

SELECCIÓN DE ESCRITOS SOBRE INTELIGENCIA ARTIFICIAL

"INTELIGENCIA ARTIFICIAL: ALGUNOS ASPECTOS DE SU IMPACTO"

Centro de Estudios en Tecnologías Inteligentes (CETI)

SERIE CONTRIBUCIONES COMPILADAS N° 6



ANCBA 2022

Ferreri, Juan Carlos

Inteligencia artificial : algunos aspectos de su impacto / Juan Carlos Ferreri ; compilación de Juan Carlos Ferreri. - 1a edición especial - Ciudad Autónoma de Buenos Aires : Academia Nacional de Ciencias de Buenos Aires, 2022.

CD-I, iBook

ISBN 978-987-537-171-2

1. Inteligencia Artificial. I. Título.
CDD 006.301

Fecha de catalogación: 3/2022

Academia Nacional de Ciencias de Buenos Aires

Av. Alvear 1711, 3º piso – 1014 Ciudad de Buenos Aires – Argentina

El presente trabajo se encuentra disponible sólo en versión electrónica

www.ciencias.org.ar

correo-e: info@ciencias.org.ar

La publicación de los trabajos de los Académicos y disertantes invitados se realiza bajo el principio de libertad académica y no implica ningún grado de adhesión por parte de otros miembros de la Academia, ni de ésta como entidad colectiva, a las ideas o puntos de vista de los autores.

ISBN 978-987-537-171-2

SISTEMAS INTELIGENTES Y ETICA

Antonio A. Martino*

*"La ética es la diferencia entre lo que tienes derecho de hacer y lo que es correcto hacer".
Potter Stewart*

"La ética es en origen el arte de recomendar a otros los sacrificios requeridos para la cooperación con uno mismo." Bertrand Russell

"Lo correcto es correcto, aunque todos los condenen, lo incorrecto es incorrecto, aunque todos lo aprueben" Charles Spurgeon

Sumario:

1. Algunos hechos

- 1.1 Los autos autónomos
- 1.2 Reconocimiento facial
- 1.3 Armas autónomas
- 1.4 Noticias falsas
- 1.5 Las máquinas reemplazan trabajo humano
- 1.6 Pandemia y pasos verdes
- 1.7 Sesgos de conocimiento

2. Los sistemas inteligentes

- 2.1 La inteligencia artificial
- 2.2 Los riesgos de la I.A.

3. La ética

- 3.1 Ética y moral
- 3.2 Siempre que aparecen riesgos se vislumbra la ética

4. Documentos poniendo límites

- 4.1 La Unión Europea
- 4.2 El Vaticano.
- 4.3 El Parlamento Europeo (PE)
- 4.4 La visión de los expertos: los 23 principios de la IA de Asilomar
- 4.5 La OCDE
- 4.6 La Unesco
- 4.7 Libro blanco de Inteligencia artificial de la UE

* Profesor emérito de las universidades de Pisa, Italia y Salvador, Argentina, miembro de la Academia Nacional de Derecho y Ciencias sociales de Córdoba Miembro asociado del Center for Artificial Intelligence and Cognate learning of the University of Greenwich

5. Ética y estética

6. Conclusiones

1. Algunos hechos

1.1. Los autos autónomos¹⁸ de marzo de 2018, a las 21.59 hs en Tempe, Arizona Elaine Herzberg fue embestida y muerta por un auto autónomo de prueba, de la firma Uber, conducido por los sistemas desarrollados para ese fin, con una conductora de seguridad llamada Rafaela Vásquez.

El automóvil disponía al momento del accidente, de cámaras con visión de toda la periferia, múltiples sensores, GPS, ubicación en mapas electrónicos, radar para la determinación de la velocidad relativa de cada objeto en movimiento, del LIDAR (Light Detection and Ranging) que permite una visión completa 300m de distancia y a 360°, con imágenes tridimensionales con muy alta definición debido a la utilización de láser. Todo esto controlado y comandado por un sistema inteligente de conducción.

La tecnología recibía los distintos datos y los procesaba con sus algoritmos, calculando continuamente y elaborando predicciones de las distintas alternativas de ocurrencia en los siguientes pocos segundos, en base a los objetos, personas, señalización y reglas de tránsito vigentes. El auto se desplazaba a 69 Km/hora, el sistema de detección de imágenes percibe 6 segundos antes de la colisión un objeto que no identifica en el medio de la calzada. Elaine Herzberg vestía un abrigo negro y llevaba a pie su bicicleta cargada.

El sistema detecta, un objeto no reconocido, luego interpreta un auto y finalmente una bicicleta, todo esto en concordancia con nuevas y sucesivas imágenes. Los sistemas de identificación necesitan varios miles de muestras de imágenes de un determinado objeto para aprender a identificarlo¹.

El algoritmo encuentra un objeto que desconocía, porque no estaba comprendido dentro los que habían aprendido, Estamos entonces en presencia de lo que en inteligencia artificial se denomina un “falso positivo”, es decir no identifica (falso) un objeto persona (verdadero). Los sistemas de reconocimiento de imagen expresan el resultado como una probabilidad que sea un objeto determinado, por ejemplo 95% casi certeza o un dudoso 15%.² Cuatro segundos después de la detección de un objeto no identificado, el sistema decide frenar para evitar el daño posible en el caso que la consistencia del objeto sea contundente, esto ocurre a 1,3 segundos de la colisión, la velocidad al momento del impacto fue de 63 Km/hora.

¹Un niño con unas pocas fotografías puede detectar enseguida una bicicleta. Es la diferencia entre un humano y una maquina: el primero se ubica en un ambiente y reconoce cosas porque puede ponderar, la máquina puede ejercer muchos razonamientos muy rápidamente, pero carece de ponderación. Para decirlo en modo simple: la maquina puede ser racional, nunca razonable.

²El programador debe evaluar cuál es el límite razonable de aceptación de la identificación, el caso del auto autónomo si el nivel de reconocimiento de la imagen es muy alto, cualquier objeto no reconocido hará que el sistema frecuentemente frene y hasta se detenga, según las circunstancias, haciendo que el vehículo realice una conducción torpe y brusca.

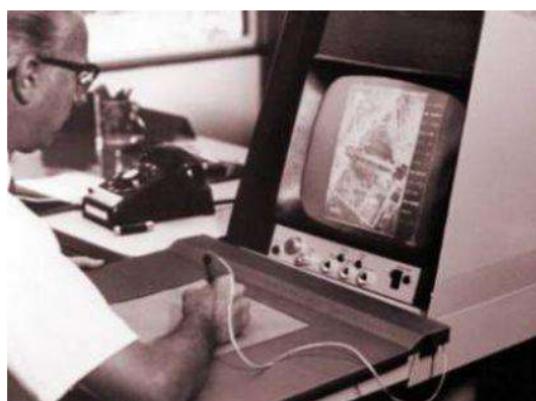
Esta larga descripción relata la concientización de la aparición de los autos autónomos: Tiene lugar cuando hay un trágico accidente. Es una constante en la vida humana: Percibimos las cosas cuando nos procuran una gran alegría o un enorme miedo. Pero hay también una segunda consecuencia tal vez más importante de la humana condición: En esos momentos de commoción y riesgo comenzamos a preguntarnos si es justo que algo así suceda, si está bien o está mal y aquí aparece la dimensión ética.

1.2 Reconocimiento facial

En los años 60 del siglo pasado Woodrow Bledsoe (1921-1995) y el dispositivo, la tableta RAND. Woodrow Wilson *Woody* Bledsoe, pionero de esta disciplina, fue además uno de los fundadores de la inteligencia artificial. Y contribuyó de forma importante al campo de reconocimiento de patrones, un aspecto determinante para la técnica que desarrolló en la década de los 60 y que entonces se llamó "reconocimiento facial de máquina".

El procedimiento consistía en clasificar una serie de retratos fotográficos de diferentes personas que previamente debía digitalizar. Aún sorprende la herramienta de la que se valió para realizarlo: La tableta RAND, un dispositivo electrónico bastante sofisticado para la época, abuelo de las actuales tabletas gráficas, que disponía de un lápiz que se comunicaba con una cuadrícula de 10 x10 pulgadas a través de impulsos electromagnéticos.

El programa de reconocimiento de escritura a mano que usaba la tableta se denominaba GRAIL (Graphical Input Language) y permitía dibujar formas y texto que el software suavizaba y renderizaba en un monitor en tiempo real. Lo más innovador era que incluso reconocía gestos y era posible eliminar, mover o cambiar el tamaño de los elementos. Bledsoe aprovechó estas funciones para marcar las características faciales mediante coordenadas, como la ubicación de la línea del cabello, los ojos o la nariz. Una vez digitalizados, estos registros numéricos se asociaban al nombre del individuo y se guardaban en una base de datos. El proceso inverso consistía en buscar en el fichero aquella cara que coincidiera en rasgos faciales, a partir de una imagen desconocida.



Tableta Rand en una demostración RAND.ORG

El sistema adolecía de las limitaciones técnicas de la época. El propio Bledsoe reconocía la incapacidad de las máquinas para distinguir incluso al mismo individuo en dos fotografías diferentes o desde distintos ángulos o iluminaciones. No obstante, supuso la primera piedra de un método que fue desarrollándose en los siguientes años gracias al progreso de los algoritmos y la capacidad de las computadoras.

El método tiene sus raíces en el siglo XIX: el oficial de policía francés Alphonse Bertillon creó un método para identificar criminales en función de sus características físicas. Había clasificado 11 tipologías y a partir de retratos fotográficos estandarizados, tomaba notas de sus rasgos por escrito.

Un siglo después, en 1994, el “padre fundador” de la identidad digital, Joseph Stick, planteó que las computadoras podían procesar información biométrica. No estaba equivocado.

La forma que tiene el software de identificar una cara en concreto es a través de un conjunto de referencias o puntos concretos (suelen ser aproximadamente 68) cuya distancia entre ellos y configuración son distintos para cada persona, confiriendo *patrones únicos como una huella digital*.

También existen otros métodos más modernos de identificación como el análisis de textura superficial que mapea y cataloga la textura de la piel, como si cartografiase cada poro y cada arruga de la cara.

La historia de esta tecnología está atada a China, uno de los primeros países en utilizar el reconocimiento facial para patrullar las calles. Y cuando decimos que es sorprendente nos referimos a este tipo de datos: más de 300 millones de cámaras de seguridad detectan, en China, caras individuales entre multitudes

El *reconocimiento facial* se ha convertido en los últimos años en un área de investigación activa que abarca diversas disciplinas, como procesado de imágenes, reconocimiento de patrones, visión por ordenador y redes neuronales. Involucra a investigadores de áreas de informática como a neurocientíficos y psicólogos. Se podría considerar también dentro del campo de reconocimiento de objetos, donde la cara es un objeto tridimensional sujeto a variaciones de iluminación, pose, etc., y ha de ser identificada basada en su proyección 2D (excepto cuando se utilizan técnicas 3D).

El objetivo de un sistema de reconocimiento facial es el siguiente: dada una imagen de una cara "desconocida", o imagen de test, encontrar una imagen de la misma cara en un conjunto de imágenes "conocidas", o imágenes de entrenamiento. La gran dificultad añadida es la de conseguir que este proceso se pueda realizar en tiempo real. El sistema identificará las caras presentes en imágenes o videos automáticamente. Puede operar en dos modos: 1. *Verificación o autentificación de caras*: compara una imagen de la cara con otra imagen con la cara de la que queremos saber la identidad. El sistema confirmará o rechazará la identidad de la cara. 2. *Identificación o reconocimiento de caras*: compara la imagen de una cara desconocida con todas las imágenes de caras conocidas que se encuentran en la base de datos para determinar su identidad.

Se utiliza principalmente en sistemas de seguridad para el reconocimiento de usuarios. En estos sistemas se utiliza un lector que define las características del rostro, y cuando este solicita el acceso, se verifica comparando los datos obtenidos con la base de datos. Sin embargo, estos sistemas no son útiles a largo plazo ya que, a medida que pasan los años, los rasgos faciales varían y al solicitar el acceso ya no coinciden con la imagen en la base de datos. Para solucionar este problema se puede utilizar un algoritmo que interprete el paso de los años, aunque igualmente sigue sin ser del todo fiable), o bien, renovar frecuentemente la base de datos.

También se utiliza en aplicaciones de interacción persona-ordenador, en gestión multimedia, y en software como Google's Picasa, Apple iPhoto, Sony's Picture Motion Browser (PMB), Facebook y Asus Smart Logon.

Una aplicación de reconocimiento facial futura se basa en establecer esta técnica a nivel de usuario. Por ejemplo, en un supermercado o en un establecimiento pequeño se podría llevar a cabo un control sobre quien abre la caja registradora mediante un reconocimiento facial previo, de esta manera también se pueden evitar intentos de robo ya que, al no reconocer el rostro, la caja permanecería cerrada. Un caso más extremo sería en los cajeros automáticos donde, para poder operar, fuese necesario un reconocimiento facial en vez del actual PIN.

Las aplicaciones que ya existen de reconocimiento facial se aplican en China para tener en todo tiempo un control absoluto sobre los ciudadanos, esto ¿está bien? ¿O es un gran reto para nuestras libertades?

Que haya una permanente visión de nuestros actos ¿es un atentado contra la privacidad a la cual, hemos aprendido, tenemos derecho?

Fuera del ámbito de la seguridad, ya vimos cómo la inteligencia artificial aplicada a este ámbito asomaba en nuestras redes sociales como algo inocente y original allá por el 2010, cuando Facebook lo incorporó para reconocer las caras de nuestros amigos en las fotografías que subíamos y las etiquetas que le proporcionábamos. Su uso se extendió rápidamente y hoy se encuentran en muchos de los smartphones y aplicaciones que utilizamos a diario.

Las polémicas sobre el uso de la biometría se han alzado principalmente al descubrirse que algunas empresas y organismos utilizaban la información recopilada para otro propósito diferente del autorizado. Este asunto ha estado en el centro del debate sobre ética y privacidad desde el 2001, propiciado a su vez por un vacío legal en la aplicación de las nuevas tecnologías de la información.

El Sistema de Reconocimiento Facial de la INTERPOL (IFRS) almacena imágenes faciales enviadas por más de 160 países, lo que la convierte en una base de datos única en el ámbito policial. Este sistema, puesto en marcha a finales de 2016 ha logrado identificar a más de 650 delincuentes, prófugos y desaparecidos. Pero su web ya advierte que la calidad de las imágenes es un aspecto esencial y que aquellas que solo posean una resolución media o baja pueden no conseguir o influir negativamente en la precisión de la búsqueda.

En los últimos meses, las controversias sobre su uso han aumentado, impulsadas en parte por el movimiento Black Lives Matter³, por lo que diferentes organismos han empezado a recular acusando excesivo control y posible promoción del racismo y la injusticia social. Amazon o IBM dieron un paso atrás. San Francisco ya se había convertido el pasado año en la primera gran ciudad de EE UU en prohibir a todas las agencias locales, entre ellas la policía, el uso de técnicas de reconocimiento facial, por miedo a que se las utilice para discriminar.

En nuestro país hubo muchos casos de detenciones erróneas, como uno en Bahía Blanca en 2019. No es sólo en Argentina que hay organizaciones que se oponen a este sistema. A mediados del año pasado, por ejemplo, hubo un caso en Estados Unidos que despertó fuertes polémicas. Esto reavivó la preocupación sobre los riesgos de sesgo en la inteligencia artificial, en medio de una ola de protestas contra el racismo y la violencia policial.

Según la Unión Americana de Libertades Civiles (ACLU), que presentó una denuncia en su nombre el 24 de junio de 2020, "aunque es el primer caso conocido, probablemente no sea la primera persona que ha sido detenida e interrogada erróneamente sobre la base de un problema de reconocimiento facial".

Para Joy Buolamwini, fundadora del grupo activista Algorithmic Justice League, el caso está revelando "cómo el racismo sistémico puede codificarse y reflejarse en las inteligencias artificiales (IA)". De hecho, bajo la presión de grupos como la poderosa ACLU, Microsoft, Amazon e IBM anunciaron a principios de junio que limitarían el uso de sus herramientas de análisis facial por parte de las fuerzas del orden. De hecho, el sistema es cuestionado en todo el mundo: está probado que hay sesgos raciales en esta tecnología.

En ciudades como San Francisco se prohibieron estas técnicas de reconocimiento facial para identificar a criminales. Alemania lo vetó para su policía y en Francia los colegios no pueden usar este sistema. En las antípodas de estas medidas están países como China e Inglaterra. En la Policía Metropolitana londinense implementó desde el año pasado la tecnología de reconocimiento facial en las calles. El debate está abierto porque los derechos civiles pueden verse fuertemente vulnerados.

1.3. Armas autónomas

La tarde del 2 de enero de 2020 dos misiles alcanzaron en los exteriores del aeropuerto de Bagdad a un vehículo en el que viajaban el comandante de una milicia iraquí Abu Mehdi al Muhandis, y Quasem Soleimani, líder de la Fuerza Quds, unidad élite de la Guardia Revolucionaria Islámica. Fuentes oficiales de Washington confirmaron horas más tarde la autoría del ataque (Guimón, 2020, 3 de enero). Este golpe militar fue significativo no solamente porque se produjo en un momento de máxima tensión política entre Irán y Estados Unidos, sino también por su modus operandi. Primero, el ataque fue ejecutado con un dron MQ-9 Reaper, un vehículo aéreo pilotado a remoto desde bases aéreas que se encuentran normalmente lejos de la zona de operación. Segundo, la zona de la operación, el

³Movimiento surgido en EEUU frente a abusos contra la comunidad afroamericana y que significa literalmente "la vida negra importa".

aeropuerto de Bagdad se encuentra en un país, con el que Estados Unidos oficialmente no está en guerra. Tercero, las víctimas del ataque fueron líderes militares a los que Estados Unidos consideran enemigos que, según las mismas fuentes, estaban preparando ataques contra el personal diplomático y militar estadounidense (IBID.). Se trataba, por tanto, de un ataque selectivo.

No es este el origen de las armas autónomas, sino que es uno más de una larga serie que Estados Unidos está ejecutando desde la primera década del nuevo milenio en el exterior, principalmente en países de Oriente Medio y forma parte de la doctrina de guerra preventiva, cuyo artífice es la anterior Secretaria de Estados de Estados Unidos, Condoleezza Rice.

Desde la publicación del informe del Christof Heyns, Relator Especial sobre ejecuciones extrajudiciales, sumarias o arbitrarias explica en su informe, del 9 de abril de 2013, se han multiplicado las voces que advierten de los peligros que podrían manifestarse en el futuro, cuando los drones adquieran tanta autonomía que podrían elegir los objetivos y ejecutar los ataques sin la intervención del hombre. El Secretario General del Consejo de Seguridad Naciones Unidas, advierte en su informe Protección de civiles en conflictos armados de la necesidad de que se regule sistemas de armas autónomas letales (SAALs en adelante) en cumplimiento con el Derecho Internacional Humanitario. En la actualidad, ya son nueve países que están desarrollando armas autónomas letales. Siete de ellos son: Estados Unidos, China, Rusia, Reino Unido, Francia, Israel y Corea del Sur⁴. A estos hay que añadir Irán, cuyos drones se han encontrado en los ataques a las instalaciones de Saudi Aramco, y Turquía que está usando sus drones en Libia. Este dato revela la importancia del área y no ha pasado de forma desapercibida en el mundo académico y la opinión pública.

Las preguntas que se plantean son ¿Quién es el responsable en el uso de armas autónomas letales y de las bajas por un ataque? ¿Qué ocurre cuando un arma actúa de forma tan autónoma que para un comandante ya no es posible frenar el ataque? ¿Cómo afectarían armas autónomas en los ámbitos del Derecho Internacional Humanitario y el Derecho de Conflictos Armados?

Todos los sistemas manifiestan cierta probabilidad de fallar, y si no existe ningún control - sea otro sistema o el hombre – las consecuencias pueden ser dramáticas. Uno de los ejemplos más notorios es el del Teniente Coronel Stanislav Petrov quien evitó una guerra nuclear el 26 de septiembre de 1983, cuando confió más en su intuición en lugar del sistema alerta temprana del que era responsable. El nuevo satélite Oko había interpretado los reflejos de la luz solar en las nubes como el lanzamiento de misiles intercontinentales. Petrov decidió contrastar la información del sistema con la vigilancia aérea y sus superiores en lugar de iniciar el contraataque que exigía el protocolo de acción ¿Qué hubiera ocurrido si un sistema automatizado hubiera actuado sin la intervención humana? No es probable que en un escenario de estas envergaduras el hombre esté fuera del bucle, pero ¿Quién puede garantizar que esto no cambie cuando un conflicto se escale?

¿Quién puede garantizar que un país no despliegue estas armas en el modo de autonomía supervisada para responder con más rapidez a un ataque masivo de enemigo?

⁴ Report: Kill the idea of 'killer robots' before they kill us, Center for a New American Security, ápid Aznar, 2020, p. 26.

Críticos de las armas autónomas alegan también una posible mayor predisposición para la entrada en conflicto, debido al menor riesgo para las personas.

La razón que alegan muchos críticos es la ingente cantidad de ejecuciones extrajudiciales que han estado ejecutando las fuerzas de seguridad estadounidenses durante los últimos años en Oriente Medio. El informe de Heyns, antes citado, estaba motivado precisamente por este tipo de ataques selectivos.

Estados Unidos llevó a cabo en diez años más de 300 ataques aéreos contra combatientes talibanes y cuadros de Al Qaeda en zonas fronterizas. Pero estos ataques no fueron operaciones militares stricto sensu. Al contrario, se trataba de una operación encubierta de la Agencia Central de Inteligencia. Según un recuento de la New America Foundation, entre 2004 y marzo de 2013, los ataques con drones acabaron con la vida de entre 1535 y 2660 militantes de talibanes⁵. La misma fuente señala que solamente el dos por ciento de las víctimas fueron cuadros de Al Qaeda o de otros grupos terroristas, mientras más de 80 por ciento eran militares rasos o colaboradores que apoyaban la yihad de forma no combativa (IBID).

En febrero de 2020, después de 15 meses de consultas con expertos de la industria, la academia, el gobierno y el público americano, el Departamento de Defensa adoptó cinco principios éticos para la inteligencia artificial (DoD, 2020). Estos principios forman parte de la Iniciativa Americana de Inteligencia Artificial que trata de promover el uso innovador de la IA, mientras protege las libertades civiles, la privacidad y los valores. El análisis de la estrategia nacional de defensa se basa en el resumen ejecutivo, porque el documento entero no está disponible en la página del departamento de defensa americano.⁶

Los principales marcos regulatorios bilaterales y multilaterales para la regulación de armas se crearon cuando las armas aún no disponían de la autonomía actual. Su objeto de regulación eran armas químicas (Convención sobre Armas Químicas, de 1993), armas biológicas (Convención sobre las Armas Biológicas de 1975), armas nucleares (Tratado de No Proliferación Nuclear de 1968 y el Tratado sobre Fuerzas Nucleares de Rango Intermedio de 1987) y armas en el espacio (Tratado sobre el Espacio Ultraterrestre de 1967). De fechas más recientes son la Convención sobre la prohibición de minas antipersonales, de 1997 y el Tratado de comercio de armas de 2014. Ninguno de estos tratados regula sistemas de armas autónomas. Una excepción sería el Régimen de Control de Tecnología de Misiles, de 1987, que regula al menos la exportación de vehículos aéreos no tripulados.

⁵ New America Foundation, apud Montoya, 2014

⁶ Según este documento, la inteligencia artificial debe cumplir los siguientes Principios: 1. Responsable: El personal del Departamento de Defensa será responsable del desarrollo, el empleo y el uso de las capacidades de la IA, manteniendo niveles apropiados de juicio y precaución. 2. Equitativo: El departamento procurará minimizar sesgos no intencionados en las capacidades de IA.

3. Trazable: El personal involucrado deberá disponer el entendimiento apropiado de la tecnología y de los métodos operacionales que se pueden aplicar en las capacidades de la IA. Estas incluyen metodologías transparentes y auditables, las fuentes de datos, el diseño de los procesos y de la documentación.

4. Fiable: Las capacidades de la IA tendrán usos bien definidos y explícitos, y la seguridad y eficiencia de estas capacidades. Estas serán objeto de pruebas de seguridad durante todo su ciclo de vida. 5. Gobernable: Con este principio, el departamento se reserva el derecho de desconectar o desactivar sistemas desplegados que muestran un comportamiento no intencionado

Un sistema de armas, por muy sofisticado y autónomo que sea, no puede ser un agente moral porque no tiene ningún vínculo emocional con los hombres. Es por este motivo por el que la mayoría de los expertos rechazan la idea que un sistema deba tener la opción de decidir sobre vida o muerte. Las implicaciones jurídicas se refieren al respeto de algunos principios claves en el Derecho Internacional Humanitario, en concreto los principios de distinción, proporcionalidad y precaución. Existen varios niveles donde se toman decisiones.²⁴ Especialmente en la planificación de un ataque existe la posibilidad de respetar estos principios. Un sistema no funcionará nunca con plena seguridad. Esta es la principal razón por la que el hombre siempre debería mantener el control sobre el sistema, reivindicación que repiten numerosos críticos de los sistemas de armas autónomas. Los dos incidentes descritos en este capítulo muestran el peligro cuando un sistema malinterpreta una información e induce a activar un protocolo de contraataque.

1.4. Noticias falsas

El periódico neoyorquino The Sun en 1835 informaba de seres que habitaban la Luna. La noticia causó un enorme impacto en EEUU gracias a tres factores: la aparición de las prensas de alta capacidad, la caída del precio de los periódicos (la penny press), y la llegada de los nuevos medios de transporte que superaban la velocidad de los caballos por primera vez en la historia: los trenes y los barcos de vapor. Esos factores ayudaron a difundir a gran escala una información falsa y sensacionalista bajo el disfraz de una noticia verdadera. Es lo que hoy denominamos fake news

La propaganda busca convencer, ser eficaz, y para eso puede recurrir a todo tipo de instrumentos, desde el arte hasta el cine, los pasquines o las redes sociales. Las noticias falsas, una de las ramas de la propaganda, son diferentes: buscan engañar, crear otra realidad. La preocupación por la forma en la que estos engaños cuajan y por los mecanismos a través de los que se crean y se multiplican no es nueva.

Las noticias falsas envenenan el mundo de las noticias y provocan desequilibrios morales.

El cuestionamiento del *ethos* comunicacional e informativo —o sea, de las normas, principios y valores que fundamentan las formas de comunicación y la información en una determinada sociedad— gira actualmente en torno a temas tan variados como la privacidad, la propiedad intelectual, el acceso libre al conocimiento, el derecho a la expresión en las redes digitales, la censura, las nuevas definiciones de género, la identidad digital, las comunidades digitales, el plagio digital, la sobrecarga informativa, la brecha digital y el control social digital

El filósofo italiano Luciano Floridi distingue entre una «ética de la comunicación global» (*ethics of global communication*) y una «ética de la información global» (*global information ethics*)⁷ la primera se refiere al diálogo y las prácticas consensuales en la interacción entre diversas culturas y generaciones. Pero tratándose de una postura meramente pragmática deja de lado, según Floridi, preguntas más básicas como, por ejemplo: ¿cuáles son en este caso los principios éticos que se toman como base de dicho

⁷Luciano Floridi, *Sextus Empiricus, The Recovery and Transmission of Pyrrhonism*. Oxford: Oxford University Press, 2002.

diálogo y en qué se fundamenta dicha elección? o ¿existe una «macroética» en el sentido de un cierto tipo de consecuencialismo o deontologismo o contractualismo? ⁸

1.5 Las maquinas reemplazan trabajo humano

Microsoft ha decidido sustituir a la mayoría de los periodistas que tenía contratados para producir noticias para el sitio web MSN.com y en su lugar emplear robots. La medida supondrá el despido de unos 50 redactores, aunque la compañía ha señalado que también se quedarán otros muchos profesionales en plantilla.

Entre las funciones que desempeñaban los periodistas que serán sustituidos estaba la de organizar las noticias en la portada de MSN, elegir los titulares y las imágenes que acompañan el contenido, y muchas otras. Sin embargo, a partir de finales de junio estas funciones de producción de noticias las realizarán sistemas automatizados utilizando la Inteligencia artificial.

La compañía no ha precisado si hará más cambios en otros departamentos que también impliquen sustituir el trabajo desempeñado por seres humanos por robots. Microsoft ha asegurado que adopta esta decisión tras evaluar la evolución de su negocio y la necesidad de reducir las inversiones para abaratizar costes. De todas maneras, la firma aclara que los cambios no se vinculan con la pandemia de COVID-19.

La decisión ha provocado las críticas de algunos periodistas afectados, que consideran «desmoralizante» que una máquina pueda reemplazarlos y realizar su trabajo. Además, hay que tener en cuenta que la Inteligencia Artificial debe pasar por un proceso de formación y aprendizaje para poder sustituirles y esto podría llevar a la generación de contenido inapropiado en el portal de noticias.

Las estimaciones con respecto a la cantidad de trabajos que se destruirán con la llegada de las máquinas (robots, inteligencia artificial o cómo le quieran llamar) son variadas, pero la gran mayoría apuntan a una pérdida importante que afectará principalmente a aquellos puestos que requieren trabajo físico.

El Foro Económico Mundial (WEF) acaba de publicar su informe '*Future of Jobs 2018*',⁹ donde se lanzan con cifras que podrían sonar igual de alarmantes, pero que en esta ocasión nos plantean un escenario un poco más alentador. Y es que sí, el estudio pronostica que se perderán empleos, pero afirman que se crearán nuevas funciones que requerían un mayor grado de especialización.

Un ejemplo concreto de conflicto entre las visiones económicas y éticas: la muy debatida cuestión de la pérdida de puestos de trabajo inducida por la automatización, incluida la IA. Un aspecto crucial, especialmente en el futuro, que puede conducir a un aumento sustancial de las desigualdades, a menos que se adopten medidas, principalmente políticas, pero no sólo, para evitarlo.

⁸Luciano Floridi, *The Blackwell Guide to the Philosophy of Computing and Information*. (a cura di) Oxford: Blackwell, 2003.

⁹ http://www3.weforum.org/docs/WEF_Future_of_Jobs_2018.pdf

"¿Es correcto introducir nuevas tecnologías que conduzcan a la pérdida de puestos de trabajo?", pregunta el representante de la Oficina Nacional de Investigación Económica de Estados Unidos. Y trata de responder: "los economistas podrían estar tentados a responder inmediatamente 'sí' a mi pregunta. Pueden observar que en un mercado que funciona bien, los salarios reflejan perfectamente el valor social del trabajo; si a un nivel determinado de salarios, una empresa encuentra conveniente innovar de manera que ahorre costes de mano de obra, libera mano de obra para emplearla en otras actividades más útiles para la sociedad.

En la práctica, si un determinado trabajo ya no es necesario, se vuelve obsoleto, pierde su utilidad y, por lo tanto, su valor, por lo que es mejor no hacerlo más y cambiarlo. Después de todo, siempre ha funcionado así. "Pero también hay que tener en cuenta que la pérdida de puestos de trabajo es socialmente más cara de lo que sugiere una visión eficiente del mercado".

No es sólo el mercado el que puede decidir

1.6 Pandemia y permisos verdes

El diciembre del 2019 aparecieron los primeros casos de Covid 19 la epidemia abrazo 2020 y lo que va del 2021 y no parece tener un final cercano. Esto ha modificado nuestro modo de vivir, aislados, distanciados, con barbijos, vacunados y ahora con obligación de presentar un certificado que se llama *Green pass*. Para viajar los países europeos han creado el *European Union digital Passenger Locator Form (EUdPLF)*¹⁰ pero se lo exige también en las escuelas, en los medios de transportes, en los lugares de comidas y en el trabajo. Esto está dando lugar a aguerridas disputas morales con intervenciones de jefes políticos (el Premier italiano Draghi dijo que no vacunarse es suicida y promover no vacunarse a los demás una instigación al suicidio y hasta el Papa quien proclamo que vacunarse es un acto de amor).

La discusión seria es sobre el pedido de certificación para realizar actividades como trabajar: no solo los docentes han protestado sino también las centrales obreras sobre todo porque después que lo solicitaran las empresas públicas, ahora también los privados piden certificados para entrar en las fábricas o lugares de trabajo.

En Francia se discute porque es obligatorio para los gendarmes, pero no para la policía y en general si este requisito no vulnera la libertad de los ciudadanos para disponer de las terapias sobre su propio cuerpo.

1.7 Sesgos cognitivos.

Trabajando con máquinas hemos descubierto que transmitimos nuestros sesgos a los programas que creamos, y por ende son prejuiciosos. Pero no somos los únicos que nos ocupamos del problema. En psicología se ocupan de sesgos cognitivos y los efectos psicológicos que dificultan procesar la información y llegan a tergiversar interpretaciones y a juicios inexactos, es la irracionalidad que distancia al *Homo sapiens* del *homo economicus*. Se han hecho estudios y se han podido verificar alrededor de 200 errores que cometemos.

¹⁰ Permiso con formas estrictas que se obtienen en la plataforma de la UE.

Pero ¿el *Homo sapiens* no es la especie cuya capacidad intelectual lo lleva a la cima de la naturaleza? Parece que muchos de los errores de sesgos cognitivos son conductas adaptativas que surgieron cientos de miles de años atrás y que fueron útiles hasta no hace mucho. Cada cambio importante en la cultura llevo a cambios en nuestros sesgos cognitivos. Y si esto fuese así todos tendríamos más o menos los mismos sesgos. Pero no es el caso: el antropólogo Joseph Henrich publicó un libro¹¹ donde los más extraños del mundo son los Weird (western educando, industrializado, rico y democrático) ¿Te identificas por tu profesión o logros, más que por tu red familiar? ¿Cultivas tus atributos y metas únicas? Si es así, tal vez usted es RARO: criado en una sociedad que es occidental, educada, industrializada, rica y democrática. A diferencia de la mayoría de los que han vivido alguna vez, las personas WEIRD son altamente individualistas, inconformistas, analíticas y orientadas al control. ¿Cómo las poblaciones WEIRD llegaron a ser tan psicológicamente peculiares? ¿Qué papel jugaron estas diferencias en nuestra historia y qué significan para nuestra identidad colectiva?

Los sesgos son diferentes según la cultura y si los Weird están en un extremo hay otros que están el opuesto y obviamente intermedios. Con un ejemplo: Ud. Acompaña una amiga en auto, doblan por una calle casi desierta a mayor velocidad de la permitida, aparece un transeúnte y lo atropellan, la pregunta es si quien conduce tiene derecho a pedir al amigo que mienta con respecto a la velocidad del auto. Los estadounidenses, canadienses, australianos y escandinavos en un 90 % dicen “no”. Venezuela, Corea del Sur o Rusia, el 50 % dice “si”. No hay un dato seguro, pero en Argentina estaríamos en un 65 % de “no”. El Weird tiene mejor economía, educación y gobierno, pero son prejuiciosos con respecto a los que no son como ellos. Los opuestos tienen mala economía, gobiernos despóticos, menor educación y menor tasa de ahorro, pero son menos prejuicios para con los demás.

Y todo esto tiene consecuencias morales: los que son más estrictos moralmente tienen más prejuicios y esto se transmite también a los programas que crean. En el reconocimiento facial hay problemas cuando la persona no es de tez blanca.

2. Sistemas inteligentes

2.1. La inteligencia artificial

Hay muchas definiciones de I.A. Una de ellas, la que se va afirmando, es decir proporcionar a una computadora capacidad de aprendizaje y aplicación de modelos estadísticos no programados explícitamente. Otra dice Programa de computación diseñado para realizar determinadas operaciones que se consideran propias de la inteligencia humana, como el autoaprendizaje.¹²

Prefiero hablar de sistemas inteligentes que es aún más amplio pero que voy a definir en modo sencillo como aquellos que presentan un comportamiento externo similar en algún aspecto a la inteligencia humana o animal. Se caracterizan por su capacidad para representar, procesar y modificar de forma explícita conocimiento sobre un problema, y para mejorar su desempeño con la experiencia. Esto les permite resolver problemas

¹¹ Joseph Hernrich *The WEIRDest People in the World*, Penguin Libros Ltd, 2021

¹² Una más simple e irónica dice “habla, pero no entiende, oye, pero no comprende, mira, pero no ve”

concretos determinando las acciones a tomar para alcanzar los objetivos propuestos, a través de la interacción con el entorno y adaptándose a las distintas situaciones.

Sistemas inteligentes se remiten a “pensar” pero el vocablo en el Diccionario de la Real Academia española tiene siete casos que van desde formar o combinar ideas o juicios.

Examinar mentalmente algo con atención, opinar, tener intención de hacer algo, formar en la mente un juicio, recordar, tener en consideración algo o a alguien.¹³ A esto podría añadirse recordar, deducir, querer, inducir, odiar, etc. Pero un sistema inteligente hace más que pensar: recoge información, llegado a una conclusión puede realizar acciones tales como dar órdenes para que se ejecuten acciones, a veces tan complejas, como dirigir el tránsito en Los Ángeles.

El aumento de la inteligencia es la capacidad de integrar nuestras facultades biológicas naturales de razonamiento con otras nuevas basadas en los avances culturales y tecnológicos.

Hasta algo tan fundamental como hablar y entender el lenguaje no es enteramente biológico. Desde el discurso, la escritura, la impresión y las comunicaciones electrónicas, el proceso ha sido largo e importante. Ha incrementado enormemente nuestra capacidad para recoger, analizar y actuar sobre la información.

El humano es un ser lingüístico: y por eso el Evangelio de San Juan comienza con 1, 1 primer versículo. «En el principio era el Verbo, y el Verbo era con Dios, y el Verbo era Dios». ¹⁴

La teoría del razonamiento tiene una larga tradición en la lógica y ésta ha sido siempre la manera primera de abordar los problemas. El próximo salto cuántico en la mejora de las herramientas que nos permiten recoger datos sobre el mundo va a ocurrir a través de interfaces cerebro-máquina de varios tipos. Ahora se están probando y luego se aplican casi exclusivamente a personas con discapacidades, Pero al igual que ocurrió con el reconocimiento de voz por ordenador, estas interfaces se convertirán rápidamente en herramientas indispensables que todos queremos utilizar.

Cuando un mundo agitado te expone a grados de adaptación requeridos más allá de tus habilidades naturales, tienes dos opciones: rendirte, porque no eres capaz de superar esos límites o abrazar las adaptaciones a través de medios tecnológicos. A través de estos medios tecnológicos, puede extender el rango de su adaptabilidad al mundo cuya aceleración está aumentando. Como siempre se trata de un tema de fines y medios. No debes sentirte privado de tu capacidad de participar activamente sin ningún poder, puedes sentirte como un protagonista en el mundo del mañana, pero debes medir cuales son los riesgos y si moralmente estás de acuerdo.

2.2 Los riesgos de la IA

¹³ <https://dle.rae.es/pensar>. Personalmente no creo que se pueda decir de una máquina que piensa. Pero obviamente hay otras posiciones.

¹⁴ La frase «el Verbo» (una traducción de la palabra griega «Logos»)

Todo cambio significa tomar riesgos y cuanto más audaz e importante y omnicomprensiva es una actividad mayor es el riesgo. En los últimos 18 años hemos tenido más cambios importantes que todos los que tuvimos anteriormente y eso supone una enorme cantidad de riesgos.

Esto es importante para comprender que ha cambiado mucho el mundo y desgraciadamente es solo cuando aparecen los peligros que los humanos nos ocupamos de la ética.

En 2017 en California (EEUU). En el inmaculado centro de conferencias de la ciudad costera de Asilomar, más de un centenar de expertos en inteligencia artificial (IA) de distintos lugares del globo están reunidos bajo una premisa aparentemente sencilla: analizar e indexar de qué forma ésta es beneficiosa para el ser humano. Los profesionales invitados que cubren áreas diversas tienen tres días por delante para debatir cómo ha de desarrollarse la IA, podría decirse que en el futuro. Pero éste, más que nunca, es sinónimo de ya, aquí y ahora.

Porque hay algo en lo que convergen intelectuales divergentes alrededor del mundo: la base sobre la que se asentará la inteligencia artificial, y también lo que supondrá para la humanidad, se está gestando mientras escribimos este artículo.

El futurista alemán Gerd Leonhard, es autor de un libro cuyo título no deja lugar a dudas: *La tecnología contra la humanidad, el choque entre el hombre y la máquina*: "Hoy es el momento en que construimos nuestro futuro y este hoy es el mayor reto al que se ha enfrentado el género humano hasta el momento, la posibilidad de una especie de superhombre que se convierte en Dios no tiene precedentes"¹⁵.

Con nuestra eterna búsqueda de libertad y felicidad. Antes de que sea demasiado tarde, hemos de detenernos y hacernos las grandes preguntas: ¿Cómo podemos acoger la tecnología sin convertirnos en ella? Cuando esto ocurra —gradualmente, luego súbitamente— la era de las máquinas creará el mayor hito de la vida humana sobre la Tierra. Estamos en uno de los últimos mapas morales que tendremos conforme la humanidad.

3. La ética

3.1 Ética y moral

La palabra ética proviene del griego ethikos (“carácter”). Se trata del estudio de la moral y del accionar humano para promover los comportamientos deseables. Una sentencia ética supone la elaboración de un juicio moral y una norma que señala cómo deberían actuar los integrantes de una sociedad. “moral” viene de mores, costumbre y se podría ensayar una distinción entre las dos diciendo que la ética es la parte teórica que se ocupa de los comportamientos humanos aceptados o permitidos (es prescriptiva) mientras que la moral se ocupa de cómo estos principios son aceptados, incorporados o evolucionan en un determinado contexto histórico y social (es descriptiva).

¹⁵ Gerd Leonhard, *La tecnología contra la humanidad, el choque entre el hombre y la máquina*, Kindel unlimited, 2017.

Por esa razón puede haber una ética universal, aunque las morales concretas sean particulares. Esta es una buena razón para mirar con respeto la tentativa de la Unesco de hacer para fines del año próximo un texto que sirva para poner los límites éticos a cualquier sistema inteligente.

La ética es parte de la filosofía y esta es una disciplina distinta a todas las demás porque obliga a quien la practica seriamente a exponerse fuera de todos los conocimientos ya adquiridos. La ética es una manera de sentirse bien con las propias acciones para con uno mismo y para con los demás. Como dice Kant no es una receta de felicidad, sino que es un camino para encontrarla o, como dice Bertand Russell “La ética es en origen el arte de recomendar a otros los sacrificios requeridos para la cooperación con uno mismo.”

Si alguien preguntase como resumir rápidamente las posiciones más famosas en materia ética creo que podría hacerse una breve lista que incluyera Sócrates, Platón, Aristóteles, el cristianismo. Kant, K. Marx, Pragmatismo, Utilitarismo y el Existencialismo, pero las más importantes para nosotros son dos, casi extremas: Emanuel Kant que sostiene “«Obra sólo según aquella máxima por la cual puedes querer que al mismo tiempo se convierta en ley universal. Obra como si la máxima de tu acción pudiera convertirse por tu voluntad en una ley universal de la naturaleza» (AA IV: 421”¹⁶

El otro es Charles Ross quien dice que tenemos un deber *prima facie* de ayudar a los demás, otro de mantener nuestras promesas, otro de devolver los actos de amabilidad anteriores y otro de no defraudar a las personas que confían en nosotros. Y habla de deberes *prima facie* porque nuestros deberes cambian con la información que obtengamos¹⁷

Por cierto, que hay posiciones intermedias pero estas dos nos dan el alfa y el omega: comportarse conforme a un canon universal e invariable, Kant dice que es un imperativo categórico porque a diferencia de los preceptos hipotéticos que son condicionales y que tienen que ver con condiciones de lugar, tiempo y espacio éste es accesible a cualquier ser humano inteligente y tiene vigor siempre. Ross, por el contrario, considera que todo deber moral está sujeto a las condiciones y es más a las informaciones que tenga el sujeto agente de dichas condiciones.

Hoy con nuevas condiciones de vida, con el mundo digital y la Pandemia Byung-Chul Han: "Nosotros, felices esclavos de la pandemia digital" En una entrevista concedida en estos días el filósofo sudcoreano afirma que estamos tan entrampados en la trama de las redes sociales y la I.A. Que estamos contentos en la Pandemia porque tenemos Internet¹⁸ "La Red nos controla y nos guía como si fuera una dictadura". Describe nuestra "sociedad de supervivencia" en la que aquellos que no pueden permitirse el lujo de aislarse corren un mayor riesgo de enfermarse y donde uno está obsesionado con salvar su vida, olvidando lo que hace que valga la pena vivir. El miedo crece, el estado de emergencia es visto como la nueva normalidad y los llamados "hombres fuertes" que se inspiran en modelos autocráticos como el chino son los que se benefician de ello. China corre ahora el riesgo de convertirse en

¹⁶ Emanuel Kant, Fundamentación de la metafísica de las costumbres, No Books, Buenos Aires, 2003, versión original Grundlegung zur Metaphysik der Sitten, Riga, 1785.

¹⁷ David Ross, Fundamentos de Ética, Eudeba, Buenos Aires, 2003. El original inglés es de 1930

¹⁸ https://rep.repubblica.it/pwa/robinson/2020/10/30/news/byung-chul_han_noi_schiavi Felici_della_pandemia_digitale_-272446064/?ref=nl-rep-a-bgr

un modelo de régimen de vigilancia impulsado por una falta de humanidad alimentada por nuestro miedo a morir.

3.2. Siempre que aparecen riesgos se vislumbra la ética

Cada vez que los humanos nos embarcamos en nuevas acciones o somos llevados a ella por acontecimientos externos y aparecen los riesgos surgen los problemas éticos. Hay una película que recomiendo que se llama *¿Cuánto vale una vida humana?*¹⁹ La película cuenta un hecho real provocado a partir de la caída de las Torres gemelas el 11 de septiembre de 2001.

Luego del ataque terrorista hubo muchos muertos y desaparecidos. Pasado el estupor y en medio del dolor de los familiares de sobrevivientes comenzó a vislumbrarse otro problema: el número de los muertos, desaparecidos, auto arrojados al vacío (*Falling Man*) mutilados, heridos y gente que continuó muriendo por enfermedades producidas por el incendio y posterior derrumbe. Se habló de más de mil cien víctimas directas a las que se agregaron tantas otras con el paso del tiempo. EEUU es un país con muchas causas legales y con muchos abogados²⁰.

Se calculaba unas seis mil familias con víctimas que podían demandar a las compañías aéreas y colapsar la economía norteamericana por años. El presidente Bush decide crear un fondo. Solo en el primer mes después del ataque, las pérdidas en general para la economía sumaron 123.000 millones. Las ayudas económicas a las aerolíneas fueron calculadas en 15.000 millones de dólares y el Congreso aprobó un paquete de emergencia por 40.000 millones de dólares. Decide crearse El Fondo de Compensación de Víctimas y se designa a un abogado Kenneth Feinberg para que actúe como árbitro en la fijación de las indemnizaciones y se le fija un plazo de 2 años para que recoja al menos el 80 % de los damnificados quienes aceptando la compensación se comprometen a no iniciar juicio.

Se desatan toda suerte de ambiciones abogadiles y grupos que quieren ir a juicio para obtener un enorme botín: un abogado lidera a los grupos fuertes que quieren una compensación mayor porque son más importantes y un grupo liderado por un abogado más idealista que espera algo más justo. Hay que agregar las víctimas de los otros atentados al Pentágono y al avión que no alcanzo su objetivo.

El resultado es que finalmente el 96 % de los damnificados acepto el resarcimiento del fondo. Se descomprime así la maquina explosiva que amenazaba la economía norteamericana, pero se sigue discutiendo si la solución encontrada fue aceptable moralmente. Pero el fondo debió continuar funcionando porque los damnificados posteriores aparecieron con toda suerte de enfermedades debido a la exposición o inhalación de gases durante el incendio y derrumbe²¹. El fondo original funcionó desde diciembre de 2001 hasta junio de 2004. El Fondo de Compensación para Víctimas inicial recibió 7.408 solicitudes para reclamos de muerte y daños personales. El fondo concedió indemnizaciones en 5.560 de esos casos.

¹⁹ Netflix *¿Cuánto vale una vida humana?*

²⁰ La American Bar Association habla de 1.330.000 abogados, pero es solo una de las Asociaciones de abogados. Según una comparación hecha en Japón hace unos años en EEUU había 11 abogados por cada ingeniero y en Japón 10 ingenieros por cada abogado.

²¹ Enero de 2009 — La oficina del médico forense dictamina que Leon Heyward, quien murió el año anterior de linfoma y enfermedad pulmonar, es una víctima de homicidio porque quedó atrapado en la nube de polvo tóxico justo después del colapso de las torres.

El 2 de enero de 2011, el presidente Barack Obama firma la Ley James Zadroga de Salud y Compensación del 11S de 2010, reabriendo y ampliando el alcance del Fondo de Compensación para Víctimas.

Al principio se intentó encontrar una fórmula que tuviese en cuenta la edad, la profesión, los ingresos anuales y otros indicadores económicos, pero finalmente se fue imponiendo la idea de una suma única. Curiosamente las victimas pedían respeto, pero necesariamente debía traducirse en una suma económica.²²

Que se logren estos acuerdos es vital para la economía, pero ¿son justos? ¿Tienen que ver con una función moral del problema?

Se trata de un problema de números: seis mil damnificados no pueden ser comparados con un accidente por grande que sea. Se dice que en moral el número no cuenta, aprendí que en política si y comienzo a pensar que también en materia moral cuanta. En el fondo, que las victimas reciban rápidamente ayudas para atender sus curas inmediatas o para resarcir un daño sufrido en breve tiempo o en tiempo dilatados es también una consideración ética.

4. Documentos poniendo límites

La difusión de los sistemas inteligentes ha sido percibida y ya países, asociaciones, entidades universales y regionales han comenzado a redactar y difundir documentos de alarma y orientación sobre la necesidad de límites éticos.

4.1 La Unión Europea

Fue la primera en plantear el problema de la inteligencia artificial no sólo desde el punto de vista ético (punto de vista que ya ha seguido el Reino Unido en 2016), sino también desde el punto de vista reglamentario, especialmente desde los mecanismos de imputación de la responsabilidad civil.

En su resolución de 16 de febrero de 2017, que contiene recomendaciones a la Comisión Europea sobre las normas de derecho civil relativas a la robótica, en las que también se tienen en cuenta los sistemas de inteligencia artificial, el Parlamento Europeo ya había indicado como temas de atención los relativos a las repercusiones éticas y sociales, subrayando que el desarrollo de la robótica y la inteligencia artificial debe tener por objeto integrar las capacidades humanas y no sustituirlas.

Esta resolución, si bien tiene por objeto proporcionar directrices en el ámbito civil, ofrece muchas ideas para comprender el alcance de los problemas jurídicos y sociales que la inteligencia artificial trae consigo.

²²Siendo director del *Istituto per la Documentazione Giuridica* del Consejo Nacional de investigaciones italiano un grupo de investigadores creo en los años 80 del siglo pasado un Robot, que en realidad era un programa de software, para el cálculo de los daños provocados por accidentes viales. Se lo llamo en italiano *Automa infortunistico*. Al principio fu aceptado con mucha desconfianza, pero ya en los años 90 comenzó a ser consultado por abogados, fiscales y jueces y termino siendo para fines de siglo y comienzo del actual un elemento determinante pues ambas partes lo consultaban y determinaban una cifra razonable según la cual con algo más o algo menos se lograba un acuerdo.

Desde el punto de vista de la responsabilidad de los agentes inteligentes, el desafío surge cuando consideramos que el criterio tradicional de imputabilidad está vinculado a una conducta del agente sujeto ¿pone una inteligencia artificial su conducta autónoma en el mundo físico? Sin embargo, los sistemas de inteligencia artificial parecen recordar los principios de la responsabilidad del productor, que deben adaptarse al hecho de que esos sistemas asumen una capacidad de decisión autónoma, y podrían también implicar a sujetos distintos del productor únicamente, como los programadores o los que elaboran los algoritmos de decisión.

Otro criterio de imputación de responsabilidad podría encontrarse también en la culpa en la vigilancia de la persona que utiliza el sistema, cuando puede tener conocimiento de la adopción de decisiones erróneas por parte del sistema.

El documento del Parlamento Europeo enumera una serie de puntos de atención y también sugerencias para una correcta solución normativa de los mismos.

Desde un punto de vista jurídico, están apareciendo trabajos sobre algunos casos concretos: La IA y la competencia, la IA y los derechos de autor directos, la IA y las armas letales, la IA y los coches que se conducen solos, la IA y el mercado financiero, la IA y el reconocimiento facial, la IA y el mercado laboral, la IA y la asistencia sanitaria.

4.2 El Vaticano

El pasado 28 de febrero se firmó en el Vaticano un documento que enfatiza los aspectos éticos de la inteligencia artificial y que coloca otra pieza en la base de la regulación. Firmaron representantes de las grandes empresas de distribución de datos Amazon, Google, Facebook, Microsoft, Apple, representantes vaticanos, del gobierno italiano y de la Unión Europea. Concretamente firmaron Paola Pisano, Ministra de Innovación Tecnológica de Italia, Monseñor Vincenzo Paglia, Presidente de la Academia Pontificia para la Vida, el Director General de Fao Dongyu Qu y los presidentes de Microsoft e IBM. David Sassoli, Presidente del Parlamento Europeo, también estuvo presente en el evento²³

4.3. El Parlamento Europeo (PE),

Tras una propuesta de borrador (junio 2016) y un informe razonado (febrero 2017), el PE ha aprobado un informe sobre Robótica en el que se establece un Código Ético de Conducta²⁴.

La propuesta de resolución del PE establece que es necesario establecer “un marco de guía ético para el diseño, producción y uso de los robots” que sirva de complemento a las diferentes recomendaciones puramente legales que se realizan. Es decir, profundizar en una nueva disciplina que aborde los problemas éticos de la Inteligencia Artificial, la “roboética”.

²³[https://www.wired.it/attualita/tech/2020/02/28/intelligenza-artificiale-etica-vaticano-microsoft-ibm/?refresh_ce=](https://www.wired.it/attualita/tech/2020/02/28/intelligenza-artificiale-etica-vaticano-microsoft-ibm/?refresh_ce=24)

²⁴ P8_TA (2017)0051 Normas de Derecho civil sobre robótica Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL)

La idea de fondo es que los estándares éticos deberían ir dirigidos a la humanidad – esto es, los diseñadores, productores y usuarios de los robots– y no tanto a los robots en sí mismos. Como indica la profesora Nathalie Nevejans, responsable del informe encargado por el propio PE, no hay que confundir la ética en la robótica con la ética en las máquinas, es decir, una ética que obligue a los propios robots a adherirse a reglas éticas. Existen varios principios fundamentales que han sido recogidos por la resolución que incluyen la protección de la dignidad humana, la privacidad, la libertad, la igualdad de acceso o los efectos sociales, entre otros.

Proteger la privacidad y el uso de datos: especialmente cuando avancen los coches autónomos, los drones, los asistentes personales o los robots de seguridad.

Protección de la humanidad ante el riesgo de manipulación por parte de los robots: Especialmente en ciertos colectivos –ancianos, niños, dependientes- que puedan generar una empatía artificial.

Evitar la disolución de los lazos sociales haciendo que los robots monopolicen, en un cierto sentido, las relaciones de determinados grupos. Igualdad de acceso al progreso en robótica: Al igual que la brecha digital, la brecha robótica puede ser esencial. Restricción del acceso a tecnologías de mejora regulando la idea del transhumanismo y la búsqueda de mejoras físicas y/o mentales

4.4 La visión de los expertos: los 23 principios de la IA de Asilomar

En febrero de 1975 un grupo de genetistas se reunió en un pequeño pueblo de California, Asilomar, para decidir si su trabajo podría destruir el mundo. Estábamos al inicio de la ingeniería genética y la manipulación del ADN, y de esa reunión surgieron una serie de principios y un estricto marco ético para la biotecnología.

Cuatro décadas después –organizado por el Future of Life Institute– otro grupo de científicos se reunió en el mismo lugar y con el mismo problema. Pero esta vez se trataba de analizar las posibles consecuencias de la Inteligencia Artificial. La idea de fondo fue clara y compartida: un profundo cambio está llegando y afectará a toda la sociedad y las personas que tengan algún tipo de responsabilidad en esta transición tienen tanto una gran responsabilidad como la oportunidad de darle la mejor forma posible, se plantean una serie de principios agrupados en tres consideraciones generales: a) Principios relativos a la investigación; b) Ética y valores; y c) Temas a largo plazo.

1) Meta de la investigación: el objetivo de la investigación de la IA no debería ser crear inteligencia sin dirigir, sino inteligencia beneficiosa. 2) Financiación de la investigación: la inversión en IA debería ir acompañada de fondos para investigar en asegurar su uso beneficioso, incluyendo cuestiones espinosas sobre ciencias de la computación, economía, legislación, ética y estudios sociales. 3) Enlace entre ciencia y política: debería haber un intercambio constructivo y sano entre los investigadores de IA y los legisladores. 4) Cultura de la investigación: una cultura de cooperación, confianza y transparencia debería ser fomentada entre los investigadores y desarrolladores de IA. 5) Evitar las carreras: los equipos que estén desarrollando sistemas de IA deberían cooperar activamente para evitar chapuzas en los estándares de seguridad. 6) Seguridad: los sistemas de IA deberían ser seguros a lo

largo de su vida operativa, y verificables donde sea aplicable y posible. 7) Transparencia en los fallos: si un sistema de IA causa daño debería ser posible determinar por qué. 8) Transparencia judicial: cualquier intervención de un sistema autónomo en una decisión debería ir acompañada de una explicación satisfactoria y auditabile por parte de una autoridad humana competente. 9) Responsabilidad: los diseñadores y desarrolladores de sistemas avanzados de IA son depositarios de las implicaciones morales de su uso, mal uso y acciones, con la responsabilidad y oportunidad de dar forma a dichas implicaciones. 10) Alineación de valores: los sistemas de IA altamente autónomos deberían ser diseñados para que sus metas y comportamientos puedan alinearse con los valores humanos a lo largo de sus operaciones. 11) Valores humanos: los sistemas de IA deberían ser diseñados y operados para que sean compatibles con los ideales de dignidad humana, derechos, libertades y diversidad cultural. 12) Privacidad personal: la gente debería tener el derecho de acceder, gestionar y controlar los datos que generan, dando a los sistemas de IA el poder de analizar y utilizar esa información. 13) Libertad y privacidad: la aplicación de la IA a los datos personales no puede restringir de forma poco razonable la libertad, real o sentida, de las personas. 14) Beneficio compartido: las tecnologías de IA deberían beneficiar y fortalecer a tanta gente como sea posible. 15) Prosperidad compartida: la prosperidad económica creada por la IA debería ser compartida ampliamente, para el beneficio de toda la Humanidad. 16) Control humano: los seres humanos deberían escoger cómo y si delegan decisiones a los sistemas de IA para completar objetivos escogidos previamente. 17) Sin subversión: el poder conferido por el control de sistemas de IA altamente avanzados debería respetar y mejorar, más que subvertir, los procesos sociales y cívicos de los que depende la salud de la sociedad. 18) Carrera armamentística: debería ser evitada cualquier carrera armamentística de armas autónomas letales. 19) Capacidad de precaución: al no haber consenso, deberíamos evitar las asunciones sobre los límites superiores de las futuras capacidades de la IA. 20) Importancia: la IA avanzada podría representar un profundo cambio en la historia de la vida en la Tierra, y debería ser planificada y gestionada con el cuidado y los recursos adecuados. 21) Riesgos: los riesgos asociados a los sistemas de IA, especialmente los catastróficos o existenciales, deben estar sujetos a planificación y esfuerzos de mitigación equiparables a su impacto esperado. 22) Automejora recursiva: los sistemas de IA diseñados para auto mejorarse recursivamente o auto replicarse de una forma que pudiera llevar al rápido incremento en su calidad o cantidad deben estar sujetos a unas estrictas medidas de control y seguridad.

4 .5. La OCDE

La OCDE y los países socios han adoptado formalmente hoy el primer conjunto de directrices de políticas intergubernamentales sobre Inteligencia Artificial (IA), y convenido en someterse a unas normas internacionales que velen por que el diseño de los sistemas de IA los haga robustos, seguros, imparciales y fiables.

Los 36 países miembros de la OCDE, junto con Argentina, Brasil, Colombia, Costa Rica, Perú y Rumanía han suscrito hoy en París los Principios de la OCDE sobre la Inteligencia Artificial en el marco de la Organización, con el lema “*La transición digital al servicio del desarrollo sostenible*”.

Los Principios, elaborados a partir de las orientaciones proporcionadas por un grupo de expertos integrado por más de 50 miembros procedentes de gobiernos, instituciones académicas, el mundo empresarial, la sociedad civil, organismos internacionales, la

comunidad tecnológica y sindicatos, comprenden cinco principios basados en valores para el despliegue responsable de una IA fiable y cinco recomendaciones en materia de políticas públicas y cooperación internacional. Su objetivo es guiar a los gobiernos, organizaciones e individuos para que, en el diseño y la gestión de los sistemas de IA, prioricen los intereses de las personas, así como garantizar que quienes diseñen y gestionen sistemas de IA respondan de su correcto funcionamiento.

«La inteligencia artificial está revolucionando nuestra forma de vivir y trabajar, y ofrece unas ventajas extraordinarias a nuestras sociedades y economías. Ahora bien, también plantea nuevos desafíos y siembra incertidumbre y preocupaciones de carácter ético. Compete, por tanto, a los gobiernos asegurarse de que el diseño de los sistemas de IA respete nuestros valores y leyes, de forma que las personas puedan confiar en que su seguridad y privacidad serán objeto de una consideración prioritaria», manifestó el Secretario General de la OCDE, Ángel Gurría. «Estos Principios constituirán un referente global para una IA confiable, de modo que las oportunidades que brinda redunden en los mejores resultados para todos», terminó de presentarlos²⁵

4.6 La Unesco

El documento que más impresiona es el de Unesco que luego de proponer un borrador al cual se podía intervenir hasta el 31 de julio de 2020 se prevé una revisión por zonas geográficas que en dos años determinará primer instrumento normativo mundial sobre la ética en la inteligencia artificial en forma de recomendación, de conformidad con la decisión adoptada por la Conferencia General de la UNESCO en su 40º período de sesiones, en noviembre de 2019.

Expertos de 155 países, sociedad civil (mediante una encuesta mundial en línea), organismos de las Naciones Unidas, grandes empresas mundiales, como Google, Facebook y Microsoft, así como el sector académico, desde la Universidad de Stanford hasta la Academia de Ciencias de China, pudieron compartir sus impresiones y enriquecer las conclusiones del proyecto.

Este proyecto de recomendación se ha enviado recientemente a los 193 Estados miembros de la UNESCO y será objeto de una serie de negociaciones con miras a su adopción definitiva en la Conferencia General de la UNESCO en noviembre de 2021.²⁶

Será muy importante tener un sistema universal de criterios morales para juzgar las nuevas tecnologías aplicando el criterio que puede haber una ética universal basada en criterio morales universales compartidos.

La estructura del documento es de gran interés porque establece por un lado los valores que deben ser protegidos, los principios que deben ser seguidos y las políticas

²⁵<https://www.oecd.org/centrodemexico/medios/cuarentaydospaisesadoptanlosprincipiosdelaocdesobreinteligenciaartificial.htm#:~:text=Cuarenta%20y%20dos%20pa%C3%ADses%20adoptan%20los%20Principios%20de%20la%20OCDE,Par%C3%ADas%2C%20mayo%202022%20de%202019.&text=%C2%ABLa%20inteligencia%20artificial%20est%C3%A1%20revolucionando,a%20nuestras%20sociedades%20y%20econom%C3%ADas>

²⁶ <https://es.unesco.org/news/unesco-da-gran-paso-primer-instrumento-normativo-etica-ia>

concretas que se pueden adoptar para lograr los fines perseguidos. Esta estructura se va a ver reflejada en otros instrumentos que se han dado posteriormente, como el de Colombia.

4.7 Libro blanco de Inteligencia artificial de la UE

Lleva como subtítulo “sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza” es de febrero de 2020.

Dice “la Comisión respalda un enfoque basado en la regulación y en la inversión, que tiene el doble objetivo de promover la adopción de la inteligencia artificial y de abordar los riesgos vinculados a determinados usos de esta nueva tecnología. La finalidad del presente Libro Blanco es formular alternativas políticas para alcanzar estos objetivos; no aborda ni el desarrollo ni el uso de la inteligencia artificial para fines militares. La Comisión invita a los Estados miembros, a otras instituciones europeas y a todas las partes interesadas, como la industria, los interlocutores sociales, las organizaciones de la sociedad civil, los investigadores, el público general y demás personas con interés en la materia, a que presenten sus opiniones con respecto de las opciones que se muestran a continuación y a que contribuyan a la futura toma de decisiones de la Comisión en este ámbito.

Establece siete requisitos esenciales contemplados en las directrices del grupo de expertos de alto nivel, a saber: • acción y supervisión humanas; • solidez técnica y seguridad; • gestión de la privacidad y de los datos; • transparencia; • diversidad, no discriminación y equidad; • bienestar social y medioambiental; • rendición de cuentas.

Se ocupa de los riesgos que produce la IA y distingue entre riesgos bajos y altos riesgos, cosa muy elemental pero que vale la pena reivindicar: no es lo mismo usar un Sistema automático como el parquímetro para relevar una prolongación indebida del tiempo de un estacionamiento y la sentencia en un juicio penal con muchos imputados, varios delitos y particulares circunstancias.

Se hace cargo de un tema común entre informáticos y organizadores y casi desconocido para el gran público: los prejuicios (sesgos) que acompañan la toma de decisiones y como ellos se transmiten a los sistemas inteligentes que creamos²⁷

5. Ética y estética

La ética tiene como objeto de estudio la moral y la acción humana, mientras que la estética es la rama de la filosofía que tiene por objeto el estudio de la esencia y la percepción de la belleza como aproximación al arte. Lo ético y estético están de muchas formas relacionados en el quehacer diario de las personas, pero el problema surge cuando en su

²⁷Los prejuicios y la discriminación son riesgos inherentes a toda actividad social o económica. La toma de decisiones de las personas no es ajena al error ni a la subjetividad. No obstante, en el caso de la IA, esta misma subjetividad puede tener efectos mucho más amplios, y afectar y discriminar a numerosas personas sin que existan mecanismos como los de control social que rigen el comportamiento humano³⁵. Puede suceder también que el sistema de IA «aprenda» mientras está funcionando. En tales casos, cuando los resultados no puedan preverse ni anticiparse en la fase de diseño, los riesgos no se deberán a fallos en el diseño original del sistema, sino más bien a las repercusiones prácticas de las correlaciones o de los modelos que reconozca el sistema en un gran conjunto de datos.

última definición de lo que es ético y estético, surge la hipocresía, la doble vara de medir y, sobre todo, manipulación mediática. Si este cóctel se adereza con la ideológica, surge un juicio popular que distingue a las personas y las clasifica, de tal modo que el buen actuar de unos se ve pésimo en otros, y todo ello desde una pretendida superioridad moral ideológica.

Así por ejemplo lo que para unos es plagiar para otros es enriquecer una obra previa, lo que para unos es estar próximos a una determinada ideología política para otros es un ejercicio de pragmatismo social y proximidad a la realidad social, lo que para unos es un error garrafal para otros es una pequeña falta excusable, lo que para unos es una broma para otros es un chiste machista, etc. Resulta asfixiante esta asimetría valorativa conductual contra la cual es muy difícil defenderse y superar. Hace poco escribí un artículo sobre la famosa frase de la mujer del César, la cual debe parecer además de ser, en el que adelantaba que la cuestión estriba en que «parecerlo» es tan subjetivo y maleable que hace que lo leve pasa por ser lo más grave, y viceversa, de tal suerte que el parecer popular sea una suerte de canon estético que se decide en lugares del poder real, que hoy por hoy es el que incide en la conformación de la opinión pública.

En la antigua Roma ya el rumor y la mentira conducían a la turba contra los cristianos acusándolos del incendio de la ciudad, y en la Rusia zarista el asesinato de Alejandro II dio lugar a una dura represión contra los judíos al atribuirles la responsabilidad del magnicidio. Resulta increíble lo fácil que es deshonrar a alguien hoy en día, y lo difícil que es defenderse de los infundios y mendacidades. Cuando para zaherir a alguien es necesario acudir a la estética, pocos argumentos se tienen. Miente, que algo queda, la vieja máxima de Goebbels es el mejor retrato que existe del infundio, que a menudo se conforma con sembrar la duda sobre el difamado e ir socavando poco a poco su prestigio personal. En España utilizamos la expresión «cuando el río suena es porque agua lleva», muy usada por los difamadores para otorgar una falsa veracidad al infundio apelando a su mera existencia, aunque sean ellos mismos los que lo han creado. Todos cometemos errores, más con desigual suerte en sus consecuencias, las cuales están determinadas por lo que parece frente a lo que es.

La estética es la manifestación externa de un sentimiento, la ética el fundamento razonado y racional de los sentimientos. La ética se manifiesta a través de la estética. La estética puede ser cambiante, sus valores y apreciaciones son subjetivas y varían según la percepción del momento y contexto histórico del individuo, la ética permanece como un valor de la Humanidad.

Wittgenstein aduce que “ética y estética son lo mismo”, o literalmente “son Uno” (*sind Eins*). Lo dice en el contexto en que afirma que “la ética es trascendental”. Y esa trascendentalidad de la ética (y de la estética, por tanto), que también se enuncia de la lógica, remite a un “sujeto” que, sin embargo, no está más allá de los límites del mundo, sino que se determina como “un límite del mundo”. Lo trascendental es, por tanto, el límite (y el sujeto como “sujeto” de ese límite, o “sujetado” a dicho límite).

Lo ético, como lo estético, al decir de Wittgenstein, es “inexpresable”. No pueden formarse “proposiciones” al respecto. La diferencia entre ética y estética, en la medida en que son “lo mismo”, es muy sutil. Wittgenstein cita a Schiller: “Seria es la vida, alegre el arte” (*Wallensteins Lager*, Prólogo). La obra de arte sería el objeto (lo que por tal se entiende en el *Tractatus*) “visto *sub specie aeternitatis*”. A la inexpresividad y silencio de lo ético se corresponde “lo mismo” en referencia a lo estético.

“Voy a describir la experiencia de asombro ante la existencia del mundo diciendo: es la experiencia de ver el mundo como un milagro” (añade Wittgenstein en su impresionante *Conferencia sobre ética*). Y añade: “Me siento inclinado a decir que la expresión lingüística correcta del milagro de la existencia del mundo -a pesar de no ser una proposición en el lenguaje- es la existencia del lenguaje mismo”.

A diferencia del modo estético “el modo científico de ver un hecho no es el de verlo como un milagro”.

Esta gran reflexión de Wittgenstein sobre el “silencio” de lo ético (y por tanto de lo estético) debe ser retenida. Algo hay “inexpresable” que impide ajustar la “experiencia” ética (y estética) al criterio según el cual este autor determina lo que denomina “proposición” (que es la expresión manifiesta del pensamiento). Pero quizás convenga fecundar esta reflexión tan exigente del *Tractatus* con la idea plural y compleja de los múltiples “juegos lingüísticos” (acordes con sus “mundos de vida”) de que habla este gran filósofo en su obra última.

Puede, pues, entonces pensarse si no es lícito determinar alguna suerte de expresión lingüística que “proponga”, con plenitud de sentido, algo relativo a esas experiencias “trascendentales” que hace el “sujeto” en el ámbito de lo ético (y de lo estético, destacar “una” proposición (ético-ontológica), la proposición que guía y orienta al sujeto (fronterizo) en relación a su obligación, imperativa, por realizar su propia condición (de lo que resultaría, por lo demás, la “vida buena”).

La expresión simbólica tiene la peculiaridad de permitir una mostración, en el objeto, o en ciertos episodios del mundo, que, mediada por el decir o hacer simbólico, permita también que “lo ético” (y en consecuencia la proposición referida, y todo el orden de experiencia de la libertad que funda), resuene. Kant concibió esa resonancia cuando dijo que la belleza era un “símbolo moral”, y que el modo de exposición simbólico, en el que indirecta y analógicamente se exponía “lo trascendental” (y por ende también lo ético, o el uso ético de las Ideas de la razón), se distinguía de la exposición esquemática que permite la conjunción de intuición y concepto para la producción de conocimientos.

La estética, la imagen, provoca asombro. La recreación en la elaboración de la imagen, la belleza, afirma Juan Pablo II en su carta a los artistas, debe provocar asombro ante la sacralidad de la vida y del ser humano, ante las maravillas del universo. Una estética bella se sustenta en el respeto y admiración por la vida y la naturaleza; y, precisamente por eso, la comunicación de estos procesos se hace necesaria y edificante.

Entender la imagen como texto implica la necesidad de interpretar la comunicación visual. La lectura visual en procura de un contenido que va más allá de la estética tradicional es una extensión de los procesos de comunicación. Los cánones griegos de la estética pretendían normar los aspectos externos de sus obras, caracterizadas por su profundidad y contenido. El elogio de la estética a lo largo de la historia la ha apartado de esa vinculación con el contenido. Asumir y retomar nuevamente esa vinculación supone

entender no sólo la estética como texto y contenido; sino aceptar que la ética, reflejada en el contenido, forma parte de los criterios estéticos. La imagen, al proporcionar un texto, presenta las dimensiones ética y estética de la comunicación. Estas dimensiones no se asemejan a la antigua diferenciación forma y fondo, más bien generan una nueva concepción: estética y ética en la forma y estética y ética en el fondo. Resulta importante retomar las palabras que Juan Pablo II dirigió a los artistas a este respecto cuando afirma que la belleza salvará al mundo, es una invitación a gustar la vida y a soñar el futuro. La belleza se convierte en un valor de vida basado en lo bueno, en lo ético. La belleza ayudará a los hombres en su afán de construir un Reino más justo. Esta concepción está ya expresada en el Génesis, cuando se manifiesta la belleza como la expresión visible del bien.

La nueva creatividad parte de la función social que los medios de comunicación, la comunicación en general, asumen. El lenguaje audiovisual moderno debe fortalecer los procesos de educación liberadora. Es necesario para el crecimiento personal y social del ser humano desarrollar una conciencia crítica. La comunicación ética debe desvirtuar el rol pasivo al que se ha enclastrado al receptor para dotarlo de las herramientas necesarias para entender los lenguajes audiovisuales. Si se pretenden renovar los procesos creativos impulsando éstos desde la ética, es necesario rescatar al receptor como lector crítico de la estética.

Reinventar los procesos interactivos en los cuales el receptor se siente involucrado contribuye a fortalecer la conciencia crítica. Para lograr esta relación dinámica se hace imperioso recuperar el diálogo y el debate. La participación del receptor como integrante activo del proceso obliga no sólo a tenerlo en cuenta, sino a formarlo dentro de un sentido crítico capaz de asumir sus propias decisiones. Crear este ámbito de diálogo requiere un ejercicio ético de los procesos de comunicación, pero además implica una estética atractiva que acerque y facilite el mensaje. Este reto, construir una comunicación ética y estética que ayude a desarrollar la conciencia crítica del ser humano, se convierte en el nuevo horizonte de la comunicación evangelizadora.

La comunicación audiovisual actual ha conseguido construir un mundo ficcional paralelo al real. Los sentimientos y las sensaciones que trasmite esta vivencia ficcional son tan fuertes o más que los del mundo real. Muchos de los sueños sociales, de las aspiraciones personales que orientan a las personas nacen y se forjan dentro de estas claves. Es necesario reconocer que las ficciones se han convertido en fuente de conocimiento y enseñanza importante. Han superado su limitante de entretenimiento para convertirse en un espejo de la vida. La verosimilitud que mueve sus historias la asemeja a una ventana con vistas a la realidad. Si las historias que cuentan son válidas, permiten al espectador crecer en sus vivencias humanas. Si las historias buscan más otros intereses como la rentabilidad antes que el servir lograrán que el espectador construya un mundo paralelo ficcional que se mueve por unos parámetros no reales.

Este mundo creado por la ficción fruto de la cultura audiovisual más comercializable confunde al público. Hace creer que la superación del ser humano se sustenta en procesos quasi mágicos, en golpes de suerte que lo pueden convertir en "presidente por accidente".

Como conclusión, y retomando la visión propositiva de la ética, se deben establecer dos líneas de trabajo:

Las instituciones formadoras deben redoblar sus esfuerzos por formar agentes sociales críticos. Es necesario crear un sentido de lectura frente a la realidad audiovisual capaz de rescatar los valores de vida frente a los valores que prometen el éxito fácil. Educar a los receptores para que sean conscientes de su papel activo a la hora de construir sociedad redundará en la mejora de los procesos de desarrollo. No sólo se beneficiará la comunicación al dirigirse a un receptor más exigente. Si los comunicadores forman unas personas con sentido crítico, la democracia, la cultura y la vida en general se consolidarán bajo los soportes de la dignidad humana.

Los productores deben recuperar la unión ético-estética de sus propuestas. Plantear trabajos audiovisuales valorados en virtud de su rentabilidad social, del conocimiento y de las pautas de vida que difunden. De este modo, el receptor podrá descubrir modelos de superación sustentados en la vivencia en valores, modelos que le permitan desarrollar sus dimensiones humanas sin mellar su dignidad.

Me contaba Luis Jiménez de Azua ²⁸ que cuando le pidieron la renuncia al profesor de estética los franquistas, renunció también el profesor de ética pues dijo que no podía haber una ética sin una estética.

6. Conclusiones

En los momentos de riesgo de la vida humana, siempre aparece la ética. Tendría que ser más duradera, pero ahí está seguro. En los últimos 18 años se han producido más novedades en el mundo debido a las nuevas tecnologías que podemos decir que hay un mundo nuevo, y ese mundo nuevo requiere nuevamente una visión ética. Por eso el tema que parecía escondido ahora toma una dimensión universal y vivimos preguntándonos hasta donde podemos llegar con nuestros adelantos técnicos.

De algo estoy convencido: los humanos saben qué está bien y qué está mal. No de una manera categórica e instantánea, pero no necesitan consejeros externos²⁹.

¿Qué es el mal? ¿Y su antítesis, el bien —o será al revés—? ¿De dónde procede el mal? ¿Cuándo se originó? ¿Qué sentido tiene? ¿Por qué tanto dolor y sufrimiento aparentemente inútiles? ¿Es que carece de límites espaciales y temporales el mal? ¿Por qué parece, en ocasiones, cebarse con los más justos y los inocentes —véase si no el caso paradigmático de Job—? ¿Qué implicaciones o consecuencias se derivan del mal? ¿Cómo erradicarlo de una vez por todas de la vida humana? ¿Es, acaso, esto último posible? Estas y otras muchas cuestiones de gran calado contenidas en el problema han agitado por centurias las mentes y vidas de insignes filósofos —y, por qué no decirlo, de todo ‘homo viator’ que ha transitado alguna vez por la faz de la Tierra—, que han intentado encontrar respuestas —en ocasiones de lo más variopintas, todo sea dicho— a las causas y consecuencias posibles planteadas por la diatriba de la presencia del mal en la naturaleza y en el núcleo de la vida misma.

²⁸ Profesor de derecho penal en la Facultad de Derecho de la Universidad de Buenos Aires que había sido vicepresidente del Congreso de los Diputados y representante de España ante la Sociedad de Naciones. Durante la dictadura franquista se exilió en Argentina

²⁹ Inclusive hay un diccionario soviético sobre el tema que dice “Se entiende por bien lo que la sociedad (clase dada) considera moral y digno de imitar. El mal tiene un significado contrario: lo inmoral, digno de condena. La interpretación metafísica de estas categorías se caracteriza por las búsquedas de los fundamentos eternos e inmutables de bien y mal”.

Estas preguntas atraviesan toda la historia de la humanidad y se la puede rastrear en cualquier cultura, incluyendo la religión. En el nuevo testamento el mal tiene un nombre: el diablo y cuentan las escrituras que lo tentó a Jesús en el Desierto. Sin embargo, esa identificación con el Maligno no es universalmente aceptada. Por lo poco que sé no aparece en el budismo y dentro de la misma iglesia católica es ejemplar la posición de algunos pensadores, por ejemplo, Leibniz: tanto en la Teodicea³⁰ con en páginas posteriores como *La presencia de Dios en el mundo*. Leibniz no creía en el Diablo, por lo tanto, define la presencia de Dios en el mundo como bondad. Y ¿la maldad? Ah, eso es la ausencia de Dios en el mundo. Y de un plumazo elimino al Diablo, no solo agrega que esos temimos pueden ser representados numéricamente con 1, bondad y 0 ausencia de Dios en el mundo.

Pero va más lejos, dice que con esos dos términos puede escribir cualquier cosa y para demostrarlo comienza a escribir con 0 y 1 los ocho primeros números de la sucesión fundamental algo así como 00001101010111100, 10000110101011001, 0110000010100, etc. Uds. Dirán, pero eso es el sistema binario, si, lo es. Y Leibniz se da cuenta de la fuerza de expresión de esos términos tanto que le escribe a un Jesuita que está en China³¹

El jesuita queda maravillado pues el sistema de Leibniz puede ser comparado con el IChin compuesto por 64 hexagramas con rayas completas y rayas quebradas. Los dos sistemas son binarios y de algún modo reflejan la manera de proceder de nuestra mente tanto en occidente como en oriente.

Ambos le escribieron al Papa para pedirle que autorice la enseñanza de la religión católica en China a través del I Chin. Probablemente Clemente XI quien sentencio que la religión católica debía enunciarse en latín.

Siendo tanto la moral (mores) como la ética (ethos) dos palabras que significan costumbre necesariamente tienen que adaptarse a los tiempos. Pero no es que haya una gran diferencia en el transcurso de la historia. Cierta la moral de un *pithecantropus erectico* es distinta de la del cosmonauta Yuri Gagarin, porque el mundo a su alrededor es distinto, pero si se va a escavar que es lo que está bien y que lo que está mal se pueden enumerar las acciones que luego las religiones habrían de establecer como virtuosas o pecaminosas.

La ética tiene que ver con valores. Nuestra vida puede ser concebida como una sucesión de elecciones (conscientes o inconscientes) que la van conformando. Una cualidad del valor es que no es un sustantivo sino un adjetivo no es independiente se habla de la belleza de una mujer o de un cuadro³², otra cualidad del valor es su polaridad: siempre que se habla de un valor se puede predicar un disvalor correspondiente: bello y feo, bueno, malo, etc. Una tercera característica es la jerarquía, pero no del valor sino como los valores se van colocando en una organización jerárquica de valores donde a veces priman algunos sobre otros, pero esto puede cambiar con el tiempo, por ejemplo, en caso de Pandemia. Generalmente se habla de una tabla de valores con algunos muy excelsos arriba otros menos fuertes debajo y así sucesivamente. Es más fácil satisfacer el valor de urbanidad que el de heroísmo. Pero el presentar la jerarquía como una tabla le da rigidez y estaticidad a

³⁰ Los Ensayos de Teodicea (sobre la bondad de Dios, la libertad del hombre y el origen del mal) fue la única obra filosófica que Leibniz publicó en vida, curioso dejó escritas más de veinte mil páginas vio la luz en la ciudad de Ámsterdam, en francés en el año 1710 sin mención a su autor, que sólo apareció en ediciones posteriores.

³¹ La carta llegó en 27 días.

³² Husserl llama a esta falta de sustantividad "no independencia"

la metáfora. Carlos Cossio ³³ pensó en reemplazar la idea de tabla por una esfera que consiente tener no uno si no varios valores en determinados niveles que con el moverse de la esfera se irán desplazando según tiempos y lugares, pasando algunos valores a ocupar lugares superiores en detrimento de otros que se desplazan a lugares inferiores.

“Podría decirse que una doctrina ética es un conjunto sistematizado de proposiciones relativas a la vida moral, que recomiendan aprueba una serie de actividades humanas consideradas valiosas”³⁴

La necesidad de controlar las decisiones producidas por la inteligencia artificial en un momento en que las soluciones digitales a los problemas sociales se invocan en toda sociedad tecnológicamente avanzada.

En la actualidad, varios países permiten a sus cuerpos de seguridad utilizar sistemas de reconocimiento facial en sus investigaciones. Los algoritmos cruzan big data e imágenes de las redes sociales de más de 3.000 millones de cuentas de todo el mundo para identificar a los sospechosos. Recientemente, escribe la Policía Metropolitana de Londres utilizó esta tecnología para identificar a 104 sospechosos, de los cuales 102 resultaron ser errores de cálculo sistemáticos.

En los países occidentales, además, estas máquinas de aprendizaje se entrena con datos que contienen un sesgo de muestreo, es decir, que no representan con exactitud la composición étnica de la sociedad. En otras palabras, están bien entrenados para reconocer las caras de los hombres blancos, pero es mucho más probable que se equivoquen en sus predicciones con las minorías étnicas.

Por eso, el uso de sistemas de inteligencia artificial para gestionar la seguridad pública hace que un experto en derechos humanos levante las antenas: una persona puede enfrentarse a un juicio por un error de apreciación de un algoritmo.

Las maquinas son increíblemente exactas para realizar determinadas tareas que tienen que ver con el cálculo racional muchas veces mejor que un humano³⁵ pero que carece de criterios de razonabilidad y adaptación al medio y a situaciones totalmente cambiantes. Por eso es recomendable que siempre hay un humano que controle, dirija, supervise y oriente las funciones de cualquier sistema de I.A. Por más complejo que sea. En Europa, el debate está quizá más avanzado que en otros lugares. El GDPR (Reglamento General de Protección de Datos) que está en vigor en Europa prevé la posibilidad de que alguien que haya sido objeto de una decisión automatizada pueda impugnar dicha decisión. Esto es muy importante porque introduce la demanda de control humano, una demanda que está ganando terreno en muchos países y para muchas aplicaciones: financieras, para las armas autónomas, para las decisiones relativas al empleo, la carrera, la concesión de un préstamo.

³³ Carlos Cossio, tucumano, 1903 – 1987 filósofo del derecho, militante universitario reformista y profesor argentino. Dentro de sus principales obras está la concepción de la Teoría egológica del derecho.

³⁴ Antonio A. Martino *Valores ética y metaética* en A. Martino, E. Russo, L. Warat, Temas para una filosofía jurídica, Cooperadora de Derecho y ciencias sociales, Buenos Aires, 1974, p.119.

³⁵ Un cirujano del hospital italiano que usa Leonardo un robot para operar confesaba que la perfección de corte y la exactitud de Leonardo son incomparables con las suyas propias que dependen de estados de ánimo, situaciones sociales y aun simpatías o antipatías. Pero Leonardo no opera solo es accionado por un médico.

La matemática e informática estadounidense Cathy O'Neil ya nos había advertido de estas cuestiones en su libro³⁶. La desigualdad de género y la discriminación de las minorías étnicas ya existen en nuestra sociedad. Se necesitan sistemas de inteligencia artificial dirigidos por un algoritmo que sea demasiado rígido a la hora de aplicar lo que ha aprendido para no introducir más. Hay que promover un debate ético lo más amplio posible y reconocer su importancia. Existe una página web del gobierno australiano en la que es posible hacer propuestas que serán evaluadas por un comité de ética con vistas a redactar un documento sobre la ética de la inteligencia artificial a finales de este año. Amén de las recomendaciones que está preparando la Unesco, con sentido universal y del cual hemos escrito *supra*.

Se dice que los sistemas de inteligencia artificial que analizan los datos de comportamiento, tanto en línea como fuera de ella, hacen transparente a la ciudadanía, revelando sus hábitos y deseos inconscientes. Sin embargo, un llamamiento debe ir en la dirección opuesta, la de hacer que los sistemas de procesamiento de datos sean transparentes para los humanos: las decisiones de las máquinas no deben ni pueden tomarse en una caja negra de circuitos y bits, inaccesible al control humano. El proceso de toma de decisiones debe ser comprensible y explicable, de lo contrario sería como confiar nuestro destino a un oráculo impenetrable.

La Pandemia está siendo un acelerador constante y también nos ha impulsado a mayor uso de servicios a distancia y de I.A. Por un lado, es un bien que haya sucedido cuando ya estábamos preparados a eso, de lo contrario las Administraciones públicas y privadas habrían quedado paralizadas, imposibilitada la educación, anuladas la interacción activa, pero esa precipitación en el uso de tecnologías que aún no están suficientemente desarrolladas y mucho menos testeadas nos ha hecho embocar tiempos de improvisación con prácticas de error y acierto que han resultado, muchas veces peligrosas y otras catastróficas.

No puede prescindirse del costo humano que significa implantar sistemas de I.A. en la sociedad: en general las grandes transformaciones de la vida humana (la rueda, el vapor, la electricidad, la imprenta, etc.) llevaron siglos y varias generaciones, Internet es de 1992, la aplicación de grandes datos (Big data) de 2017 y muchos otros cambios significativos se dan en decenios o menos. Esto provoca no solo la desaparición de lugares de trabajo para los humanos sino también esfuerzos inconcebibles para estar al par con los tiempos. No hay porqué deshumanizar la I.A.

Los sistemas inteligentes están acelerando nuestros tiempos y cambiando el mundo que nos rodea: personalmente no creo que haya otra moral, pero no se puede desconocer que hay otro mundo y, por lo tanto, la ética debe explicarse y expresarse en este nuevo mundo. Un mundo nuevo acelerado por los cambios tecnológicos como nunca lo fue y a la sociedad ir adaptándose. En el último cuarto del siglo pasado todo se aceleró de manera vertiginosa: la aparición de Internet significó un antes y un después. Se puede pensar que eso sucedió porque tanto la teoría como la práctica estaban preparadas y por supuesto que es así, siempre es así, pero la aceleración que se desató fue incommensurable con cualquier cambio del pasado.

³⁶ Cathy O'Neil *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Penguin Books Ltd, 2017

Pero no todas son malas noticias: las nuevas metodologías permiten controlar mucho más lo que hace una Administración, favorece la participación, extiende la dimensión horizontal de la burocracia que fue vertical desde Luis XIV, crea chatbots de información para el ciudadano (91), alerta sobre la controlabilidad de lo que se decide y se hace. Nacen ciudades inteligentes. ¿El gran público queda indiferente? Siempre el gran público fue indiferente. Inclusive la rebelión de las masas de orteguiana memoria fue simplemente para que los muchos obtuvieran los objetos exteriores de los pocos no para que cogobernaran. Pero los medios están, se trata de ponerlos en movimiento. Además, grandes instituciones supranacionales, estados nacionales, grandes corporaciones, iglesias, hasta la Unesco, se han percatado de los riesgos de una sociedad con sistemas inteligentes incontrolados y han comenzado a formular acuerdos, a redactar principios o pactar soluciones y vigilancias, a redactar un documento que sirva como guía mundial para vigilar el uso moral que se hace de los mismos. Es de estos días la noticia que la comisaria europea de Asuntos Digitales y Antimonopolio, Margrethe Vestager, dijo que los gigantes de la tecnología tendrán que abrir sus archivos de anuncios a los reguladores e investigadores. Es probable que la medida impacte en empresas como Alphabet, propietaria de Google, Amazon, Apple y Facebook. Sabemos que el medio no es el mensaje, pero lo limita, lo condiciona, lo coloca en otro contexto, lo puede manipular, tergiversar. Y entonces no basta con saberlo, no basta con recurrir a las soluciones éticas del pasado, no porque no sirvan sino porque hay que contextualizarlas y ello requiere que se afronte la nueva sociedad que se ha implantado.

Las ciudades digitales han incorporado una cantidad de funciones desde la multiplicación de las cámaras que permiten un mayor control contra la delincuencia hasta chatbot, como el Boti de la ciudad de Buenos Aires que da todo tipo de información sobre las cosas que normalmente se necesitan: estacionamientos, farmacias abiertas, el restaurante más próximo, la comisaría o el hospital. Y sistemas de resolución pacífica de conflictos totalmente digitalizados, como existe en muchas ciudades, incluyendo Buenos Aires³⁷. Además, un *Buenos Aires elige* que permite sugerir temas que si obtienen un cierto número de votos son necesariamente tratados por la Legislatura de la ciudad como un proyecto cualquiera. Esto es la ciudadanía puede presentar proyectos de ley y ser aprobados o al menos tratados por la Legislatura.

La economía tiene que rendirle cuentas a la moral, no es posible que solo las leyes del mercado decidan la falta de trabajo de millones de personas, la hambruna y la angustia de muchos. Con cada situación, con cada mundo que se va creando con los sistemas inteligentes es necesario plantearse hasta donde las acciones que se generan y las consecuencias que derivan pueden traer felicidad o dolor a los seres humanos que serán implicados por las mismas porque, en definitiva, como decía Kant: la moral no es realmente la doctrina de cómo hacernos felices, sino de cómo debemos hacernos dignos de la felicidad.

Se trata siempre de un tema de valores de fines y medios y la frase “el fin justifica los medios” históricamente carece de sentido.³⁸ Sigue que si se quiere un fin fuertemente se termina aceptando los medios, pero el medio puede polucionar, envenenar de tal manera el fin de hacerlo inaceptable. Por eso se debe tener en cuenta desde el punto de vista moral

³⁷ Con un dato interesante: el año pasado las causas de litigio por relaciones familiares disminuyó de un 37% en parte debido a este sistema de resolución automática o digital.

³⁸ Es curioso la frase “el fin justifica los medios” atribuida a Nicolo Macchiavelli en el *Príncipe*, en realidad es un comentario que hizo Napoleón, luego de su lectura.

que cuando se acepta un fin los medios para llegar a él no deben tornarlo indeseable. Torturar a un maestro para aprender todo lo que él sabe es un ejemplo muy claro y burdo, pero la vida real está llena de claros oscuros, es allí donde se debe indagar sobre los medios y no basta con decir “no lo sabía”.³⁹

Cualquier persona razonable puede averiguar los peligros que implica cualquier uso de cualquier cosa. Es tan obvio esto que se empiezan a distinguir entre riesgos pequeños, medianos y grandes como quiso Alan Kay quien el autor de la famosa frase “la mejor manera de predecir el futuro es inventándolo”. La discusión sobre el uso de un medio que tiene poco riesgo pero que podría este aumentar, si los beneficios son enormes, pero no se trata solo de un tema de cantidad.

Grandes instituciones supranacionales, estados nacionales, grandes corporaciones, iglesias, hasta la Unesco, se han percatado de los riesgos de una sociedad con sistemas inteligentes incontrolados y han comenzado a formular acuerdos, a redactar principios o pactar soluciones y vigilancias, a redactar un documento que sirva como guía mundial para vigilar el uso moral que se hace de los mismos.

Sabemos que el medio no es el mensaje,⁴⁰ pero lo limita, lo condiciona, lo coloca en otro contexto, lo puede manipular, tergiversar. Y entonces no basta con saberlo, no basta con recurrir a las soluciones éticas del pasado, no porque no sirvan sino porque hay que contextualizarlas y ello requiere que se afronte la nueva sociedad que se ha implantado.

A Al-Juarismi⁴¹ “le preguntaron por el valor del ser humano y así respondió. Si tiene ética, entonces su valor es 1. Si además es inteligente, agréguele un cero y su valor será 10. Si también es rico, súmele otro cero y será 100. Si por sobre todo eso, es además, una bella persona, agréguele otro cero y su valor será 1000. Pero si pierde el 1, que corresponde a la ética, perderá todo su valor, pues sólo quedarán los ceros”.

Sin valores éticos ni principios sólidos, no quedará nada. Solamente delincuentes corruptos y malas personas.

³⁹ Un trabajo extraordinario sobre el tema lo realizó Giuliano Pontara, *Se il fine giustifichi i mezzi*, Il Mulino, 1974, con un prologo de Norberto Bobbio.

⁴⁰ Marshall McLuhan ‘, Understan- ding Media: The Extensions of Man, 1964 “el medio es el mensaje”

⁴¹ Abu Abdallah Muḥammad ibn Mūsā al-Jwārizmī, conocido generalmente como al-Juarismi, y latinizado antiguamente como Algorismi, fue un matemático, astrónomo y geógrafo persa. Fue astrónomo y jefe de la Biblioteca de la Casa de la Sabiduría de Bagdad, alrededor de 820