a la población eliminando las trampas burocráticas que impiden que el gasto social llegue a los más necesitados. Vid: Tondani, Davide, Universal basic income and negative income. Two different ways of thinking redistribution in The Journal of Socio-Economics 38-2 (2009), pp. 246-255.

Ahora bien, aunque el ingreso universal garantizado no solucionaría necesariamente el problema de la concentración de la riqueza, toda vez que la rentabilidad del capital aumentará con la aplicación de tecnología a los medios de producción, mientras que las rentas del trabajo disminuirán con la escases del mismo, si tiene la fuerza suficiente como para que tal concentración no sea penosa para la población.

Es posible que un mundo donde el hambre no nos obligue a trabajar, algunos dediquen sus esfuerzos a actividades socialmente nocivas, pero creemos que ésta no será la tendencia general. Las aristocracias del Antiguo Régimen no tenían necesidad alguna de dedicarse a trabajar para vivir, pero no por ello se transformaron en una banda de criminales. Más bien, gracias a su gusto y mecenazgo nació la Ilustración. Si el privilegio de liberarse de las cadenas del hambre se extendiese a toda la humanidad, una revolución intelectual y social profunda podría llevarnos hacia el jardín del que un día fuimos expulsados.



BIBLIOGRAFÍA

- AAVV, Report of the Commission on Technology (1966).
- Abbot, Ryan, The Reasonable Computer, The George Washington Law Review.
- Alarie, Benjamin, Niblett, Anthony, Yoon, affect the practice of law, University.
- Ashley, K., Artificial Intelligence and Legal (Cambridge 2017).
- Barton, Benjamin H., Bibas, Stephanos, fewer lawyers, and the future of law (New York).
- Bayern, Shawn, "The Implications of Regulation of Autonomous Systems" Euro.
- Bhat, Aditya, Autonomous Vehicles: A International Journal of Engineering (2017).
- Blakin, Jack M., The Path of Robotics Law.
- Broussard, M., Artificial Unintelligence World, (MIT Press-Kindle, Cambridge).
- Buolamwini, J., Gebu, T, Gender Shade Commercial Gender Classification in search 81 (2018).
- Casey, Anthony J., Niblett, Anthony, Self Journal, 66-4 (2016).
- Chopra, Samir and White, Laurence F., Agents (University of Michigan Press).
- Colin, Rule, Online Dispute Resolution (Cisco).
- Davenport, Thomas H., Artificial Intelligence Review, On AI, Analytics and the Publishing-Kindle, 2019, Boston, MA.