

ÍNDICE

Introducción <i>Antonio A. Martino</i>	3
---	---

¿Por qué la ética establece criterios para los sistemas inteligentes? <i>Antonio A. Martino</i>	7
--	---

CAPÍTULO PRIMERO

Ética y sistemas inteligentes <i>Antonio A. Martino</i>	9
--	---

Semblanza de Antonio A. Martino <i>Antonio A. Martino</i>	48
--	----

CAPÍTULO II

La importancia de un marco jurídico adecuado a la ética y los valores en IA <i>Bibiana Beatriz Luz Clara</i>	50
---	----

Principios de inteligencia artificial y ética por diseño <i>Daniel Peña Valenzuela</i>	59
---	----

“Inteligencia artificial”, ética y administración de justicia <i>Carlos Gregorio de Gràcia</i>	65
---	----

El derecho humano a las decisiones jurídicas humanas <i>Emilio Suñé Llinás</i>	83
---	----

La inteligencia artificial para prevenir la corrupción <i>Romina Florencia Cabrera</i>	98
---	----

Diritti umani e intelligenza artificiale: la regolamentazione internazionale dei sistemi di arma autonoma letale (Laws) <i>Natascia Arcifa</i>	106
---	-----

De la revolución de la información: ¿qué ética para la sociedad de la red? <i>Valentina Grazia Sapuppo</i>	119
---	-----

Algunas reflexiones sobre la ética de los desarrolladores y diseñadores de sistemas inteligentes
Elisa Palomino Angeles.....131

¿Qué límites éticos se plantean hoy un hacedor de programas inteligentes?
Horacio Granero.....140

¿Ética de la tecnología o ética de los productos tecnológicos?
La IA enfrenta sus límites y los de los seres humanos
Enzo Maria Le Fevre Cervini.....142

CAPÍTULO III

Los sistemas inteligentes: una oportunidad para el ámbito ético
Sebastián Chavarría L.....144

Principios de inteligencia artificial y ética por diseño
Daniel Peña Valenzuela.....149

Blockchain como elemento de gobernanza, confianza, ética y autorregulación de la IA
Rafael Y. Cuartas Báez.....155

Consideraciones constitucionales acerca de la política y gestión de las nuevas tecnologías en el proceso judicial peruano
Jorge Isaac Torres Manrique.....159

Panorámica del curso “ética y sistemas inteligentes”
Amir García Villalpando.....168

CAPÍTULO IV

Ética en IA: Dimensión epistemológica y ontológica
Daniela López De Luise.....203

Introducción

Este ciclo es una de las primeras actividades del CETI Centro de Estudios sobre Tecnologías Inteligentes de la Academia Nacional de Ciencias de Buenos Aires.

Nació de una pregunta genuina del presidente de la Academia, Ing. Juan Carlos Ferreri ¿es posible que nadie haya pensado en los límites que deben tener la nueva tecnología, en particular la IA desde el punto de vista ético?

Dado que me ocupo de filosofía desde siempre pues he sido miembro del Instituto de Filosofía del Derecho de la UBA, luego profesor de Filosofía del Derecho en la Universidad de Belgrano y en la de Haute Normadie, pareció natural que me ocupara del tema.

Para ser sinceros con mucho escepticismo pues pensé que los mayores de 50 años no sabían nada de IA y los menores nunca habían sido ocupados por la ética. Y me equivocaba. Primero porque rápidamente encontré cuatro documentos internacionales que trataban del tema, segundo porque contrariamente a mis dudas para hallar relatores los encontré mucho más rápidamente de lo que creía, tercero porque ha sido un éxito resonante hemos tenido, llegamos a tener 190 personas conectadas: 100 en el zoom de la Academia y el resto en el Youtube de la Academia.

Pensamos con la directora del CETI, Daniela Lopez de Luise, que había que encontrar un día en la semana para dedicarlo al tema y el miércoles, en medio de la semana nos pareció ideal queríamos comenzar en agosto y el primer miércoles de agosto era el 5 y contando todos los miércoles hasta fin de año teníamos 21. Necesitábamos 21 disertantes que además fueran de diferentes áreas del conocimiento y de diferentes países y aun continentes, si fuera posible.

Y nos lanzamos a la búsqueda, siempre con desconfianza, a punto tal que tomé yo la primera conferencia pensando que si no llegábamos a los 21 podíamos intercalarnos nosotros mismos como relatores. Como dicho antes nos equivocábamos, hubo 36 relatores (o mejor dicho 37 y uno suprimido pues le dio el Covid).

Esto significaba que no solo habíamos cubierto los 21 miércoles, sino que nos sobraba gente. Entonces pensamos en poner dos los miércoles que fuese necesario.

Pensamos en una fórmula que fuese simple y atractiva: el relator tenía unos 30 minutos para exponer y esperábamos otros 20/25 para preguntas del público. Tengo experiencia en materia de conferencias en línea pues he organizado todos los años congresos con ese formato. Por lo tanto, sé que es muy peligroso abrir los micrófonos pues el participante antes de hacer una pregunta prefiere una larga introducción con su pensamiento que, a veces se transforma en una contra relación y la experiencia me indica que si se le pide una pregunta escrita ésta es aceptablemente breve y ceñida al tema en discusión. Además, permite al coordinador eliminar las preguntas repetidas y separar las impertinentes y ordenarlas dejando a los que preguntan dos veces, después de los que lo hacen una vez sola.

La directora del CETI, pensó que limitarnos a la IA era una fórmula perdedora y sugirió “sistemas inteligentes” más acorde con nuestro formato.

Así arrancamos el 5 de agosto 2020 con una conferencia a mi cargo ¿Por qué la ética establece criterios para los sistemas inteligentes? Donde debía exponer unos 30 minutos y luego habría otros 20 /25 de pregunta coordinados por Daniela Lopez de Luise. Por las dudas pensamos que la reunión podía extenderse durante 90 minutos, comenzando siempre a las 17,30 hora argentina, diferentes horarios en América y casi media noche en Europa.

El relator podía compartir pantalla para que su conferencia fuese atractiva con exposición e ilustraciones. Me excedí unos minutos de lo previsto, pero la sorpresa apareció con las preguntas. No solo eran muchas, sino que eran muy inteligentes y pertinentes lo que me obligaba a una tarea inmensa de pensamiento frente a las cámaras. Tuvimos que interrumpir a las 19 hs. como previsto habiendo usado una hora y media de tiempo.

Nunca había tenido una respuesta tan interesante por parte del público, pero también es cierto que era la primera vez que el tema a tratar era filosófico, pues la ética es una parte de la filosofía.

Esta experiencia nos hizo reflexionar sobre la importancia de la participación del público y ya en la segunda conferencia teníamos dos oradores: Enzo Maria Le Fevre, joven en edad, pero uno de los más informados sobre el tema pues trabaja directamente en la Comisión de la Unión Europea y Horacio Granero, un anticipador de sistemas inteligentes en el derecho en la Argentina. Le Fevre era un orador desde Bruselas y en la coordinación tuve que hacer malabares para otorgar unos 50 minutos a cada uno, con preguntas incluidas y terminar a la hora prevista con preguntas que no pudieron ser contestadas. Los oradores lo hicieron posteriormente y pudimos enviar las respuestas.

El 19 de agosto fue la prueba del nueve porque el famoso jurista español Emilio Suñé Llinás, trató *Un “decision making” ético para sistemas inteligentes* y la conocida informática argentina Lidia Seratti, se ocupó nada menos que de *Hacia una racionalización del alcance de la algorítmica de la IA y de los desafíos éticos y regulatorios impuestos por su irrupción sin límite ni fronteras*.

El ciclo estaba consolidado no solo por la calidad de los oradores sino por el interés y la calidad del público que se notaba en las preguntas cada vez más concretas, más precisas y más interesantes.

Ya estaba vigente el slogan “si es miércoles debe ser ética” y el interés se notaba no solo en la cantidad de participantes conectados sino también en la cantidad de personas que visitaban la versión grabada que apareció todos los miércoles en el Youtube de la Academia minutos después que terminara la conferencia.

Muchas personas hicieron posible la calidad y la precisión del ciclo, entre ellas Juan Cejas el responsable del funcionamiento de la plataforma que habilitaba 15 minutos antes la pantalla para estar seguros que el relator podía coligarse y participar pantalla si tenía ilustraciones para mostrar. Eso fue fundamental el 25 de noviembre cuando el prof. Emérito de la Universidad de San Pablo, Tercio Sampajo Ferraz (Jr.) trató el tema ¿Los robots mienten? Tercio, de quien soy amigo desde hace más de cincuenta años no podía entrar en la plataforma y pudimos resolverlo solo luego de una estrecha y larga conversación por WhatsApp.

Karina Libano cuidó que todo lo relativo al ciclo que fuese escrito estuviese siempre al día y Rodolfo Mesturini que los flyers, atractivos, anunciando cada conferencia estuviesen la semana antes para hacer la difusión.

No puedo analizar todas las conferencias dado que están en vivo y en directo en el sitio Youtube de la Academia. Vale solo recordar que, de los cinco relatores italianos, cuatro hablaron en español, algunos por primera vez como Valentina Sappo, otros recuperándolo de un pasado en Colombia como Giampaolo Maria Ruotolo. En italiano habló solo Natascia Archifa quien nos trajo el tema de la cyberguerra, despertando tanto interés que tuvimos que traer a los profesionales, es decir dos marinos uno Pablo Daniel Sorrentino, responsable de la cyber seguridad en la Armada argentina y Néstor Antonio Dominguez, ya retirado quienes hablaron de *Una posible defensa ante la amenaza de una ciberguerra. Algunas consideraciones macroéticas*.

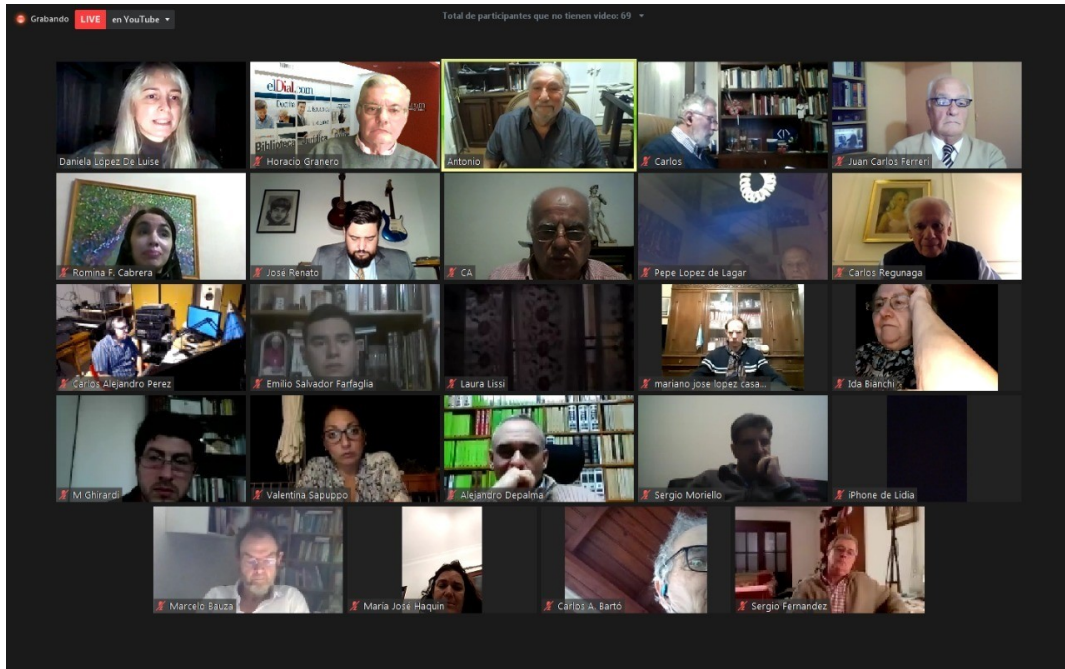
El 26 de agosto Luis Kun desde EEUU nos deslumbró con todo lo que hay en los aspectos biológicos de la inteligencia artificial. Tuvimos un presidente de Academia, Armando Andruet, de la Academia de Derecho de Córdoba, quien trató el tema de la ética en la función judicial. Un sacerdote Néstor Ramos: La inteligencia y la persona en Rome Call for Ethics, el tratado que se firmó en Roma en febrero del 2020 y una cubana Yadira Liza: *Las deficiencias actuales en el tratamiento del bias en Machine learning*, actualmente haciendo un doctorado en Canadá.

Tuvimos toda clase de problemas que afortunadamente pudimos superar siendo todo *on line*, Yarina Amoroso desde Cuba no pudo conectarse con la plataforma y tuvo que hacerlo a través de Carlos Rafael Rodriguez, retransmitiendo desde Rusia.

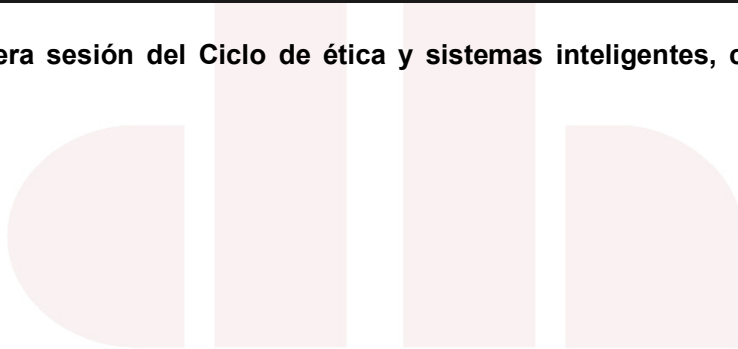
Y lo más importante: las preguntas que fueron satisfaciendo necesidades muy variadas y profundas desde el primer día y que le han dado una riqueza inusual al ciclo. Eso pueden apreciarlo, con paciencia en las 21 conferencias grabadas.

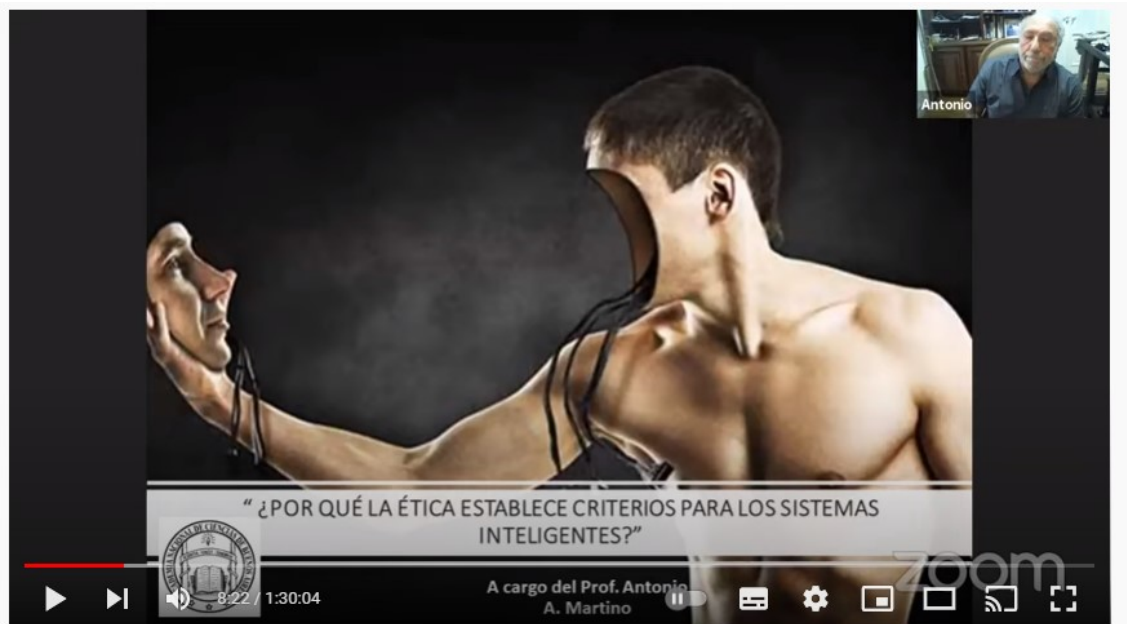
El material escrito que presento tiene la función de profundizar lo que se puede ver en los videos y dar algunos datos adjuntos que siempre vienen bien en un tema tan vasto y con una actualidad que no sospechábamos.

El libro está dividido en capítulos. El primero está compuesto mi trabajo sobre la ética y los sistemas inteligentes y mi semblanza. El segundo contiene las importantes notas que han hecho otros relatores del ciclo. El tercero trae la visión que la coordinadora del CETI, Daniela Lopez De Luise, hace del ciclo desde una perspectiva ontológica y gnoseológica.



Primera sesión del Ciclo de ética y sistemas inteligentes, con los asistentes





Primera conferencia del Ciclo "Sistemas Inteligentes y Etica" - CETI - ANCBA - 05/08/20

¿Por qué la ética establece criterios para los sistemas inteligentes?

Por Antonio A. Martino

Se supone que sabemos de IA ¿y la ética? Es una materia filosófica y requiere curiosidad y capacidad de asombro. Estando con los otros hay maneras correctas de comportarse e incorrectas. Esa es la ética. **Ética** proviene del griego **ethos**, etos que significa carácter, costumbre. En este sentido, la ética es considerada como una teoría del carácter y de las costumbres.

Creo que todos sabemos cómo comportarse éticamente y la mejor prueba es una **cola**. Y es una regla social que nos dice cómo comportarnos. Nos permite juzgar el comportamiento de los demás. Y esperamos que otros en las mismas condiciones hagan lo mismo.

La IA ha entrado en toda nuestra vida desde el celular hasta los drones que matan en la cyberguerra. ¿Y la ética? Lucha entre el imperativo categórico kantiano y los deberes prima facie de Ross (hay otras posturas). El imperativo para robots dice "actúa de tal manera que los efectos de tu acción sean compatibles con la vida humana".

Ya hay varios documentos sobre el tema: 1) Reglas éticas para la inteligencia artificial. El pasado 28 de febrero se firmó en el Vaticano; 2) Directrices éticas de la U. E. para una IA confiable; 3) Principios éticos de la Oede, y 4) La Unescoha emprendido un proceso de **dos años** para elaborar el **primer instrumento normativo mundial** sobre la **ética de la inteligencia artificial** en forma de recomendación.



La ética se aplica a **conductas**. En general debemos tomar decisiones. Y actuamos en automático, solo cuando hay problemas serios debemos tomar decisiones lógicas, analíticas y deliberadas necesitamos. Nuestro actuar tiene sesgos cognitivos: un peso desproporcionado a favor o en contra. Y construimos programas con nuestros sesgos (o prejuicios).

Después del accidente de un auto tripulado Tesla en mayo de 2017 en Florida apareció el tema del dilema del tranvía de Filippa Foot. Y luego un juego ético del MIT (*moral machine*) www.moralmachine.net.

A las 3 leyes robóticas de Asimov, se le ha agregado una cuarta: El robot debe siempre respetar la continuidad de la especie humana, aunque esto colisione con las tres reglas anteriores.

La IA se ha usado por Cambridge analítica para influenciar elecciones: Brexit y la de EEUU el año pasado. Los 4 grandes Facebook, Apple, Amazony Google fueron sancionados en USA por comportamiento predominante. En un mundo leve rápido exacto visible y múltiple. La forma automática de tomar decisiones se complica. De allí la necesidad de reglas éticas para un mundo transformado en particular por las nuevas tecnologías.



CAPÍTULO PRIMERO

ÉTICA Y SISTEMAS INTELIGENTES

Por Antonio A. Martino¹

Proposición 7 de Baruch Spinoza, *Ética more geométrico*:

“El orden y conexión de las ideas es el mismo que el orden y conexión de las cosas.

Quinta parte: Del poder del entendimiento o de la libertad humana Paso, por fin, a esta última parte de la *Ética*, que trata de la manera de alcanzar la libertad, es decir, del camino para llegar a ella. En esta parte me ocuparé, pues, de la potencia de la razón, mostrando qué es lo que ella puede contra los afectos, y, a continuación, qué es la libertad del alma, o sea, la felicidad”.

Wittgenstein carta a Ludwig von Fricker: “Mi libro consiste de dos partes: la aquí presentada, más lo que no escribí. Y es justamente esta segunda parte la más importante. Mi libro traza, por así decirlo, desde el interior los límites de la esfera ética, y estoy convencido que esta es la ÚNICA forma rigurosa de trazar sus límites. En pocas palabras, considero que ahí en donde tantos otros hoy peroran, me las arreglé para poner todo bien en su sitio callándome sobre ese asunto”.

Proposición 2.04 del *Tractatus*: La totalidad de los eventos subsistentes es el mundo
Proposición 6.421. Es claro que la ética no puede formularse. La ética es trascendental (ética y estética son una sola cosa)

Introducción

Lo primero que va dicho es que los sistemas inteligentes están llegando a casi todas las partes de nuestra vida y como han crecido en forma vertiginosa van creando mundos diferentes. Como toda invención o descubrimiento humano importante los sistemas inteligentes son peligrosos y por ello es necesario encontrar límites morales a su uso.

La inteligencia artificial se ha desarrollado en modo estrepitoso en los últimos años no solo desde el punto de vista teórico sino también práctico. No hay parte de las actividades humanas que no hayan sido alcanzadas por ellas, desde los teléfonos celulares hasta las armas de la guerra cibernética.

Hay muchas definiciones de IA. 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.

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 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.

¹ Miembro de Academia Nacional de Derecho y Ciencias Sociales de Córdoba, Profesor emérito de las Universidades de Pisa, Italia y Salvador, Argentina.

Sistemas inteligentes se remiten a “pensar” pero el vocablo en el Diccionario de la Real Academia española tiene siete acepciones 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 alguien². A esto podría añadirse recordar, deducir, querer, inducir, odiar.

En este artículo vamos a dejar el tema de los sistemas inteligentes a lo que ya el lector conoce, con algunas precisiones y algo parecido haremos con la ética.

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.

Incluso 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 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 Elon Musk habla de la misión de Neuralink³, se refiere a la necesidad de mejorar nuestra capacidad, de aumentar nuestro ancho de banda de procesamiento de información, especialmente en un mundo cada vez más definido por la inteligencia artificial.

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.

La inteligencia artificial se ha puesto de moda desde el punto de vista teórico y desde el punto de vista práctico y nadie quiere quedar fuera de ella. Ello supone algunos beneficios en el sentido que habiendo mucha producción hemos mejorado mucho nuestro conocimiento del tema dado que todos quieren hacer IA, pero también alguna dificultad sea porque se llama IA a cosas no lo son, y con este afán de

²<https://dle.rae.es/pensar>. Personalmente no creo que se pueda decir de una máquina que piensa.

³*Neuralink de Elon Musk: el último avance del multimillonario empresario en su plan de conectar nuestros cerebros a computadoras*, 2018/20, www.bbc.com/mundo/noticias-53955394.

llenar todo con IA impide que algunos casos que pueden resolverse en modo más sencillo se lo haga de manera complicada.

Las modas traen complicaciones. Como todos la quieren, muchos productores que no la tienen visten de IA lo que tienen y lo venden. Actualmente, por esa razón, se asocia a la inteligencia artificial productos que, interesantes en sus campos no tienen nada que ver con la IA. El primero es el blockchain. Más allá del uso siempre más frecuente de “blockchain” en forma de “buzzword” desde un punto de vista técnico se refiere a registros compartidos y distribuidos o “privados/centralizados” los registros están constituidos por nudos validados por técnicas de hash y criptografía. Una de las propiedades notables es la no modificabilidad de los “elementos/transacciones” que forman parte de la cadena de bloques del registro. Que no se puedan modificar porque está de por medio el hash es algo así como el descubrimiento del agua tibia.

Las únicas reales diferencias entre las varias tecnologías para construir un blockchain se pueden resumir en el hecho que la distribución de los nudos (miners) es pública y libremente distribuida o privada/centralizada/controlada. O sea, cualquiera puede ser el minero o solo algunos. Obviamente se pueden “imprimir” monedas virtuales, notarizar contratos y transacciones o validar una bella producción musical o un film porno o cadenas de procesos y/o contratos entre empresas o dentro de una empresa.

El haber podido superar (dicen) la amenaza de la computación cuántica y su aplicación en prácticamente cualquier actividad humana lo hace muy atractivo. El verdadero límite de esta tecnología es el enorme dispendio computacional legado a la validación de bloques pues más se alarga la cadena, más se gasta.

Dicho esto, las competencias comunes, sobre todo de tipo criptográfico en un especialista en tecnología blockchain y uno de IA son iguales a cero.

Otro tanto puede decirse de una tecnología que se hará muy famosa próximamente que es el 5G, tecnología de redes que poco tiene que ver con la IA.

Y que podemos decir de la ética. La ética es una parte de la filosofía. Para hacer filosofía se requiere capacidad de asombro⁴. Ética proviene del griego *ethos*, que significa carácter, costumbre. En este sentido, la ética es considerada como una teoría del carácter y de las costumbres.

Detengámonos un momento en esta capacidad de asombro: no es sorpresa, ésta puede tenerla cualquier animal. El asombro en el Diccionario de la Real Academia⁵ es gran admiración y extrañeza y aun susto y espanto. Extrañarse es ponerse fuera y hay algo de eso en el asombro: se pone uno fuera, fuera de todo lo que le es familiar y por eso lo del susto o espanto. Y esta es una capacidad solo de la especie humana que le permite ir más allá de lo conocido de lo familiar y exponerse fuera a lo desconocido.

⁴ Por eso recomiendo a quienes quieran adentrarse en la historia de la filosofía el libro de Jostein Gardener, *El mundo de Sofía*, Madrid, Siruela, 2008.

⁵ <https://dle.rae.es/asombro>.

Es interesante notar que asombro trae también susto y espanto y esta es una característica de lo que en psicología se llama trauma que se define como un evento que amenaza profundamente el bienestar o la vida de un individuo. El trauma, la ruptura aparecen cuando uno abandona los conocimientos familiares y se aventura, exponiéndose, a saberes nuevos.

Y filosofía viene de amor a lo que no se conoce, pero se quiere tener, uno quiere lo que no tiene y quiere tener: conocimiento. Pero el conocimiento es peligroso y de allí también que pueda producir susto o espanto desde el asombro. No olvidemos que el árbol prohibido en el Paraíso era el del conocimiento⁶. Por eso la ética que es una actividad filosófica supone que estemos extrañados, asombrados, fuera de los conocimientos familiares, expuestos.

La pandemia ha traído muchas cosas, no todas malas. Pero el distanciamiento social es ciertamente entre los perniciosos, ya que el hombre es un ser “naturalmente político”. El distanciamiento físico evita muchas efusiones afectivas y distancias mucho más. En esta distancia nace una nueva ética: no diferente y adversa a la tradicional, sino más precisa, más detallada y que prevé nuevas circunstancias. En una ética de deberes *prima facie* las nuevas circunstancias cambian los deberes. También la difusión de las redes es un nuevo elemento que debe ser considerado en lo social y en las costumbres.

Byung-Chul Han: “Nosotros, felices esclavos de la pandemia digital”. En una entrevista concedida en estos días el filósofo sudcoreano afirma que estamos tan atrapados en la trama de las redes sociales y la IA que estamos contentos en la Pandemia porque tenemos Internet⁷. “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 un modelo de régimen de vigilancia impulsado por una falta de humanidad alimentada por nuestro miedo a morir.

2. Inteligencia artificial y riesgo

Lo cierto es que la inteligencia artificial llena casi todas las actividades humanas y comienza a mostrar sus riesgos.

Voy a obviar los riesgos que algunos autores atribuyen a la IA, como ser que sus algoritmos rápidamente sobrepasaran (si no lo han hecho ya) a la inteligencia

⁶El árbol del conocimiento del bien y el mal(en hebreo, אֵץ טוֹב וְרָע *Etz ha-da'at tov va-ra*) en la historia del Jardín del Edén, en Génesis 2-3.

⁷https://rep.repubblica.it/pwa/robinson/2020/10/30/news/byung-chul_han_noi_schiavi_felici_della_pandemia_digitale_-272446064/?ref=nl-rep-a-bgr.

humana porque, en general son debidos a miedos⁸. El robot provoca miedo y evoca miedos ancestrales. Desde el *golem* de la tradición hebrea, que aparece citado una sola vez en el Antiguo Testamento, esto es, en Salmo 139, verso 16, allí se dice “inconcluso o esbozado me vieron tus ojos. Tus ojos vieron mi embrión, y en tu libro se escribieron todos los días que *me* fueron dados, cuando *no existía* ni uno solo de ellos”. Ese embrión, según la interpretación, es el *golem*, pero para otros Adán. En hebreo, idioma original del Antiguo Testamento, se coloca el origen de la figura del *golem*, de donde la *Qabbal* (tradición) busca sus significados⁹.

En los orígenes hay una conexión etimológica entre Adán, creado por Dios y la tierra, concebida como metonimia en su componente de polvo. Adán, antes de que llegue su creador, el perfeccionamiento espiritual que lo vuelve una figura animada y dotada de intelecto, debidamente humano, es Adamah, tierra, *golem*. *Golem* (*golmi*) es una forma adjetival que en la Biblia aparece una sola vez en el salmo citado¹⁰.

De allí que las lecturas del *Midrash* cumplen la conjugación entre el aspecto humano (y paradigmático) y el aspecto no humano, inorgánico y material.

El estado golémico se atribuye al mismo Adán. Es importante percibir esta diferencia: en el estado orgánico, primitivo, hay materia, pero todavía no está el hombre, este va a venir luego, insuflado por el hálito divino. Y esta misma diferencia se cree ver en los robots, que son materia (que en algunos casos puede tener forma humanoide), pero le falta la acción que aparece luego a través de programas especiales, entre los cuales existen también los programas de aprendizaje. Y eso hace temer a quienes no conocen la robótica y tampoco los límites de IA que ese robot poco a poco se independice y llegue a ser más inteligente que el ser humano y termine dominándolo¹¹.

Indico un libro como representante de esta tendencia el de Yuval Noha Harari *21 lecciones para el siglo XXI*¹², el autor es un historiador israelí que ha escrito por lo menos dos libros que han sido *best sellers*, pero de inteligencia artificial no sabe mucho.

⁸ Los seres humanos tenemos miedos ancestrales. Hay 6 que son universales: a la muerte, a la falta de autonomía, al hambre, a la soledad, a las pandemias, a la mutilación física o psíquica. También a la muerte.

⁹ Dentro de estos cuatro métodos de comprensión de la Torá, existen un sin número de posibles vías de entendimiento. Por ejemplo: Hay muchas maneras diferentes de entender la Torá según Peshat. Es por eso que hay muchos comentaristas de la Torá que se concentran en Peshat (Rashi, Iben Ezra Rashbam y muchos más) y ellos estarán muy a menudo (al parecer, la mayoría) de acuerdo sobre el significado de un verso. De hecho, según la Cábala hay 600.000 formas de entender Peshat.

¹⁰ Algunas de las referencias pueden verse en el excelente libro *Dal Golem al Cyborgs*, de Bárbara Henry, Salomone Belforte, Livorno, 2016.

¹¹ Hay en ciencia ficción un relato de unos robots muy sofisticados que van perfeccionando un modelo y, cuando llegan a la Tierra, satisfechos por su perfección lo dejan: han creado al hombre.

¹² Yuval Naoah Harai, *21 lecciones para el siglo XXI*, 2018, Bs. As., Penguin Random House Grupo Editorial.

El Golem es un monstruo creado por el Rabi Levi en Praga y de allí hay mucha literatura y cine dedicad a él. A veces es un monstruo, a veces sirve para salvar a los judíos del gheto. Además del Golem hay otros monstruos como Frankenstein. Todos seres inventados.

Los signos son improntas que modelan la realidad a su imagen. El Golem es construido y vuelto vital del correcto uso de las letras (otot) del Nombre Divino. El análisis de los signos puede hacerse con un código y esta palabra aúna informáticos, juristas, lógicos, biólogos y lingüistas.

Del Golem al ciborg, ajedrecista, tramite el expediente de reconducir la variada tipología de vida antropomórfica y artificial a la “mera” fabricación de autómatas: el robot indica solo la máquina que exonera del trabajo manual (trabajar, en checo *roboti*).

El maridaje entre orgánico e inorgánico aparecen en las discusiones social y culturalmente acreditadas en el pasado de Occidente antes de Mary Shelley (autora de Frankenstein) y las inquietudes góticas del romanticismo de Byron¹³.

Sí, vale la pena ocuparse de los peligros serios que han aparecido ya y que están siendo tratados; podemos llamarlos riesgos y son debidos al hecho que se trata de una metodología nueva a la cual la sociedad no ha tenido tiempo de habituarse. Empecemos por el uso de armas autónomas: Las armas programadas para matar suponen un grave riesgo en el futuro de la IA. No solo son peligrosas por que puedan convertirse en autónomas completamente, cosa que veo poco probable pero sí que se la utilice sin mucha supervisión para hacerlas más eficientes. Son una gran tentación para cualquier vigía que quiere dominar el espacio circundante todo el tiempo y puede poner el arma en automático, lo que quiere decir que actúa sin control humano.

La manipulación de la sociedad ya apareció sea en el caso de las redes sociales pueden ser una gran fuente para obtener información sobre cualquier persona, y además de utilizarse para el marketing y brindar anuncios específicos para cada persona, como el caso de Facebook. Con el escándalo de Cambridge Analytica en relación con las elecciones presidenciales de EEUU y el Brexit en Reino Unido, se pudo ver el enorme poder que puede dar tener datos para manipular a las personas, con lo cual la IA al ser capaz de identificar algoritmos y datos personales, puede ser extremadamente peligrosa¹⁴.

La invasión de la privacidad para oprimir socialmente: es posible “seguir” las pistas de un usuario en la red, y utilizar mucha información invadiendo su privacidad. Por ejemplo, en China se utilizará información como el reconocimiento facial de

¹³ Algunas de estas observaciones pude comprobarlas en el excelente libro Dal Golem al Cyborgs, de Bárbara Henry, Salomone Belforte, Livorno, 2016.

¹⁴ La resolución de la Cámara de representantes norteamericana del 7 de octubre de este año, con sus 415 páginas es una fuente valedera del riesgo y de lo que razonablemente harán las autoridades.

cámaras y la forma en la que se comportan, si fuman, o si ven muchos videojuegos para el sistema de crédito social¹⁵.

Divergencia entre nuestros objetivos y los de la producción de IA: Si nuestros objetivos no coinciden con los de la máquina, las acciones que pedimos llevar a cabo pueden terminar en un desastre. Por ejemplo, mandar una orden a la IA de llevarnos a un lugar lo más rápido posible, pero sin especificar que tiene que respetar las reglas de tránsito para no hacer peligrar las vidas humanas. O cualquier otra acción humana que comporta riesgos y que los humanos deberíamos calcular cuando hacemos el cálculo general de la acción.

Discriminación. Como las máquinas de IA pueden recoger tu información, analizarla y rastrearla, también pueden utilizar esta información en tu contra. Por ejemplo, una compañía de seguros puede negarte el seguro debido a las veces que las cámaras tienen imágenes de ti hablando por teléfono, o una persona que aspire a un puesto de trabajo puede perder la oportunidad debido a su baja red social de contactos en Internet. Sucede además que los humanos tenemos sesgos (prejuicios) y que los pasamos a las máquinas sin darnos cuenta. Un peligro extraordinario consiste en el aprendizaje de los programas de IA. Lo primero que hacemos es enseñarle a recoger toda la documentación existente sobre el tema (que no es solo escrita, sino fotos y videos) y aprender de allí. Pero como los humanos nos comportamos discriminatoriamente los programas aprenden a discriminar.

Dicho brevemente, si un invento humano es importante necesariamente será también peligroso, casi podría decirse que para medir la importancia de una invención humana hay que medir su peligrosidad. La desintegración del átomo y la fuerza que produce e Internet podrían decirse las dos invenciones más importantes del siglo XX, bueno son las más peligrosas. Un tema que le sigue es ¿podemos pensar una sociedad sin energía atómica? O ¿podemos pensar una sociedad sin Internet? Creo que la respuesta es negativa lo que comporta asumir los riesgos de la dos.

Los sistemas inteligentes no van a volver atrás, por lo tanto, la solución no es demonizarlos como si no existieran o fuéramos a desterrarlos sino ver cuáles son los riesgos que crean, diferenciarlos pues algunos son menores y otros mayores, colocarlo en los justos contextos, estudiar las medidas para contrarrestarlos y tener un ojo vigilante a una sociedad siempre más compleja, siempre más cruzada por todas las líneas de las nuevas tecnologías.

3. El ser humano

Para Aristóteles el hombre es un animal naturalmente social¹⁶. Estando con los otros hay maneras correctas de comportarse e incorrectas. Esa es la ética.

¹⁵ Va dicho que el sistema chino es una forma autoritaria de gobierno, una feroz dictadura. Sobre dictadura ver Martino, Antonio A., *Osservazioni sulla definizione di "dittatura"*, Il Politico, XLII, n° 2, Bologna, 1978.

¹⁶ El hombre es un animal naturalmente político.

Creo que todos sabemos cómo comportarse éticamente y la mejor prueba es una cola. Es una regla social que nos dice cómo comportarnos. Nos permite juzgar el comportamiento de los demás Y esperamos que otros en las mismas condiciones hagan lo mismo. Basta enunciar su nombre para convencer a alguien que está haciendo lo incorrecto: “esto es una cola”.

Pero no es tan fácil saber en cada caso cual es el comportamiento ético o moral (moral viene del latín mores, costumbre). Hay muchas discusiones sobre la diferencia entre moral y ética, en general no me convencen y por razones prácticas aquí las vamos a considerar sinónimos.

Justamente porque no es fácil distinguir el comportamiento moral hay muchas teorías éticas.

La primera distinción que conviene hacer es entre ética y metaética: mientras la primera se ocupa de las reglas morales la segunda se plantea el tema de saber qué tipo de enunciados son los éticos y aquí hay dos grandes corrientes metaéticas: los cognitivistas, quienes sostienen que los enunciados éticos son susceptibles de ser verdaderos o falsos y por lo tanto puede demostrarse una verdad ética y quienes sostienen que ello es imposible, los no cognitivistas metas éticos. El que no se crea que un enunciado ético sea demostrable empíricamente no quiere decir que haya una fe menor en cumplirlo. Simplemente no se lo considera demostrable por verdad y falsedad.

Casi todas las ideas éticas que provienen de la religión son cognitivistas metas éticos porque sostienen que la obligatoriedad deriva del hecho que algún Dios lo ordenó y como creen, consideran que es empíricamente demostrable.

Entre las teorías éticas hay algunas que han conseguido una gran difusión. Por ejemplo, la teoría kantiana del imperativo categórico.

Dice Emanuel Kant en *Fundamentación de la metafísica de las costumbres* “Obra sólo según aquella máxima por la cual puedas 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)¹⁷. Y lo considera un imperativo categórico porque es una norma que no tiene ningún criterio de atenuación desde el punto de vista de las condiciones que se verifican en el momento de aplicarlo.

Obviamente es muy rígido y muy duro y muchos autores han tratado de desmentirlo o de falsificarlo. El caso más interesante es el del Marqués de Sade con su *Filosofía del tocador*¹⁸. Vale la pena leerlo porque en la introducción el autor dice explícitamente que ha escrito ese libro para falsificar la teoría kantiana e iniciar a los franceses en la cultura del libertinaje. Pero en el transcurso de la obra, donde tratan de inculcar a Eugenia a divertirse y considerar a los demás como objetos o medios,

¹⁷ Kant, Emanuel, *Fundamentación de la metafísica de las costumbres*, Bs. As., No Books, 2003, versión original *Grundlegung zur Metaphysik der Sitten*, Riga, 1785.

¹⁸ Francois de Sade, *La filosofía nel boudoir ovvero i precettori immorali. Dialoghi per l'educazione delle fanciulle*, Garzanti, Milano 2004. Título original *La Philosophie dans le boudoir ou Les instituteurs immoraux*, Paris, 1795.

nunca como fines, uno de sus maestros confiesa que ha sido muy grata la enseñanza porque vio que Eugenia gozaba. Pero eso, depravado o no, es contrario a los postulados del autor y reafirma la teoría kantiana de tratar a los demás como fines, nunca como medios: si el maestro estaba contento porque Eugenia gozaba se tuerce la finalidad primordial de pervertir, tratando al otro como un medio y renace la idea que el otro ha obtenido un beneficio y por lo tanto fue tratado como un fin y no un medio.

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.

Entre los que han tratado de aliviar la dura ley kantiana hay muchos y Jean Jacob Rousseau, lo intenta en la Nueva Eloísa¹⁹.

El siglo pasado se caracterizó por el progreso científico, tecnológico y médico, entre otros, pero también por una inusitada crueldad que se tradujo en la pérdida de millones de vidas y una falta de libertad por culpa de numerosas dictaduras.

Los movimientos políticos de dicha época presumían de argumentar desde la lógica de la razón y se escudaban en interpretaciones interesadas de intelectuales como Friedrich Nietzsche o Martin Heidegger. Tanto el leninismo, como el fascismo y nazismo se aferraron a pensadores como estructura legitimadora, pero no son autores de una política destructiva. Hubo una lectura sesgada. En realidad, se puede sacar citas de todo.

Con la obsesión por el racionalismo y la falta de sensibilidad, el valor de la vida humana se depreció hasta tal punto que algunas personas dejaron de ser consideradas como ciudadanas. Este fenómeno derivó en la creación de guetos, campos de concentración y gulags. Esa deshumanización se traduce en un “tú no eres de los nuestros” y, en su forma más radical, “tú no eres humano”. Es vital reclamar la dignidad de las relaciones interpersonales.

Un pensador que ha envenenado la política y por consecuencia la moral es Karl Smith, creador de la disyuntiva “amigo - enemigo”²⁰. Estado es definido como el estatus político de un pueblo organizado sobre un territorio delimitado. En la época actual, el Estado se encuentra entremezclado con la sociedad, al punto que lo político no admite ser definido a partir de lo estatal. Por eso, para llegar a una definición de lo político se requiere el descubrimiento y la fijación de una distinción específica a la cual sea posible referir las acciones y los motivos políticos. Esa distinción es la distinción de “amigo y enemigo”, la cual determina la esencia de lo político, y desde este criterio se puede determinar lo político y lo apolítico. En la medida que no es

¹⁹ Rousseau, Jean J., *La nouvelle Eloise*, Paris, 1761. “Ce n'est pas assez, croyez-moi, que la vertu soit la base de votre conduite, si vous n'établissez cette base même sur un fondement inébranlable”.

²⁰Norberto Bobbio se carteaba con Karl Smith desde que lo había conocido en Alemania en 1933. Una vez me pregunto qué me parecía y le dije que estaba en desacuerdo a lo que él retrucó “el liberalismo obliga a escuchar a todos”. Solo ahora me doy cuenta que no tenía razón porque el “todos” no puede incluir a gente que obra de mala fe.

derivable de otros criterios, ella corresponde, para la política, a los criterios relativamente autónomos de otros ámbitos humanos: bueno y malo para la moral, bello y feo para la estética, etcétera²¹.

Gracias a esa distinción aparecen “ellos” y “nosotros” y en sistemas populistas sirve para eliminar toda otra opción que no sea con el régimen o contra él, sin dejar lugar a otras elecciones. Se es fascista o disfattista, comunista o anticomunista, castrotrista o contrera.

En el pasado siglo también apareció un autor inglés Charles Ross quien en su libro *Fundamentos de Ética* 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 obtengamos²². En una cola lo elemental es respetar el orden de llegada, pero si aparece un señor con problemas cardíacos, nuestro deber cambia y es él quien debe ocupar el primer lugar en la fila. Algo parecido sucede con la llegada de una señora embarazada²³. La idea de los deberes *prima facie* es muy útil en nuestro tiempo donde permanentemente se van cambiando las condiciones del mundo en que vivimos: hay nuevos datos, hay nuevos deberes morales.

4. La llegada de los sistemas inteligentes

Con la llegada de los sistemas inteligentes hay situaciones nuevas. Se dice que en los últimos dos años se han hecho tantos avances como en los 18 anteriores y estamos hablando solo de este siglo. La velocidad del cambio tecnológico no tiene parangón en la historia de la humanidad y esto obliga a pensar una moral mucho más prolija y dedicada a cada una de las nuevas situaciones en las cuales se ha puesto a los humanos. No es que haya cambiado la moral, ha cambiado y mucho, el mundo.

El progreso tecnológico y la masa extraordinaria de datos a disposición impulsó gobiernos, académicos y la sociedad civil a promover una reflexión sobre los desafíos éticos ligados a la gran progresión de sistemas inteligentes y comenzaron a surgir directivas y acuerdos.

Tenemos en claro los datos y la tecnología son herramientas utilizadas por los Estados para automatizar, predecir, identificar, monitorear, detectar, atacar y castigar. Actualmente son tantísimos, pero se destaca cómo deberían utilizarse princi-

²¹ Smith, Karl, *El concepto de lo político*, Madrid, Alianza Editorial, 2005. La obra original es de 1932.

²² Ross, David, *Fundamentos de Ética*, Bs. As., Eudeba, 2003. El original inglés es de 1930.

²³ Una cosa curiosa es que los deberes *prima facie* de Ross dieron lugar a especulaciones de lógica difusa en campo de las ciencias duras. Y esto es debido a que todas las lógicas para consistentes disminuyen el valor de la negación. Lo opuesto a 1 no necesariamente es 0, lo opuesto a verdadero no necesariamente es falso, sino que hay alternativas intermedias.

palmente para apoyar la protección social y los sistemas de bienestar y garantizar el bienestar.

a. Un caso importante

El domingo 18 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 identificarlo²⁴.

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%²⁵. 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.

²⁴ Un niño con unas pocas fotografías puede detectar enseguida una bicicleta. Es la diferencia entre un humano y una máquina: 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 máquina puede ser racional, nunca razonable.

²⁵ El 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 reconocidos hará que el sistema frecuentemente frene y hasta se detenga, según las circunstancias, haciendo que el vehículo realice una conducción torpe y brusca.

Los sistemas inteligentes disponen de cantidad de información muy superior a la que tiene el ser humano, pero también tienen sus limitaciones por lo que exigen cuantificar los riesgos sin olvidarnos que se trata de pruebas encaradas para aprender de la realidad y corregir. En los dos años transcurridos desde el accidente, muchos de estos problemas fueron resueltos.

Se pudo verificar luego que la conductora de seguridad Rafaela Vásquez iba mirando un programa de TV en su celular, y recién levanta la vista en el último segundo antes del impacto, de acuerdo con la cámara interior del vehículo. Cuando hacen la autopsia de la persona fallecida, Elaine Herzberg, determinaron que su sangre contenía cristal y marihuana, además de una larga historia de drogas en su vida. El Estado de Arizona había permitido en el 2015 la circulación de autos autónomos de prueba, sin mayores indicaciones, es decir ni siquiera mencionaban la necesidad de un conductor de seguridad. Por esa razón las empresas se establecieron en ese Estado como centro de desarrollo y pruebas. Al momento del accidente la empresa Uber contaba con 400 autos autónomos en prueba circulando en Tempe. Un dato: el Gobernador del Estado de Arizona había manifestado antes del accidente, que las empresas podían tomar como empleados ex convictos. Rafaela Vásquez lo era porque había sido condenada a 5 años de prisión por intento de robo. Uber se encontraba en una situación complicada por una disputa con Google que le costó 245 millones de dólares.

Pido disculpas por el exceso de datos, pero son interesantes para mostrar como un cúmulo de circunstancias muy difíciles de conocer con anterioridad pueden dar un rasgo fundamental a un hecho que luego puede ser considerado como disparador de reflexiones éticas que no existían hasta ese momento.

A partir de ese hecho comenzaron las preguntas que tal vez habría que haberse hecho antes del experimento sobre la eticidad de comportamiento y volvió al ruedo un viejo dilema que Filippa Foot había enunciado en los años sesenta del siglo pasado: el dilema del tranvía. Se escribieron libros como *Moral Machine* y hasta el famoso MIT (Instituto Tecnológico de Massachusetts) compuso una serie de 15 juegos sobre el famoso dilema que se puede ver (y jugar) en www.moralmachine.net²⁶.

b. La cosa se pone más peligrosa

Los vehículos aéreos no tripulados (UAV) son ahora una herramienta central de antiterrorismo. Más específicamente, los vehículos aéreos no tripulados armados se han demostrado altamente valiosos para el ejército de los Estados Unidos al llevar a cabo asesinatos selectivos en regiones remotas del Oriente Medio y el Sub Sahara, África.

Los progresos recientes y en curso en la esfera de la inteligencia artificial (IA) abren la puerta a varias nuevas aplicaciones militares, particularmente en la asisten-

²⁶ Hay muchas críticas y escándalos a la manera de jugar de *moral machine* pero se olvida: 1) que es un juego, 2) muchas de las críticas vienen de la visión meta ética de quien la formula, por ejemplo, pensando que los enunciados éticos pueden ser demostrados como verdaderos o falsos.

cia del proceso de toma de decisiones de los operadores de vehículos aéreos no tripulados militares que llevan a cabo asesinatos selectivos a distancia. Entre otras aplicaciones, el aprendizaje automático, a través de las redes neuronales y el aprendizaje profundo, puede ayudar a la identificación (y quizás incluso la selección) de los objetivos.

Las armas autónomas letales tienen algoritmos cognitivos que permiten identificar, seleccionar y eliminar un objetivo (humano o no) sin necesidad, o más bien con escasa, presencia humana. Drones, misiles de crucero, bombarderos y tanques están ya dotados de sensores, cámaras y radares que permiten el control remoto²⁷.

Va dicho que varios países, comenzando por EEUU tienen un sistema de espía permanente a través de satélites y que, apoyándose en estructuras autorizadas expresamente en el suelo de aliados, como Australia, le permiten verificar qué sucede en casi todo el mundo a un nivel de acercamiento digno de películas²⁸.

En los Estados Unidos, desde 2001 existe una autorización para el uso de la fuerza militar dirigida desde espionaje satelital. Las convenciones que rigen el uso de la fuerza empezando por la de Ginebra de 1949, no regulan las diversas aplicaciones militares que con medio de IA podría llevar a cabo en la lucha contra el terrorismo. Se sabe un arma o un sistema de armas hoy se usa con fines antiterroristas, mañana en una guerra contra otro país, en algún momento como sistema de represión nacional.

Y estos usos son importantes en teatros de operaciones lejanos donde solo puede llegarse con medios modernos de observación y de transporte de armas, así como armas mismas que pueden operar sin necesidad de tripulación a distancias enormes y sin el riesgo de vidas amigas. Esto depende de la precisión de la observación y del ataque sucesivo.

De particular atractivo para los EEUU, así como otras naciones que recurren a los UAVs, es la capacidad de “actualizar la violencia a distancia”²⁹ puede matar sin poner las botas en el suelo. Una operación clave.

La ventaja de esta realidad es que las vidas de los operadores de UAV no están en peligro de extinción.

Hay factores principales explican por qué los vehículos aéreos no tripulados están particularmente favorecidos cuando se trata de ataques dirigidos. En primer lugar, ya que los operadores encuentran se alejan del campo de batalla convencional y no tienen que temer por sus vidas, pueden flotar los UAVs sobre un lugar específico hasta que su objetivo previsto es donde tiene que estar para un ataque letal exitoso³⁰.

²⁷ Por ejemplo, el Phalanx (MK15) US- Navy.

²⁸ De hecho, hay una serie que habla del tema: Pine Gap de Netflix.

²⁹ 17/6/13, en línea: Institución Brookings www.brookings.edu/articles/why-drones-work-el-caso-de-las-armas-de-eleccion-de-los-lavaderos [perma.cc/Z9YC-W7FP].

³⁰Byman, Daniel L., *Por qué funcionan los drones: El caso del arma elegida por Washington*.

Lo importante es que la falta de un piloto humano en la cabina permite más la precisión de tales ataques porque un UAV puede ser volado a una altitud menor que su equivalente tripulado. En altitudes más bajas, los pilotos convencionales cada vez más vulnerables a los sistemas de defensa tierra-aire³¹, lo que significa que deben volar rápidamente sobre su objetivo para evitar ser disparados hacia abajo si deciden volar bajo o permanecer a una mayor altitud y más lejos de su objetivo. En ambos casos, un ataque realizado por un hombre es probable que el avión sea menos preciso que el de un UAV.

En segundo lugar, no tener un piloto en la cabina también significa que los UAVs pueden permanecer en el aire por largo tiempo períodos de tiempo en comparación en comparación con los aviones tripulados que tienen que aterrizar cuando la resistencia de sus pilotos se ha agotado. Por ejemplo, uno de los pilotos de los EEUU, los UAVs más usados –el Predator– pueden volar hasta 35 horas sin tener que tocar tierra³². Esto permite a múltiples operadores tomar turnos consecutivos y asegurar que un objetivo no se pierda porque un piloto necesite dormir. Esto también asegura una presencia casi permanente en el cielo. Tercero, el láser contemporáneo. Municiones guiadas, con las que se equipan los vehículos aéreos no tripulados armados, ahora permiten que un misil golpee su objetivo dentro de un margen de cinco pies de radio de la base de un régimen de gobierno para la proyección de la fuerza a través de AI...de los vehículos aéreos no tripulados asistidos.

Previendo estos riesgos que aumentan con las armas no tripuladas se han realizado diversas conferencias como la de Buenos Aires en 2015 donde se publicó una carta abierta de problemas éticos ligados al uso bélico de la IA, en 2016 se creó un grupo de expertos gubernamentales y en 2019 se han activado conversaciones diplomáticas sobre el tema.

La CCW es la conferencia internacional sobre la revisión de armas convencionales para tratar los riesgos y efectos traumáticos excesivos que las nuevas armas ponen sobre la población civil y regularla a través de un derecho internacional humanitario. Se está tratando de aprobar un alegato adicional a la CCW, convocar una conferencia de expertos y aprobar una Robot Law o ley sobre el uso ético de sistemas inteligentes.

El art. 36 del Protocolo adicional a los acuerdos de Ginebra obligan a los Estados a revisar si el uso de armas que producen o adquieren respetan las reglas del derecho internacional. Se empieza a considerar que las decisiones de herir, matar o traumatizar seres humanos no debe ser aprendido por sistemas autónomos bélicos.

Se va afianzando la idea que en la ciberguerra también el control humano significativo sea la piedra angular del todo el desarrollo.

Un nuevo programa de investigación de DARPA está desarrollando interfaces cerebro-computadora que podrían controlar “enjambres de zánganos, operando a la velocidad del pensamiento”. ¿Y si tiene éxito?

³¹ James Der Derian, *Virtuous War: Mapping the Militar-Industrial-Media-Entertainment-Network*, Nueva York, Routledge, 2009.

³² Alcides Eduardo dos Reis Peron, *The “Surgical” Legitimacy of Drone Strikes? Issues*.

¿Cómo las sospechas de espionaje amenazan a la ciencia transfronteriza?

Una empresa de inteligencia advierte que China está explotando a los científicos cuánticos occidentales con fines militares. Las pruebas son escasas, pero las tensiones están aumentando. Siendo un sistema político opaco es difícil saber exactamente qué están haciendo y dónde quieren llegar.

El ejército de los EEUU no tiene parangón en su capacidad de proyectar el poder en todo el mundo, y eso no va a cambiar.

Toda tecnología que cambia el juego es absorbida por los militares y la IA no es una excepción. Hasta cierto punto, máquinas como los drones utilizados en la guerra ya tienen una gran autonomía. Y eso plantea serias cuestiones éticas.

Esta semana, dice Weekend Reads, estamos explorando la inteligencia artificial militar y cómo se está utilizando tanto ahora como en el futuro. Y nos preguntamos: ¿qué pueden o deben hacer los investigadores para evitar que su trabajo sea convertido en un arma?³³.

En su reciente libro David Kelcullen³⁴ pregunta cómo y qué han aprendido los oponentes de Occidente durante el último cuarto de siglo de conflicto. Aplicando una combinación de teoría evolutiva y detallada observación de campo, explica lo que pasó con las “serpientes”—amenazas no estatales incluyendo terroristas y guerrilleros— y los “dragones”—competidores estatales como Rusia y China—. Explora cómo los enemigos aprenden en condiciones de conflicto y examina cómo el dominio occidental sobre una forma de guerra. Los agentes no estatales pueden ahora acceder a niveles de precisión y sistemas de armas letales que antes sólo estaban a disposición de los gobiernos.

El tema se lo planteaban ya en los años 70 en Alemania con las brigadas rojas y la banda Baader Meinhof: ¿era una nueva forma de guerra? Sí, pero desde las hordas bárbaras, pasando por Güemes, Joseph Broz, Tito que escribe un libro de guerra de guerrilla, Ernesto Guevara que teoriza sobre la guerrilla urbana, el tema no era nuevo.

Si los costos son bajos, cualquier grupo puede armarse de ciberguerra y hacer terrorismo. El ciberespacio es un dominio, tanto físico como virtual, muy complejo. Allí se desarrollan actividades de creación, procesamiento, almacenamiento, intercambio y visualización de información usando software y hardware interconectado. Lo constituyen tanto la Internet como todas aquellas redes aisladas que se utilizan con finalidades particulares.

A diferencia de los dominios terrestre, aeroespacial y marítimo —donde se pueden desarrollar operaciones militares— el ciberespacio, es de características arti-

³³La Weekend Reads, revista militar especializada, publica en su número del 31 de octubre 2020 *Military artificial intelligence*, donde hay artículos sobre el tema desde el miedo a futuros terroristas con sistemas inteligentes hasta el programa del ejército norteamericano de uso de la lectura (o averiguación) de la mente.

³⁴David Kelcullen, *The Accidental Guerrilla: Fighting Small Wars in the Midst of a Big One*, Chicago, 2016 y en el más reciente. *The Dragons and the Snakes: How the Rest Learned to Fight the West*, 2020.

ficiales. Por ejemplo, en los espacios tradicionales, aún si se interrumpiera el suministro de energía eléctrica o nadie estuviera en esos lugares, ese ambiente seguiría existiendo. Hemos consultado un militar que se ocupa del tema y sostuvo que los ciberataques son prácticamente diarios³⁵.

Cambian las armas que se pueden usar y las despreocupaciones morales en aras de un resultado. *Nec novum sub sole*. Lo que tal vez falte y es encajar todo esto en el tema general de la guerra y sus grandes tratados: Carl von Clausewitz, Sun Tzu, Robert McNamara, para citar 3 excepcionales.

c. Dando la cara

Los métodos de reconocimiento facial dedicados a vigilancia masiva tuvieron que esperar hasta la segunda década del siglo XX, en parte porque los primeros intentos mostraron que los resultados estaban lejos de ser los deseados. La primera vez que se utilizó en un acontecimiento de grandes proporciones fue en la Super Bowl de 2002, que resultó, en general, un fracaso. Numerosos falsos positivos mostraron que la tecnología no estaba aún preparada para las grandes multitudes. A partir del 2015, las fuerzas policiales del Reino Unido empezaron a probarla también para eventos públicos en vivo, pero un informe realizado por Big Brother Watch encontró que estos métodos seguían devolviendo resultados hasta un 98% imprecisos.

Una de las trabas para que funcione correctamente entre grandes multitudes puede ser dificultad de obtener una imagen de calidad. 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 dicha página se especifica que “lo ideal sería disponer de una fotografía de pasaporte conforme a la norma OACI, ya que se trata de una imagen frontal completa de la persona con iluminación homogénea en el rostro y un fondo neutro”.

Pero 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

³⁵ Aníbal Luis Intini, jefe del Comando de Ciberdefensa del Estado Mayor Conjunto argentino.

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.

En los últimos meses, las controversias sobre su uso han aumentado, impulsadas en parte por el movimiento Black Lives Matter³⁶, 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 EEUU 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.

d. Las noticias falsas

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 informacional, 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*)³⁷ 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 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?³⁸

Estas preguntas son las que Floridi intenta responder con lo que él llama “ética de la información global”. Todo aquello que soporta la segunda. Ética multicultural y una sociedad viva en las redes que trata de evitar el sufrimiento, así como toda forma de destrucción de los entes, puesto que todo ente, por el hecho mismo de ser, tiene derecho a una forma específica de respeto. *Ens et bonum convertuntur*, afirmaba el adagio escolástico. Floridi llama a esta ontología mínima “ontocéntrica”, siendo aparentemente más radical que, por ejemplo, la “biocéntrica” y la “antropocéntrica”.

³⁶ Movimiento surgido en EEUU frente a abusos contra la comunidad afroamericana y que significa literalmente “la vida negra importa”.

³⁷ Floridi, Luciano, *Sextus Empiricus, The Recovery and Transmission of Pyrrhonism*, Oxford, Oxford University Press, 2002.

³⁸ Floridi, Luciano, *The Blackwell Guide to the Philosophy of Computing and Information* (a cura di) Oxford, Blackwell, 2003.

La ontología ontocéntrica coloca en su centro no al agente sino al paciente de la acción, incluyendo los entes no vivientes, extensión que constituye una tesis muy loable y de gran relevancia ecológica actual, pero de difícil aplicación en los casos concretos. Floridi se opone tanto a una teoría metafísica que afirme algo respecto del ser de los entes –lo que sería una forma de “imperialismo ontológico”– como a un mero relativismo que es incapaz de promover una interacción eficaz a nivel global en referencia a problemas que afectan a todas las culturas. Esta “ética de la información global” no quiere imponer una jerarquía de valores comunes, sino permitir que estos se expresen en situaciones concretas con toda su situacionalidad.

Floridi opta por una ontología ligera y horizontal (*lite, horizontal ontology*) como condición de posibilidad de interacción pragmática entre culturas, las cuales, vistas en su densidad vertical o gruesa (*thick cultures*), son a menudo irreconciliables. La diferencia entre culturas ligeras y gruesas puede relacionarse con la distinción del filósofo estadounidense Michael Walzer entre argumentos éticos finos (*thin*) y gruesos (*thick*), según se los analice o no en su profundidad cultural³⁹.

Esta posición de Floridi es muy atractiva y útil a nivel pragmático, pero tiene el peligro de ignorar la necesidad de un análisis intercultural grueso o vertical.

Según el ético holandés Philip Brey, un diálogo ético intercultural tiene que tomar en serio las diferencias interculturales. Brey usa el concepto de ética de la información en el sentido ordinario de cuestiones ético⁴⁰. Desafíos teóricos y prácticos de la ética intercultural de la información y casos relacionadas con las TIC pero incluyendo también la ética de la computación (*computer ethics*), así como la de los medios de comunicación de masas (*media ethics*) y el periodismo.

Hay que distinguir entre un relativismo moral descriptivo y uno normativo llamado también meta ético. Este último, aplicado a la ética de la información, encara la pregunta de si existen conceptos y principios con validez universal o bien si la ética de la información es culturalmente relativa. Brey defiende la necesidad de un relativismo descriptivo en la ética de la información, dado que, si estas diferencias no existen, las discusiones en torno al relativismo se volverían abstractas.

La forma en la que se utilizan los sistemas cerrados no está contribuyendo a proporcionar una plataforma multicultural para el debate, sino que más bien permite que se restablezcan las viejas morales –más locales– de comunicación entre individuos. Por otro lado, los sistemas abiertos como The Well (y en cierta medida, Messenger de MSN), permiten que la gente sea más libre de discutir los temas que ellos mismos escojan, a su modo y con la profundidad que decidan.

En cualquier caso, nadie puede entender cómo serán utilizados realmente dichos sistemas por los usuarios finales, ya que –como se argumentará más adelante– su uso dependerá de factores que involucran a un individuo o a un grupo muy limitado de personas, hablamos de ética privada. Un ejemplo de ética a este nivel es

³⁹ Walzer, Michael, *Pensar políticamente*, Barcelona, Paidós, 2010.

⁴⁰ Brey, Philip - Søraker, Johnny H., *Philosophy of technology and engineering sciences*, Elsevier, 2009.

el aborto, cuando se debe tomar una decisión en torno a la supervivencia de un individuo en particular.

La responsabilidad corresponde únicamente a ese individuo que deliberadamente decide dentro de ese dilema; el escrutinio sobre dicha deliberación es realmente muy problemático, ya que son aquellos que analizan el dilema desde fuera quienes deben ponerse en los zapatos del otro individuo y conocer cada uno de los hechos y las contingencias emocionales en las que el otro se encuentra. Por otro lado, la dimensión pública de la ética involucra más a muchas personas y no solo grupos de ellas. De hecho, las decisiones públicas tienen un impacto generalizado no solo para las generaciones presentes, sino muy probablemente para generaciones futuras también.

Las decisiones públicas dependen de valores generales mediante los cuales queremos, por así decirlo, diseñar una buena sociedad. La discusión en torno a estas decisiones (p.ej., Facebook) son instancias de lo que Karl Popper llamó “sociedades abstractas”. De acuerdo con Popper: podríamos concebir una sociedad en la cual las personas prácticamente nunca se conozcan cara a cara, donde todas las transacciones de negocios sean llevadas a cabo por individuos aislados que se comunican a través de escritura mecanografiada o a través de telegramas, y que se desplazan en autos motorizados⁴¹.

Hay aspectos moralmente negativos muy fuertes en torno a esto, pero Popper ve también los beneficios: por ejemplo, la posibilidad de nuevas relaciones personales, independientes del lugar de nacimiento o de su ubicación.

Popper entendía que las sociedades modernas son en muchos aspectos sociedades abstractas; sin embargo, la realización más próxima a dichas sociedades es posible solo a través de Internet. Hay una diferencia fundamental entre los Sistemas Mediados por Computadora, tales como The Well y Facebook. La primera es un tipo de sistema construido de abajo hacia arriba, una sociedad cerrada se caracteriza por una jerarquía rígida inmutable, de hecho.

En ese mismo sentido las redes sociales diseñadas de arriba hacia abajo (Facebook y Twitter) son como las sociedades cerradas: al parecer no tienen reglas compartidas, que tienen las mismas funciones que los tabúes que son incuestionables puesto que son impuestos por quienes diseñaron el sistema. Sus valores fundamentales, basados en el credo dominante del fundamentalismo de mercado y el interés personal como bendición moral se refuerzan con la red: no es posible una apertura hacia otras culturas o individuos.

Las sociedades abiertas son exactamente lo contrario a las sociedades cerradas; son sociedades sensibles y tolerantes, donde los mecanismos y procedimientos se dice que son transparentes, flexibles y siempre negociables. Los países occidentales son sociedades abiertas, Facebook es una herramienta exitosa para el cambio, pero peligrosa por la dependencia que crea.

⁴¹Karl Popper, *Conjeturas y Refutaciones: el Crecimiento del Conocimiento Científico*, revisada por el autor en 1972, Barcelona, Paidós, 1981, Título original: *Conjectures and Refutations. The Growth of Scientific Knowledge*, publicado en inglés por Routledge & Kegan Paul, 4ª ed, Londres, 1972.

Con la promesa de libre consumo, Facebook ha ayudado a países cerrados y subdesarrollados –como los del norte de África– a deshacerse de sociedades jerárquicas y pasadas de moda mediante el cambio de una sociedad tribal a una basada en el individuo. Esa es la principal razón por la cual en países desarrollados y corruptos –aunque sean sociedades abiertas– Facebook no ha funcionado; estas ya eran sociedades individualistas, líquidas (en el sentido explicado por el filósofo social Zygmunt Bauman)⁴² en las cuales el código moral está dominado por los valores del mercado. En estas sociedades, únicamente herramientas comunitarias desarrolladas de abajo hacia arriba –como The Well– tienen más probabilidades de funcionar. Internet, en otras palabras, nos libera de los impedimentos de las sociedades a las que pertenecemos, es posible que sea por esta misma razón que podamos escapar de la influencia nefasta de este mismo ideal. Con el fin de obtener una comprensión cultural y moral real entre las personas de todo el mundo, necesitamos un cambio comunitario, posible gracias a sistemas como The Well, para que ayude cambiar realmente los sistemas de comunicación enfocados en el individuo por sistemas de colaboración.

Es interesante la resolución de octubre del 2020 de las autoridades escolares de Cataluña erradicando de las escuelas el uso de Facebook y otras redes sociales por no tener suficiente control de las noticias que en ellos aparecen. Y esto rige ya para el año lectivo en curso.

e. ¿Y el trabajo?

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.

“¿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”.

⁴² Zygmunt, Bauman, *Modernidad líquida*, México, Fondo de Cultura Económica, 2003.

No es sólo el mercado el que puede decidir. En esencia, a menudo vale la pena que los empresarios y las empresas innoven (también) en términos de costos, mucho menos para los trabajadores que se quedan en casa, pero más generalmente para la sociedad en su conjunto, que tiene que intervenir con servicios públicos y dinero.

En los últimos decenios, aunque el crecimiento económico general ha continuado, la distribución de los beneficios económicos ha sido cada vez más desigual. En los Estados Unidos, la mitad inferior de la población, compuesta principalmente por trabajadores poco cualificados, apenas ha experimentado un aumento de los ingresos después de la inflación. Durante el mismo período, los ingresos reales de los ricos se duplicaron, los de los megamillonarios se triplicaron y los de los multimillonarios se cuadruplicaron.

A pesar de las diferentes tesis desde ambas perspectivas, los economistas y los éticos de la innovación pueden estar de acuerdo en varios puntos. En primer lugar, es conveniente garantizar que los trabajadores que pierden su empleo reciban ayuda no sólo en términos monetarios, por ejemplo, con medidas contra el desempleo, sino también en términos del valor más amplio que la sociedad atribuye a sus pérdidas. Después de todo, estos trabajadores han sido sacrificados en aras de la eficiencia económica: cuando sus puestos de trabajo fueron trasladados o cancelados, fueron el daño colateral para permitir que la economía adoptara procesos de producción más eficientes. Esas ayudas no económicas comienzan con formación para los nuevos desocupados y su familia, el acceso a lugares de lectura y recreación, el acceso a una alimentación sana y sobre todo la cobertura médica mínima para su salud y la de su familia.

f. Los periódicos

Este año el ganador del premio más interesante es en la categoría “innovación tecnológica al servicio del periodismo”. Ganó el proyecto realizado por un periódico canadiense, The Globe and Mail, cuyo sitio web ha sido dirigido por un sistema inteligente para elegir los artículos durante algún tiempo y los lectores nunca se han quejado de ello; juran, de hecho, que ni siquiera lo notarían. Digámoslo mejor: la disposición del 99% del contenido del Globe and Mail la decide Sophi, un motor de inteligencia artificial que lo actualiza según lo que hacen los lectores en tiempo real. Puede parecer el fin del periodismo y la victoria de los ordenadores, pero sólo puede parecerlo en una lectura muy superficial.

Sophi no escribe nada, simplemente se ocupa de la gestión de contenidos, de la administración de contenidos; ha sido capacitada durante algunos años por los periodistas de Globe and Mail para comprender qué temas son pertinentes, qué temas no son adecuados para la primera página, cuándo un artículo no reciente merece ser sacado, e incluso qué mezcla debe haber en las fotos de cada pantalla para reflejar los diferentes grupos étnicos de Canadá utilizando un software de reconocimiento facial. Los periodistas no han contribuido a que este sistema sea despedido sino, por el contrario, a hacer un mejor producto y al mismo tiempo tener más tiempo para hacer encuestas o entrevistas o análisis de calidad. Obviamente también en Globe and Mail hay un verdadero editor, que se reserva el derecho de elegir sólo el

puesto principal: su nombre es David Walmsley y cree que el trabajo de los periodistas es encontrar y escribir grandes historias, el resto lo pueden hacer los algoritmos.

Esta innovación no viene por casualidad. En 2015, cuando la crisis se agudizaba, el editor, en lugar de recortar o impulsar supuestos contenidos virales, creó un laboratorio de innovación, el Lab351, con la tarea de apoyar las mejores innovaciones propuestas por los propios periodistas; para ello, además del dinero para cada proyecto, contrató a ingenieros informáticos y analistas de datos. “Tienes que gastar dinero si quieres ganar dinero” es una de las frases que repiten en el Globe and Mail cuando cuentan su historia de éxito. Pero tienes que gastarlo bien. Incluso en el periodismo, la tecnología no es un enemigo, no es una alternativa a los humanos. Bien utilizada, la tecnología debe darnos tiempo para volver a ser humanos.

Obviamente Sophi está sujeta a todos los riesgos de sesgos que significa preparar un sistema inteligente para que actúe en tareas humanas, más en esta tan delicada de elegir las noticias y diferenciarlas en principales y accesorias. Ya hay un problema de noticias falsas (*fake news*). Los periodistas deben cubrir o descubrir cuáles son los sesgos que provoca la recolección de datos de preferencia sociales y decidir si deben ser corregidos y como.

5. 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.

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.

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.

a. 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 evento⁴³.

b. 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 Conducta⁴⁴.

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”.

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

⁴³ www.wired.it/attualita/tech/2020/02/28/intelligenza-artificiale-etica-vaticano-microsoft-ibm/?refresh_ce=

⁴⁴ 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)).

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.

Enumeración de los principios sancionados por el Parlamento Europeo:

- Proteger a los humanos del daño causado por robots: la dignidad humana.
- Respetar el rechazo a ser cuidado por un robot.
- Proteger la libertad humana frente a los robots.
- 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.

c. 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 auditable 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.

d. 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 en París los Principios de la OCDE sobre la Inteligencia Artificial en el marco de la Reunión del Consejo de Ministros 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, Angel 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 presentarlos⁴⁵.

⁴⁵www.oecd.org/centrodemexico/medios/cuarentaydospaísesadoptanlosprincipiosdelaocdesobreinteligenciaartificial.htm#:~:text=Cuarenta%20y%20dos%20pa%C3%ADses%20adoptan%20los%20Principios%20de%20la%20OCDE,Par%C3%ADs%20mayo%202022%20de%202019.&text=%C2%ABLa%20inteligencia%20artificial%20est%C3%A1%20revolucionando,a%20nuestras%20sociedades%20y%20econom%C3%ADas.

e. 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 40° 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 2021⁴⁶.

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 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.

f. 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 • diver-

⁴⁶ <https://es.unesco.org/news/unesco-da-gran-paso-primer-instrumento-normativo-etica-ia>.

idad, 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 creamos⁴⁷.

g. Marco ético para la inteligencia artificial en Colombia

Es interesante ver lo que establece el marco del país latinoamericano sobre el tema: habla de riesgos del uso de sistemas inteligentes, principios que son más o menos los que ya hemos expuesto y herramientas entre las cuales aparecen auditoría de algoritmos, limpieza de datos, evaluación de la legitimidad, definición y gestión de riesgos, modelos de gobernanza para asegurar la ética en el uso de la IA. No es particularmente original, pero lo ponemos a la atención porque es el primero en América Latina en ocuparse en forma sistemática del tema.

6. Economía y ética

Siempre hay una justificación económica para un desastre social, pero los economistas serios no están de acuerdo. Lo que no se puede negar es que el índice Gini ha tocado los límites más altos⁴⁸.

Observando la miseria creada por la pérdida de puestos de trabajo, la ética puede verse tentada a responder “no” a la pregunta formulada supra. Pueden ver los daños y sufrimientos tangibles infligidos a los trabajadores despedidos y observar que no es ético imponérselos a los trabajadores, quienes no pueden apreciar los efectos a largo plazo del progreso económico en el bienestar humano.

⁴⁷ Los 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 humano. 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.

⁴⁸El coeficiente de Gini es una medida de la desigualdad ideada por el estadístico italiano Corrado Gini. Normalmente se utiliza para medir la desigualdad en los ingresos, dentro de un país, pero puede utilizarse para medir cualquier forma de distribución desigual.

No sólo eso, cuando los mercados son tan perturbadores, puede ser socialmente indeseable que sean la única guía de la toma de decisiones humanas. Por eso hay que hacer compromisos entre la ética y el mercado.

Las pérdidas de puestos de trabajo causadas por la automatización suelen reducir la demanda de mano de obra y los ingresos de los trabajadores, y pueden aumentar las ganancias de los empresarios y los accionistas. Los resultados pueden generar una distribución más desigual de los ingresos, que la sociedad percibe como menos deseable. “No hay ninguna razón teórica para creer que el libre mercado dirigirá los esfuerzos de innovación hacia las innovaciones más deseables socialmente”, señalan Zachary, Bthune y Korinek⁴⁹.

El primer teorema fundamental de la riqueza en la economía, comúnmente llamado “teorema de la mano invisible”, establece que, bajo ciertas condiciones idealizadas, el mercado generará una distribución eficiente de los recursos existentes en la economía. Sin embargo, este teorema no se aplica al progreso tecnológico, por lo que el mercado puede impulsar la innovación en la dirección equivocada.

Si las innovaciones generan importantes perturbaciones sociales que la “mano invisible” del mercado no puede manejar, entonces los innovadores tienen el deber moral de internalizar los efectos perturbadores que generan; las cuestiones de la eficiencia del sistema y la distribución de los ingresos no pueden separarse. Sin embargo, es demasiado común que los perdedores del progreso tecnológico sean abandonados a su suerte, o que reciban un apoyo mínimo de los ganadores o de la sociedad para compensar sus pérdidas.

Los economistas y los éticos de la innovación pueden estar de acuerdo en varios puntos. Vale la pena garantizar que los trabajadores que pierden su empleo reciban ayuda no sólo en términos monetarios, sino también con políticas contra el desempleo y atendiendo a las exigencias del desempleado y su familia que exceden el campo económico. Preservar a los trabajadores y sus familias es una forma de evitar daños colaterales del mejoramiento de la producción de bienes con nuevas tecnologías. El mejoramiento de la vida, en general, no puede, no debe traer aparejado exclusivos reducidos a la miseria. Es una película ya muchas veces vista que hay que evitar.

7. Siempre teniendo cuidado

Si las normas son demasiado débiles, existe un riesgo demasiado grande de que nuestros derechos y libertades se vean socavados: Esto se aplica actualmente a todas las aplicaciones de la inteligencia artificial, que hasta ahora sólo se han basado en principios y valores éticos no vinculantes. Con esta nueva legislación, Europa tiene la posibilidad de adoptar un marco jurídico para la IA con normas claras. Necesitan instrumentos sólidos para proteger los derechos fundamentales y la democracia.

⁴⁹Bethune, Zachary A. - Korinek, Anton, *Covid-19 Infection Externalities: Trading Off Lives vs. Livelihoods*, NBER Working Papers 27009, National Bureau of Economic Research Inc., 2020.

En los últimos años, los gobiernos de todo el mundo se han ocupado de crear comités, consejos y grupos de expertos especiales para debatir la ética de la inteligencia artificial: el Grupo de Expertos de Alto Nivel sobre Inteligencia Artificial nombrado por la Comisión Europea, el Grupo de Expertos sobre la IA en la Sociedad de la Organización de Cooperación y Desarrollo Económicos (OCDE), el Consejo Asesor sobre el Uso Ético de la Inteligencia Artificial y los Datos de Singapur y el Comité Selecto sobre Inteligencia Artificial de la Cámara de los Lores del Reino Unido o, en los Estados Unidos, la administración Obama elaboró una Hoja de Ruta para la Política de la IA.

La ciudad de Nueva York, que fue una de las primeras en crear un consejo especial llamado “Grupo de Tareas del Sistema de Decisión Automatizada”, tardó más de un año en acordar incluso una definición de la toma de decisiones automatizada. Hasta la fecha, todavía no ha logrado elaborar un panorama general de todos los sistemas de IA que se utilizan en la ciudad.

Al mismo tiempo, muchas empresas han aceptado con gratitud el libre paso a la autorregulación estableciendo juntas de ética, escribiendo directrices y patrocinando investigaciones en temas como el sesgo algorítmico o la “imparcialidad en la inteligencia artificial”. No hace falta molestar a Marshal McLuhan para saber que el medio condiciona el mensaje, desde ese punto de vista no hay medios neutros.

El debate sobre el cuestionamiento moral a la acción de megaempresas que manejan el 75% de los datos que hoy puede ser mucho y ha sido usada como pantalla para distraer la atención del daño que están infringiendo a la sociedad y de los extraordinarios beneficios que obtienen con prácticas monopólicas.

En los últimos años, el sector empresarial ha estado construyendo una “lavadora de ética” tras otra. Facebook ha financiado a los especialistas en ética de la IA en la Universidad Técnica de Múnich en Alemania, mientras que Google, SAP y Microsoft han adoptado principios rectores y códigos éticos.

Sin cuestionar estas supuestas buenas intenciones, es obvio que el enfoque en los debates éticos ha sido desde hace mucho tiempo una manera eficiente para la industria de ganar tiempo y evitar una dura regulación. Con demasiada frecuencia, confiar en la ética y la autorregulación ha sido insuficiente para que las empresas no rindan cuentas, si no se dispone de una aplicación estricta y de mecanismos de supervisión independientes proporcionados por un marco institucional.

En los últimos años hemos visto abundantes pruebas que ponen de relieve el hecho de que nuestras leyes actuales son insuficientes para proteger contra la discriminación, cuando la carga de la prueba sigue recayendo en la víctima, que puede ni siquiera ser consciente de que está siendo discriminada por un algoritmo, y cuando la responsabilidad por los daños no siempre puede establecerse a lo largo de la compleja cadena de suministro de un sistema de inteligencia artificial.

Las personas que tienen más probabilidades de ser discriminadas por los sistemas de inteligencia artificial también tienen más probabilidades de no tener los medios financieros o la confianza en sí mismas para entablar una demanda de resul-

tado incierto. Las herramientas automatizadas de toma de decisiones pueden a menudo exacerbar el racismo y el sexismo anclado en nuestras sociedades⁵⁰.

Piensen en el sistema COMPAS que se utilizó en los Estados Unidos para predecir si los acusados volverán a cometer delitos –se descubrió que discriminaba a los acusados negros–. En un estudio reciente del Instituto Alan Turing de Londres y la Universidad de Oxford, un equipo de investigación demostró que las leyes actuales en Europa son insuficientes para proteger a las personas contra el daño causado por los algoritmos defectuosos.

Hay muchos ejemplos que ilustran que un enfoque de derecho no vinculante a través de la ética, la autorregulación y la responsabilidad social de las empresas fracasa drásticamente casi siempre. Por ejemplo, The Intercept reveló en 2018 que Google estaba desarrollando una versión censurada de su motor de búsqueda para el gobierno chino, en violación directa de sus principios éticos de IA.

No será fácil para los responsables políticos europeos redactar una ley que vaya al grano, que contenga medidas específicas, que siga el ritmo de la tecnología y que no cause daños colaterales a nadie.

Pero este desafío también es una oportunidad en Europa. Tienen la oportunidad de ser el primer continente en poner a la humanidad en el centro de las políticas digitales y tener en cuenta que poderosos intereses intentan impulsar la “agenda ética”. Necesitamos derechos fuertes y ejecutables para los usuarios de los futuros sistemas de inteligencia artificial porque tienen una influencia cada vez mayor en todos los aspectos de las vidas de sus ciudadanos.

Por ello, el proceso legislativo debe ser lo más participativo e inclusivo posible. La Comisión de la UE acaba de celebrar una consulta pública con el fin de preparar su propuesta para el enfoque europeo de la IA. En teoría, los ciudadanos podrían responder, pero el proceso de consulta de la Comisión de la UE no es, lamentablemente, muy accesible, debido a que se le pide que cree una cuenta antes de poder responder a un cuestionario que contiene una larga lista de preguntas muy técnicas.

Por consiguiente, los encargados de la formulación de políticas de la UE deberían encontrar otros foros y hacer participar proactivamente a los grupos más afectados por la IA y con menos probabilidades de sentarse a la mesa donde se toman las decisiones. Esto también ayudaría a aumentar la transparencia y a crear confianza en la toma de decisiones de la UE.

Sólo así podrán asegurarnos de que adoptan normas que hagan aplicables derechos y libertades en virtud de los Tratados y la Carta, en lugar de adoptar simplemente una orientación ética en nombre de “principios y valores”.

Es de señalar la existencia de una roboética, primero como disciplina universitaria y ahora prevista directamente por la Unesco. Es curioso, pero a las 3 leyes

⁵⁰ Ver el informe de Juan G. Corvalán y Enzo María Le Fevre Cervini, *¿El sistema de IA discrimina las personas?* Hecho en el Laboratorio de Ialab, www.ialab.com.ar, IA confiable <https://ialab.com.ar/wp-content/uploads/2020/07/Evaluacion-IA-confiable-.pdf>.

robóticas de Asimov se agregó una cuarta: “El robot debe siempre respetar la continuidad de la especie humana, aunque esto colisione con las tres reglas anteriores” y esto por dos razones: la primera es que algunos temen que la inteligencia artificial supere a la natural y la segunda porque, de cualquier modo, en un mundo de mucha tecnología es mejor defender y proteger a los seres humanos que vendrá, o sea defendamos la especie. Que es una ley de Darwin adaptada a los desarrollos tecnológicos.

Se van perfilando éticas consecuencialísticas, basadas en resultados, que tienden a proteger el bienestar de los usuarios y éticas deontológicas basadas en los procesos de toma de decisiones donde el valor moral es el parámetro para decisiones individuales o colectivas que atañen el bienestar de los demás.

Una cuestión fundamental es si los procesos de decisiones se pueden trasladar en todo o en parte a sistemas algorítmicos que usan miríadas de datos. La idea que se va imponiendo es no limitar el uso de sistemas inteligentes sino calcular, detectar y perfeccionar el modo de prever los riesgos, encontrar formas universales de calificarlos y medirlos y mitigarlos a través de un uso serio y eficaz de sistemas de derechos humanos.

Los sistemas de inteligencia artificial deberían empoderar a los seres humanos y ser resistentes y seguros. Capaces de asegurar respeto total por la privacidad y la protección de datos, y mecanismos adecuados de gobierno de datos. Los modelos de negocio de datos, los sistemas expertos y la IA deben ser transparentes y garantizar diversidad, no discriminación y equidad, evitando el sesgo injusto. También es necesario que los sistemas de IA beneficien a todos los seres humanos.

El desarrollo tecnológico ha dado un gran salto adelante en este siglo, pero no ha sido seguido por su desarrollo social, político e informativo. Las fantasías, por un lado, la ignorancia no inocente y los miedos atávicos, por el otro, están dando una considerable contaminación de información. Todas nuestras vidas se ven afectadas por la IA, desde los teléfonos móviles hasta los vuelos espaciales, pero cinco compañías poseen el 75% de todos los datos del planeta. Nos encontramos con nuevas situaciones y no tenemos ni ética ni cultura, ni siquiera normas para tratarlas. Y es urgente hacerlo.

8. Ética y estética

Siempre pensé que había una relación entre ética y estética, pero siempre me resultó difícil explicar por qué. Recuerdo la anécdota que contaba Giménez de Azua, era que cuando el régimen había echado al profesor de estética el de ética dijo “entonces yo también debo irme”.

Revisando el *Tractatus* de Wittgenstein en la proposición 6.421 dice “ética y estética son la misma cosa”⁵¹. Este argumento tiene dos partes. La primera es nega-

⁵¹ Ludwig Wittgenstein, *Tractatus logico-philosophicus* proposición 6.421 “Está claro que la ética no resulta expresable. La ética es trascendental” (Ética y estética son una y la misma cosa.) *Ethik und Aesthetik sind Eins*.

tiva, y se desprende directamente de la operación de someter el lenguaje humano al canon de esa suerte de lenguaje ideal que se exhibe a través de la semántica que Wittgenstein pone en marcha en el *Tractatus*, la horma de su atomismo lógico, mientras que la segunda, de carácter afirmativo, es deudora de una distinción fundamental que subyace al tratamiento del complejo lingüístico, la distinción entre decir (*sagen*) y mostrar (*zeigen*)⁵².

Con respecto a “trascendental” parece que Wittgenstein se remite a la idea de Kant en la *Crítica de la razón pura*, cuando ya en la Introducción dice textualmente: “Llamo trascendental todo conocimiento que se ocupa, no tanto de los objetos, como de nuestro modo de conocerlos, en cuanto que tal modo ha de ser posible a priori” (A12/ B25).

Pero no solo los filósofos del pasado nos ayudan en esta tarea, también algunos del presente como Marcus Gabriel, quien sostiene que si bien la reducción puramente cibernética de la mente humana es rechazable por muchas razones (por ejemplo, la tontería de que nuestro cerebro nos manipula y por tanto no hay libre albedrío), también lo es minimizar la importancia del cerebro para la mente. Basta con reparar en las devastadoras consecuencias de la carencia de empatía, del Alzheimer o de la transmisión hereditaria de patologías como la esquizofrenia⁵³.

“El arte es aquella super inteligencia temida por los críticos de las tecnologías digitales”⁵⁴ y defiende la de que hoy es el arte quien controla al poder y no al revés. Según él vivimos en una época estética, por lo que se encuentra presente en todas partes, como herramienta de persuasión o disfraz; fusionado con el diseño, en la publicidad y en las gráficas de las páginas web, películas, ropa.

El catalán Alex Hinojo formula críticas al tecno consumismo y dice que con la multiplicidad de fuentes lo que hemos logrado son las noticias falsas *fake news*, mientras que del lado de la producción se ha mantenido un nivel y una moral gracias al pedido de trazabilidad “La ética ha exigido trazabilidad a los sistemas de producción. Hoy, cualquier producto del supermercado incorpora un número de lote que nos permite, en caso de incidencia, detectar su origen. En este caso, la tecnología ha servido para mejorar la calidad de los sistemas de producción y de distribución, así como del servicio al consumidor”⁵⁵.

En su libro *On the Existence of Digital Objects*⁵⁶ el chino Yuk Hui, se apoya en sus antecedentes como ingeniero en computación y programador para investigar entidades digitales como virus de computadora, videoclips, algoritmos y redes. En el prólogo de dicho libro, el filósofo francés Bernard Stiegler describe el pensamiento

⁵² Ver también la cita de la carta a Ludwig von Fricker puesta en el acápite.

⁵³ Markus, Gabriel, *Yo no soy mi cerebro. Filosofía de la mente para el siglo XXI*, Barcelona, Pasado y Presente, 2016.

⁵⁴ Markus, Gabriel, *El poder del arte*, Bs. As., Roneo, 2019.

⁵⁵ Hinojo, Alex, *Hacia una nueva ética informativa*, CCCBLab, 16/1/20, <http://lab.cccb.org/es/hacia-una-nueva-etica-informativa>.

⁵⁶ Yuk Hui, *On the Existence of Digital Objects*, vol. 48, Minnesota, Electronic Mediations, 2020.

de Hui como un “generoso y abierto entorno teórico para la exploración de la experiencia humana en conexión con la infoesfera”.

En este libro Hui, realiza un examen filosófico de los objetos digitales y su esquema de organización creando un diálogo entre Martin Heidegger y Gilbert Simondon, que Yuk Hui contextualiza dentro de la historia de la informática. ¿Cómo pueden entenderse los objetos digitales según la individualización y la individuación? persigue esta pregunta a través de la historia de la ontología y el estudio de los lenguajes de marcas y las ontologías de la web; investiga la estructura existencial de los objetos digitales dentro de sus sistemas y entornos. Con este enfoque relacional hacia los objetos digitales y los sistemas técnicos, el libro aborda la alienación, descrita por Simondon como la consecuencia de ver erróneamente las técnicas en oposición a la cultura. Y llega a la cosmotécnica.

Y hablando de la cosmotécnica “Para los griegos, ‘cosmos’ significa un mundo ordenado. Al mismo tiempo, el concepto apunta a lo que está más allá de la Tierra. La moralidad es ante todo lo que concierne al reino humano. La cosmotécnica, como yo la entiendo, consiste en la unificación del orden moral y el orden cósmico a través de actividades técnicas. Si comparamos Grecia y China en la antigüedad, descubrimos que tienen una comprensión muy diferente del cosmos y también concepciones muy diferentes de la moralidad. El arbitraje entre ellas tiene lugar de diferentes maneras con diferentes tecnologías. Una cosmotécnica del tipo de la tianxia no es posible en un tiempo que ya no tiene una concepción de ‘Cielo’, como la tuvo la gente del pasado. Como otras grandes naciones, China tiene satélites orbitando la tierra. Los cielos se han convertido en un lugar secular, utilizado por los humanos, y no pueden jugar un papel como poder legitimador moral”⁵⁷.

Hay relaciones entre la estética y sistemas inteligentes: artistas que, gracias a sus conocimientos matemáticos e informáticos y a su relación con los grandes centros de computación de Google, NVIDIA, etc., exploran las posibilidades estéticas y creativas que ofrece la IA e incluso crean sus propios algoritmos, produciendo obras originales, a menudo muy escénicas. Este es el caso de la *Historia latente* de Refik Anadol presentada en @fotografiska en Estocolmo, para la cual el artista ha reunido un millón de datos fotográficos de materiales de archivo: “A través de la exploración de las memorias fotográficas de los últimos 150 años, esta exposición pretende investigar y reimaginar la memoria colectiva, las capas ocultas de la historia y la conciencia de una ciudad [Estocolmo] que de otra manera podría ser reinterpretada mal o quedar a oscuras”⁵⁸.

La inteligencia artificial juega un papel crucial en el ecosistema cultural mundial. Aconseja lo que debemos ver, oír, leer y comprar. Decide cuántas personas verán nuestro contenido compartido. Nos ayuda a tomar decisiones estéticas cuando creamos contenido digital. En la producción cultural, la IA se ha utilizado durante

⁵⁷ Edwin Lo, *Entrevista: sobre Tecnodiversidad: una conversación con Yuk Hui*, 27 de julio de 2020, Seminario de Tecnologías Filosóficas, <http://philosophyandtechnology.network/3939/entrevista-sobre-technodiversity-una-conversacion-con-yuk-hui/>.

⁵⁸ Refik Anadol: “*Melting Memories*” *esculturas digitales en movimiento*, exposición en San Francisco, 1019.

mucho tiempo para producir trailers de películas, álbumes de música, artículos de moda, diseño de productos y de páginas web, arquitectura, etc. Lev Manovich es un creador y al mismo tiempo un crítico que se ocupa de los límites que pone la estética a los sistemas inteligentes⁵⁹, nos presenta un marco sistemático para ayudarnos a pensar en los usos culturales actuales y futuros de la IA. Y nos ofrece la posibilidad analítica de entender los medios, el diseño y la estética en la era de los sistemas inteligentes.

El arte se está planteando temas de conocimiento y como las nuevas cogniciones llevan a otra estética o por lo menos a una estética que tiene en cuenta los sistemas inteligentes y va observando sus límites. Algo que hace Gianluca Consoli y escribe sobre estas experiencias⁶⁰.

La inteligencia artificial con los actuales sistemas de aprendizaje automático y aprendizaje profundo, está ganando cada vez más interés por parte de los artistas que se dedican a investigar las relaciones entre el arte, la ciencia y la tecnología, por lo que importantes centros de exposición también están proponiendo y discutiendo sus obras: Entre los que están en evidencia, el Barbican Center de Londres presenta AI: 1) More than Human, 2) el MAK de Viena Künstliche Intelligenz & Du, 3) la Kate Vass Galerie de Zurich Automat und Mensch, 4) la Transfer Gallery de Nueva York Forging the Gods y, finalmente 5) el HEK de Basilea Entangled Realities, Living with Artificial Intelligence.

Es muy difícil obtener una idea acabada de lo que está sucediendo dado que del conjunto de propuestas es difícil, pero generalizando, sin embargo, es posible identificar dos líneas de tendencia algo opuestas: de una parte los artistas usan los sistemas inteligentes para realizar obras que consideran más actuales de las que otros artistas que no usan tales medios obtienen y del otro lado muchos artistas consideran que algunos resultados de la IA son feos, poco atractivos y deben ser reformulados. En ambos bandos la ética tiene que ver con la estética.

En una entrevista que saldrá en la edición de noviembre de 2020 de Icon, Pedro Almodóvar, el cineasta confiesa que le parece “un milagro” llegar a los cines con una experiencia cinematográfica en este momento de tantas dificultades. Se refiere a la película “La voz humana” sobre el texto de Cocteau “Al principio, mi idea de la exhibición era otra. Pero después de la proyección en Venecia, los distribuidores nos llamaron para estrenarla en cines de España y del extranjero, y sólo puedo decir que es maravilloso que suceda y más aún en un momento como éste”. Y agrega: “El algoritmo me aterroriza y me horroriza”. Vaya uno a saber qué entiende Almodóvar por algoritmo.

Dicho de otro modo, arte y moral, ética y estética dialogan sobre los sistemas inteligentes y van dibujando recorridos posibles y también límites.

⁵⁹ Lev Manovich, *L'estetica dell'intelligenza artificiale: Modelli digitali e analitica culturale*, Numerus Kindle Edition, Luca Sossella editore, 2020.

⁶⁰ Consoli, Gianluca, *Arte e cognizione. Rapporti tra estetica e intelligenza artificiale*, Milano, Bulzoni, 2006.

9. Conclusiones

Sinceramente creo que los humanos tenemos una noción bastante clara de lo que está bien y lo que está mal de la vida en sociedad y además de su propia vida. Dicho de otro modo, que la moral social y la ética individual han acompañado al hombre desde su aparición en la Tierra.

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. Ciertamente, la moral de un *pithecanthropus erectico* es distinta de la del cosmonauta Yuri Gagarin, porque el mundo a su alrededor es distinto, pero si se va a excavar 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.

En el Antiguo Testamento se establecen las reglas de las religiones monoteísta y Dios tiene su oponente: Satán⁶¹. En el nuevo testamento se llamará el Diablo y se le aparece a Jesús, como tentación en el desierto.

Satán o el diablo no aparecen, por lo que se, en el budismo ni en hinduismo y hasta algunos creyentes famosos tuvieron problemas con esta figura. Leibniz⁶², cuando escribe sobre la presencia de Dios en el mundo, dice que es fácil verla: la bondad, pero cuando quiere encontrar la maldad dice simplemente que es la ausencia de Dios en el mundo. A punto tal que puede indicar lo mismo apelando a números. La presencia de Dios en el mundo 1 y su ausencia 0. Y agrega que con esos dos elementos puede nombrar cualquier cosa y comienza con la sucesión fundamental de los primeros ocho números que se convierten en una tira de 0 y 1⁶³.

Leibniz de este modo se libera del diablo en el cual no creía.

Pero sin tropezarnos con el diablo, la vida virtuosa y la maldad (o ignorancia en el budismo) parecen acompañar al hombre desde su aparición en Tierra. Es cierto que algunos autores han concebido y expresado las máximas morales en modos muy estrictos como Kant y su imperativo categórico y otros de una manera más contextualizada como David Ross con sus deberes *prima facie*. Pero ahí estamos: la moral parece no cambiar mucho, lo que cambia es el mundo y entonces debemos adecuar la ética al mundo nuevo.

Esto es lo que nos sucede con los sistemas inteligentes: personalmente no creo que haya otra moral, pero no se puede desconocer que hay otro mundo y por lo

⁶¹ *Satán* es una palabra hebrea que significa “acusador o adversario, alguien que resiste”. El término se utiliza 19 veces en el Antiguo Testamento, 14 de las cuales se encuentran en Job 1 y 2. También se menciona a *Satán* en 1 Crónicas 21:1, Salmos 109:6 y Zacarías 3:1-2.

⁶² Gottfried, Wilhelm, Freiherr von Leibniz, *Ensayos de Teodicea. Acerca de la bondad de Dios, la libertad del hombre y el origen del mal*, Madrid, Abada Editores, 2019.

⁶³ El lector dirá, “pero eso es el sistema binario” y tiene razón. A tal punto que le escribí a un jesuita que enseñaba en China sobre el tema y este le contestó alborozado que era una manera de explicar a los chinos el catolicismo pues el sistema binario coincidía con el IChin compuesto por 64 hexagramas. Ambos le escribieron al Papa quien sentenció que la religión católica debía enunciarse en latín.

tanto la ética debe explicarse y expresarse en este nuevo mundo. Un mundo nuevo acelerado por los cambios tecnológicos como nunca lo fue. Antes los cambios tecnológicos duraban algún siglo y permitían 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 inconmensurable con cualquier cambio del pasado. La intuición que la lógica era sintáctica acercó la lógica que se estaba acelerando con la matemática y con la informática⁶⁴.

Llevábamos veinticinco siglos, desde que Aristóteles sostuvo que la lógica se ocupaba de los enunciados verdaderos o falsos⁶⁵ y liberamos a la lógica de la semántica y la hicimos sintáctica como los programas de computación.

Los nuevos medios han potenciado algo que el hombre hizo siempre: imaginar realidades distintas y contarlas como un relato. Tal vez en torno a los fuegos de las cavernas, seguramente en el mito de la caverna de Platón, y se han multiplicado en modo vertiginoso.

Hoy los sistemas inteligentes están en todas partes desde del celular al dron y están transformando el modo de tratar con el banco, con la Administración pública, la Justicia, los transportes y nuestras relaciones con los otros y nosotros mismos.

No hay tiempo para reflexionar porque como bien había visto Ítalo Calvino, nuestro siglo esta signado por la velocidad⁶⁶ pero hay que actuar y ver si todo lo que se está haciendo tiene respeto por el otro, por la sociedad, por la humanidad, por el clima de la tierra.

La pandemia aceleró los procesos sociales, mostró las debilidades de los sistemas económicos, sociales y políticos; encerrando a la gente las hizo enfrentar con sus peores miedos: hambre, soledad, limitaciones, robots, carencias materiales y afectivas, pestes. Y nos privó de la esencialidad social que hablaba Aristóteles.

Se nos han venido encima los automóviles autónomos, las armas cibernéticas, *las fake news*, los reconocimientos faciales y la posibilidad de ubicar a cualquiera en cualquier lugar, la escasa defensa de nuestra privacidad de datos, las manipulaciones que se pueden hacer desde una red social o desde un sistema de voto, las manipulaciones con libros, películas y noticias, la facilidad con la cual se pueden limitar libertades con las nuevas tecnologías y con los encierros. A los viejos populismos del siglo XX se agregan los nuevos con rostro no tan duro, pero tanto o más violentos que aquellos, se deteriora la democracia en una época de crisis democrática.

⁶⁴ Alchourrón, Carlos E. - Martino, Antonio A., *Logic Without Truth*, "Ratio Juris", vol. 3, 1990.

⁶⁵ Aristóteles, *De Interpretations. In The Works of Aristotle*, Trans. and ed. W.D. Ross, vol. 1. Oxford, Oxford University Press, 1923.

⁶⁶ Calvino, Ítalo, *Lezioni americane: sei proposte per il prossimo millennio*, Milano, Oscar Mondadori, 1985. Las seis propuestas son: levedad, rapidez, exactitud, visibilidad, multiplicidad, consistencia.

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, 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. Incluso 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.

Nace una ética global, aparecen morales ontocéntricas, biocéntricas y antropocéntricas y es necesario informarse, reflexionar, tomar partido. Se tiene clara conciencia que nuestras decisiones no solo afectan nuestra vida y nuestro futuro inmediato sino también el largo plazo y las próximas generaciones y es por eso que, a las tres reglas clásicas de Asimov, se agrega una cuarta en la cual el robot debe cuidar de no afectar la perdurabilidad de la especie humana, como el mal tratamiento del ambiente o la afectación de los recursos naturales.

Se evidencian los sesgos con los cuales obramos y con los cuales creamos los sistemas inteligentes. Es fácil descubrir que cualquier regla de aprendizaje para un sistema inteligente (robótico o no) contiene las bases para hacer filtrar decisiones de mayorías prejuiciosas por las cuales es necesario hacer intervenir personas de culturas diferentes para descubrir los prejuicios anudados en la mayoría local.

Nace la identidad digital y es necesario velar sobre ella⁶⁷.

A las viejas tradiciones filosóficas se agregan las nuevas que acentúan la necesidad del multiculturalismo y la imperiosa necesidad de enfoques holísticos sin

⁶⁷El decreto del DL 16 julio 2020, n° 76 de simplificación digital italiano que comprende la carta d'identità digitale.

necesidad de perder la especificidad y la profundidad de los conocimientos puntuales. Recordando a Calvino: multiplicidad, visibilidad, exactitud y consistencia. Occidente va quedando chico y hay que hacer las cuentas también con la cultura oriental que nuestra cultura toma siempre como “excentricidad” tal vez provechosa, como demuestra Jung, pero excéntrica.

La activación de las neurociencias demuestra sin lugar a dudas que el que medita tiene mayor conectividad de sinapsis que quien no lo hace y entonces la cultura tiene que ver directamente con la salud y con la inteligencia no solo de los sistemas sino del hombre mismo.

Hasta los artistas reclaman un ámbito ético a su labor sobre todo si se sirven de sistemas inteligentes y reivindican algo que siempre estuvo en la cultura flotando con dificultades de expresión: la ética está indisolublemente ligada a la estética.

Un joven acude a una clínica para un chequeo rutinario; en esa clínica hay cinco pacientes esperando trasplantes de órganos. Para poder vivir, dos de ellos necesitan un pulmón, otros dos necesitan sendos riñones y el quinto requiere un corazón. Curiosamente el joven que ha ido a hacerse el chequeo tiene el mismo grupo sanguíneo que ellos, lo que le convierte en el donante idóneo. ¿Es lícito matar a una persona para salvar a otras cinco?⁶⁸

En 2009, durante un experimento en el Laboratorio de Sistemas Inteligentes en la Ecole Polytechnique Fédérale de Lausanne en Suiza, robots programados para cooperar entre sí (en la búsqueda de un recurso beneficioso y evitar uno venenoso) finalmente aprendieron a mentir a cada uno otro en un intento de acumular el recurso beneficioso. Un problema en este caso puede haber sido que los objetivos eran “terminales” (es decir, en contraste, los motivos humanos últimos típicamente tienen la cualidad de requerir un aprendizaje interminable).

Entre asustados, ignorantes, exaltados, angustiados, delirantes, habrá que concluir que “algunas personas temen que la inteligencia artificial nos haga sentir inferiores, pero entonces cualquiera de nosotros tendría que sentir complejo de inferioridad al mirar una flor”⁶⁹.

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.

⁶⁸ Johnson, Judith Jarvis, *The Trolley problema*, The Yale Law Journal, 1985.

⁶⁹ Kay, Alan, *The Computer “Revolution” Hasn’t Happened Yet!* Educom Conferences, 1998. Alan Kay es también el autor de la famosa frase “la mejor manera de predecir el futuro es inventándolo”.

Se trata siempre de un tema de fines y medios y la frase “el fin justifica los medios” históricamente carece de sentido. Sucede 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 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 claro oscuros, es allí donde se debe indagar sobre los medios y no basta con decir “no lo sabía”.

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. Se puede admitir la discusión sobre el uso de un medio que tiene poco riesgo pero que podría aumentar, si los beneficios son enormes, pero no se trata solo de un tema de cantidad⁷⁰.

Para terminar nada mejor que la paradoja de Fermi: él se preguntaba cómo era posible que cuando se buscaban vidas inteligentes estos siempre aparecían en la Tierra cuando era estadísticamente improbable que no hubiese vida inteligente en el vasto universo. Su hipótesis era que cuando la tecnología va más rápido de la razón, la vida inteligente se destruye a sí misma. Tomémoslo como una advertencia.

Semblanza de Antonio A. Martino

Por Antonio A. Martino

Soy sustancialmente un maestro. En 1955 me recibí de Maestro normal nacional en la escuela n° 2 de Buenos Aires Mariano Acosta. Tuve la suerte de ser aceptado como maestro suplente ya el año siguiente y en el mes de marzo y luego gané el concurso como maestro titular y ejercí la profesión hasta 1962 cuando ya estaba recibido de abogado en la Universidad de Buenos Aires.

En 1963 Ambrosio Lucas Gioja me pidió que me integrara a su Instituto de Filosofía del Derecho (que hoy lleva su nombre) y que dictara clases como Profesor Ayudante y años después me alentó a que me presentara como candidato a Profesor Adjunto en la Facultad de Derecho y gané en 1969. Seguí dando clases en la Universidad de Buenos Aires y también en la de Belgrano donde me pidieron en 1970 que fuera Profesor Titular de Filosofía del derecho.

Es importante que se reflexione sobre el hecho que para enseñar es necesario, primero, aprender. Así que pase mi vida estudiando y haciendo mi doctorado cuando las Universidades argentinas retomaron esa posibilidad. Y no alcanza con tener los extraordinarios profesores que tenía en Argentina, me presenté a una beca del gobierno italiano para estudiar con Norberto Bobbio, el más famoso filósofo del

⁷⁰Si descalificar el magnífico trabajo del Mit para hacer el juego de la *Moral Machine* y sin desprestigiar la intuición de la creadora del dilema del tranvía, a veces, el tema de las cantidades es humanamente intolerable.

derecho italiano entre 1982 y 1983 (las clases en Europa van de octubre a octubre) y la gané. Hice mi perfeccionamiento en Italia y volví a Argentina a continuar con mis clases en las tres universidades donde enseñaba, pues ínterin había sido nombrado Profesor Titular de Filosofía del Derecho en la Universidad de Morón.

En enero de 1976 me vi obligado a abandonar el país y volví con mi viejo maestro Bobbio a Turín, en agosto de ese año me presenté en un concurso de Ciencia Política en la Facultad homónima de la Universidad de Pisa y gané y enseñé durante 31 años esa materia en Pisa. El Consejo Nacional de Investigaciones italiano me confió la dirección de un Instituto de esa prestigiosa institución desde 1982 hasta 1993 y pude organizar congresos de importancia internacional gracias a mi maestro finlandés Georg H. von Wright, alumno y sucesor de Wittgenstein en la cátedra de filosofía de la Universidad de Cambridge.

Jubilado en la Universidad de Pisa en el 2007 como profesor emérito, volví a Buenos Aires a dirigir la Maestría en Ciencia de la Legislación que traía de la Universidad pisana. Y seguí enseñando aun cuando en 2012 me nombraron también profesor emérito de esa universidad.

Desde 1992 soy académico correspondiente de la Academia Nacional de Derecho y Ciencias Sociales de Córdoba. Desde el 2008 soy miembro del Instituto de Filosofía Política e Historia de las Ideas Políticas de la Academia Nacional de Ciencias Morales y Políticas y desde el 2009 del Instituto de Derecho Constitucional y desde este año formo parte del CETI (Centro de Estudios de Tecnologías Inteligentes de la Academia Nacional de Ciencias de Buenos Aires).

Siempre enseñé o participé de entidades de enseñanza y estudio. Soy un lector empedernido, a los 11 años leí la *Summa Teológica* de Santo Tomás de Aquino y me fascinó ese método de colocar un problema, decir que escribieron otros sobre el tema y luego dar su opinión. En la pubertad conocí a quien luego sería mi mejor amigo: Rafael Campanella, un italianito que vino a vivir aquí. Juntos leímos a los existencialistas empezando por Kierkegaard, pero quien nos marcó fue Ortega y Gasset con su manera de contar la realidad y analizarla.

Sin duda la filosofía analítica ha sido un punto central en mi formación y en mis estudios y siguiéndola encontré la lógica y con ella la sintaxis. La sintaxis fue el gran instrumento que me hizo descubrir que tanto en oriente como en occidente se razona de la misma manera sea con el sistema de Leibniz o los hexagramas del I Chín. Mirando la sintaxis descubrí que la lógica no es acerca de verdad o falsedad como decía Aristóteles (que es semántica) sino solo sintaxis, no dice nada sobre el mundo, pero permite hacer cálculos, escribir fórmulas y programas de computación, escribí *Lógica sin verdad* y allí tropecé con la inteligencia artificial.

Siempre me fascinó la filosofía y estudié derecho pues mi padre me preguntó de qué viviría si estudiaba filosofía y letras. Pero rápidamente me hice un filósofo del derecho, que en el fondo es un filósofo. Todos mis maestros fueron filósofos.

La filosofía exige capacidad de asombro, por eso es recomendable para los niños y para quienes no les asusta dejar de lado los conocimientos que le son familiares y que le dan confianza por el entorno. O se es un niño o un aventurero intelectual que no tiene idea donde irá a parar.

Y dentro de la filosofía, que es más una actitud que un conocimiento porque de entrada se dice que quiere saber (filo-sofía) o sea que no se sabe. Hay algo de irresponsable en la filosofía por eso siempre me gustó la frase *fortuna juvat audaces* (la fortuna ayuda a los audaces). Y creo que en mi caso me ayudó pues me largué a hacer cosas más grandes de mis capacidades.

Pero no es irresponsabilidad de no hacerse cargo de los propios actos. Creo que todos nos debemos hacer cargo de lo que hacemos y no vale decir “no lo sabía”. Debiste haber pensado y calculado.

Hemos liberado una fuerza extraordinaria con el desarrollo de los sistemas inteligentes, tenemos la obligación de averiguar hasta donde podemos llegar sin hacer más daño que beneficio.



CAPÍTULO II

LA IMPORTANCIA DE UN MARCO JURÍDICO ADECUADO A LA ÉTICA Y LOS VALORES EN IA

Por Bibiana Beatriz Luz Clara⁷¹

Introducción

⁷¹ Doctora en Derecho Procesal por la Universidad de Salamanca, Magister en Derecho de Internet, vicepresidente de la Federación Iberoamericana de Asociaciones de Derecho e Informática (FIADI), Docente e Investigadora de UFASTA, UCAECE y UCASAL. Profesora de la Maestría en Legaltech de la Universidad de Salamanca. Autora de libros y capítulos de libros y conferencista internacional.

El ser humano está tomando en el entorno de la nueva economía un papel activo, buscando nuevos desarrollos basados en la investigación. Dentro de estos desarrollos nos encontramos aquellos que integran Inteligencia Artificial (IA), buscando mayor eficacia en los procesos, y el mejoramiento de la calidad de vida de la sociedad toda. Las investigaciones se han multiplicado en los últimos tiempos, y es por ello que es preciso analizarlas y estudiar el marco normativo adecuado para que todo marche en el sentido esperado.

El hombre debe siempre ocupar el lugar central y ser quien realice los desarrollos para que nada se salga de su cauce, cimentado en la ética y las previsiones jurídicas, para que podamos hablar de verdadera evolución. Sus desarrollos deben ser para beneficiar a la humanidad iluminándola sobre las nuevas posibilidades y no para acarrear nuevos problemas.

Dentro de las ciencias de la computación la IA constituye una categoría, cuya finalidad es diseñar sistemas inteligentes y autónomos, basados en el conocimiento humano.

1. La inteligencia artificial

El concepto de IA se aplica a los sistemas que manifiestan un comportamiento inteligente, capaces de analizar su entorno y pasar a la acción con cierto grado de autonomía, con el fin de alcanzar objetivos específicos.

Tiene un objetivo complejo: la percepción de su entorno; la obtención de datos; elegir entre varias opciones la que sea óptima para alcanzar el objetivo establecido; adapta su conducta según como el entorno se ve afectado por sus acciones anteriores.

Big data es el alimento de la IA. Sistemas que a partir de algoritmos aprenden y reaccionan en consecuencia con autonomía y que puede implicar a todas las áreas de la sociedad. Datos que pueden predecir lo que va a ocurrir (sistemas predictivos). Nuestros deberes cambian según la información de la que disponemos. Así los deberes éticos también cambian. La ética se refiere al hombre que es quien desarrolla los sistemas. Pues los prejuicios se pueden transmitir al software, creando sesgos cognitivos sobre razas, ideas políticas, sexos, etcétera. Pero ¿qué pasaría con máquinas que desarrollan su próxima generación? Las órdenes que se les indican deberían ser compatibles con la vida humana.

Riesgos e impactos de la IA nos llevan a preguntarnos por la responsabilidad civil contractual y extracontractual que tenemos que considerar dentro del ámbito civil, pero también pensar en el ámbito militar con todas sus posibilidades y consecuencias. La idea es poder desarrollar el máximo potencial para obtener beneficios, minimizando los riesgos fundados en la explicabilidad y transparencia.

Se destaca la importancia de poner límites a la IA basados en la dignidad del ser humano y los valores. Si existe un riesgo alto deberían instrumentarse sistemas de aprobación y autorización previa para que el marco ético quede garantizado.

La ética marca las líneas básicas generales y el derecho debe brindar el marco jurídico específico, que mantenga en resguardo los derechos humanos que son inalienables.

A día de hoy cuando queremos realizar alguna consulta interactuamos en muchas oportunidades con asistentes virtuales que nos responden, dan recomendaciones y realizan acciones en base a los datos que obtienen o les proporcionamos. Estos sistemas nos realizan sugerencias de películas, videos, series y música en base a nuestros gustos. Preparan publicidades a nuestra medida de acuerdo a las búsquedas que anteriormente hemos realizado.

Utilizando IA los sistemas pueden obtener y almacenar los datos y en base a ellos sacar conclusiones o tomar decisiones, replicando o imitando la inteligencia de las personas o bien actuando o desarrollando un pensamiento racional en busca del mejor resultado posible con cierto grado de autonomía e incluso con la posibilidad de tener autoaprendizaje.

Los avances y desarrollos para obtener mayor almacenamiento y procesamiento de datos, permitirán su mayor autonomía. La finalidad de la IA debe ser mejorar nuestra calidad de vida facilitando nuestras tareas, para que podamos dedicarnos a otras de mayor relevancia.

La faceta importante que hay que considerar es que la IA trae consigo riesgos que pueden derivarse de defectos, errores o sesgos en la programación, de la incidencia de terceros, que atenten contra la privacidad, intimidad, dignidad, salud o patrimonio, de las personas, pudiendo estar la responsabilidad ser atribuida entre sus distintos actores como el programador, el fabricante, el propietario, el usuario, etc. entre quienes también operará la solidaridad, haciendo que quien tenga que responder luego pueda reclamar a los demás de la cadena.

Encontramos que la IA puede ser:

Inteligencia Artificial general: Es aquella donde las máquinas tienen plena capacidad para replicar procesos intelectuales humanos, como, por ejemplo, jugar videojuegos o reconocer imágenes concretas en fotografías.

Inteligencia Artificial específica: la cual utiliza técnicas como la minería de datos, aprendizaje profundo o aprendizaje de redes neuronales artificiales.

De acuerdo a las estimaciones del Instituto de Investigaciones de la Multinacional McKinsey⁷², del 4 de setiembre de 2018, en el año 2030 la tecnología de la IA tiene un alto potencial para contribuir a la actividad económica global mediante sus herramientas y técnicas, aunque puede impactar seriamente en el sector laboral. Esto provocaría que en el proceso de transición un 20% de la población económicamente activa a nivel mundial sea sustituida por alguna clase de máquina inteligente, y que al menos un 70% de las compañías podrían haber adoptado al menos un tipo de IA. En

⁷² Mc Kinsey & Co., *Apuntes desde la frontera de IA: Modelado del impacto de la IA en la economía mundial*, www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-modeling-the-impact-of-ai-on-the-world-economy/es-cl#.

la carrera de la IA los adoptantes tardíos obtendrán menos ventajas que sus competidores que la adoptaron tempranamente.

Por todos estos motivos la regulación de la Inteligencia Artificial es indispensable, pues ello garantizaría una transición digital ordenada y al mismo tiempo brindaría certeza jurídica a la sociedad que pudiera verse afectada por ella.

De acuerdo al desarrollo de los países, hoy es posible utilizar la inteligencia artificial en la vida cotidiana, mediante robots, asistentes de voz, drones, chatbots, automóviles autónomos o aplicaciones del internet de las cosas. Esto trae como implicancia que en un futuro no muy lejano existirán conflictos sociales, económicos, legales, educativos, laborales e incluso médicos y de salud, que hace necesaria una adecuación del marco jurídico para contemplar dichas situaciones.

Algunas de las críticas, respecto del uso de la IA se basan en que podría utilizarse con fines que vayan contra la ética o ilegales en diferentes tareas, como espionaje industrial, fines militares, tráfico de datos de personas, entre otros fines diversos y oscuros.

La norma ISO sobre inteligencia artificial⁷³, la define de forma un poco vaga, con la finalidad de no limitarse, como: *“la capacidad de una unidad funcional para realizar funciones que generalmente están asociadas con la inteligencia humana, como el razonamiento y el aprendizaje”*.

2. IA en la Unión Europea

En el ámbito de la Unión Europea, se propone el cumplimiento de estos objetivos a través de la aplicación de normas internacionales, de la creación de un código de conducta, de la revisión de las reglas de responsabilidad vigentes y de la evaluación del impacto social que la IA pueda llegar a tener.

El Consejo Europeo, el Comité Económico y Social Europeo y el Comité de las Regiones, entre otras, enunciaron el 25 de abril de 2018, las siguientes propuestas:

- Potenciar la capacidad tecnológica e industrial de la UE e impulsar la adopción de la IA en todos los ámbitos de la economía, tanto en el sector privado como en el público. Las inversiones en investigación e innovación y en la mejora del acceso a los datos.
- Prepararse para las transformaciones socioeconómicas que origina la IA, fomentando la modernización de los sistemas de educación y formación, favoreciendo el talento, previendo los cambios en el mercado laboral y prestando apoyo a las transiciones que se operen en él y la adaptación de los sistemas de protección social.
- Garantizar el establecimiento de un marco ético y jurídico apropiado, basado en los valores de la Unión y en consonancia con la Carta de los Derechos Fundamen-

⁷³“ISO 10218-1:2011, *Robots and robotic devices - Safety requirements for industrial robots*”, “ISO 13855, *Safety of machinery, Positioning of safeguards with respect to the approach speeds of parts of the human body*” y la “ISO/TS 15066:2016 *Robots and robotic devices - Collaborative robots*”.

tales de la UE. Incluye una próxima directriz sobre la interpretación de las actuales normas en materia de responsabilidad por productos defectuosos y un análisis pormenorizado de los retos emergentes, así como la cooperación con las partes interesadas, en el seno de una alianza europea de la IA, para elaborar directrices éticas en la materia.

Para minimizar los riesgos de la IA debe tenerse en cuenta:

- 1) la bioética, la transparencia e inteligibilidad para el ser humano, la responsabilidad social;
- 2) buscar el bien común, no hacer daño;
- 3) el respeto por la autonomía del ser humano;
- 4) la igualdad de trato para todos;
- 5) supervisión humana; sistemas seguros y robustos probados en forma independiente para evitar daños masivos;
- 6) informar si los algoritmos dieron los resultados esperados o no.

Se vincula así la ética al diseño de los productos tecnológicos para hacer de este mundo un lugar mejor. La trazabilidad de los algoritmos, la supervisión humana y la rendición de cuentas son fundamentales, antes durante y terminado el desarrollo para poder verificar las pautas de seguridad para su adecuado funcionamiento. Un sistema robusto para la prevención de eventos dañinos consecuencia de errores humanos como intencionales. La IA nos brinda mayor potencia y poder sobre los datos que pueden ser utilizados para: automatizar, predecir, identificar, monitorear, atacar, castigar, detectar. Soluciones ágiles, confiables y sencillas para volver a los sectores público y privado más eficientes. La idea que debería imperar es la de la cooperación hombre máquina, y no su reemplazo, para asegurar el futuro de la IA, que permita ir en el sentido de los ODS de la ONU y su agenda 2030 para una industria, economía y sociedad sustentables.

Entre las cuestiones más importantes que deben regularse, destacan la posible adquisición de personalidad jurídica de los robots programados con Inteligencia Artificial, así como su responsabilidad civil en caso de daños y perjuicios. La regulación de la IA garantizaría una transición digital ordenada y al mismo tiempo brindaría certeza jurídica a la sociedad que pudiera verse afectada por ella aún de manera indirecta.

Día a día irán apareciendo muchos nuevos productos tecnológicos y es difícil hacer reglas para productos aun desconocidos, aunque en principio, podríamos basarnos en las reglas éticas que tenemos.

Debido a que ya es posible utilizar la inteligencia artificial en la vida cotidiana mediante robots, asistentes de voz, drones, chatbots, vehículos autónomos o aplicaciones de IoT, eso implica que en un futuro cercano su uso planteará problemas sociales, económicos, legales, educativos, laborales e incluso médicos y de salud, por ello debe pensarse en el marco adecuado de regulación, aunque es difícil debido a que cada país tiene tiempos, formas y modos diferentes de crear legislación, aunque podría seguirse alguna ley modelo.

El tema de la responsabilidad jurídica debe quedar aclarado en breve pues la IA operando a través de distintos dispositivos puede realizar en muchos casos las mismas tareas que las personas, pero desde un nivel de programación, por lo que la responsabilidad jurídica obviamente debe especificarse de forma distinta.

Es posible que la falta de regulación de Inteligencia Artificial, provoque demoras en la investigación y desarrollo. Es fundamental determinar parámetros de actuación y fijar límites, condiciones y responsabilidades dejando en claro si las máquinas inteligentes pueden considerarse jurídicamente responsables por sus acciones ya sea de forma parcial o total.

Si la IA es capaz de interactuar en su entorno obteniendo mayores datos y adquiriendo conocimientos que aquellos para los que fue programada originalmente, si de este modo adquieren mayor autonomía, capacidad de decisión y ejecución de acciones, entonces deberán establecerse claramente a quien corresponden las responsabilidades por los eventos dañinos que puedan producirse y que afecten a las personas o a sus bienes.

La misión de las normas jurídicas debe ser resguardarnos de los riesgos que la IA pueda generar, optimizar sus beneficios brindando pautas claras para resolver los conflictos legales y de seguridad que puedan surgir, promoviendo un desarrollo y utilización técnicamente sólida basada en el respeto por los derechos de las personas, favoreciendo que los desarrolladores y testers sean diversos para evitar sesgos cognitivos que vulneren la dignidad humana por discriminación.

Las directrices de la Unión Europea remarcan la necesidad de centrar el marco normativo en los seres humanos y en el respeto de sus derechos fundamentales, su autonomía, privacidad y seguridad, haciendo foco en evitar que se afecten los grupos sociales que son más vulnerables y/o que se encuentren en una situación de asimetría de poder o información.

Recogen el principio de no discriminación y remarcan la importancia de la diversidad de los equipos a la hora de desarrollar, aplicar y probar los productos basados en IA.

Se busca afianzar la aplicación de los principios de explicabilidad y transparencia, lo que traduce en la necesidad de que exista información clara sobre las posibilidades del sistema, que su funcionamiento sea trazable, susceptible de ser explicable y de fácil auditoría, disminuyendo así el riesgo de sesgo y error.

Las directrices remarcan la importancia de adoptar acciones de educación y formación para quienes estén expuestos a la IA, y de que se respete el derecho a informarse sobre la lógica detrás de las decisiones que la IA adopte en base al tratamiento de datos personales.

Ha habido propuestas legislativas fundadas en la premisa de que son las personas las responsables por las decisiones tomadas por la IA, a la vez que se han realizado propuestas de interpretación de las normas vigentes sobre responsabilidad por productos para aplicarlas a la IA y hasta se ha mencionado la posibilidad del pago de impuestos destinados a reducir el impacto social.

El 19 febrero de 2020 se presentó en Bruselas el libro Blanco de la Inteligencia Artificial. Tiene un enfoque orientado hacia la excelencia y la confianza, buscando lograr mayor eficiencia, y rápido desarrollo. Sus pilares fundamentales son:

a) Lograr la excelencia a lo largo de toda la cadena de valor implementando soluciones adecuadas, mediante incentivos a la innovación e investigación.

b) Crear un marco normativo sobre la IA, centrada en el ser humano, velando por el cumplimiento de las normas, con la finalidad de generar confianza.

c) Agilidad y seguridad en el manejo de los datos, proveer información para la mejor toma de decisiones para mejorar la vida de los ciudadanos.

La IA ofrece beneficios y desafíos, pues si bien puede agilizar y brindar ingentes oportunidades, mantenimiento predictivo y seguridad, por otro lado, puede traer el riesgo de intromisión en la privacidad, falta de transparencia en la toma de decisiones, y algún tipo de discriminación.

Para que pudiera responsabilizarse a la IA por sus acciones debería tener una personalidad jurídica propia. ¿Tendríamos que pensar en la necesidad de una personalidad electrónica? ¿así como existe un domicilio electrónico y otras opciones electrónicas?

Pero teniendo en cuenta el grado de avance tecnológico actual, por el momento puede parecer apropiado acudir a una eventual clasificación de la IA como objeto peligroso, a la aplicabilidad de la responsabilidad por daños causados por productos defectuosos, o a los conceptos de responsabilidad objetiva, in vigilando, entre otros, y así hacer responsables a las personas detrás de la IA por los daños a terceros: esto es, a los fabricantes, propietarios o a los usuarios, lo que está alineado a lo que actualmente se propone en el marco de la Unión Europea.

3. IA en Argentina

En la Argentina, más allá de ciertas disposiciones que apuntan al estudio y desarrollo de propuestas de marcos regulatorios o a la promoción de buenas prácticas en términos de IA, la materia no ha sido específicamente regulada y entonces se nos presentan distintas opciones:

1) No hacer nada y observar permaneciendo expectante y luego actuar, aunque puede que sea tarde.

2) Encontramos a quienes sostienen la necesidad de regular jurídicamente la inteligencia artificial de manera específica ya que la falta de regulación podría entorpecer su desarrollo, pero la regulación excesiva también, implicando tal vez un daño más grande a la innovación del que se pretende prever.

3) Al ser de suma complejidad lograr que nuevas normas abarquen todos los aspectos de la materia, sumado a que el proceso de elaboración de las normas va habitualmente muy por detrás de los avances tecnológicos, con el consecuente riesgo de que los continuos avances vuelvan obsoleta a la eventual nueva norma en poco tiempo. Es por ello que debemos plantearnos seriamente si es eficiente tomar ese camino. Tal vez sería mejor un “sandbox” regulatorio que establezca un entorno

controlado e ir regulando de acuerdo a los desarrollos y actividades, porque se requiere experimentar primero para luego regular.

4) Por otro lado, encontramos la postura que sostiene que puede resultar más conveniente, si es posible, encuadrar las situaciones que surjan con motivo del uso de la IA en las normas jurídicas ya existentes aplicando los principios generales vigentes del derecho contractual, de las normas sobre propiedad intelectual, responsabilidad y privacidad vigentes como las de nuestro Código Civil y Comercial, las normas de la Ley de Defensa del Consumidor 24.240 o la Ley de Protección de los Datos Personales, entre otras.

Por ejemplo, en materia de IA se suele recomendar reconocer la responsabilidad de tanto los fabricantes como de quienes comercialicen productos con IA y en nuestro ordenamiento jurídico, y en el marco de relaciones de consumo, ello lo podríamos lograr a través de lo prescripto por los arts. 2⁷⁴ primera parte, y 40⁷⁵ de la Ley de Defensa del Consumidor, por cuanto reconocen responsabilidad al productor, fabricante, proveedor, vendedor, entre otros, por los daños resultantes de la prestación del servicio, como así también del vicio o riesgo de la cosa.

Similarmente, frente al avance de la IA se suele hacer énfasis en que las personas tengan acceso libre a los datos generados o procesados por sistemas de IA y en ese sentido encontramos la posibilidad de ejercer el Derecho de Información prescripto por el art. 13⁷⁶ de la Ley Nacional de Protección de los Datos Personales 25.326.

Las posturas de búsqueda de encontrar puntos de contacto pueden provocar que se genere riesgo de restricción al usar moldes viejos para instituciones o modelos nuevos lo que puede generar fricción.

Sin embargo, para llevar a cabo lo antedicho, es necesario que los operadores jurídicos estemos capacitados e informados sobre los alcances actuales y potenciales de la IA.

En la medida en que se cuente con suficiente capacitación y formación, que los principios generales vigentes sobre responsabilidad, privacidad, propiedad inte-

⁷⁴Art. 2°: Proveedor. “Es la persona física o jurídica de naturaleza pública o privada, que desarrolla de manera profesional, aun ocasionalmente, actividades de producción, montaje, creación, construcción, transformación, importación, concesión de marca, distribución y comercialización de bienes y servicios, destinados a consumidores o usuarios. Todo proveedor está obligado al cumplimiento de la presente ley”.

⁷⁵Art. 40. “Si el daño al consumidor resulta del vicio o riesgo de la cosa o de la prestación del servicio, responderán el productor, el fabricante, el importador, el distribuidor, el proveedor, el vendedor y quien haya puesto su marca en la cosa o servicio. El transportista responderá por los daños ocasionados a la cosa con motivo o en ocasión del servicio. La responsabilidad es solidaria, sin perjuicio de las acciones de repetición que correspondan. Sólo se liberará total o parcialmente quien demuestre que la causa del daño le ha sido ajena”.

⁷⁶Art. 13 (Derecho de Información). “Toda persona puede solicitar información al organismo de control relativa a la existencia de archivos, registros, bases o bancos de datos personales, sus finalidades y la identidad de sus responsables. El registro que se lleve al efecto será de consulta pública y gratuita”.

lectual y otras ramas estén dotados de suficiente flexibilidad, se podrán aplicar en materia de IA a través de una interpretación concordante con nuestra realidad social. Esto nos podrá ahorrar las complejidades del camino que implica el dictado de nuevas normas específicas sobre IA.

Actualizar las normativas y definir si el concepto de idóneo que se aplica puede ser utilizado respecto de un software, pero dejando aclarado que siempre debería ser asistido por el control humano.

5) La propuesta de cambio de paradigma regulatorio: aborda la transversalidad de la tecnología en la sociedad y en la industria en particular.

Se generan cambios tan grandes y repentinos, que se requiere reinventar el derecho en lo ligado a IA. No pueden seguirse normas rígidas y generales, sino que mediante estrategias flexibles sobre principios y objetivos adecuados que vayan acompañando la innovación, junto a reglas éticas para que nada se salga de su cauce.

Se trata de aproximaciones multidisciplinarias en diferentes capas, y de principios de neutralidad tecnológica.

Se requiere delinear el curso de los pasos a seguir para encarar la regulación de la IA junto estándares éticos.

Busquemos lograr el desarrollo de una IA centrada en la persona a quien tiene que ayudar, beneficiar y proteger.

La actuación en forma conjunta entre instituciones regionales será beneficiosa para poder regular con bases responsables y éticas y con información de calidad.

Bibliografía

Bourcier Daniele, *Inteligencia artificial para Europa*, <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52018DC0237&from=ES>.

García, Víctor M., *Inteligencia artificial: su regulación y desafíos legales*, <https://medium.com/@abogadovicgarcia/inteligencia-artificial-su-regulaci%C3%B3n-y-desaf%C3%ADos-legales-segunda-y-%C3%BAltima-parte-e13ced0994e3>.

Marco jurídico internacional de la Inteligencia Artificial, www.uasb.edu.ec/web/ciberderechos/marco-juridico-ia.

Merchán Murillo, Antonio, *Ventajas y riesgos de la Inteligencia Artificial*, www.muyinteresante.es/tecnologia/articulo/ventajas-y-riesgos-de-la-inteligencia-artificial-651449483429.

Repond, Pablo, *Inteligencia artificial y su marco normativo*, <https://abogados.com.ar/inteligencia-artificial-y-su-marco-normativo/25187?tech=1>.



Principios de inteligencia artificial y ética por diseño

Por Daniel Peña Valenzuela⁷⁷

⁷⁷Abogado de la Universidad Externado de Colombia con maestría en derecho de la Universidad de Londres (LLM en materia Corporativa y Comercial) y en Grenoble, Francia (Propiedad Intelectual y Nuevas Tecnologías) así como estudios de profundización en Derecho Internacional en La Haya y en Arbitraje Internacional en la Cámara de Comercio Internacional. Profesor Titular de la Universidad Externado de Colombia en las Facultades de Derecho y de Finanzas y Relaciones Internacionales. Director de Investigación en Comercio Electrónico en la misma Universidad. Árbitro y Panelista de nombres de dominio de internet del Centro de Arbitraje y Mediación de la Organización Mundial de la Propiedad Intelectual OMPI. Árbitro activo de las Cámaras de Comercio de Bogotá y Medellín. Miembro de la Comisiones de Propiedad Intelectual y de Negocios Electrónicos y Telecomunicaciones de la Cámara de Comercio Internacional CCI de París. Socio Fundador de la firma Pe-Ciclo “Sistemas Inteligentes y Ética”

Introducción

El día 23 de septiembre de 2020, en el Ciclo de Ética en Sistemas Inteligentes de la Academia Nacional de Ciencias de Buenos Aires, se presenta la conferencia “Principios de Inteligencia Artificial y Ética por Diseño”. Este escrito comprende las principales temáticas analizadas teniendo como dos ejes fundamentales la reflexión ética, es decir, respecto de los valores o axiología que definen un deber ser y la inteligencia artificial como sistemas de información que permiten la predicción de conductas mediante el aprendizaje profundo de máquinas o dispositivos.

La inteligencia artificial potenciada por el *Deep Learning* y la robótica generan simultáneamente temores y esperanzas. La historia del miedo a la “creación” (humana) de inteligencia tiene un comienzo remoto incluso la mitología griega muestra al atrevido Prometeo robar el fuego de los Dioses y tener un elemento (soplo!) vital. Muchos siglos después, Mary Shelley con su personaje ficticio Frankenstein aproximó el mito del soplo creador a un repulsivo monstruo –al menos así es en el imaginario popular y de las películas–.

Esta aproximación conflictual entre los seres humanos y la inteligencia creada de manera artificial sugiere la continuidad del miedo ancestral al desborde tecnológico. Es el primer pilar que justifica el análisis prospectivo ya que debemos tener la capacidad de construir la hoja de ruta del futuro posible o probable.

Otras aplicaciones de la inteligencia artificial representan una mayor cercanía con el sentido de lo humano y su desarrollo muestra armonía con actividades cotidianas. Tomamos arbitrariamente el ejemplo de Alexa (desarrollo inteligente y virtual en Echo) así como otros asistentes virtuales y botnets representan una versión más agradable y seductora de la inteligencia artificial. El asistente de voz escucha al usuario y realiza distintas acciones según lo que se le pida. Puede ofrecer a sus usuarios sus propias opiniones sobre programas de televisión, series o su bebida y comida favorita. El asistente va más allá de su función de motor de búsqueda por voz y son sujetos interactivos con los que se puede entablar una conversación (cada vez más sofisticada).

A favor de los beneficios que trae Inteligencia Artificial se puede referir el experimento de Google, para identificar tejidos cancerosos y daños en la retina. Una combinación de flujo de datos de los pacientes provenientes de los teléfonos inteligentes y dispositivos portátiles permiten un diagnóstico médico automatizado a partir de una descripción de síntomas, detectar rasgos de comportamiento que sugieran que está deprimido e identificar si tiene un riesgo especial de enfermedad cardíaca. La agregación de datos también facilitará encontrar otras personas con enfermedades similares y ver cómo respondieron a varios tratamientos.

ña Mancero Abogados. Autor de libros sobre el tema de Comercio Electrónico, entre otros: *Aspectos Legales de Internet y Comercio Electrónico* (2001); *Contrato de diseño, desarrollo y hosting de un sitio web* (2003); *El contrato por medios electrónicos* (2003); *La Sociedad de la Información Digital* (2007); *Responsabilidad civil en la era digital* (2007); *Aspectos legales de la Computación en la nube* (2013); *Software libre y software propietario* (2013); *Responsabilidad de los proveedores de servicios de internet en materia de propiedad intelectual* (2013); *De la Firma Manuscrita a las Firmas Electrónica y Digital* (2016).

De todas las nuevas tecnologías parece que la inteligencia artificial es la más retadora desde el punto de vista del monopolio humano respecto de la capacidad de análisis, creación, aprendizaje, memoria, intuición y sensibilidad. Muchas son las preguntas recurrentes en ambientes académicos y empresariales ¿Se reemplazará al ser humano en algunas dimensiones como la laboral? ¿Desaparecerán algunas profesiones y oficios? ¿Las máquinas estarán fuera del control humano? ¿Hasta dónde llegarán en su capacidad de pensar? ¿Las máquinas sólo aprenderán o serán intuitivas y creativas? ¿La masificación de las aplicaciones de la inteligencia artificial favorecerá o perjudicará a las economías emergentes que han centrado su modelo de desarrollo en las materias primas y los bajos salarios para los trabajadores?

La ética puede ser parte de la solución frente a la imprevisibilidad de la fuerza de la inteligencia artificial. Sin embargo, proponemos la prospectiva como metodología para establecer la relación entre derecho y máquinas pensantes. Cada cultura y sociedad tiene una aproximación diferente a la relación hombre-máquina, no es la primera vez en la historia que se experimenta esa fricción. Los automóviles reemplazaron a los carruajes de caballos, pero aún hoy en día la potencia se mide por decenas o cientos de caballos de fuerza (algunas utilizan los cientos de caballos de fuerza para abusar de la velocidad y poner en riesgo a otros inocentes ciudadanos con piques en las carreteras).

La masificación de los computadores personales hacía prever la desaparición de miles de empleos rutinarios, pero más bien se convirtió en una herramienta necesaria para secretarías, digitadores y operarios. Se crearon profesiones nuevas y empleos por doquier.

Frente a la inteligencia artificial no podemos ser reactivos porque la humanidad podría estar en riesgo en sus características más eminentes. Tampoco podemos ser preactivos porque no sabemos realmente cómo prepararnos para un cambio esperado. Más bien debemos ser proactivos para definir cómo controlamos los cambios, definimos los límites éticos y determinamos si queremos complementarnos con las máquinas y potencializar nuestras cualidades o más bien ser reemplazados en nuestra esencia.

1. Problema jurídico planteado

La pregunta central formulada fue: ¿podemos/debemos construir agentes/sistemas inteligentes éticamente conscientes?

Existen dos visiones respecto de este tema:

a) La *ética por regulación* (regida por los principios de prevención y precaución), en donde la legislación regula y pone límites a los sistemas de inteligencia artificial existentes y

b) La *ética por diseño*, la cual busca incluir dentro del diseño mismo del sistema de inteligencia artificial (en sus algoritmos) el componente ético, es decir, que este queda embebido dentro del mismo sistema de inteligencia artificial.

De esta manera, el principio de prevención parece no ser tan eficiente frente al aprendizaje autónomo de las máquinas y el cada vez más complejo control efectivo

vo sobre la toma de decisiones, la aplicación del principio de precaución parece demasiado limitativa y prohibitiva de los desarrollos aplicados de la inteligencia artificial por ejemplo en los vehículos autónomos, en los asistentes personales y en los bots para asesoría financiera automatizada. El diseño preestablecido en los sistemas de inteligencia artificial puede ser una alternativa. Según el Principio 5 sobre inteligencia artificial de la OECD: *“Las organizaciones y las personas que desarrollen, desplieguen o gestionen sistemas de IA deberán responder de su correcto funcionamiento en consonancia con los principios precedentes. Este principio ratifica la responsabilidad de los diseñadores, encargados e inversionistas en los sistemas de inteligencia artificial”*.

2. ¿Qué es la ética por diseño?

Es un concepto originado en la privacidad *by design*, de la protección de datos, en donde la toma de decisión de la Inteligencia Artificial está embebida en el diseño mismo de ésta en sus algoritmos de manera tal que las decisiones que este tome involucren un componente ético.

a. Importancia de la ética por diseño

Una de las razones por las cuales la ética por diseño se está convirtiendo en esa tercera vía, al lado de la regulatoria y la preventiva o de precaución, es porque involucra principios, reglas y normas del propio código (utilizando la expresión del profesor Lawrence Lessig) porque implica:

1) *Accountability (responsabilidad demostrada)*. La rendición de cuentas implica la justificación frente a terceros interesados sobre la toma de decisiones en una organización. a) ¿Qué y cómo fue el origen o causa de la decisión que se tomó? b) El algoritmo de inteligencia artificial puede incorporar dentro de sí las funcionalidades que permitan detectar la manera como se decidió adoptar un curso de acción y tener una trazabilidad de dicha decisión. De esta manera, la ética por diseño va de manera más amplia pues no sólo abarca el recaudo de datos personales sino la explicación *ex post* respecto de cualquier toma de decisiones por un sistema de inteligencia artificial y permite tener una trazabilidad de estos procesos de toma de decisión. El principio 3 de la OECD sobre Inteligencia Artificial establece: *los sistemas de IA deben estar presididos por la transparencia y una divulgación responsable a fin de garantizar que las personas sepan cuándo están interactuando con ellos y puedan oponerse a los resultados de esa interacción*.

2) *Responsabilidad (liability)*. a) Capacidad de los sistemas de no sólo tomar la decisión sino también de que, en caso de causar un daño, de responder por esa decisión. Lo anterior implica, incorporar una capacidad de examen de las decisiones del sistema de Inteligencia Artificial y una capacidad de respuesta del mismo; b) Esto permite que la cadena de responsabilidad pueda ser definida y permite funcionalidades en la Inteligencia Artificial que permitan discernir sobre sus decisiones y de responder sobre las consecuencias que éstas generan; c) La ética por diseño se trata precisamente de anticiparse desde el diseño mismo de la inteligencia artificial a los

posibles daños que ésta pueda causar en su proceso de toma de decisiones y que la funcionalidad misma permita corregir los errores cometidos por estos sistemas.

3) *Transparencia.*a) Este elemento permite el activismo respecto de los sistemas de Inteligencia Artificial, en el sentido de que al dejar una trazabilidad del proceso de toma de decisiones del sistema, esto permite que el sistema tenga un proceso de auto-aprendizaje y, a su vez, cumpla con su deber de información frente a terceros interesados; b) El deber de información muchas veces choca con el derecho a la confidencialidad, pero si pensamos en la transparencia y en la capacidad de responder por sus errores, es preferible que esté incorporado en el diseño de la inteligencia artificial y no a través de los mecanismos de precaución y prevención.

b. Problemas o cuestionamientos de la ética “by design”

La ética por diseño representa una convergencia entre el concepto de diseño (*legal design*) y la arquitectura del sistema jurídico (sistema normativo tradicional), con el fin de ajustar la inteligencia artificial a los sistemas legales, pensar los procesos jurídicos de manera tridimensional.

Las aplicaciones de la ética por diseño deben complementarse con una interactividad humana. El algoritmo siempre debe tomar una decisión ética, pero siempre que se refiera a conceptos éticos le concedemos a la inteligencia artificial la capacidad de tomar decisiones complejas. En ese sentido, la interacción hombre-máquina debe ser más intensa en los casos límite en los cuales la decisión implica un juicio moral. La inteligencia artificial ha dejado de ser puramente rutinaria y mecánica y los sistemas de decisión automatizado van a tener una repercusión moral a medida que van avanzando.

De cualquier manera, el sistema de ética por diseño también tiene algunos problemas/cuestionamientos, tales como: 1) ¿Es o no viable? la viabilidad de este sistema *ex ante* es una alternativa más sostenible en ese *trade off* entre la regulación y la prohibición; 2) definir cuál es la norma correcta a aplicar y el valor es complejo. Pero este es el gran problema del derecho: axiología frente a lo normativo; 3) cómo convertirlo en una decisión práctica que tenga efecto en estándares internacionales y locales de diseño y puesta en el mercado de los sistemas de inteligencia artificial.

3. Conclusiones

a) En conclusión, la razón de incorporar el sistema de ética por diseño en la toma de decisiones de la inteligencia artificial depende de un raciocinio de anticipación respecto de los efectos potencialmente nocivos de las aplicaciones prácticas de este nuevo paradigma tecnológico.

b) La idea de la ética por diseño es anticipar sin prohibir, es decir, generar un sistema técnico-jurídico de responsabilidad *ex ante* y generar mayor información y mayores opciones de control sin necesidad de afectar el desarrollo de los sistemas de inteligencia artificial.

c) El principio de prevención parece no ser tan eficiente frente al aprendizaje autónomo de las máquinas y el cada vez más complejo control efectivo sobre la toma de decisiones, la aplicación del principio de precaución parece demasiado limitativa y prohibitiva de los desarrollos aplicados de la inteligencia artificial.

d) El algoritmo de inteligencia artificial puede incorporar dentro de sí las funcionalidades que permitan detectar la manera como se decidió adoptar un curso de acción y tener una trazabilidad de dicha decisión.

e) El deber de información muchas veces choca con el derecho a la confidencialidad, pero si pensamos en la transparencia y en la capacidad de responder por sus errores, es preferible que esté incorporado en el diseño de la inteligencia artificial y no a través de los mecanismos de precaución y prevención.

Referencias

Brynjolfsson, Erik - McAfee Andrew, *The Second Machine Age. Work, Progress and Prosperity in a Time of Brilliant Technologies*, 2014.

Comisión Europea, *Libro Blanco sobre la inteligencia artificial: Un enfoque europeo orientado a la excelencia y la confianza*, 2020, https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_es.pdf.

El Tiempo, *La inteligencia artificial vive su gran momento en Colombia*, 2019, www.eltiempo.com/tecnosfera/novedades-tecnologia/inteligencia-artificial-en-colombia-entrevista-con-andrew-ng-416822.

Grupo Bancolombia, *Capital Inteligente*, "Inteligencia Artificial: el aprendizaje de las máquinas al servicio de las personas", 2018, www.grupobancolombia.com/wps/portal/empresas/capital.

iProup, *Por qué L'Oréal utiliza inteligencia artificial para realizar diagnósticos personalizados de la piel*, 2019, www.iproup.com/innovacion/2888-tecnologia-realidad-virtual-automatizacion-LOreal-utiliza-inteligencia-artificial-para-realizar-diagnosticos-de.

Llinares, Fernando M., *Inteligencia artificial y justicia penal: más allá de los resultados lesivos causados por robots*, "Revista de Derecho Penal y Criminología", 2018, p. 87 a 130.

Lessig, Lawrence, *Code and Other Laws of the Cyberspace*, New York, Basic Books, 2001, p. 34 a 89.

Navas Navarro, Susana y otros, *Inteligencia Artificial*, "Tecnología Derecho", 2017, p. 23 a 71.

OECD, *Artificial Intelligence Principles*, 2019, www.oecd.org/going-digital/ai/principles, París.

Peña Valenzuela, D., *Inteligencia Artificial, explicabilidad y la toma de decisiones*, Blog de Derecho de los Negocios, 2019,

<https://dernegocios.uexternado.edu.co/comercio-electronico/inteligencia-artificial-explicabilidad-y-la-toma-de-decisiones>.

Schneider, U. (prod.), *Límites éticos para la inteligencia artificial*, Alemania, Deutsche Welle documentales, 2019.

Schwab, Klaus, *The Fourth Industrial Revolution*, 2018, p. 25 a 47.



Inteligencia artificial, ética y administración de justicia

Por Carlos Gregorio de Gràcia⁷⁸

1. Introducción

⁷⁸ Actualmente es investigador y consultor en el Programa Justicia Abierta del Ministerio de Justicia y Derechos Humanos, de Argentina, y Director de Investigación del Instituto de Investigación para la Justicia.

En los últimos años hemos sido sorprendidos por aplicaciones informáticas cuya característica es poder desarrollar actividades que normalmente eran realizadas por seres humanos. Se las llama “Inteligencia Artificial” (o IA) –genéricamente– pero incluyen cosas diferentes, y parecen ser exitosas en cierto tipo de tareas.

Algunos ejemplos son: los vehículos autónomos⁷⁹, aplicaciones capaces de hacer diagnósticos médicos⁸⁰ o hacerse cargo del pronóstico del clima⁸¹. Un elemento reiterativo en algunas de estas aplicaciones son las “redes neurales artificiales”, pero no todas las que se adornan con la etiqueta de “Inteligencia Artificial” las usan. Su impacto más sorprendente es que logran desarrollar algunas tareas “mejor que los humanos”.

Las primeras valoraciones éticas de estos ejemplos coinciden en que la “Inteligencia Artificial” basada en redes neurales tiende a operar como una caja-negra, y por ende resulta casi imposible desarrollar algunos aspectos habitualmente esperados de la inteligencia humana como: adicionar una fundamentación, transparentar las opciones descartadas, y quedar sometida a control y auditoría.

a. Redes neurales artificiales

El uso tradicional de las computadoras (Gráfico 1) supone aprovechar su potencia de cálculo por medio de una secuencia de instrucciones (programa de cómputo) que se aplican a los datos de entrada. Al final de estas operaciones se obtienen los resultados.

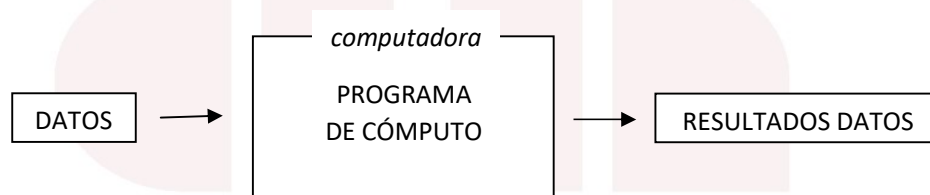


Gráfico 1. Modelo computacional

⁷⁹ Los vehículos autónomos (capaces de conducir en la vía pública e interactuar con otros vehículos conducidos por humanos) se caracterizan por poseer muchos sensores; incluyen imágenes que son procesadas conjuntamente para identificar todas las señales y circunstancias del tránsito de vehículos necesarias para llegar a un destino, sin accidentes.

⁸⁰ Por solo mencionar un ejemplo, puede verse H. A. Haenssle & al., 2018. El sistema consiste en una aplicación que analiza fotografías de algún trastorno de la piel (e.g. lunares) y es capaz de establecer la probabilidad de que se trate de un melanoma.

⁸¹ Específicamente, Google está desarrollando *Nowcast*, que utiliza una red neural convolucional –un tipo de algoritmo de aprendizaje– cuya arquitectura es particularmente propicia para el análisis de imágenes; es “libre de física”, o sea, esto significa que la red neural aprende solo de sus datos de entrenamiento y no incorpora el conocimiento sobre cómo funciona la atmósfera; todo lo que tiene que seguir son los patrones que identifica en las imágenes que la alimentan (ver: Jason Hickey, *Using Machine Learning to “Nowcast” Precipitation in High Resolution*, 13/1/20). Esto muestra un cambio en el paradigma del conocimiento. De alguna forma con la fuerza bruta computacional (procesamiento y memoria), junto con datos brutos (*big-data*), se resolverían los problemas sin necesidad del conocimiento científico. Esto no es necesariamente malo, solo chocante; ¿por qué usar la investigación científica, una metodología lenta y cara, cuando hay una más barata disponible?

La expresión redes neurales artificiales aparece asociada a la idea de emular los procesos del sistema nervioso central de los vertebrados. Si bien las redes neurales operan dentro de una computadora (Gráfico 2), el procedimiento difiere del “programa” de cómputo tradicional.

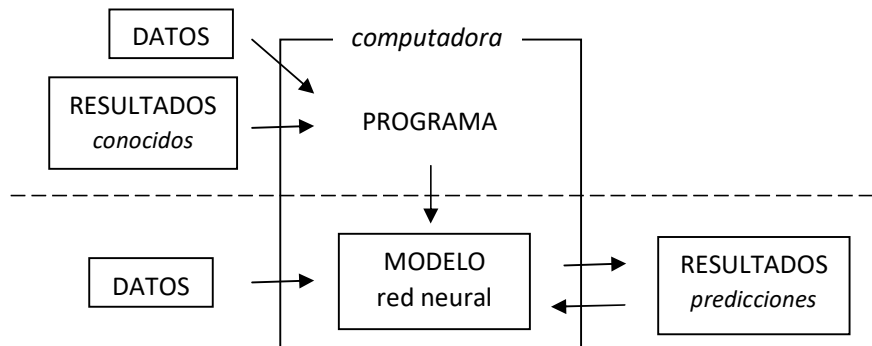


Gráfico 2. Esquema de una computadora que soporta una red neural artificial

Las redes neurales consisten en un conjunto de unidades de procesamiento (nodos) que simulan neuronas y están interconectadas (en forma de grafo conexo) a través de un conjunto de “pesos” de una manera que permite que las señales viajen a través de la red en paralelo.

Los nodos (neuronas) son elementos informáticos simples. Acumulan información de otras neuronas por medio de una suma ponderada. Luego, se alimenta el sistema con una gran cantidad de casos con un resultado conocido, y los pesos de las conexiones entre las neuronas se cambian mediante un algoritmo de entrenamiento diseñado para minimizar el error total del sistema.

Una diferencia importante en comparación con los sistemas tradicionales basados en reglas o estadísticas es la capacidad de las redes neurales de realizar este tipo de aprendizaje. Al comienzo de un proceso de aprendizaje una red neural artificial no contiene información explícita.

Una red entrenada ha generado un modelo (estadístico) que está representado por todas las matrices de pesos entre las neuronas. Esta característica se llama generalización y permite a una red neural predecir casos que nunca antes se han presentado al sistema. Las redes neurales artificiales han demostrado ser útiles para predecir varios tipos de eventos. Se pueden modelar y pronosticar relaciones especialmente complejas, no lineales y que dependen del tiempo.

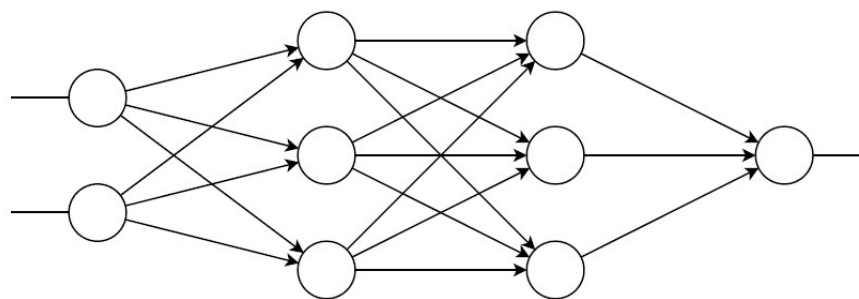


Gráfico 3. Esquema de una red neural multicapa

En todo este proceso –en términos generales– tienen lugar principalmente tres operaciones: función de activación⁸², descenso del gradiente⁸³, y retro-propagación⁸⁴. Una red neural está formada por numerosas neuronas (nodos) y sus conexiones. Se distingue también tres tipos de neuronas: neuronas de entrada, salida y ocultas.

b. Analogía y paralelismo histórico

Existe cierta analogía entre las redes neurales y el cálculo infinitesimal, y al mismo tiempo, un paralelismo histórico sobre como evolucionaron ambas concepciones⁸⁵.

Para fines del siglo XVII el cálculo diferencial e integral era aún un conocimiento informal⁸⁶, utilizado, pero no dotado de fundamentación teórica. Por esta

⁸² Una neurona puede describirse por su estado de activación en un momento determinado. Además, una función de activación define cuál será el nuevo estado, que a su vez se compone de la activación existente y las nuevas que fluyen. Las señales de un nodo de entrada a una neurona se multiplican cada una por un peso específico variable. Los resultados del cálculo (una transformación lineal) de los nodos internos deben imaginarse como una especie de resultado intermedio determinante para el funcionamiento de la red.

⁸³ El objetivo del proceso de aprendizaje es minimizar gradualmente el error entre el resultado esperado y el resultado real. El error cuadrático medio se utiliza habitualmente como medida del error total en la matriz neural.

⁸⁴ La señal de corrección primero se retroalimenta a la última capa oculta y se cambian las ponderaciones correspondientes. Luego, el error se informa al siguiente turno. Esto se hace hasta la capa de entrada.

⁸⁵ Para resaltar la importancia histórica del “cálculo” Steven Strogatz (2019) en su libro –que reaviva la polémica entre Newton y Leibniz– dice: “Sin cálculo, no tendríamos teléfonos celulares, computadoras ni hornos de microondas. No tendríamos radio. O la televisión. O ultrasonido para mujeres embarazadas o GPS para viajeros perdidos. No hubiéramos dividido el átomo, desenredado el genoma humano o puesto astronautas en la luna. Puede que ni siquiera hubiéramos tenido una Declaración de Independencia”.

⁸⁶ Los problemas comunes que dieron origen al cálculo infinitesimal comenzaron a plantearse en la época clásica de la antigua Grecia (siglo III a.C.), pero no se encontraron métodos sistemáticos de resolución sino hasta el siglo XVII, gracias a los trabajos de Isaac Newton y de Gottfried Wilhelm

razón Isaac Newton no publicó sus primeros escritos que intentaban darle sustento. Newton temía a las críticas y el desprestigio de apoyar una metodología considerada –en ese entonces– críptica, al estilo de la alquimia. Eso lo llevo a solo tratar el tema en cartas con amigos. Leibniz se le adelantó y publicó sus ensayos, Newton enfureció y reclamó que sus escritos privados eran anteriores.

Cuando en el desarrollo de redes neurales se define que todo cuanto ocurre en los nodos (neuronas) sea lineal (suma de las entradas ponderada por sus pesos), se está estableciendo una gran similitud con los fundamentos filosóficos que influyeran en el pensamiento de Leibniz. Él se refería al principio de “continuidad de universo”, que se resume en la expresión “el universo no da saltos”⁸⁷. Si es así, cualquier observación en el tiempo o el espacio puede aproximarse eficientemente por una línea recta (al menos en el entorno de un punto o instante). De esta forma el cálculo infinitesimal puede modelar un comportamiento no lineal por medio de líneas tangentes, en cada punto o instante. Las redes neurales utilizan un modelo lineal (local en las neuronas artificiales) que contribuye a modelar uno no-lineal en el cálculo final de la relación datos-resultados.

En el pensamiento de Leibniz el otro gozne filosófico que puede haber influido en la fundamentación del cálculo infinitesimal es el argumento “del mejor de todos los mundos posibles”, y su consecuencia inmediata: podemos explicar las leyes físicas mediante criterios optimales. Precisamente el “descenso del gradiente”, que opera en el aprendizaje de las redes neurales, es un método optimal⁸⁸.

La retro-propagación que utilizan las redes neurales, no tienen relación con el pensamiento filosófico de Leibniz, pero tampoco guarda similitud con el cerebro de los vertebrados. A falta de analogía, no sorprende que sea tan contemporáneo como el enfoque de “prueba y error”⁸⁹.

Leibniz. Ambos –paralelamente– sintetizaron los conceptos y métodos usados por sus predecesores en lo que hoy se denomina “diferenciación” e “integración”. Desarrollaron reglas para manipular las derivadas (reglas de derivación) y demostraron que ambos conceptos eran inversos (teorema fundamental del cálculo).

⁸⁷ Está claro que Leibniz no había oído sobre la teoría del *Big Bang*, ver: Eric Kincanon *Leibniz & the Big Bang: Leibniz could have predicted the Big Bang in 1715*, 121 *Philosophy Now*, 2017. Existen, por tanto, elementos contextuales relevantes, por ejemplo: la eficacia de la IA se ha acelerado después del *Big Data*. Entonces ¿qué perspectiva es más adecuada para valorar la IA, y qué contextos considerar? La respuesta es, todas las perspectivas que sea posible encontrar: histórica, cultural, psicológica, etc. Un ejemplo en este sentido es el libro de John Brockman (2015), que es como una encuesta abierta a más de 190 pensadores, con muy disímiles puntos de vista y contextos, que dan una respuesta en menos de dos páginas a la pregunta: *¿Qué opina sobre las máquinas que piensan?* Obviamente fue criticado por incluir “mucho redundancia”, pero también elogiado por dar cabida a muchas ideas originales o ensayos que presentan un ángulo único.

⁸⁸ Leibniz publicó entre 1684 y 1686 sus ensayos sobre el cálculo infinitesimal, uno de ellos se titula *Sobre una geometría oculta y el análisis de los indivisibles y los infinitos*, tampoco su nombre completo aparece en estos ensayos (firmó solo con sus iniciales G.G.L.). De hecho, sus ensayos leídos con la formación matemática del siglo XXI son difíciles de entender. Desde el siglo XVII, muchos matemáticos han contribuido al cálculo diferencial y le han dado rigor matemático.

⁸⁹ Dado que este método está ligado con la inteligencia, es un método que el sociólogo, psiquiatra e investigador inglés, William Ross Ashby, señaló como necesario para la creación de inteligencias artificiales. Ashby, W. R. *Design for a Brain*, Londres, Chapman & Hall, 1960.

c. Discusión sobre las redes neurales artificiales

Si bien la “Inteligencia Artificial” basada en redes neurales se considera una metodología útil y exitosa, no existe un consenso sobre si es ciencia o *pseudo-ciencia*. Es frecuente ver cada día nuevas publicaciones que formalizan –tanto como les es posible– la metodología y sus principios (Petersen 2020). Sin embargo, proliferan las opiniones en los *blogs*⁹⁰; en unos critican las redes neurales y las denominan como un procedimiento meramente estadístico, otros desestiman la necesidad de una fundamentación rigurosa, ya que ven que sus desarrollos son exitosos y funcionan (o sea son capaces de resolver problemas con menor error que grupos de profesionales especializados).

Petersen (2020), afirma “Debe quedar claro que no podremos proporcionar una respuesta integral a porque los métodos de aprendizaje profundo funcionan de manera especial. Por otro lado, veremos que este enfoque nos permite hacer declaraciones rigurosas que proporcionan explicaciones e intuición de por qué ciertas arquitecturas de redes neurales son preferibles a otras”.

d. Percepción artificial

Del análisis de los ejemplos mencionados en la introducción es posible acotar un poco el problema. La expresión “Inteligencia Artificial” es presuntuosa, ya que en la mayoría de los casos que son vistos como exitosos (la red neural) en realidad se trata de un sistema computacional capaz de “percibir” mejor que los humanos. Por ello es preferible el término “percepción artificial” tal como sugiere Gandhi (2017)⁹¹. En realidad, deberíamos referirnos simplemente como “percepción”, ya que la percepción humana es extremadamente débil, dada la fragilidad de los sentidos. La gran habilidad de la mente humana ha sido desarrollar instrumentos de percepción como: telescopio, osciloscopio, periscopio, giroscopio, microscopio y cualquier otro *-scopio*.

Si nuestra preocupación es la ética, deberíamos preguntarnos cuáles son las reglas éticas de la percepción. Este cometido es difícil porque la percepción es “libertad de información”, legalmente incluida en la “libertad de expresión”, y por tanto cualquier limitación o reproche a “percibir” debe ser muy restrictivo. Queda claro que percibir la naturaleza no puede tener muchas reglas (algunas quizás sí presentes en la ética de la investigación médica). Ciertamente las limitaciones se presentan al momento de percibir la vida o personalidad de otros seres humanos. Tor Ekeland opina que en el *common law* se considera el rostro humano como “público”⁹², aun si fuera así, no es muy ético mirar a una persona insistentemente. O sea, un límite éti-

⁹⁰ Ver, por ejemplo, *Is a machine learning a science or art? o Is machine learning a pseudo science?*

⁹¹ Sharad Gandhi, *AI is Essentially “Artificial Perception”*, 2017.

⁹² Ryan Daws, *Clearview AI lawyer: Common law has never recognised a right to privacy for your face*, 6/3/20.

co de la percepción es el respeto de la intimidad (vida privada) y los datos personales⁹³.

Sin embargo, sobre la legalidad de la percepción, y en el contexto de la administración de justicia, sorprende la reciente ley francesa que sanciona penalmente la construcción de perfiles sobre cómo deciden los jueces⁹⁴.

2. Sistema de Justicia

Antes de analizar qué características debe tener cualquier tecnología, algoritmo o sistema que reemplace –aunque sea supervisada o parcialmente– la decisión de un juez, es prudente analizar algunos ejemplos sobre qué se espera –constitucional o convencionalmente– de un juez.

André Malraux dice que “la misión del juez consiste en transformar el Derecho en Justicia”⁹⁵. Entonces, un juez aplica la ley, resolviendo controversias o sancionando los delitos, siempre en función de las pruebas que declara admisibles. Es claro que los jueces deben tomar sus decisiones, solo en las cuestiones que les competen, con independencia de cualquier otro poder constituido o no, éste es un requisito mínimo para la función judicial. Al respecto se le preguntó al doctor Jorge Bacqué–ministro de la Corte Suprema de Argentina entre los años 1985 y 1990– ¿puede abstraerse la Corte del plano político? Bacqué dijo: “*Personalmente creo que lo que tiene que hacer un tribunal es aplicar la ley, y dentro de ella los ministros pueden dar prioridad a sus convicciones morales, preferencias personales o al dolor de muelas que tenga en ese momento. Dentro de lo que la ley permite, la función del tribunal es hacer justicia*”⁹⁶.

Si acompañamos esta apreciación con los arts. 24; 75, inc. 12, y 118 de la Constitución Argentina que instituye los juicios por jurados –en particular en las causas criminales– queda claro que el sistema constitucional argentino imagina un juez técnico pero teniendo en cuenta un veredicto con sensibilidad humana. ¿Cómo se compatibiliza esta interpretación con la idea de trasladar a una máquina parte de las funciones judiciales?

También las constituciones políticas crean la garantía del juez natural (en Argentina según el art. 18, “Nadie puede ser ... sacado de los jueces designados por la

⁹³ Aun en el caso de las figuras públicas (despojadas por la jurisprudencia de casi todos sus derechos de imagen, privacidad y vida privada, incluso familiar) la legislación de California prohíbe el uso de telescopios o cámaras fotográficas con teleobjetivo para invadir los domicilios de las figuras públicas.

⁹⁴ Art. 33, Loi n° 2019-222 du 23 mars 2019 de réforme pour la justice: “Les données d’identité des magistrats et des membres du greffe ne peuvent faire l’objet d’une réutilisation ayant pour objet ou pour effet d’évaluer, d’analyser, de comparer ou de prédire leurs pratiques professionnelles réelles ou supposées. La violation de cette interdiction est punie des peines prévues aux articles 226-18, 226-24 et 226-31 du Code Pénal”.

⁹⁵ Citado por Bielsa, Rafael, *Transformación del Derecho en Justicia*, La Ley, 1993.

⁹⁶ Entrevista de Irina Hauser a Jorge Bacqué, ex Juez de la Corte Suprema: “*Hubo un exceso de la Corte para abrirle la puerta a Menem*”, “Página/12”, 2/12/01.

ley”. También el Pacto de San José de Costa Rica, art. 25, consagra el derecho de acceso a la justicia “Toda persona tiene derecho a un recurso sencillo y rápido o a cualquier otro recurso efectivo *ante los jueces o tribunales competentes*, que la ampare contra actos que violen sus derechos fundamentales reconocidos por la Constitución, la ley o la presente Convención”. Sobre este punto el Tribunal Constitucional alemán también analiza la noción del juez legítimo para definir la impropiedad de que, aun dentro del sistema judicial, actúe un juez que no es aquel al que legítimamente le corresponde el asunto⁹⁷. Sin embargo, el creciente congestionamiento de los tribunales en Brasil ha requerido acciones extremas, como el *mutirão* donde jueces de otras jurisdicciones concurren en grupo a un juzgado en problemas, para resolver causas pendientes⁹⁸. El *mutirão* es una figura creativa que se sustenta en la ponderación de los derechos fundamentales⁹⁹.

Un ejemplo más, en 1985 la Ley Francesa del 5 de julio de 1985, llamada comúnmente “*Loi Badinter*”¹⁰⁰, estuvo dirigida a resolver los problemas de congestionamiento de los tribunales debido a la gran cantidad de casos de accidentes de circulación. El art. 12 de la ley obligaba a la compañía de seguros citada en garantía a traer información adicional al proceso, en particular una oferta económica de indemnización. La introducción de esta “oferta” –junto con el art. 17– fue revolucionaria, y en pocos años los tribunales se descongestionaron¹⁰¹. En Argentina se proyectó un camino similar, pero al poco tiempo se apreció que bastaba con la publicidad de los precedentes y no era necesaria una reforma legal (o sea solo el art. 26 de la Ley Badinter); esta misma solución fue luego implementada con éxito en Chile y Eslovenia¹⁰². En estos casos no existe problema ético en admitir este tipo de

⁹⁷ Sentencia BVerfGE 4, 412 de la Primera Sala, del 20 de marzo, 1956, 1 BvR 479/55. Ver: Jürgen Schwabe, *Jurisprudencia del Tribunal Constitucional Federal Alemán*, Fundación Konrad Adenauer, 2009.

⁹⁸ Ver Regina Cœli Formisano, *Mutirão judiciário e o princípio do Juiz Natural, Justiça e Cidadania*, 2011. El *mutirão* fue cuestionado judicialmente pero, entre otros, el Supremo Tribunal Federal “*reputou legítima a possibilidade de se proferir sentença em regime de mutirão*” (STF, HC 182551-RS, 20/3/20).

⁹⁹ Ver *Mutirão “Judiciário em Dia”* que describe en detalle el funcionamiento del *mutirão*, palabra de origen tupi-guaraní para describir el auxilio gratuito que se prestan, unos a otros, los miembros de una comunidad –denominado: *andecha* (en Asturias), *auzolan* (en euskera) o *a vereda* (en Rioja)–.

¹⁰⁰ www.ijusticia.edu.ar/docs/LoiBadinter5julliet1985.htm.

¹⁰¹ El art. 17 determina que si la oferta no es aceptada por el demandante, el juez que finalmente fije la indemnización debe valorar si la oferta inicial fue razonable. Si la estima manifiestamente insuficiente debe condenar al asegurador a pagar una multa del 15% de la indemnización que establezca. Ver, Jean-Marc Sarafian, *La révolution opérée par la procédure d’offre*, L’argus de l’assurance, 2015.

¹⁰² En Argentina se accede al Sistema Judicial para Cuantificar los Daños a la Persona, en Chile al Baremo Judicial y en Eslovenia al Odmera Nepremoženjska Škoda (al ingresar seleccionar solo NEGM y luego buscar por campos “*iskanje po poljih*”).

información, ya que tiene un fundamento legal y consiste en valorar precedentes jurisprudenciales¹⁰³.

Como último ejemplo, se citan aquí algunas normas constitucionales que caracterizan la decisión judicial:

Constitución de la Provincia de Córdoba, art. 95. Las sentencias y autos interlocutorios deben tener motivación suficiente, so pena de nulidad.

Constitución de la Provincia de Chubut, art. 46. Los procedimientos judiciales, el sumario y la prueba son públicos en todos los casos salvo aquéllos en que la publicidad afecte la moral o la seguridad pública. La resolución es motivada. ... art. 60. El Estado garantiza la plena reparación de los daños causados por error judicial, sin otro requisito que su demostración.

Constitución de la Provincia de San Juan: art. 198. El Poder Judicial tiene todo el imperio necesario para mantener su inviolabilidad e independencia ante los otros poderes del Estado.

Esta lógica de la decisión judicial determina el marco normativo y ético en el que deben incluirse los sistemas inteligentes, y en particular la entrada de información en el proceso, o solo en la mente del juez. En particular si esta información es determinante para tomar una decisión. El tema se torna especialmente delicado si esta información procede de una máquina.

3. Sistemas inteligentes en la administración de justicia

Ciertamente la inteligencia artificial no es llevada a los tribunales por estar de moda ni porque se espere que sea de mejor calidad, sino por necesidad. La razón es que el modelo clásico de administración de justicia está hoy colapsado por la gran cantidad de conflictos que plantean los ciudadanos.

Con esta idea en la mira, cabe pensar que requisitos y que control ético debe tener cualquier ayuda que busquen o reciban los jueces para sus funciones naturales.

En este sentido, la irrupción de la tecnología en las decisiones judiciales ha requerido la formulación de algunas normas éticas; por ejemplo, la *American Bar Association* ha regulado cuando pueden los jueces buscar información en Internet sobre las causas en las que tienen que decidir.

La *Formal Opinion 478* del 8 de diciembre de 2017 "*Independent Factual Research by Judges Via the Internet*" establece que:

¹⁰³ Ver como argumentan los jueces usando este sistema en CNCiv, 21/11/02, "R., M. J. c/López, Gerardo y otro". Ver: Elena I. Highton & al., *Predictibilidad de las indemnizaciones por daños personales por vía de la publicidad de los precedentes*, "Revista de Derecho de Daños" 2004-3-29.

— Si el juez necesita información adicional para decidir un caso; si necesita información que influencia la decisión, ésta debe ser aportada por las partes o sus abogados.

— Si el propósito de la búsqueda es corroborar o desacreditar hechos, o llenar una laguna factual en el proceso, la búsqueda es éticamente inapropiada.

— Si el propósito de juez es buscar información general o ilustrativa para entender mejor un asunto que no está directamente vinculado a un caso pendiente o inminente: la búsqueda es apropiada.

— Si el juez quiere obtener más información sobre una parte o sobre la materia relativa a un caso pendiente o inminente: sin tienen consecuencias factuales en la definición del caso, ellas deben ser sometidas a proceso contradictorio.

Varios principios éticos atinentes a la función judicial se desequilibran al momento de buscar o admitir –conscientemente o inconscientemente– información determinante en la toma de decisiones (fuera de los causes naturales/procesales). En las decisiones judiciales pueden identificarse al menos dos caracterizaciones; las decisiones que se toman de acuerdo a derecho y las que se toman de acuerdo a ciencia. En las que se toman de acuerdo a ciencia están –por ejemplo– apreciar si el reo se presentará en el juicio, si reincidirá, si está él en riesgo o pondrá en riesgo a otros, o en el fuero de familia cuales son los riesgos de dar la tenencia de un niño a los padres o a otras personas. La justicia penal de adolescentes está repleta de este tipo de decisiones: ¿está el adolescente infractor *a priori* en condiciones de rehabilitarse socialmente? ¿Cuáles son medidas más apropiadas que le ayudaran a reinserirse? Este tipo de decisiones perturban al juez porque ninguna norma jurídica le dice que hacer, así que podrá abrazarse a cualquier ayuda que reciba, ya sea de un perito o de un sistema inteligente, sin preguntar mucho con qué fundamento.

La adopción de medidas cautelares en libertad mientras se espera el juicio, es un tema de gran debate (donde se pondera la libertad, las garantías procesales y la seguridad pública, y se atiende –además– a los altos costos y estigmas de la encarcelación). En este caso el error judicial tiene siempre graves consecuencias, pero el sesgo de incrementar la privación de libertad para reducir los riesgos es ciertamente peor. En varias jurisdicciones de los EEUU se está utilizando cierta IA para ponderar los factores de riesgo de los acusados que buscan la libertad bajo fianza (Arizona, Kentucky, Ohio, Alaska and New Jersey, y ciudades como Chicago, Houston, Pittsburgh y Phoenix) usan el sistema PSA –*Public Safety Assessment*– que es un “cálculo actuarial” (así lo denominan los autores) que ayuda a los jueces a predecir si el acusado comparecerá en el juicio, y si cometerá un nuevo delito mientras duren las medidas cautelares. Para algunos jueces esto es un cambio revolucionario, para quienes ya ven en la IA la *singularidad tecnológica*, esto es prácticamente cero. El balance de esta experiencia es que está sometida a un control de eficacia, a un mo-

nitoreo permanente en cuanto a sus resultados, y entonces el sistema aprende de cada error¹⁰⁴.

Las primeras objeciones éticas de estos ejemplos es que la IA no debería operar como una caja-negra¹⁰⁵, la información que deviene o facilita una decisión debe estar acompañada por una fundamentación, transparentar las opciones que descartó, y quedar sometida a un control *a posteriori* que asegure un aprendizaje de los errores.

Con este mismo ánimo se han desarrollado varios documentos contentivos, como por ejemplo en la Unión Europea la *European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment*, CEPEJ (2018).

En síntesis, esta carta enfatiza los siguientes principios:

- Especial respeto de los derechos fundamentales;
- Prevenir cualquier discriminación entre individuos o grupos de individuos;
- Calidad y seguridad;
- Transparencia, imparcialidad y equidad: que los métodos de procesamiento de datos sean accesibles y comprensibles, y que faciliten las auditorías externas;
- Asegurar que los usuarios sean actores informados y que controlen las elecciones realizadas¹⁰⁶.

Algunos sistemas crean más dudas que adhesiones; es el caso de *COMPAS– Correctional Offender Management Profiling for Alternative Sanctions*– que es una herramienta de gestión que les da a los jueces una estimación de la probabilidad que un acusado reincida en un delito. El principal problema con este sistema es que los algoritmos que usa son secretos comerciales, y no pueden ser examinados por el público ni las partes afectadas, situación que podría ser una violación del debido proceso.

Es muy interesante analizar la decisión del Tribunal de Apelaciones de Wisconsin en “Wisconsin v. Loomis”¹⁰⁷, que rechazó la apelación de Eric Loomis quien fue condenado a seis años de prisión con fundamento en este *software*. Loomis alegó que la sentencia viola el derecho del acusado al debido proceso porque la caja negra evita que impugne la validez científica y la precisión de dicha prueba. La ma-

¹⁰⁴ Ver: *Public Safety Assessment: how it works* y en particular como se hace el monitoreo permanente en <https://advancingpretrial.org/psa/research/>.

¹⁰⁵ Esto es un problema, porque “el concepto de IA tiene una desventaja fundamental en comparación con los modelos físicos: la IA está aprendiendo el comportamiento de un sistema a partir de los datos como una ‘caja negra’, mientras que los modelos físicos tienen como objetivo simular todas las características de las predicciones” (extraído de Peter Dueben, *Can Artificial Intelligence predict the weather?*, 506 *The Parliament Magazine*, 2019, p. 47).

¹⁰⁶ Dice Margaret Levi (2018) “Las máquinas que piensan crean la necesidad de regímenes de rendición de cuentas –que aún no hemos diseñado– y de responsabilidad social (es decir, humana) por las consecuencias que aún no hemos previsto”.

¹⁰⁷ <https://epic.org/algorithmic-transparency/crim-justice/Loomis-Brief.pdf>.

yoría del tribunal sostuvo que “si se usa adecuadamente con conocimiento de las limitaciones y precauciones, la consideración de un tribunal de circuito de la evaluación de riesgo COMPAS en la sentencia no viola el derecho del acusado al debido proceso”. El tribunal afirma que “en la medida en que la evaluación de riesgos de Loomis se basa en sus respuestas a preguntas y datos disponibles públicamente sobre su historial criminal, Loomis tuvo la oportunidad de verificar que las preguntas y respuestas enumeradas en el informe de COMPAS fueran correctas”.

En “State v. Samsa”, 859 N.W.2d 149 se dice: “COMPAS es simplemente una herramienta disponible para un tribunal al momento de la sentencia, y un tribunal es libre de confiar en algunas partes de la evaluación, mientras rechaza otras partes”.

En otro caso similar, “People v. Chubbs”, el *software* en cuestión era *TrueAllele casework*, también un secreto comercial. El tribunal revocó la orden del tribunal superior de Los Ángeles que obligaba a *TrueAllele* a exhibir el código fuente – secreto comercial– ordenándole la modifique negando el acceso.

Algunos argumentos esgrimidos son a favor de *no revelar el secreto comercial* ya que los jueces deben proteger la inversión que hicieron los desarrolladores, y que si se revela la competencia los anularía. Otros argumentos son tan simples como *otros Estados también lo usan*¹⁰⁸. El software COMPAS va a contrapelo de los desarrollos de IA. Mientras la IA es fundamentalmente “percepción artificial”¹⁰⁹, COMPAS percibe los hechos a partir de un cuestionario que se le hace al reo, o sea no hay una búsqueda de información pertinente en un conjunto confuso o complejo que requiera mejorar la percepción humana. COMPAS parece más bien un modelo matemático que estima probabilidades a partir de los datos de un cuestionario.

Es claro que la situación ideal es que el código fuente sea transparente y de acceso público, sin embargo, habría que evaluar si es viable que los defensores ensayaran formas de instruir a los acusados para lograr resultados favorables¹¹⁰. La decisión legislativa no se hizo esperar. En marzo de 2019 el Estado de Idaho introdujo cambios en el procedimiento criminal §19-1910, ch. 258, sec. 1, p. 768, estableciendo que “las herramientas de evaluación de riesgos previa al juicio estarán abiertas a la inspección, auditoría y pruebas públicas”. También en febrero de 2020 se presentó en la legislatura del Estado de Michigan un proyecto (*Bill No. 5464*) para reformar en el mismo sentido el código de procedimiento criminal.

Surgen entonces varias preguntas para completar los principios éticos:

¿la IA debe dar –además de un resultado– una fundamentación?

¹⁰⁸ También es posible que sean reticentes a mostrar el código fuente porque en realidad éste sería tan simple que no justificaría su costo. Ver: Ed Yong, *A popular algorithm is no better at predicting crimes than random people: the COMPAS* (2018). Cabe preguntar: *¿tienen estos jueces el imperio necesario para garantizar su independencia de un proveedor externo?*

¹⁰⁹ Ver Sharad Gandhi, *op. cit.* Nota 91.

¹¹⁰ COMPAS usa los datos del registro de antecedentes criminales, y Dory Reiling (2020) acota: “la herramienta tiene sus defectos. Al utilizar datos del pasado, sobrestima sistemáticamente la reincidencia entre los acusados afroamericanos en comparación con los estadounidenses caucásicos”.

Dada la costumbre de los votos disidentes en las decisiones colegiadas ¿debe ofrecer la IA también las alternativas más relevantes, para que el juez elija?

¿Debe exigirse que la IA tenga evaluación previa de riesgos? ... o sea una validación. ¿Toda IA debe tener monitoreo *a posteriori*? ... o sea seguimiento permanente y a largo plazo de las decisiones que se toman.

¿Debe estar clara la responsabilidad por una decisión errada? O sea que no dé lugar a una responsabilidad diluida.

¿Debe la IA ser el resultado de minimizar los riesgos antes que maximizar los resultados? En materia penal los riesgos no son comparables con los resultados. William Blackstone enunció en 1769 “*para la ley, es preferible que diez culpables escapen a que un inocente sufra*”¹¹¹ ... ¿captan los procedimientos de aprendizaje automático este principio?, ... ¿quién dice como ponderar los riesgos en relación con los beneficios?

La información que usa/obtiene la IA para su aprendizaje ¿es admisible de acuerdo a las reglas de la prueba? ¿su estructura es la misma que la de los litigios que alimentan el sistema de precedentes?

Cuando la IA ha generado su propio aprendizaje “no supervisado” es probable que no sea posible recuperar todo el proceso para una auditoria. ¿será suficiente una validación sobre su desempeño?

Estévez & *al.* (2020) describen la experiencia de Prometea—acotando— “Se descartaron explícitamente tecnologías como las redes neuronales por la falta de justificación del resultado provisto, y por funcionar como cajas negras, sin posibilidad de explicar el resultado de una decisión” ... “La característica distintiva de Prometea es que realiza la recomendación de la sentencia en base a un *conjunto de reglas definido previamente por los abogados* y almacenadas en el sistema. Es por esto que Prometea *funciona como un algoritmo de caja negra (esto es, no se desconoce cómo es su lógica o cómo funciona)*, sino que es completamente trazable y permite explicar de una manera clara y sencilla cómo se arriba a una determinada propuesta de solución”. Además, ofrecen una serie de definiciones de “Inteligencia Artificial” en las que no es necesario que se base en el uso de redes neurales¹¹².

Sobre los “sistemas inteligentes” se puede decir que funcionan, que son eficientes, más baratos que la investigación científica y que muchos de ellos han sido incorporados a la vida cotidiana sin mayores problemas (al igual que el *mutirão*); so-

¹¹¹William Blackstone, *Commentaries on the Laws of England, Book the Fourth, Chapter the Twenty-Seventh: Of trial, and conviction*, *359 (1769).

¹¹²Estévez & *al.*, 2020, p. 20 citan una serie de definiciones de “Inteligencia Artificial” dirigidas a mostrar que no son necesarias las redes neurales: Bellman (1978) “actividades que asociamos con el pensamiento humano, actividades como la toma de decisiones, la resolución de problemas, el aprendizaje”; Kurzweil (1990) “El arte de crear máquinas que realizan funciones que requieren inteligencia cuando las realizan personas”; Winston (1992) “El estudio de los cálculos que hacen posible percibir, razonar y actuar”; Nilsson (1998) “se ocupa del comportamiento inteligente de los artefactos”. Prometea es un sistema desarrollado para el Ministerio Público Fiscal de la Ciudad Autónoma de Buenos Aires.

bre la inteligencia artificial quizás podemos preguntarnos si ¿existe o no existe? y probablemente nunca tengamos una respuesta definitiva¹¹³. Sobre la “percepción artificial” no hay dudas, las máquinas tienen mayor capacidad de percepción que los humanos quienes estamos totalmente condicionados por nuestra biología¹¹⁴, un automóvil autónomo dispone de muchísimos sensores (como telémetros, radar, visión infrarroja, entre otros) que son un salto cualitativo en cuanto a percibir.

Pasar de “sí funcionan” a que sean aceptables exige cambios en algunas normas de la administración de justicia. Se parte de que existe un “estado de necesidad”; ya se han intentado todo tipo de medidas, aumentar el número de jueces, jurisdicciones simplificadas, ... mega-juzgados con muchos jueces (Costa Rica), ... y los tribunales siguen congestionándose y la administración de justicia sigue retardada. Acceso, celeridad, son algunas de las razones que se deben ponderar, como el “juez natural” y otros principios. Si se implementa IA es probable que las demandas deban ser estructuradas (basadas en completar formularios con categorías cerradas) y que las sentencias también.

Los sistemas inteligentes se ocuparán de la gestión de las causas (eso no es un problema), y de proponer decisiones a las partes o al juez. La *Ley Badinter* ofrece dos ejemplos a seguir. Las partes podrían inicialmente recibir una posible decisión del sistema inteligente (art. 12) y llegar a un acuerdo, ahorrando tiempo y costas; y también pueden no acordar y pedir la intervención de un juez humano, que debería estar facultado para incentivar el uso del sistema inteligente, sancionando a la parte que bloqueo el acuerdo, si este era razonable (art. 17).

La posibilidad de un recurso con una decisión judicial tradicional siempre deberá existir para satisfacer todas las garantías. Para viabilizar esta instancia es necesario que los sistemas inteligentes tengan la transparencia necesaria.

4. Argumentos para concluir

a) La primera complicación es encuadrar todos los desarrollos que se llaman inteligencia artificial, ya que se ha convertido en un tipo de “apetencia” o “aspiración”; ya sea por estar de moda o por la disponibilidad de fondos para financiamiento. De allí que cualquier cosa que se hace con una computadora tiene chance de ser “Inteligencia Artificial”. Esto se complica por no tener la “Inteligencia Artificial” una metodología clara, única y por ser en la práctica un arte que pocos logran alcanzar con éxito.

b) Razonar –ética o legalmente– en torno a la “Inteligencia Artificial”, en particular la que se funda en redes neurales, es complejo. La caja negra existe¹¹⁵, aun

¹¹³ Esto explica la tendencia actual de reemplazarla por una expresión más sobria, como “sistemas inteligentes”. No obstante, la distinción entre sistemas inteligentes y IA es aún controvertida.

¹¹⁴ Según Jakob von Uexküll cada organismo biológico tiene su propio y único “*umwelt*” (mundo circundante), donde la conciencia y la percepción son una función de la naturaleza del propio organismo (Uexküll, *A Foray into the Worlds of Animals and Humans*, 2010).

¹¹⁵ *Google’s submission*, 2020, p. 21, dice: “es una falacia que los sistemas de IA sean cajas negras. Con suficiente esfuerzo y las herramientas adecuadas, es posible obtener una idea de por
Ciclo “Sistemas Inteligentes y Ética” 78

cuando la opacidad sea un tema de costos o tiempos para auditar; en este caso es definitivamente mucho más complejo. En la práctica habría que repetir todo el proceso y auditar los datos que alimentaron el aprendizaje, pues, aunque muchos desarrollos funcionan satisfactoriamente, nunca está claro –para todos los públicos, o puntos de vista– como se llega a esos resultados.

c) Otro aspecto que perturba una opinión apaciguada es la “singularidad tecnológica”¹¹⁶. Aun cuando no sea plausible tal hipótesis, muchos tienen la sensación que los aspectos ocultos, no explícitos, casi mágicos, que se predicen de la “Inteligencia Artificial” podrían sobrepasar las prevenciones. Uno de los elementos más interesantes del documental *The Social Dilemma* –producido por Netflix– es la idea que ya estamos dentro de la singularidad tecnológica, y que la peculiaridad es que muchos esperaban un ROBOT dictador esclavizando humanos, cuando en realidad se trata de un puñado de humanos, que –usando IA– han logrado manipular a cientos de miles de personas¹¹⁷.

d) La variable tiempo también agrega complejidad. No es igual evaluar y aceptar una metodología como el cálculo infinitesimal, en un intervalo de al menos 250 años, que una opción como la “Inteligencia Artificial” que va pisando los talones de los profesionales (que ven amenazada su competencia –basada en conocimientos– y les asusta tener que adaptarse a una biblioteca ubicua y sabionda que resulta mejor que ellos), y la sociedad (que ve resultados económicamente atractivos y al mismo tiempo observa atónita episodios inimaginables, como el ocurrido con *Cambridge Analytica*)¹¹⁸.

e) Un aspecto adicional –preocupante– es que la protección de los datos personales no termina de cuajar en la sociedad moderna, muchos renuncian a sus datos personales en un solo *click*, sin pensar, ansiosos y embelesados por las primicias de la conectividad en red. Por otra parte, las leyes específicas de protección de datos no logran asentarse, deben más bien adaptarse o convivir con el poder de los *large entropy reducers*¹¹⁹. Al mismo tiempo las autoridades de protección de datos en

qué cualquier sistema de inteligencia artificial se comporta de cierta manera” ... “Existe el riesgo de que los usos innovadores de la IA se impidan inadvertidamente al exigir que los sistemas de IA cumplan con un ‘estándar de oro’ de explicabilidad que supere con creces lo requerido por los enfoques establecidos sin IA (incluidos los basados en humanos)”.

¹¹⁶ Ryan Daws, *Musk predicts AI will be superior to humans within five years*, AINews, 28/7/20.

¹¹⁷ *The Social Dilemma*, es una recopilación de entrevistas a un conjunto de profesionales arrepentidos de haber trabajado para empresas de tecnología, mayormente gestoras de redes sociales.

¹¹⁸ Pablo Guimón, *Cierra Cambridge Analytica, la empresa clave en el escándalo de Facebook: la pérdida de clientes y los altos costes jurídicos relacionados con la filtración de datos personales desde la red social han provocado la decisión de cerrar la compañía*, “El País”, 3/5/18.

¹¹⁹ Según Ohm, p. 1760, 2010, “Los grandes reductores de la entropía son entidades que acumulan bases de datos masivas que contienen tantos enlaces, entre tantos tipos dispares de información, que representan una parte significativa de las bases de datos que arruinarían a cualquier persona [chantajear, acosar, difamar, encasillar o discriminar]. En la práctica, esta categoría incluye grandes agencias de crédito como Experian, TransUnion y Equifax; o proveedores de búsqueda en Internet como Google, Microsoft y Yahoo”.

muchos países no son más que una oficina pública (administrativa), que se encuentra bombardeada por las nuevas estrellas del momento: el *big data*, los datos abiertos, la transparencia, la minería de datos, los perfiles personales, la video-vigilancia.

5. Conclusiones

Es esperable que la IA tenga un impacto especialmente preocupante sobre el conocimiento humano. Es cierto que el método científico no es la única forma de conocimiento; pero sin duda la IA cambiarán la ecuación final de la adquisición de conocimiento. Es mucho más barata y rápida, que la investigación científica y probablemente más efectiva. El problema es que se renuncia a “saber porque” y se acelera “sabiendo cómo” (conocimiento artificial); un nuevo equilibrio con efectos difíciles de pronosticar¹²⁰.

Los niveles de congestión e ineficiencia de la administración de justicia llevarán de alguna forma a incorporar IA. La solución que parece más viable –sin cambios constitucionales– es una forma de conciliación entre las partes guiada por un sistema inteligente, capaz de procesar los hechos y proponer soluciones dando también fundamentos¹²¹. Sin embargo, para que esto funcione –al igual que la *Ley Baidinter*– deben crearse –prudentemente– incentivos económicos que desalienten la solución de los conflictos por jueces humanos¹²².

Resulta necesario perfeccionar una ética de la percepción (como mejorar y adecuar la legislación específica), en particular cuando se recolecten y procesen datos personales. Existe al presente mucha tensión entre las normas de protección de datos personales y los desarrollos tecnológicos, pero ciertamente es más un tema de conciencia pública que de las normas.

La IA será sin duda un nuevo desafío para la sociedad humana sobre cómo administrar el futuro. Es evidente que algunos contextos anteriores son lamentables: como la esclavitud, el uso de recursos naturales –como el petróleo o el litio–, o desarrollos científicos –como los aerosoles–, muestran que los intereses económicos re-

¹²⁰ El conocimiento científico incluye una hipótesis, una historia experimental, una fundamentación; es decir no es solo el conocimiento aplicable, es mucho más. El “conocimiento artificial” es más rápido, más barato y probablemente igual de útil, pero no es igual.

¹²¹ Aun en el caso en que el juez solo busque en un sistema inteligente opciones, o asesoramiento, sobre como decidir un caso, mantener la independencia sería difícil. Un gesto ético –en esta situación– supondría compartir con los abogados de las partes las sugerencias que obtuvo del sistema inteligente.

¹²² Esta solución –como cualquier otra que obvie la decisión de un juez humano– tiene un costo para la creación del derecho. En los últimos años el valor de la jurisprudencia como fuente de derecho está creciendo aún fuera del *common law*, probablemente porque estamos en una sociedad más compleja, con perturbaciones tecnológicas y por la dificultad de una reacción legislativa rápida. La jurisprudencia se alimenta de conflictos reales resueltos por jueces reales, si éstos no llegan al menos a la primera instancia, recrudescerán las nuevas líneas jurisprudenciales. Una alternativa es que el sistema inteligente perciba que para algunos conflictos –por su tipicidad– sea conveniente derivarlos a una decisión por parte de un juez humano. Ver: Mary Ann Glendon, *The Sources of Law in a Changing Legal Order*, 17 Creighton Law Review, 1984, p. 663 a 698.

basan la prudencia y ciegan sobre cómo imaginar los escenarios futuros. El *Llamado desde Roma por una ética de la Inteligencia Artificial (Rome Call for an AI Ethics)* – entre otros–es insistente en la necesidad de una IA que afirme la magnificencia de los seres humanos y que garantice cooperativamente un futuro para las generaciones más jóvenes.

En este futuro, el escenario más débil es para los trabajadores, ya que se les obligará a re-adaptarse, re-definirse y re-construirse, prácticamente solos y en tiempos brevísimos. Los problemas ya conocidos de los marginados laborales han sido enfrentados con programas de reeducación; los resultados son aún poco satisfactorios (Yu-Qian Zhu & al., 2020).

Finalmente, todo se centra en la vigencia de los Derechos Humanos¹²³, que aquí enfatizamos concebidos como una protección para los seres humanos ante cualquier forma o concentración de poder. Que los GAFAM tengan hoy ubiquidad sin que nadie los regule en su totalidad (las leyes antimonopolio no parecen ser las más efectivas) parece ser un factor determinante¹²⁴. El poder informal y económico que han alcanzado supera la media de los países. Todo este escenario hace pensar que será bastante difícil una regulación efectiva.

Referencias

- Asúnsolo Morales, Carlos R., *Los derechos humanos como límites al poder económico en el contexto de globalización*, Bogotá, “Anuario de Derecho Constitucional Latinoamericano”, año XXIII, 2017, p. 21a 37.
- Brockman, John (ed.), *What to think about machines that think: today’s leading thinkers on the age of machine intelligence*, Harper Perennial, 2015.
- CNIL, *Le Conseil d’État demande au Health Data Hub des garanties supplémentaires pour limiter le risque de transfert vers les États-Unis*, 14/10/20.
- Estévez, Elsa - Linares Lejarraga, Sebastián - Fillottrani, Pablo, *Prometea, transformando la administración de justicia con herramientas de inteligencia artificial*, Banco Interamericano de Desarrollo, 2020.
- Google, *Artificial intelligence at Google: our principles*.
- Google’s submission, Consultation on the white paper on AI - a European approach, 28/5/20.

¹²³ Los enunciados éticos son de gran ayuda –pero aun así, son muy poco participativos–, ayudan en la medida que crean conciencia y promueven los desarrollos legislativos; finalmente será necesario un consenso global sobre las normas regulatorias.

¹²⁴ “GAFAM”, acrónimo utilizado para referirse a Google (Alphabet), Amazon, Facebook, Apple y Microsoft y similares.

- Haenssle Holger A., *Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists*, "Annals of Oncology", n° 29, 2018, p. 1836 a 1842.
- Ko, Jina - Baldassano, Steven N. - Po-Ling Loh - Kording, Konrad - Litt, Brian - Issadore, David, *Machine Learning to detect signatures of disease in liquid biopsies - a user's guide*, Lab Chip, 2018, 18(3), p. 395 a 405.
- Lehoux, Pascale - Hassane, Alami - Mörch, Carl - Lysanne, Rivard, Robson, Rocha - Hudson, Silva, *Peut-on innover de façon responsable en temps de pandémie? L'intelligence artificielle et le numérique face au Sars-Cov-2*.
- Nadler, Jerrold (chairman), *Investigation of competition in digital markets*, U.S. House of Representatives, 2020.
- Ohm, Paul, *Broken promises of privacy: responding to the surprising failure of anonymization*, "UCLA Law Review", vol. 57, 2010, p. 1701, <https://ssrn.com/abstract=1450006>.
- Park, Joshua, *Your honor, AI*, "Harvard International Review", 3/4/20.
- Red Iberoamericana de Protección de Datos, *Recomendaciones Generales para el Tratamiento de Datos en Inteligencia Artificial*, 2019.
- Reiling, Dory, *Courts and Artificial Intelligence*, 11(2) International Journal for Court Administration, 2020, p.8.
- Rome Call for an AI Ethics, 28/2/20.
- Rudin, Cynthia - Wang, Caroline - Coker, Beau, *The age of secrecy and unfairness in recidivism prediction*, Harvard Data Science Review, 2.1, 2020.
- Sarafian, Jean-Marc, *La révolution opérée par la procédure d'offre*, L'argus de l'assurance, 2015.
- Strogatz, Steven, *Infinite Powers: how calculus reveals the secrets of the universe*, Houghton Mifflin Harcourt, 2019.
- Swetz, Frank J., *Mathematical Treasure: Leibniz's Papers on Calculus - Integral Calculus*, Convergence (Mathematical Association of America), 2018.
- Yu-Qian Zhu - Corbett, Jacqueline - Yi-Te Chiu, *Understanding employees' responses to artificial intelligence*, Organizational Dynamics, 2020.



El derecho humano a las decisiones jurídicas humanas

Por Emilio Suñé Llinás¹²⁵

Declaración de Derechos del Ciberespacio¹²⁶ (Emilio Suñé, 2008)

Artículo 15. Dignidad de la persona.

1. La dignidad del ser humano, en su significado más profundo de inviolabilidad de la persona, es un metaderecho, presente en buena parte de los Derechos Fundamentales, incluidos algunos de los aquí enunciados y como tal metaderecho debe ser protegido incluso frente al consentimiento del propio afectado.

¹²⁵ Catedrático de Derecho Informático, Universidad Complutense de Madrid.

¹²⁶ Suñé Llinás, E. (coord.), *La Constitución del Ciberespacio*, México, Porrúa, 2015.

2. Se prohíbe que cualquier decisión que afecte a una persona humana se adopte exclusivamente por medio de sistemas automatizados, sin intervención de otra persona humana con potestad decisoria para cada caso concreto.

1. La denominada justicia robótica

Parafraseando los criterios de la Curva de Gartner, estamos en fase de expectativas tan sobredimensionadas como exageradas hasta rayar en lo insólito, en el ámbito de la inteligencia artificial asociada a la robótica. Lo importante –prefiero quedarme con lo positivo– es que tanto en el sector privado, como también en el sector público existe ya conciencia de la importancia de adoptar sistemas inteligentes en ámbitos jurídicos, tanto en el Poder Ejecutivo, como en la esfera del Poder judicial, conforme veremos acto seguido.

El problema es que estamos en un mundo de expertos de última hornada, empeñados en una absurda carrera por descubrir la pólvora. Ello nos sitúa ante un cambio constante de palabras, para expresar realidades añejas como si fueran nuevas: que si *Uberización del Derecho*, antes *E-Derecho*, ahora *Legal Tech*, ¡como si el mundo anglo-norte-americano fuera un referente en Derecho Informático, mayor que nosotros mismos! Otra cosa es que ellos sean los más avanzados en innovaciones tecnológicas, pero no necesariamente en Derecho Informático, precisamente porque a ellos les interesa la desregulación, que les permite afianzar mejor su dominio tecnológico y económico. En todo caso, son multitud de expresiones cambiantes, para llegar con cada palabra a lo mismo; es decir, al Derecho Informático (Informática Jurídica y Derecho de la Informática) en el que algunos somos expertos desde principios de los 80, como en mi caso, o antes, como en el del Prof. Martino. Incluso el *Derecho de las Nuevas Tecnologías*, que es una expresión razonable, se ha empleado para eso (por cierto, la actual expresión *Legal Tech* no me suena muy diferente de la añeja *Derecho de las Nuevas Tecnologías*). De hecho, con esta última expresión, introducida en el nuevo Plan de Estudios del Grado en Derecho UCM, no se ha buscado otra cosa sino repartir entre cinco departamentos el Derecho Informático, que fue mi equipo quien introdujo en la Universidad Complutense, dentro del Área de conocimiento de Filosofía del Derecho a caballo entre los 80 y los 90, afirmando sin pestañear que Derecho de las Nuevas Tecnologías es otra cosa distinta del Derecho Informático, cuando lo único que hay son matices: puede ser ligeramente más, al incluir algunas cuestiones, sobre todo de Derecho de la Biotecnología en la vertiente regulativa del estricto Derecho Informático y bastante menos, al prescindir de la vertiente tecnológica del Derecho Informático, es decir, de la Informática Jurídica que, desde el principio, ha englobado tres cuestiones básicas: Bases de Datos Jurídicas, Informática Jurídica de Gestión, e Inteligencia Artificial aplicada al Derecho. Lo que realmente se está jugando con la magia de las palabras, son cuestiones mucho más terrenales de poder, en este caso de poder académico.

Acabo de referirme a la IA aplicada al Derecho, como una cuestión *clásica* de la Informática Jurídica, que ya existía cuando inicié mis estudios de dicha disciplina, en 1983. Y muchos la pretenden presentar como si fuese una cuestión surgida en el presente. La introducción de la IA en la resolución de casos litigiosos –la Informática Decisional Jurídica– es algo que, entre los expertos que lo son de verdad, es bastan-

te añeja y tradicionalmente no ha planteado grandes debates. No se trata de que un *juez robot* inteligente sustituya al operador humano. De hecho, no hace falta, no hay ventaja alguna en eso, ni siquiera en términos de productividad hay una ventaja apreciable. Siempre se ha tratado de tener sistemas lógicos e inteligentes de apoyo a procesos de decisión jurídica, que antaño nadie dudaba que habría de ser adoptada finalmente por un ser humano. ¿Tanto se reduce la productividad, al establecer una razonable supervisión humana, o lo que se pretende es dejar el Derecho y la propia justicia en manos de algoritmos supuestamente neutrales que alguien pueda manejar a su libre albedrío, o lo que es lo mismo, de forma arbitraria?

Plantearse un juez estrictamente robótico en los años 80 y hasta final de siglo se veía con escándalo, hasta el punto de que personas que, por su conocimiento de la IA, tenían claro que nos hallábamos ante sistemas muy sofisticados de apoyo inteligente a las decisiones jurídicas, que serían en última instancia humanas, no se atrevían a decir que los entonces *sistemas expertos* jurídicos tenían una potencialidad de adoptar por sí mismos decisiones jurídicas, por más que no se plantease utilizarlos de esta manera, sino, como antes se decía, como sistemas super sofisticados de apoyo lógico a las decisiones que, en última instancia, seguirían siendo humanas. Esto lo vi en la Tesis Doctoral de la Prof^a Nuria Cuadrado¹²⁷, sobre la materia, que me correspondió dirigir, y que se presentó en 1999, para ser publicada algo después, en 2004. La doctoranda, fue muy cautelosa en su exposición, para evitar herir hipotéticos prejuicios de los miembros del Tribunal, que sin embargo habían sido muy bien escogidos. La mayoría de ellos habían escrito sobre Derecho Informático, hasta el punto de que mi Maestro y Director de Tesis, el Profesor Villar Palasí, animó a expresarse a la todavía doctoranda: *Atrévase usted a decir lo que piensa, que estos sistemas son potencialmente decisivos*. Potencialmente... porque nadie se planteó nunca sustituir al operador jurídico humano, por la sencilla razón de que lo que se puede, no es lo que se debe; porque una cosa es lo que *son* las cosas y otra lo que *deben ser* (ser y deber ser) y, en última instancia, porque no hay que confundir el poder con el Derecho.

A día de hoy parece que los que seguimos diciendo lo mismo que a finales del pasado siglo, somos poco menos que ignorantes retrógrados, frente a mentes iluminadas, que a veces ni siquiera son juristas, que confunden sus particulares intereses –y a fin de cuentas lo que les gustaría–, con lo que realmente son las cosas.

2. La justicia robótica en el procedimiento administrativo

Los robots pueden utilizarse no sólo en el ámbito del Poder Judicial, sino también en procedimientos administrativos. Es ya un clásico la circulación vial, donde ya se emplean desde hace años y ni siquiera son robots inteligentes, como en el caso de esta factoría de ingresos que son las multas por exceso de velocidad.

¹²⁷ Cuadrado Gamarra, N., *Aplicación de los Sistemas Expertos en el Campo del Derecho*, Madrid, Servicio de Publicaciones de la Facultad de Derecho de la Universidad Complutense, 2004.

También Hacienda –otra factoría de ingresos– se plantea incorporar *funcionarios robots* en su ámbito propio, con idéntico objetivo de incrementar los ingresos del Estado.

En el supuesto paradigmático de la seguridad vial, me extenderé un poco más, para recordar el art. 15.2 de mi Declaración de Derechos del Ciberespacio (2008).

Se prohíbe que cualquier decisión que afecte a una persona humana se adopte exclusivamente por medio de sistemas automatizados, sin intervención de otra persona humana con potestad decisoria para cada caso concreto.

Mayor mérito –al ser anterior– tiene la primera Directiva Europea de Protección de los Derechos Humanos, la Directiva 95/46 (1995), que dice:

Artículo 15. Decisiones individuales automatizadas.

1. Los Estados miembros reconocerán a las personas el derecho a no verse sometidas a una decisión con efectos jurídicos sobre ellas o que les afecte de manera significativa, que se base únicamente en un tratamiento automatizado de datos destinado a evaluar determinados aspectos de su personalidad, como su rendimiento laboral, crédito, fiabilidad, conducta, etc.

Inciso: No se olvide que dicho artículo aplica *el derecho humano a las decisiones jurídicas humanas* a un ámbito específico, cual es la protección de datos personales; pero si prescindimos de estas referencias específicas a su propio ámbito, observaremos que no estamos sino ante una concreta especificación del principio que después expresé en el art. 15 de mi Declaración de Derechos del Ciberespacio.

2. Los Estados miembros permitirán, sin perjuicio de lo dispuesto en los demás artículos de la presente Directiva, que una persona pueda verse sometida a una de las decisiones contempladas en el apartado 1 cuando dicha decisión:

a) se haya adoptado en el marco de la celebración o ejecución de un contrato, siempre que la petición de celebración o ejecución del contrato presentada por el interesado se haya satisfecho o que existan medidas apropiadas, como la posibilidad de defender su punto de vista, para la salvaguardia de su interés legítimo; o

b) esté autorizada por una ley que establezca medidas que garanticen el interés legítimo del interesado.

En el Derecho Europeo vigente, corresponde citar el Reglamento 2016/679, también de Protección de Datos Personales, que es de aplicación directa en todos los Estados de la UE. El conocido como Reglamento General de Protección de Datos o RGPD, que se refiere a la cuestión de las decisiones individuales automatizadas en los arts. 21 y 22, sobre todo en este último:

Artículo 22. Decisiones individuales automatizadas, incluida la elaboración de perfiles.

1. *Todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar.*

2. *El apartado 1 no se aplicará si la decisión*

a) es necesaria para la celebración o la ejecución de un contrato entre el interesado y un responsable del tratamiento;

b) está autorizada por el Derecho de la Unión o de los Estados miembros que se aplique al responsable del tratamiento y que establezca asimismo medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado, o

c) se basa en el consentimiento explícito del interesado.

3. *En los casos a que se refiere el apartado 2, letras a) y c), el responsable del tratamiento adoptará las medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado, como mínimo el derecho a obtener intervención humana por parte del responsable, a expresar su punto de vista y a impugnar la decisión.*

4. *Las decisiones a que se refiere el apartado 2 no se basarán en las categorías especiales de datos personales¹²⁸ contempladas en el artículo 9, apartado 1, salvo que se aplique el artículo 9, apartado 2, letra a) o g), y se hayan tomado medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado.*

El art. 22 del RGPD es un precepto poco aplicado hasta el presente y ello a pesar de que, como acabamos de ver, no representa una innovación de fondo del ordenamiento de la UE, en materia de protección de datos personales, pues en términos semejantes estaba ya redactada la Directiva 95/46, salvo en un extremo importante y es que la Directiva 95/46 ponía límites a las decisiones individuales automatizadas cuando se trataba de obtener perfiles personales, o al menos esta era la interpretación más plausible de la misma, lo que eventualmente podría limitar los supuestos en los que sería de aplicación. En cambio, el nuevo RGPD no vincula necesariamente las decisiones individuales automatizadas a cuestiones de perfil personal. Sin necesidad de otro tipo de interpretaciones, en su tenor literal queda claro que la prohibición de las decisiones individuales automatizadas, va más allá de las estrictas cuestiones de perfil personal.

Los operadores jurídicos, incluidos los jueces, tradicionalmente han sido bastante legos en Derecho Informático, por la sencilla razón de que no ha estado presente en su formación como juristas; al menos hasta hoy, cuando con la irrupción social de la IA asociada a la robótica, a tan gran escala que no puede ser ignorada por más tiempo, junto con la consiguiente fase de expectativas sobredimensionadas inherente a la curva de Gartner, preceptos como el referido se han puesto ya en primer plano, sobre todo cuando son cruciales para resolver casos, como el que se

¹²⁸ Se refiere a los ordinariamente conocidos como *datos sensibles*.

plantea en la última parte de este artículo, en relación con los algoritmos de Uber, para controlar a sus conductores.

Siempre me pregunté qué hubiera sucedido, de haberse invocado este precepto frente a sanciones de los consabidos *radares recaudatorios* de la DGT en España –o situaciones análogas de otros países de la UE– en los que con el pretexto legal o reglamentario, de que no es posible que un agente pare al supuesto infractor, para comunicarle la asimismo supuesta infracción y correspondiente denuncia, en la práctica lo que sucede es que la información del radar es transmitida telemáticamente a un centro de proceso de datos, que en España radica en León, donde por procedimientos a su vez telemáticos, sale ensobrada la notificación de la denuncia al afectado, sin intervención humana de ningún tipo. Y ¡encima! nos hallamos ante un sistema informático tradicional. Ni siquiera ante un proceso inteligente de datos. ¿No es eso una decisión individual automatizada, basada en meros pretextos normativos (como que no exista la posibilidad de parar al conductor), con una finalidad que pocas veces se compadece con la seguridad vial y casi siempre con un afán meramente recaudatorio? Probablemente no sea tanto a la mayor parte de los conductores, a quienes haya que *poner freno*, sino a las prácticas viciadas de la DGT, y a fin de cuentas al abuso o desviación de poder, que se produce tanto en España, como en buena parte de los países de la Unión Europea, que por si fuera poco colaboran entre sí, para la cobranza de tales *sanciones recaudatorias*.

Es más, se infringe también el art. 22 del RGPD, porque en la que en ocasiones he denominado *deriva autoritaria* de los sedicentes Estados democráticos de nuestros días, se penaliza la posibilidad de recurrir, pues si se paga la multa en un plazo que en España es de 20 días, sin recurrirla, el importe se reduce en un 50%, mientras que si es recurrida no hay ninguna posibilidad de reducción. Lo que se presenta como bonificación es en realidad un impedimento al recurso, pues es exactamente lo mismo bonificar un 50% si se paga y no se recurre, que incrementar un 100% si se recurre. Es pura *magia de las palabras*, con la conclusión, radicalmente opuesta a la pedagogía democrática y del *Rule of Law*, de que el súbdito que se atreve a ser ciudadano y en consecuencia ejercita sus derechos, es penalizado por ello, para que interiorice que cada día más la condición de ciudadano se aproxima a la retórica, mientras que la condición de súbdito es la que se acerca a la realidad.

He puesto sólo un ejemplo de un precepto como el de la prohibición de las decisiones individuales automatizadas, al que se podría sacar mucho partido en favor de las libertades ciudadanas, si fuera debidamente invocado por los particulares y aplicado por las autoridades nacionales, incluidas las competentes en materia de protección de datos personales. Sin embargo, es una cuestión poco estudiada, salvo por algún experto de gran solera, como mi amigo y compañero de largo recorrido, Antonio Anselmo Martino¹²⁹.

3. Justicia robótica o predictiva y proceso judicial

¹²⁹ Martino, Antonio A., *La decisión jurídica automatizada*, en Bauzá Reilly, M. (dir.), “El Derecho de las TIC en Iberoamérica”, FIADI, 2019.

Frente al práctico consenso de los expertos en Derecho Informático, que desde siempre veíamos la Informática Decisional Jurídica, como una herramienta super sofisticada –al ser potencialmente decisional–, de apoyo a una decisión jurídica, que finalmente adoptaría un operador humano, algunos y hasta bastantes Estados ya han introducido o se plantean introducir *jueces robot* que directamente sustituyan, en determinados asuntos, al operador humano. En otras palabras, es la justicia robótica propiamente dicha, en el ámbito del Poder Judicial, que al adoptar decisiones con criterios algorítmicos estandarizados, se denomina también *justicia predictiva*. A ella me referí en mi reciente libro *Derecho e Inteligencia Artificial*¹³⁰, en los siguientes términos:

Por centrarnos en Europa, Estonia es el abanderado de este tipo de cuestiones, lo que no es de extrañar, al tratarse de un Estado ex soviético, mal que les pese. De 1940 a 1991: 51 años, o lo que es lo mismo, generación y media. Siempre dije que hay cosas que imprimen carácter, más incluso que un uniforme... o una sotana.

Estonia, en su afán de reafirmar su identidad no soviética, ha dado un bandazo desde el comunismo a un capitalismo radical, que entre otras cosas se manifiesta en una apuesta firme por la tecnología, lo que está muy bien y de hecho es un país puntero en E-Government; pero la *ley del péndulo* no debe llegar al extremo de hacer de la tecnología el demiurgo de un mundo perfecto en el que sea posible prescindir incluso de jueces humanos, sobre todo cuando en el subconsciente de este mundo perfecto pueden pervivir *tics* relacionados con la antigua dominación soviética, a partir de los cuales acaso no se tome suficientemente en cuenta la implicación en todo ello de algo tan sagrado como los Derechos Humanos.

En Estonia, según informaciones de prensa relatadas por Joaquín Hernández¹³¹ (2019) los robots decidirán sobre asuntos de cuantía inferior a 6000 euros. Las partes se limitarían a subir la documentación a La Red, para que el juez robot decida. Por supuesto sus resoluciones podrán ser objeto de recurso antes tribunales humanos. Mas no se olvide que, de entrada, hemos prescindido de la prueba en el Juzgado de Instancia y en principio las instancias judiciales superiores tienen carácter revisor de la aplicación del Derecho, a partir de hechos previamente establecidos por el órgano de instancia. En resumen, con la posibilidad de recurso, ya hemos tranquilizado la conciencia frente a la denominada *justicia robótica*, *pero los hechos están fijados*. Y ya se sabe, *da mihi factum, dabo tibi ius*.

Francia se ha lanzado a copiar a Estonia y quiere hacer algo parecido en su proyecto de reforma de la Justicia (2018-2022), también para asuntos de menor cuantía. Se estima que las máquinas podrían resolver 2,7 millones de casos cada año.

¹³⁰ Suñé Llinás, E., *Derecho e Inteligencia Artificial. De la robótica a lo posthumano*, México, Tirant lo Blanch, 2020.

¹³¹ Hernández, J., *Los robots ahorrarían a los jueces españoles una cuarta parte de su trabajo*, 2019, www.vozpopuli.com/economia-y-finanzas/robots-ahorrarian-cuarta-parte-trabajo-jueces-espanoles_0_1260174449.html.

Es aberrante que se esté planteando tanto en Estonia como en Francia –el país más socialista de la UE, con un Estado que representaba el 57% del PIB antes de la pandemia– una *justicia robótica* para pobres y una justicia humana para ricos. Es algo que no tiene nombre y además no resuelve nada. La entrada de la IA en la Justicia, sin distinción de procesos, incrementaría en igual o mayor medida la eficiencia del aparato judicial, si se introdujese en todo tipo de procesos, reservando *siempre* la decisión final al ser humano. Lo demás es discriminatorio, contrario a la dignidad de la persona y profundamente inhumano, aparte de hoy por hoy inconstitucional en la mayor parte de los países... y además, como ya hemos visto, contrario al que hoy se suele denominar, con más o menos acierto, Derecho Fundamental a la Protección de Datos.

En España, el mismo artículo alude a una consultora tecnológica cuyo CEO – a quien no cito de pura estupefacción– normaliza sin pestañear una situación que otros vemos con asombro y rechazo: *lo que se busca es que un conflicto de 200 euros no frene la investigación de uno de 10 millones, y los robots son una vía para ello*. Ya se ve cómo queda la *justicia igual para todos*, en esta peculiar postmodernidad.

Asimismo, se pone de relieve que el Consejo General del Poder Judicial ya se ha planteado la cuestión en España, donde se estima que de los 6 millones de asuntos que se resuelven cada año, una cuarta parte serían susceptibles de *soluciones de automatización*.

El propio Joaquín Hernández, en otro artículo¹³² sobre la que denomina *justicia predictiva* desgrana las ventajas e inconvenientes de la *justicia robótica*. Entre las primeras destaca la *eliminación de la incertidumbre jurídica*, la *reducción de burocracia*, sobre todo en asuntos que *atascan* los juzgados, la *rebaja de costes* y un *incremento de las conciliaciones*, como alternativa a los juicios. Entre los inconvenientes, cita el *viraje del “Imperio de la Ley” a la “norma aplicada”*, lo que acabaría reduciendo la necesidad de argumentar las resoluciones y crearía un *efecto rebaño* a partir del no cuestionamiento de los algoritmos de IA. Obvio es decir que estos planteamientos parten de una mentalidad que nos resulta extraña a los especialistas acreditados y que es la sustitución del juez por el robot cuando, lo reitero una vez más, el planteamiento debiera ser que la IA fuese una herramienta muy sofisticada al servicio *siempre y en todo caso* de un juzgador humano. En lo que más útil le puede ser, por ejemplo un SE al juez, es en todo lo contrario de lo que se está postulando; me refiero a hacer transparente la lógica del razonamiento jurídico, que muchas veces en las resoluciones judiciales del presente –no apoyadas por la IA– la falta de tiempo del juzgador hace que estén poco o deficientemente argumentadas. Y por supuesto la productividad del trabajo judicial, que al estar simplificada por la IA, permitiría multiplicar por un factor probablemente de 2 dígitos, el número de resoluciones judiciales dictadas; pero siempre manteniendo la decisión final humana.

¹³² Hernández, J., *El Poder Judicial estudia implementar robots que predigan sentencias*, 2019, www.vozpopuli.com/espana/poder-judicial-estudia-implementar-algoritmos-predigan-sentencias_0_1256574916.html.

Otro inconveniente, que se cita en el periódico Vozpopuli y también en The-technolawgist¹³³, es que la justicia predictiva puede poner de manifiesto los sesgos de cada juez, lo que al parecer Francia se plantea prohibir, e incluso los sesgos que pueden estar presentes en los propios algoritmos de IA, que podrían afectar a las minorías y a los más desfavorecidos, pues en palabras de Cathy O'Neil, los algoritmos suelen contener *opiniones encerradas en matemáticas*.

Se conocen los prejuicios raciales; pero no es tanto esto como el escaso CI de quien diseña sistemas tan grotescos. Como lo que se acaba de publicar este mismo mes de noviembre de 2020, en el que se redactan estas líneas, sobre una IA de seguimiento video de partidos de fútbol, que en vez de centrarse en el movimiento del balón, se dedicó a perseguir la calva del árbitro.

En estas apresuradas coordenadas, que en general no se plantean a partir de una reflexión pausada, incluso algún magistrado *habla por hablar* con afirmaciones tales como: *donde el robot no puede entrar es en la apreciación de pruebas, el componente subjetivo, pero de lo otro hay una enorme cantidad de asuntos pequeños que no permiten dar respuesta a asuntos de gran envergadura*. De nuevo lo que se pone sobre la mesa es una *justicia paralela robótica para pobres*, con una frivolidad espantosa, al tiempo que se habla de lo que no se sabe. Un sistema de IA puede entrar en la apreciación de la prueba perfectamente. Que nos hallemos ubicados en un sistema probatorio no jerárquico basado en la libre valoración de la prueba por parte del juzgador, no significa que se dé pie al subjetivismo. Y eso un juez debiera saberlo. Sólo se trata de no hacer reglado algo que por su propia naturaleza es muy difícil de regular, como establecer que determinado tipo de prueba “vale más” que otra, como sucede en los sistemas de jerarquía probatoria; pero el juez puede actuar con criterios lógicos, estadísticos y de experiencia, que deben ser siempre objetivos, no subjetivos, y por supuesto formalizables en sistemas de IA.

Junto a este tipo de afirmaciones, *lanzadas al voleo*, otros magistrados tienen las ideas mucho más claras, como Raimundo Prado, que hace planteamientos razonables y certeros, cuando da por bienvenida *cualquier ayuda tecnológica que nos ayude a hacer mejor nuestro trabajo*, o cuando concluye, con un rigor jurídico impecable, que *son las personas las que juzgan, se trate del asunto que se trate, porque así lo dice la Constitución. Para que la palabra final sobre un asunto jurídico, sea el que sea, la tenga un robot y no una persona, tendría que llevarse a cabo no sólo una reforma legal, sino constitucional*. Y yo voy todavía más allá: al estar en presencia de cuestiones profundamente vinculadas a la dignidad de la persona y de Derechos Humanos básicos, casi seguro que esa reforma constitucional daría lugar a una *norma constitucional inconstitucional*, porque afecta a la arquitectura básica de cualquier constitución moderna, que son los Derechos Fundamentales.

Es la idea que expresó por vez primera Otto Bachoff¹³⁴ en una conferencia pronunciada en 1951 y reiterada en obras posteriores, en cuanto a un derecho su-

¹³³ www.thetechnolawgist.com/2019/07/18/espana-considera-la-introduccion-jueces-robot-en-la-prediccion-de-sentencias/?fbclid=IwAR12THQBHv62gtl4y-nWr70Uu4Z9YJwMfvh4yLqPKGHH-9MdMxSNrqEjigg.

¹³⁴ Bachof, O., *¿Normas constitucionales inconstitucionales?* Lima, Palestra, 2010.

pralegal, que puede estar escrito en la Constitución, como los valores superiores del art. 1 de la Constitución Española (CE) y algunos derechos o metaderechos más, como la dignidad y la libertad, e incluso los Derechos y libertades susceptibles de amparo constitucional.

A mi juicio, ninguna Constitución de un país libre puede contener disposiciones contrarias a la dignidad de la persona, o a su libre autodeterminación, porque son presupuestos de los demás Derechos Fundamentales. La mejor definición de dignidad es *inviolabilidad de la persona*, que entre otras cosas significa que una persona sólo puede ser tratada como una persona, no como un número o una cosa. Todo ello, a fin de cuentas, se resume en un aforismo: *homo res sacra homini*. Como decía en mi Declaración de Derechos del Ciberespacio (2008), la dignidad es un *metaderecho* que, en ocasiones, debe ser protegido incluso frente al consentimiento del propio afectado. Nadie puede consentir en ser esclavo, por ejemplo, porque dicho estado, que consiste en tratar a una persona como una cosa, atenta frontalmente contra la dignidad de la persona humana.

La dignidad, al igual que la libertad (libre autodeterminación de la persona) son, como se ha dicho, presupuestos de los propios derechos fundamentales. Así lo entendió el Tribunal Constitucional Alemán, en su famosa sentencia de 1983, sobre la Ley del Censo, cuando estableció el derecho a la libre autodeterminación informativa, que es el derecho fundamental a la protección de datos personales, sin necesidad de que la Constitución Alemana, la Ley Fundamental de Bonn de 1948, contemplase siquiera expresamente el derecho a la intimidad, porque éste era un derivado de la dignidad de la persona, asociada a su libertad o libre autodeterminación, en base a los cuales construía el Tribunal el Derecho Fundamental a la libre autodeterminación informativa, de manera semejante a lo que hicieron Warren y Brandeis¹³⁵, en su conocida obra *The Right to Privacy* (1890), cuando sostenían que la intimidad estaba reconocida en el *common law*, porque la jurisprudencia ya había postulado que detrás de otros derechos no legislados, como el derecho moral de autor, estaba la dignidad de la persona, que asimismo fundamentaba su derecho a la intimidad y que estaba recogida previamente por la jurisprudencia, como base de otros derechos no legislados, como el referido derecho moral de autor y más en concreto el derecho a decidir sobre la publicación de los propios escritos, cuyo entronque con la intimidad no puede ser más evidente.

Una justicia robótica independiente, no meramente auxiliar de la justicia humana, supone tratar a la persona como a un número o una cosa y violarla en lo más íntimo y sagrado, en su dignidad. *Homo res sacra homini* que, como ya se ha dicho, es también un aforismo que refleja el significado profundo de la dignidad de la persona humana. De ahí el título de la ponencia: *El Derecho Humano a las Decisiones Jurídicas Humanas*, que no es sino un Derecho Humano sólidamente enraizado en la dignidad de la persona humana.

Algo muy semejante a lo que aquí se expone es lo que expresa el art. 10 de la Constitución Española, cuando dice:

¹³⁵ Warren, S. - Brandeis, L.D., *El derecho a la intimidad*, Madrid, Civitas, 1995.

1. La dignidad de la persona, los derechos inviolables que le son inherentes, el libre desarrollo de la personalidad, el respeto a la ley y a los derechos de los demás son fundamento del orden político y de la paz social.

Nótese que la Constitución Española, al igual que yo mismo, pone la dignidad de la persona y el libre desarrollo de la personalidad, *junto a* los derechos inviolables de la persona y no *dentro de* los mismos, porque son un presupuesto de los Derechos Fundamentales, y en el fondo auténticos metaderechos, como ya se ha expresado.

Por lo demás, las normas constitucionales suelen referirse también directamente a una justicia humana. Por seguir con el punto de referencia de la Constitución Española, se citarán dos párrafos del art. 117, aparte de otros preceptos constitucionales:

Artículo 117

1. La justicia emana del pueblo y se administra en nombre del Rey por jueces y magistrados integrantes del Poder Judicial, independientes, inamovibles, responsables y sometidos únicamente al imperio de la ley.

3. El ejercicio de la potestad jurisdiccional en todo tipo de procesos, juzgando y haciendo ejecutar lo juzgado, corresponde exclusivamente a los juzgados y tribunales determinados por las leyes, según las normas de competencia y procedimiento que las mismas establezcan.

El art. 117.1 CE es diáfano, a la hora de exigir que sean personas físicas, los jueces y magistrados, quienes administren Justicia y entiendo que lo mismo sucede en el párrafo tercero de dicho artículo, cuando se refiere a Juzgados y Tribunales, a nivel orgánico, pues una supuesta *justicia robótica* no podría emanar del pueblo en modo alguno y sería incompatible con el principio democrático. Admito no obstante que la letra de la Constitución podría dar lugar a otras interpretaciones que, a mi entender, chocarían de frente con los principios nucleares de cualquier Constitución de un país libre, de cualquier Constitución Democrática, por lo que una tal interpretación no debiera ser nunca admisible y si llegara a legislarse, aunque fuera por el Poder Constituyente, a mi juicio daría lugar a *normas constitucionales inconstitucionales*, impropias de un Estado libre y democrático. Mantener al tiempo una permisón constitucional de la justicia robótica y afirmar la dignidad de la persona, en la propia constitución, es sencillamente contradictorio y habría de prevalecer el principio, el basamento constitucional que significa la dignidad, frente a la permisón de la justicia robótica. Esta es la idea de Otto Bachoff. La alternativa de prescindir de la dignidad de la persona, en el edificio constitucional, no sería aceptable en la Constitución de un país libre.

Otros artículos de la Constitución Española que podrían traerse a colación, son los siguientes, por orden de numeración:

Artículo 106

1. Los Tribunales controlan la potestad reglamentaria y la legalidad de la actuación administrativa, así como el sometimiento de ésta a los fines que la justifican.

Artículo 118

Es obligado cumplir las sentencias y demás resoluciones firmes de los jueces y tribunales, así como prestar la colaboración requerida por éstos en el curso del proceso y en la ejecución de lo resuelto.

Artículo 120

1. Las actuaciones judiciales serán públicas, con las excepciones que prevengan las leyes de procedimiento.

2. El procedimiento será predominantemente oral, sobre todo en materia criminal.

3. Las sentencias serán siempre motivadas y se pronunciarán en audiencia pública.

En este último caso, los principios constitucionales de publicidad, oralidad y pronunciamiento de las sentencias en audiencia pública, son también difícilmente compatibles con una hipotética justicia robótica.

Creo que con lo dicho debiera ser suficiente para dejar esta cuestión zanjada y si no lo fuera, estoy seguro de que en un análisis no ya de Derecho Constitucional Comparado, que sin duda sería muy semejante al del Derecho español, sino de Derecho Internacional Humanitario, hallaríamos en los más importantes Tratados Internacionales, tanto universales como regionales, principios muy parecidos a los que informan el Derecho Constitucional, por la sencilla razón de que son inherentes a cualquier sociedad libre y democrática.

4. Las decisiones jurídicas robóticas en el derecho privado

Todas estas cuestiones han entrado también en el ámbito del Derecho privado. Hay un caso significativo, que se ha planteado hace escasos días que es el de los despidos algorítmicos de Uber¹³⁶. En resumen, el caso se ha planteado cuando cuatro conductores holandeses de Uber han llevado a la compañía a los tribunales laborales, al haber sido despedidos de forma improcedente, según alegan, por un algoritmo de Uber denominado *Mastermind*.

¹³⁶ Faes, I., ¿Despido improcedente? Cuatro conductores llevan a Uber a los Tribunales en Holanda tras despedirles su algoritmo, "El Economista", 31/10/20.

www.economista.es/legislacion/noticias/10860771/10/20/Despido-improcedente-Cuatro-conductores-llevan-a-Uber-a-los-tribunales-en-Holanda-tras-despedirles-su-algoritmo.html.

Al parecer el algoritmo localiza a los conductores que indiciariamente se comportan de manera fraudulenta, por estimar que hay acuerdos fraudulentos entre conductor y pasajero, por cuestiones relacionadas con el sistema de pagos, etc.

Fue el propio algoritmo el que canceló automáticamente sus cuentas con Uber, sin que se les diera ningún derecho a defender su punto de vista (en otras palabras, *inaudita parte*), ni tampoco darles información, más allá de expresiones genéricas, según las cuales habían efectuado *viajes irregulares con intención fraudulenta*, o bien habían incurrido en un *uso inadecuado del sistema informático*. Uber ha intentado a su vez defenderse alegando *ex post facto* que las cuentas sólo fueron desactivadas después de un control humano, que efectivamente apreció un intento de fraude.

Los afectados alegan haber sido sometidos a *decisiones individuales automatizadas*, contrarias al RGPD y solicitan:

- Readmisión a su puesto de trabajo.
- Indemnización por los daños y perjuicios causados.
- Transparencia sobre el funcionamiento del algoritmo de Uber.
- Multa coercitiva de diez mil euros, por cada día que tarde Uber en aplicar la resolución judicial que resuelva el proceso laboral.

Interesa aclarar que lo de menos, a nuestros efectos, es que el juzgado considere que es un asunto laboral, o declare incompetencia de jurisdicción y derive el asunto al juez civil ordinario, porque el fondo de la cuestión no está en la naturaleza de la relación contractual, máxime cuando lo que pretenden es volver a su puesto de trabajo, sino en que la resolución de la relación jurídica (sea civil o laboral) se ha efectuado en base a una decisión individual automatizada, que infringe el RGPD y, en base a él, el que hoy se denomina Derecho Fundamental a la Protección de Datos. De estimarse esto, la reclamación prosperará, independientemente de que se la califique como civil o laboral.

El caso de los *despidos algorítmicos de Uber* es de suponer que va a ser sonado; pero esto va a ser el pan nuestro de cada día en el ámbito del Derecho Privado, mucho más allá de que afecte a la organización del Estado, como veíamos al principio, tanto en la rama administrativa, como en la judicial. De hecho, ya existe una amplísima y diversa casuística de este tipo de cuestiones en el ámbito privado que, en el propio periódico “El Economista” ponía de manifiesto José Varela¹³⁷, responsable de digitalización del sindicato español UGT. Entre los muchos casos que cita, mencionaré sólo algunos, al mero efecto de que seamos conscientes de que las

¹³⁷ Varela, J., *Algoritmos e infalibilidad: una falacia de nuestra época*, “El Economista”, 23/9/20, www.eleconomista.es/opinion-blogs/noticias/10774823/09/20/Algoritmos-e-infalibilidad-una-falacia-de-nuestra-epoca.html.

decisiones jurídicas adoptadas directamente por algoritmos, están a la orden del día en el ámbito del Derecho privado:

- Despidos en base a algoritmos que no miden bien el desempeño profesional, sin llegar al supuesto de detección de fraudes del algoritmo reseñado de Uber.
- En las plataformas digitales, el tiempo de trabajo suele realizarlo un algoritmo, que reparte turnos, tareas y descansos.
- Contratos y reclamaciones de consumo que se efectúan de forma exclusivamente robótica, a veces por llamadas sin posibilidad de contactar con un operador humano.
- Muchos algoritmos han tenido que ser retirados de determinadas decisiones por tener sesgos raciales, sexistas, o prejuiciosas con respecto a nacionales de determinados países.

Sobre todas estas cuestiones, los acontecimientos se están precipitando en los últimos días de noviembre de 2020, que es cuando se termina de redactar el siguiente escrito. Ya antes, el 25 de septiembre de 2020, la Sala de lo Social del Tribunal Supremo de España dictó sentencia, en un recurso de casación para la unificación de doctrina, en la que estableció, a mi juicio con toda la razón, que los denominados “riders”, en este caso los de la plataforma Glovo (pero es aplicable a todas los demás) son falsos autónomos y, en consecuencia, sus contratos mercantiles son, en realidad, contratos laborales en fraude de ley. Han aparecido bastantes noticias de prensa criticando la sentencia, por entender que es una carga de profundidad a la denominada economía colaborativa. Personalmente estoy 100% de acuerdo con la sentencia, que se limita a hacer justicia en un caso claro de explotación laboral. Ya está bien de ampararse en las plataformas digitales y una supuesta economía colaborativa, para llevar a cabo modelos de negocio, que se podrían establecer igualmente por empresas tradicionales y que nadie calificaría de economía colaborativa y menos todavía intentaría justificar en base a la tecnología. La utilización de plataformas digitales es accidental y no sustancial en este tipo de negocios basados en la precarización de las relaciones laborales y, a fin de cuentas, en la explotación de los trabajadores. La presencia de plataformas digitales en el sector no es sino un pretexto de modernidad, con el fin de burlar la legislación laboral, que el Tribunal Supremo de España ha hecho muy bien en aplicar.

A colación del caso de los “riders”, los sindicatos españoles UGT y CC.OO. propusieron al Ministerio de Trabajo regular y controlar la actividad de esas plataformas digitales, que con lo que acaba de suceder en Holanda, a raíz de la antes mencionada resolución de contratos de trabajo de conductores de Uber, por un algoritmo de IA, es de presumir que va a tener un alcance mucho más amplio que el de los propios “riders” y procurará controlar en general a las plataformas digitales, para que no se menoscaben con un pretexto –que es sólo pretexto– tecnológico, los derechos de los trabajadores.

Ya el día 12 de noviembre, el Ministerio de Trabajo ha confirmado que en breve establecerá un registro obligatorio de plataformas digitales, para controlar su acti-

vidad, tal como se ha publicado en *Invertia*¹³⁸, entre otros periódicos económicos y de información general. En el borrador de la ley ha incluido una Disposición Adicional, en la que se crea un registro de plataformas digitales, dependiente del propio Ministerio, en la que será obligatorio inscribirlas, si tales plataformas realizan transacciones a partir de algoritmos. Se les pedirá una amplia información, incluida la relativa a los convenios colectivos aplicables; es de suponer que a fin de controlar si sus trabajadores van a realizar sus actividades en el marco de la regulación laboral, o si no tienen intención de hacerlo.

Según la Información de *Invertia*: *También tendrán que comunicar modelo de organización del trabajo, de la descripción de los diferentes puestos de trabajo, de su número de trabajadores, y del capital inicial y procedencia de éste, incluidas las posibles ayudas económicas públicas. Asimismo, deberán informar del algoritmo aplicado para la organización de la actividad, incluyendo, como mínimo, “el pseudo código o diagrama de flujo utilizado, así como los sistemas de reputación utilizados, en su caso, y a quién se aplican”.*

La forma de obligar a que dichas plataformas se inscriban, con carácter previo al inicio de su actividad, es el cuadro sancionador que se aplicará a las que no lo hagan, con multas que oscilan entre 6.251 euros y 187.515 euros. Es un cuadro sancionador que parece el de la Ley de Infracciones y Sanciones en el Orden Social, o al menos está inspirado en ella y acaso sea demasiado poco severo para ser eficaz, frente al gran volumen de negocio de las plataformas digitales. Pienso que más bien habría que ir a un sistema de multas, calculado como porcentaje de la facturación anual de la correspondiente empresa, al modo de lo que hace el Reglamento Europeo de Protección de Datos (RGPD).

Por lo demás, el pasado martes día 17 de noviembre, se conoció que la Secretaría de Estado de Digitalización e Inteligencia Artificial ha presentado una “primera versión” de *Carta de Derechos Digitales*, que al parecer no está destinada a tener carácter normativo, sino que servirá como guía de actuación del Gobierno a la hora de diseñar su política legislativa, en el entorno digital. Tal como está formulada, se asemeja a la Declaración de Derechos del Ciberespacio, publicada por mí mismo, hace ya 12 años, en 2008, que se puede encontrar en varias webs, como la del Observatorio Iberoamericano de Protección de Datos¹³⁹, o (en español y en inglés) en el libro que más recientemente coordiné sobre *La Constitución del Ciberespacio*¹⁴⁰. Incluso hay un estudio propio sobre la Declaración de Derechos del Ciberes-

¹³⁸ *Invertia*, *Trabajo creará un registro obligatorio de plataformas digitales, que tendrán que revelar sus algoritmos*, 2020, www.elespanol.com/invertia/economia/empleo/20201112/trabajo-registro-obligatorio-plataformas-digitales-revelar-algoritmos/535447346_0.html.

¹³⁹ <http://oiiprodat.com/2013/04/21/declaracion-de-derechos-del-ciberespacio/#:~:text=%20Declaraci%C3%B3n%20de%20Derechos%20del%20Ciberespacio%20%201,medidas%20sistem%C3%A1ticas%20de%20espionaje%20en%20las...%20More%20>.

¹⁴⁰ Suñé Llinás, E. (coord.), *La Constitución del Ciberespacio*, México, Porrúa, 2015.

pacio¹⁴¹, comparada minuciosamente, artículo por artículo, con algunos instrumentos posteriores, asimismo sobre Derechos Humanos en el ciberespacio.

Por ceñirnos a la cuestión que nos ocupa, *el Derecho Humano a las decisiones jurídicas humanas*, vuelvo a poner de manifiesto el art. 15.2 de la Declaración de Derechos del Ciberespacio de 2008, de mi autoría.

2. Se prohíbe que cualquier decisión que afecte a una persona humana se adopte exclusivamente por medio de sistemas automatizados, sin intervención de otra persona humana con potestad decisoria para cada caso concreto.

De forma semejante, el numeral XXIII.2 de la Carta de Derechos Digitales, dice:

2. Las personas tienen derecho a no ser objeto de una decisión basada únicamente en procesos de decisión automatizada, incluidas aquéllas que empleen procedimientos de inteligencia artificial, que produzcan efectos jurídicos o les afecten significativamente de modo similar, salvo en los supuestos previstos en las leyes. En tales casos se reconocen los derechos a:

a) Solicitar una supervisión e intervención humana.

b) Impugnar las decisiones automatizadas o algorítmicas.

Dicha Carta de Derechos Digitales, que en general está bien elaborada, se halla al finalizar el presente escrito, en un breve período de información pública. Según la web del Ministerio de Asuntos Económicos y Transformación Digital, fue redactada por un grupo de expertos, constituido por el Ministerio el 15 de junio de 2020, no sé en base a qué criterios de selección, pues probablemente son todos los que están, pero no están todos los que son. En ella se integran compañeros de la Universidad de los que tengo la mejor opinión, al igual que una brillante antigua alumna del Máster en Informática y Derecho, que promoví y dirigí entre 1995 y 2005; pero si yo tuviera que citar dos personas relevantes en el ámbito del Derecho Informático en España, sin duda las primeras que me vendrían a la cabeza son Valentín Carrascosa y Miguel Ángel Davara, porque son expertos de largo recorrido en la materia, que no necesitan presentación, y ninguno de los dos está en el mencionado grupo. Como mi propio caso, que llevo largo tiempo publicando sobre Derechos Humanos en el Ciberespacio (Declaración incluida), más que cualquier otra persona en España (dato objetivo) y nadie me consultó si quería estar en la Comisión ministerial. Esto me indica que va siendo hora de terminar el escrito, porque esto nos lleva de regreso a lo que planteaba al principio del mismo.

La inteligencia artificial para prevenir la corrupción

Por Romina Florencia Cabrera¹⁴²

¹⁴¹ Suñé Llinás, E. - Reyes, A. - Servín, O. (coords.), *Derecho Informático e Informática Jurídica*, en la *Obra Jurídica Enciclopédica en Homenaje a la Escuela Libre de Derecho en su Primer Centenario*. México, Porrúa, 2012.

1. Introducción

Como decía el insigne jurista alemán Radbruch (1956)¹⁴³ debemos “dar a nuestra sociedad no un mejor derecho penal, sino algo mejor que el derecho penal”.

Desde finales del siglo XX, se alzan distintas voces doctrinales que postulan la necesidad de superar, ampliar o complementar el sistema judicial penal (por ejemplo, tomando de base al español), regido por el principio de legalidad y de justicia retributiva. En este sistema, los poderes públicos están obligados a actuar procesalmente cuando tienen noticia de la comisión de unos hechos delictivos, además de estar fundados en la pena y en el castigo al culpable, se pretende incluir lo que se conoce como justicia restaurativa (Subijana Zunzunegui, 2013)¹⁴⁴, ésta desarrolla una función de prevención, apuesta por una mayor atención a las víctimas, por espacios de comunicación entre víctima y victimario, y se rige en su caso, por el principio de oportunidad en el ejercicio de la acción penal.

Otra manera de prevención innovativa, es aplicar la inteligencia artificial a la tarea de investigación, combatir el crimen, y sobre todo, disminuir la corrupción en la sociedad, para contribuir al avance ético y superador de la misma.

2. Corrupción

Estos últimos años, especialmente en América Latina, los temas de corrupción han estado en la boca de la opinión pública y han sido protagonistas de diversos debates y divisiones en la sociedad. La gobernabilidad permite que estos temas dejen de ser prohibidos y permitan un sano y positivo¹⁴⁵ tratamiento para mejorar las

¹⁴² Doctorando UNLP. Academia.

¹⁴³ Radbruch, G., *Rechtsphilosophie*, 5ªed., Stuttgart, 1956, p. 269. Ruiz Sierra, Joana, “Víctima y Mediación penal”, <http://noticias.juridicas.com/conocimiento/articulos-doctrinales/10614-victima-y-mediacion-penal>.

¹⁴⁴ Véase a Subijana Zunzunegui, I.J., *El significado innovador y la viabilidad de la justicia terapéutica, restaurativa y procedimental en nuestro Ordenamiento Jurídico*, “Cuadernos penales José María Lidón”, n° 9, 2013, p. 21 a 58, que añade a la justicia restaurativa, la justicia terapéutica y la justicia procedimental. Ruiz Sierra, Joana, *Víctima y Mediación penal*, <http://noticias.juridicas.com/conocimiento/articulos-doctrinales/10614-victima-y-mediacion-penal>.

¹⁴⁵ International Council on Human Rights Policy, *La corrupción y los derechos humanos. Estableciendo el vínculo*, <https://observatoricorruptio.org/wp-content/uploads/2017/11/6.-La-corruptio%CC%81n-y-los-Derechos-Humanos.-Estableciendo-el-vi%CC%81nculo.pdf>.

Como señala Ferrajoli, “el derecho penal, aun rodeado de límites y garantías, conserva siempre una intrínseca brutalidad que hace problemática e incierta su legitimidad moral y política”. Un análisis del sistema penal en sus consecuencias revela la fría y deshumanizada mecánica del crimen y castigo.

“En efecto, en el modelo hoy imperante en nuestro sistema penal –basado en los postulados de la justicia retributiva–, ‘el mal reclama el mal’, la reparación de la ofensa, del delito, exige una pena, una sanción equivalente. La justicia tiene que ver más con la venganza que con el reconocimiento del daño causado, con el perdón, con la construcción de una sociedad más justa. La justicia retributiva se centra primaria y monotémicamente en el castigo del culpable y otorga relevancia casi exclusiva a las instituciones de control formal. La justicia restaurativa, sin embargo, valora la necesidad de

Ciclo “Sistemas Inteligentes y Ética” 99

políticas públicas de que es destinatario final el ciudadano, el soberano. La democracia la realizamos y consensuamos entre todos.

La corrupción es un asunto tratado por las instituciones financieras, agencias gubernamentales, agencias de financiación bilaterales, organizaciones internacionales, organizaciones no-gubernamentales (ONGs) y los profesionales del desarrollo. Se han comprobado empíricamente las causas de la corrupción y sus impactos en el nivel de desarrollo humano. Se han investigado las instituciones y los procedimientos administrativos. Los países han negociado y han firmado tratados internacionales contra la corrupción. El Programa Global contra la Corrupción de la Oficina de las Naciones Unidas contra la Droga y el Delito (UNODC) ha actuado como un catalizador que ayuda a los países en la implementación de la Convención de las Naciones Unidas Contra la Corrupción (CNUCC). Transparencia Internacional (TI) y otros actores sociales han creado un gran forum para la discusión y la defensa contra la corrupción en sus variadas formas; ha surgido una coalición internacional de ONGs que desafía a la corrupción “desde abajo”. Sin embargo, la corrupción sigue siendo un reto. A pesar de incontables diagnósticos de políticas, de campañas públicas para crear conciencia y de las reformas institucionales y también legales para mejorar la administración pública, las investigaciones muestran que la corrupción continúa floreciendo.

De hecho, las encuestas de opinión sugieren que el público es ahora más pesimista que antes ante la posibilidad de su erradicación. El combate contra la corrupción requiere de esfuerzos colectivos potentes, desde diversos sectores de la sociedad que actúen en forma coordinada¹⁴⁶.

3. Vínculo

Si se analizan los vínculos entre la corrupción y los derechos humanos suponiendo que, si la corrupción surge donde existe la inclinación y la oportunidad, un enfoque desde los derechos humanos puede ayudar a minimizar las oportunidades para el comportamiento corrupto y posibilitar que aquéllos que son corruptos puedan ser capturados y sancionados debidamente. Este enfoque de derechos humanos también centra la atención en las personas que por su vulnerabilidad tienden a sufrir una desventaja desproporcionada cuando son víctimas de la corrupción, brinda un

la intervención de las instituciones penales, pero insiste en procurar la corresponsabilidad de la sociedad y de todo el tejido social en la prevención y evitación del delito, así como en el tratamiento y la inserción social de los infractores. Mientras que la justicia retributiva se asienta en la idea de que la única intervención ha de ser la que se sostiene en el monopolio de la violencia del Estado, la justicia restaurativa, aun manteniéndola, reivindica el diálogo y el encuentro personal como formas saludables y no violentas de restablecer la paz quebrada por el delito” (Ríos Martín, Julián C. - Olalde Altarejos, Alberto J., *Justicia restaurativa y mediación. Postulados para el abordaje de su concepto y finalidad*, “Revista de Mediación ADR, Análisis y Resolución de Conflictos”, n° 8, <https://revistademediacion.com/articulos/justicia-restaurativa-y-mediacion-postulados-para-el-abordaje-de-su-concepto-y-finalidad>).

¹⁴⁶La corrupción y los derechos humanos. Estableciendo el vínculo.

enfoque de género y ofrece pautas para el diseño y la implementación de políticas anticorrupción.

Si se demuestra que la corrupción viola los derechos humanos puede influir en aumentar el apoyo público anticorrupción. Si los individuos son más conscientes del daño que la corrupción causa a los intereses públicos e individuales y toman conciencia del daño que pueda ocasionar hasta la menor forma de corrupción, estarán más dispuestos a apoyar las campañas y los programas para prevenirla. Esto es importante porque, a pesar de la fuerte retórica, el impacto político de la mayoría de los programas anticorrupción ha sido bajo.

La identificación de los vínculos específicos entre la corrupción y los derechos humanos puede convencer a los actores claves (funcionarios públicos, parlamentarios, jueces, fiscales, abogados, empresarios, banqueros, contadores, los medios y el público en general) para que encaren con mayor rigor la corrupción. Esto puede suceder incluso en aquellos países donde referirse a los derechos humanos puede ser un asunto sensible.

Los estándares de derechos humanos, como se establecen en los principales tratados internacionales y las legislaciones nacionales, imponen obligaciones a los Estados. Un análisis de derechos humanos específicos ayudará a identificar quiénes pueden legítimamente reclamar cuando ocurren actos de corrupción y proteger a aquéllos que sufran perjuicios a causa de ellos. Una comprensión clara de las conexiones prácticas entre los actos de corrupción y los derechos humanos puede empoderar a quienes tienen reclamos legítimos para exigir sus derechos contra la corrupción y puede ayudar a los Estados y a otras autoridades públicas para que respeten, protejan y cumplan sus responsabilidades en materia de derechos humanos a todos los niveles.

En los últimos sesenta años, luego de la adopción de la Declaración Universal de los Derechos Humanos (DUDH), se han creado muchos mecanismos para buscar la responsabilidad de los Estados y de los individuos en caso de violaciones de los derechos humanos. Por ejemplo, además de mecanismos judiciales, los informes parlamentarios son instrumentos importantes en muchos países, igualmente el monitoreo de la sociedad civil que es cada vez más amplio en muchos países. Las instituciones intergubernamentales también se han desarrollado y además de los mecanismos principales de las Naciones Unidas ahora encontramos una serie de mecanismos regionales como la Corte Europea de Derechos Humanos (CtEDH), la Corte Africana de Derechos Humanos y de los Pueblos (ACtHPR) y la Corte Interamericana de Derechos Humanos (CIDH). Es igualmente significativa la evolución de las instituciones nacionales de derechos humanos.

Cuando los actos de corrupción se vinculan a las violaciones de los derechos humanos, todas estas instituciones pueden actuar para buscar la responsabilidad en caso de actos de corrupción y así disminuir los incentivos de la corrupción. Aunque los mecanismos de derechos humanos no sustituyen a los mecanismos tradicionales contra la corrupción –principalmente, a la justicia criminal– estos mecanismos pueden dar visibilidad a los casos de corrupción, pueden forzar a un Estado a que emprenda acciones preventivas o pueden disuadir a los funcionarios corruptos de un

mal ejercicio de sus funciones. Por lo tanto, pueden ayudar a crear conciencia y tienen un efecto persuasivo.

El sistema de los derechos humanos enfatiza el deber de dar protección prioritaria a los grupos vulnerables y desventajados contra abusos. Así, los derechos humanos tienen principios transversales –en especial, los relativos a no-discriminación, la participación y la responsabilidad– que buscan empoderar a aquéllos que se encuentran en una situación de desventaja. El sistema de los derechos humanos requiere que los Estados apliquen seriamente estos principios¹⁴⁷. Las comunidades no deben ser consultadas en forma superficial, por ejemplo. Se les debe permitir y estimular una participación activa en los esfuerzos para combatir la corrupción. Desde el punto de vista de los derechos humanos los gestores de políticas públicas deben preguntarse de qué manera el diseño o la implementación de los programas anticorrupción afectarán a las personas marginadas, o víctimas de discriminación social o en desventaja en algún otro aspecto. Adherirse a los principios de los derechos humanos implica identificar y superar obstáculos (tales como diferencias lingüísticas y culturales, racismo y discriminación de género) que hacen que estas personas sean especialmente vulnerables a la corrupción. Aunque pareciera existir acuerdo en que la corrupción tiene impactos concretos específicos en los grupos más vulnerables o desventajados de la sociedad, la incorporación de criterios de vulnerabilidad¹⁴⁸ y género en el diseño de los programas anticorrupción es aún limitada y excepcional. Utilizar plenamente los derechos humanos puede ayudar a fortalecer estas dimensiones de las políticas públicas. El principio de no-discriminación puede ser, especialmente, útil como guía para alcanzar este objetivo.

De conformidad con los tratados internacionales contra la corrupción, las medidas anticorrupción deben ser compatibles con los principios de los derechos humanos y no deben afectar los derechos de aquéllos que estén implicados. Sin embargo, los tratados ofrecen muy pocas orientaciones sobre cómo los funcionarios públicos deben conciliar su compromiso para combatir la corrupción con su obligación para promover y proteger los derechos humanos. Analizar los programas anticorrupción desde una perspectiva de derechos puede ayudar a los Estados a cumplir con sus obligaciones en materia de derechos humanos en el diseño e implementación de las leyes y los procedimientos para investigar y sancionar los casos de corrupción.

4. Inteligencia artificial contra el crimen

Las recientes repercusiones de público conocimiento del caso de Cambridge Analytica, confirman que la tecnología cada día alcanza más campos del ámbito académico, político, social y laboral. Por lo tanto, es importante destacar la creciente utilización de la inteligencia artificial (IA) para comenzar a prevenir y detectar hechos de corrupción y fraude en las organizaciones.

¹⁴⁷ *La corrupción y los derechos humanos. Estableciendo el vínculo.*

¹⁴⁸ *La corrupción y los derechos humanos. Estableciendo el vínculo.*

Nuestra última encuesta sobre delitos económicos indicó que el 23% de las organizaciones comenzaron a utilizar la IA y el 35% planea su desarrollo en el corto y mediano plazo.

En los últimos años y, más aún con la vigencia de la ley de responsabilidad penal empresaria, se incrementó la necesidad de construir programas de integridad, antifraude y/o anticorrupción y el desarrollo de equipos de investigación internos y compliance, para establecer políticas y procedimientos específicos e implementar tecnología que ayude a las organizaciones a combatir la amenaza de ser víctima de fraude¹⁴⁹.

En general, el foco principal de cualquier actividad de investigación, prevención y detección, es la búsqueda de algún indicio, patrón o conducta que permita identificar un acto no ético o irregular. En esa búsqueda, está la habilidad del investigador, su capacidad de analizar diferentes fuentes de información, el desarrollo del pensamiento lateral para unir puntos que no tienen a priori ninguna conexión a la vista y la perseverancia de plantear hipótesis de cómo podría ocurrir el delito, probando luego cada una de ellas hasta finalmente descubrir los hechos. Pero hoy el investigador ya no está solo en esa desafiante tarea.

Gracias al avance de la tecnología en los últimos años y la explosión en la cantidad y diversidad de datos estructurados y no estructurados que se generan continuamente, las organizaciones más innovadoras han comenzado a poner en práctica modelos predictivos que soporten la tarea del equipo a cargo de promover la ética y transparencia, sentaron a la misma mesa al auditor interno, investigador de fraude, científico de datos, oficial de ética y cumplimiento, responsable de seguridad de la información, de sistemas, gerente comercial, de operaciones, CFO y hasta el CEO. ¿Con qué propósito? Comenzar a plantear posibles escenarios de amenazas a los que la organización puede estar expuesta y descubrir qué datos serán necesari-

¹⁴⁹Sarcuno, Andrés, *Inteligencia artificial y machine learning para prevenir fraudes en las organizaciones*, www.telam.com.ar/notas/201804/273015-inteligencia-artificial-y-machine-learning-para-prevenir-fraudes-en-las-organizaciones.html. “Las alertas no se han hecho esperar, el director de inteligencia artificial de Google, John Giannandrea, citado por el *MIT Technology Review*, dijo que lo que más le preocupaba de la inteligencia artificial en cuanto a los algoritmos de aprendizaje automático utilizados para la toma de decisiones es ‘que si les damos a estos sistemas datos sesgados, estarán sesgados’. El problema jurídicamente relevante en este tema es la proliferación del uso de sistemas de información creados y entrenados con datos sesgados en materias críticas, como la financiera, la legal, la laboral, la médica y, por qué no, el consumo. Varias empresas en EEUU han creado sistemas para identificar sesgo o discriminación en las ofertas de empleo por cuestión de género, edad o raza. En materia legal, hace poco, se conoció el estudio de Pro Publica, una corporación sin ánimo de lucro que descubrió que el *software* utilizado por el sistema judicial en EEUU para determinar el riesgo de reincidencia ‘*tenía el doble de probabilidades de señalar erróneamente a individuos de raza negra*’. No podemos cambiar la naturaleza humana, lo que sí podemos hacer es regular las acciones que los individuos despliegan en sociedad. Aparte de los ejemplos en materia legal-criminal que revisten una importancia y merecen un tratamiento mucho más profundo, desde el punto de vista del Derecho Comercial, por ejemplo, el banco solo querrá hacer préstamos a quien tiene dinero y liquidez suficiente y la publicidad querrá ser dirigida puntualmente a un público objetivo con capacidad de compra del producto, ya que esto hace parte de la naturaleza propia de un comportamiento comercial en una sociedad de consumo” (*Inteligencia artificial y ¿discriminación?* <http://blogs.portafolio.co/abogado-tic/2017/10/26/inteligencia-artificial-discriminacion>).

rios recopilar para cimentar las bases de un modelo que le permita predecir cuándo esas amenazas pueden materializarse.

Con el paso del tiempo y el registro de aquellas actividades que resultaban sospechosas, las firmas se encuentran en la posición de aplicar aprendizaje supervisado (“*machine learning*”) e inteligencia artificial, para etiquetar cada nueva transacción y predecir si es irregular o no.

Quienes hoy están abordando la aplicación de un modelo predictivo, son aquellos que hace algunos años confiaban en que la analítica avanzada les iba a permitir fortalecer sus programas antifraude y comenzaron a registrar y almacenar ciertos datos relevantes. Y es justamente allí donde está la clave, la materia prima es el historial de transacciones (o alertas) que se investigó y determinó si eran o no fraudulentas. Por lo tanto, a las organizaciones que les preocupa la corrupción y el fraude y están implementando un programa de integridad, es necesario que incluyan en su agenda la aplicación de inteligencia artificial y aprendizaje supervisado para luchar contra el flagelo y anticiparse a las amenazas del mundo actual y del mañana.

El proceso está protegido en el Sistema Interamericano de los Derechos Humanos. Cualquier menoscabo o violación del mismo supone el no acatamiento a las normas supranacionales que conforman la comunidad internacional. El Estado que vulnere dichas normativas, y que ha suscripto debidamente los instrumentos internacionales de Derechos Humanos, queda sujeto a las sanciones pertinentes dispuestas por los mecanismos del Derecho Internacional Público.

Obviamente, cualquier aplicación de las Nuevas Tecnologías, deberá respetar las garantías procesales penales básicas, consagrado el debido proceso desde la Carta Magna de 1215, y no provocar sesgos discriminatorios, por razón de raza, sexo, religión, orientación sexual, etc. La investigación criminal y la prevención de la corrupción son fundamentales, pero no pueden vulnerar derechos fundamentales, ni amenazar el tesoro más valioso que tiene una sociedad: la garantía de la libertad.

5. Conclusiones

La inteligencia artificial está revolucionando el medio en el cual la especie humana desarrolla su vida en sociedad. El avance vertiginoso de la tecnología modifica hábitos y modos de relaciones, tanto entre pares como con el medio. Surgen nuevos hechos jurídicos, nuevas legislaciones, derechos y obligaciones, y sobre todo, se replantean los valores clásicos, y hasta las creencias. El futuro demostrará si la IA logrará los cambios ambiciosos que se propone, mejorando la sociedad y logrando un mundo con corrupción cero o por lo menos disminuida, y los valores de justicia y equidad.

Gracias al avance de la tecnología en los últimos años y la explosión en la cantidad y diversidad de datos estructurados y no estructurados que se generan continuamente, las organizaciones más innovadoras han comenzado a poner en práctica modelos predictivos que soporten la tarea del equipo a cargo de promover la ética y transparencia.

La sociedad evoluciona según las demandas sociales y la dinámica misma de la civilización. Pero para que haya resultados positivos, no deben perderse los valores, inmutables, pero adaptados a la realidad, y sobre todo la ética.

El desafío consiste en ver los pro y contra de la tecnología, y adaptarlos a una sociedad con ansias de avanzar hacia modelos más evolucionados, pero no perdiendo los principios, la esencia del ser humano. Y sobre todo, garantizando las garantías procesales penales y los Derechos Humanos, sin entrar en discriminaciones.

Bibliografía

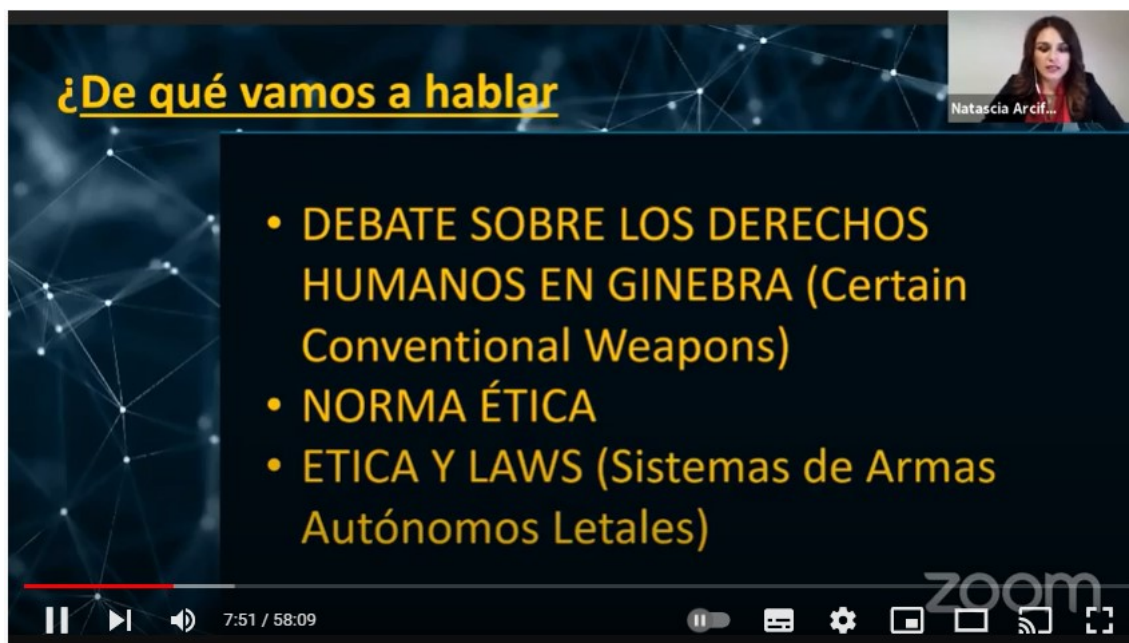
Inteligencia artificial y machine learning para prevenir fraudes en las organizaciones, www.telam.com.ar/notas/201804/273015-inteligencia-artificial-y-machine-learning-para-prevenir-fraudes-en-las-organizaciones.html.

La corrupción y los derechos humanos. Estableciendo el vínculo, International Council on Human Rights Policy, www.ichrp.org/files/reports/52/131_report_es.pdf.

Radbruch, G., *Rechtsphilosophie*, 5ªed. Stuttgart, 1956, p. 269.

Ruiz Sierra, Joana, *Víctima y Mediación penal*, <http://noticias.juridicas.com/conocimiento/articulos-doctrinales/10614-victima-y-mediacion-penal>.

Subijana Zunzunegui, I.J., *El significado innovador y la viabilidad de la justicia terapéutica, restaurativa y procedimental en nuestro ordenamiento jurídico*, “Cuadernos Penales José María Lidón”, n° 9, 2013, p. 21 a 58, que añade a la justicia restaurativa, la justicia terapéutica y la justicia procedimental.



***Diritti umani e intelligenza artificiale:
la regolamentazione internazionale dei sistema
di arma autonoma letale (laws)***

Por Natascia Arcifa

La nostra società si è sempre basata sulla supremazia dell'essere umano che acquisisce conoscenze e usa ogni strumento a sua disposizione per evolversi. Nel nuovo millennio col termine "evoluzione" si suole indicare il livello raggiunto dalle scoperte tecnologiche, consolidandosi l'idea che le macchine, eliminando gli errori, possono fare meglio degli uomini, spingendo addirittura verso il ridimensionamento dell'intervento umano.

L'intelligenza artificiale e il machine learning, ovvero l'apprendimento automatico, vengono usati ormai in ogni ambito, commerciale, governativo e aziendale, attraverso servizi che riescono a monitorare cose e persone producendo effetti nei confronti della società tutta.

In particolare, nel mondo della robotica militare l'utilizzo dell'intelligenza artificiale si è spinta, in molte Nazioni, nell'implementare un nuovo sistema di armi, programando macchine che possono identificare e colpire bersagli da soli, senza alcun input diretto dell'uomo. Già USA, Russia, Cina e Francia hanno investito miliardi nello sviluppo dell'autonomia nell'uso della forza militare.

Il progressivo utilizzo di tali sistemi solleva però molte problematiche etico-giuridiche in quanto l'attuale impianto normativo non è in grado di ricomprendere tutte

le diverse tipologie di sistemi di intelligenza artificiale esistenti e quelle future, lasciando, inoltre, un vuoto sul piano della responsabilità nelle ipotesi di danno a colui che utilizza questi nuovi strumenti o a soggetti terzi.

Considerando le leggi internazionali a tutela dei diritti umani, la critica maggiore riveste la moralità dell'azione svolta dall'arma autonoma che può intaccare la tutela dei diritti umani.

Cosa succederebbe se alcune decisioni venissero prese solo da calcoli matematici? Il controllo umano sull'arma è riferibile a norme di diritto internazionale solo indirettamente, perché, sino ad oggi, non si era posta la possibilità di impiegare armi in grado di selezionare e ingaggiare da sole obiettivi umani.

In molti si sono chiesti: cosa è più necessario, l'evoluzione tecnologica o la tutela dei diritti umani? Nessuno è in grado di poter dare una risposta personale o certa a questa domanda, ma serve confrontare i dati e conoscere le nozioni per provare a dare una risposta concreta nel vicino futuro.

1. La nozione di arma autonoma letale

Nella corrente letteratura militare e giuridica viene impiegato il termine armi autonome o, più nello specifico, armi autonome letali (c.d. lethal autonomous weapons o Laws), per indicare quelle armi che possono colpire un obiettivo militare senza l'intervento umano.

Secondo il World Economic Forum, sono sei i maggiori Stati che stanno promuovendo ricerche per sviluppare e testare armi autonome letali, quali Inghilterra, Cina, Stati Uniti, Israele, Russia e Sud Corea¹⁵⁰.

Ben 44 Stati, per la prima volta nel 2013, hanno pubblicamente discusso di questa tipologia di armi durante il primo dibattito sui diritti umani tenutosi in occasione della Convention on Certain Conventional Weapons (CCW). Inoltre, tenendo sempre presente l'obiettivo di prevenire che i civili possano essere sottoposti ad occasioni di pericolo e di evitare possibili effetti destabilizzanti, le Nazioni Unite in occasione delle CCW 2016 hanno identificato un gruppo di esperti governativi per meglio intervenire sulla gestione dei sistemi di armi autonome letali¹⁵¹.

La definizione di Laws è un elemento indispensabile per potere procedere alla regolamentazione di queste armi. La comunità internazionale ha sostenuto negli anni il dibattito¹⁵² sulla corretta qualificazione delle armi autonome, senza però trovare una definizione comune. Sono stati però selezionati gli elementi che permettono di

¹⁵⁰ Per maggiori informazioni, visita www.weforum.org/reports.

¹⁵¹ La documentazione relativa agli incontri informali ed ai lavori del Gruppo di esperti governativi è consultabile su: www.unog.ch.

¹⁵² P. Scharre - M.C. Horowitz, *An introduction to Autonomy in Weapon Systems*, Working Paper, Center for a New American Security, 2015.

definire un'arma come autonoma: a) l'autonomia rispetto al controllo umano¹⁵³; b) la selezione; c) l'attacco di un obiettivo che avviene senza alcun input; d) l'intervento umano non necessario o assente. Questi elementi si prestano a diverse interpretazioni e conseguenti definizioni.

Una definizione di arma autonoma è stata data dal Comitato Internazionale della Croce Rossa (CICR) che, quale risultato di una serie di conferenze, così scrive: "Qualsiasi sistema d'arma con autonomia nelle sue funzioni critiche. Cioè, un sistema d'arma che può selezionare (rilevare, identificare, tracciare) e attaccare (usare la forza contro, danneggiare o distruggere) obiettivi senza l'intervento umano".

Nonostante non sia stata concordata una definizione comune, sul piano tecnico lo sviluppo delle stesse non trova un effettivo rallentamento. Infatti, ad oggi, sono già stati impiegati sul campo diverse tipologie di armi¹⁵⁴, la maggior parte di

¹⁵³ L'autonomia delle armi letali autonome viene anche c.d. "autonomia operativa", cioè relativa allo svolgimento di un determinato compito. In questa prospettiva un sistema robotico può essere definito "autonomo" in base alla sua capacità di eseguire un dato compito senza richiedere alcun intervento da parte degli esseri umani. I dispositivi tecnologici che godono in questo senso di autonomia operativa formano una classe ampia ed eterogenea, poiché il repertorio dei compiti comprende tanto attività complesse, quali la navigazione stradale dei veicoli autonomi o l'esecuzione di un'operazione chirurgica, quanto compiti che consideriamo molto più semplici dalla prospettiva della loro eseguibilità da parte di agenti artificiali, come ad esempio il controllo dell'accensione e dello spegnimento di una caldaia da parte di un termostato; D. Amoroso e G. Tamburrini, I sistemi robotici ad autonomia crescente tra etica e diritto: quale ruolo per il controllo umano?, *BioLaw Journal*, Rivista di BioDiritto, 2019.

¹⁵⁴ Alcuni esempi: Missili da usare in campo aereo, anti-nave, da crociera o che agiscono sia via aria che via terra. Questi vengono raggruppati in due categorie. La prima è c.d. "go-onto-target" e raggruppa quei missili progettati per colpire un obiettivo seguendo le sue tracce, attraverso l'uso di sensori e segnali elettromagnetici. La seconda, c.d. "go-onto-location-in-space", si riferisce ad armi progettate per colpire il bersaglio situato in una specifica posizione geografica, come le bombe-GPS.

Alcune armi automatizzate sono più sofisticate e si attivano con precisione nel tempo stabilito, nonostante la scelta spetti sempre all'operatore. Ma si consideri, ad esempio, l'ipotesi dell'ingaggio di un missile air-to-air oltre il raggio visivo. Il pilota dovrà confermare l'ingaggio del missile, verso il potenziale obiettivo identificato da un computer, in base alle informazioni a lui trasmesse senza la capacità di confermare visivamente l'identità del bersaglio, ma solo sulla base di informazioni alimentate da un computer per l'identificazione di un nemico potenziale. È la c.d. "munizione vagante" che si differenzia da quella guidata in quanto non viene lanciata per un obiettivo specifico, ma per trovare e poi colpire, in autonomia e in un'area determinata, il bersaglio che rientra nei parametri di programmazione. L'operatore, dunque, conosce solo il bersaglio potenziale dell'arma.

Le forze di difesa israeliane usano il c.d. "Sentry Tech" con munizionamento in grado di coprire una Kill Zone per più di un chilometro, ed è progettato per sorvegliare e controllare eventuali sconfinamenti territoriali, tanto da essere stato posto lungo la linea di confine con la Striscia di Gaza.

La Russia ha testato diversi sistemi di combattimento autonomi e semi-autonomi, come il Kalashnikov con modulo di combattimento a "rete neurale", per il quale i suoi creatori sostengono può rendere i propri giudizi sul bersaglio senza intervento umano. Gli Stati Uniti hanno già molti programmi militari di combattimento IA, come la Sea Hunter, nave da guerra autonoma, progettata per operare in mare senza membri dell'equipaggio, guidando sé stessa. In Corea del Sud, è stata progettata una mitragliatrice in grado di identificare, rintracciare e distruggere un bersaglio in movimento a una distanza di 4 km.

esse rappresentata da robot, o sistemi autonomi in generale, per applicazioni terrestri e aeree, progettati per scopi difensivi, ma che implementano anche soluzioni offensive, seppur con un ridottissimo input umano.

2. Le argomentazioni etiche

Le regole etiche attingono alla morale e hanno una valenza universale, il loro rispetto non è assicurato tramite la coercizione da parte dello Stato, ma alcune volte vengono formalizzate all'interno di un ordinamento giuridico, come è avvenuto nella Carta Europea dei Diritti dell'Uomo, divenendo così norme giuridiche, il cui rispetto è assicurato tramite l'apparato statale e sovrastatale.

Considerazioni etiche e legali possono richiedere simili vincoli all'autonomia nei sistemi d'arma, in modo da mantenere un significativo controllo umano¹⁵⁵. Ciò è possibile attraverso la supervisione umana e la capacità di intervenire e disattivare il sistema, nonché la necessità di stabilire i requisiti tecnici di prevedibilità e affidabilità, in base agli algoritmi utilizzati. Serve anche determinare i vincoli operativi sull'utilizzo dell'arma, sul tipo di bersaglio, sull'ambiente operativo, sui tempi di funzionamento e sulla portabilità.

La questione etica fondamentale è se i principi di umanità e i dettami della coscienza pubblica possano consentire che il processo decisionale umano sull'uso della forza sia efficacemente sostituito con processi controllati dal computer e che le decisioni di vita o di morte¹⁵⁶ siano cedute alle macchine attraverso una scelta derivante dall'analisi di dati programmati.

In Europa, secondo la World Bank, c'è il potenziale per diventare il leader globale dello sviluppo dei sistemi di armi autonome letali, pubblicando già tanti documenti concernenti l'uso dell'Intelligenza Artificiale.

Francia, Germania, Slovenia e Italia stanno investendo sulla ricerca dei sistemi robotici. La Germania ha firmato un contratto per il sistema di un drone israeliano Heron TP; il Belgio e i Paesi Bassi sono in procinto di acquisire i sistemi Reaper MQ-9; l'Italia ha ottenuto il permesso, dagli Stati Uniti, di armare la sua flotta di droni MQ-9 Reaper; la Norvegia ha iniziato a formulare una regolamentazione dei sistemi autonomi.

¹⁵⁵ La documentazione relativa agli incontri informali ed ai lavori del Gruppo di esperti governativi è consultabile su www.unog.ch; United Nations, Report of the Group of Governmental Experts on "Lethal Autonomous Weapons Systems" (Laws), CCW/GGE.1/2017/CRP.1, 20/11/17, p. 7: "The importance of considering Laws (Lethal Autonomous Weapon Systems) in relation to human involvement and the human-machine interface was underlined. The notions that human control over lethal targeting functions must be preserved, and that machines could not replace humans in making decisions and judgements, were promoted. Various related concepts, including, inter alia, meaningful and effective human control, appropriate human judgement, human involvement and human supervision, were discussed"; United Nations, Recommendations to the Review Conference Submitted by the Chairperson of the Informal Meeting of Experts, november 2016, p. 1: "Views on appropriate human involvement with regard to lethal force and the issue of delegation of its use are of critical importance to the further consideration of Laws amongst the High Contracting Parties and should be the subject of further consideration".

¹⁵⁶ International Covenant on Civil and Political Rights, ICCPR, Art 6: "Ogni essere umano ha il diritto intrinseco alla vita. Questo diritto è protetto dalla legge. Nessuno sarà arbitrariamente privato della sua vita".

Gli argomenti etici sono stati formulati sia a favore che contro sistemi d'arma autonomi, riflettendo, in una certa misura, le due teorie dell'etica normativa, quella consequenzialista (focalizzata sui risultati) e quella deontologica (focalizzata sui processi).

La prima valuta la moralità di una determinata scelta alla luce della bontà delle conseguenze che ne derivano, come gli argomenti che promuovono l'autonomia delle macchine per ottenere migliori prestazioni a salvaguardia del benessere degli utenti.

L'etica deontologica, invece, si focalizza sui doveri morali quali guida per l'azione e parametro di giudizio del valore morale delle scelte individuali e collettive, considerando infatti che certe attività devono necessariamente essere condotte dall'essere umano.

Laddove l'etica consequenzialistica può motivare la concessione di una maggiore autonomia, l'imperativo deontologico la limita con la presenza del controllo umano¹⁵⁷, anche se vanno interpretati in modo contestuale, seguendo un'indagine delle norme morali rilevanti e immaginando scenari concreti.

Nei casi in cui la sostituzione dell'uomo con la macchina sia ritenuta accettabile, occorre stabilire come regolamentare il comportamento dei sistemi robotici di fronte a problemi eticamente e giuridicamente complessi e come incardinare un rapporto di responsabilità in ipotesi di eventi dannosi causati dalla macchina. L'argomentazione principale a favore di queste armi si basa sul migliore rispetto sia del diritto internazionale che dei valori etici umani, in quanto più precisi e affidabili dei sistemi di armi controllati direttamente dagli esseri umani, avendo inoltre conseguenze umanitarie meno negative per i civili¹⁵⁸.

Al contrario, gli studiosi dell'argomento¹⁵⁹, per promuovere la messa al bando delle armi autonome, hanno sostenuto, tra i vari argomenti deontologici, che le armi autonome non sarebbero in grado di assicurare il rispetto dei principi cardine del diritto internazionale umanitario, quali il principio di distinzione, il principio di proporzionalità ed il principio di precauzione. L'uso incontrollato dell'intelligenza

¹⁵⁷ D. Amoroso e G. Tamburrini, I sistemi robotici ad autonomia crescente tra etica e diritto: quale ruolo per il controllo umano?, *BioLaw Journal*, *Rivista di BioDiritto*, 2019.

¹⁵⁸ Ad esempio, droni armati telecomandati con munizioni guidate di precisione possono offrire il potenziale per una maggiore precisione e quindi meno rischio di effetti indiscriminati. Tuttavia, se le informazioni sul bersaglio sono imprecise, le pratiche di targeting sono troppo generalizzate, o le persone o gli oggetti protetti sono deliberatamente, o accidentalmente, attaccati, allora il potenziale di precisione non offre alcuna protezione.

¹⁵⁹ Kellenberger (2011: 5-6), the former president of the International Committee of the Red Cross (ICRC), expressed numerous concerns relating to the potential deployment of autonomous weapons, for example, by warning: A truly autonomous system would have artificial intelligence that would have to be capable of implementing IHL (international humanitarian law). Their development represents a monumental programming challenge that may well prove impossible. It would also raise a range of fundamental legal, ethical, and societal issues, which need to be considered before such systems are developed or deployed", *The Future of Killing: Ethical and Legal Implications of Fully Autonomous Weapon System*, M.Lark, *Salus Journal*, 2017.

artificiale può comportare la violazione dei diritti tutelati e porre rischi¹⁶⁰ di potenziale violazione dei diritti umani.

L'importanza di mantenere il libero arbitrio umano nelle decisioni di usare la forza è uno degli argomenti etici centrali per i limiti dell'autonomia nei sistemi d'arma autonome. Molti ritengono che le decisioni di uccidere, ferire e distruggere non debbano essere delegate all'apprendimento automatico delle macchine e che gli esseri umani debbano essere presenti in questo processo decisionale in modo sufficiente a preservare un legame diretto tra l'intenzione dell'essere umano e l'eventuale funzionamento del sistema di armi letali autonome. Per collegare l'agire umano e l'intenzione alle eventuali conseguenze di un attacco si fa ricorso alla prevedibilità e affidabilità nell'uso di un sistema d'arma autonomo. Tuttavia, in quanto armi che auto-innescano attacchi, i sistemi di armi autonomi sollevano interrogativi sulla prevedibilità, a causa di vari gradi di incertezza nel determinare quando, dove e perché si verificherà un attacco contro il bersaglio umano.

Secondo il Comitato internazionale della Croce Rossa (CICR)¹⁶¹, le considerazioni etiche sono parallele al requisito di un livello minimo di controllo umano sui sistemi d'arma e all'uso della forza per garantire la conformità al diritto internazionale e soddisfare le preoccupazioni. Se l'etica e la moralità umana sono meno controllate nelle relazioni impersonali con le macchine, è importante controllare se la macchina possa uccidere un uomo e come lo fa, al fine di non perdere il senso di dignità che contraddistingue il genere umano. Dunque, per preservare il libero arbitrio, da un punto di vista etico, il controllo umano deve essere "significativo", "efficace" o "appropriato" così da determinare la responsabilità morale nelle decisioni di usare la forza. Ciò richiede che sia mantenuta una connessione sufficientemente diretta e stretta tra l'intento umano dell'utente e le eventuali conseguenze del funzionamento del sistema d'arma in un attacco specifico. Spetta agli Stati stabilire limiti di autonomia di questi sistemi d'arma.

Lo sviluppo tecnologico dei Laws ha creato tanto interesse proprio perché l'autonomia riguarda le funzioni critiche della selezione e dell'ingaggio degli obiettivi nelle operazioni belliche. Tali funzioni sono considerate "critiche"¹⁶² perché la loro esecuzione: 1) è oggetto di approfondita regolamentazione nel diritto internazionale ed in particolare nella branca del diritto internazionale dei conflitti armati (o diritto internazionale umanitario); 2) è un fattore chiave ai fini della responsabilità individuale e statale; 3) implica scelte morali suscettibili di incidere, anche in misura assai profonda, su posizioni individuali eticamente rilevanti e giuridicamente tutelate¹⁶³.

Di fronte all'evoluzione dei Laws, l'idea di guerra della nostra società verrà modificata, con il rischio che potrebbe essere considerata moralmente accettabile. Le

¹⁶⁰ L. Andersen, *Human Rights in the Age of Artificial Intelligence*, www.accessnow.org, 2018.

¹⁶¹ ICRC Report - Ethics and autonomous weapon systems: An ethical basis for human control?, Geneva, 2018.

¹⁶² D. Amoroso e G. Tamburrini, I sistemi robotici ad autonomia crescente tra etica e diritto: quale ruolo per il controllo umano?

¹⁶³ Ad esempio: diritto alla vita, diritto all'integrità fisica, diritto ad una casa.

guerre deldomani potrebbero essere combattute solo da sistemi artificiali, e poiché non saranno più gli esseri umani a perdere la vita, difficilmente l'opinione pubblica si indignerà alla vista di oggetti metallici distrutti, anche se sono "macchine intelligenti".

La guerra ridotta ad un videogioco potrebbe essere accettata dall'opinione pubblica, e così di conseguenza l'evoluzione delle armi autonome, che diminuirebbe il costo delle vite umane, ma potrebbe aumentare il numero di conflitti e quindi il rischio per tutta l'umanità.

Ciò potrebbe rendere la guerra uno strumento di risoluzione delle controversie tra Stati più accettabile, producendo effetti dirompenti e impatti sociali importanti per gli uomini. Ma, allo stesso tempo, potrebbe compromettere la moralità della "guerra giusta"¹⁶⁴ in cui uno Stato entra in guerra per aiutare una parte debole.

I modelli di IA sono progettati per ordinare e classificare i risultati di ricerca e le persone, creando una discriminazione che può interferire con i diritti umani. Considerato che le traiettorie tecnologiche sono difficili da prevedere, la rimozione del controllo umano¹⁶⁵ rispetto alle decisioni di uccidere o danneggiare e le decisioni di usare la forza, porterà ad una lacuna di responsabilità morale che gli esseri umani non possono sostenere.

Allontanando ulteriormente, fisicamente e psicologicamente, gli esseri umani dal campo di battaglia, verranno rafforzate le asimmetrie esistenti e, senza limiti, si potrà ricorrere ad una violenza più semplice o meno controllata, minando la dignità umana di quei combattenti che sono presi di mira e dei civili esposti al rischio di morte e lesioni a seguito di attacchi contro obiettivi militari legittimi¹⁶⁶.

3. I prospetti giuridici

Nel 2010 il mercato per usi militari o di pubblica sicurezza dei velivoli senza pilota controllati da remoto, o droni, valeva quasi cinque miliardi di dollari. Questa è cresciuta di un quinto nel giro di un anno, attestandosi intorno ai sei miliardi di dollari su scala globale nel 2015. Vari tipi di mezzi militari controllati da remoto e senza esseri umani a bordo sono oggi in dotazione alle forze armate di numerosi Stati¹⁶⁷.

¹⁶⁴ Singer, P. W., *Robots at War: The New Battlefield*, The Wilson Quarterly, Winter, 2009.

¹⁶⁵ Per i sistemi d'arma autonomi il controllo esercitato dall'uomo può assumere varie forme e gradi in diverse fasi di sviluppo, spiegamento e uso, tra cui: a) lo sviluppo e la sperimentazione del sistema d'arma (stadio di sviluppo); b) la decisione del comandante o dell'operatore di attivare il sistema d'arma (fase di attivazione), e c) il funzionamento del sistema d'arma autonomo durante il quale seleziona e attacca in modo indipendente gli obiettivi (fase di operazione).

¹⁶⁶ Vedi, per esempio R Sparrow, *Killer robots*, Journal of Applied Philosophy, 2007; H Roff, *Killing in War: Responsibility, Liability and Lethal Autonomous Robots*, Routledge Handbook of Ethics and War: Just War Theory in the 21st Century, Routledge, UK, 2014.

¹⁶⁷ Tamburrini, Guglielmo, *Per la messa al bando delle armi autonome: ragioni morali e deliberazioni politiche*, DIETI - Dipartimento di Ingegneria Elettrica e Tecnologie dell'Informazione Università di Napoli Federico II, 2018.

Alcuni ritengono che il Laws rappresenta la terza rivoluzione del sistema bellico. A combattere la guerra del prossimo futuro potrebbero essere dunque robot, droni o sistema elettronici vari che, in modo più efficiente e veloce degli esseri umani, potranno raggiungere il loro obiettivo, ponendo l'intervento umano in secondo piano.

Basterà che gli aerei muniti da sensori computerizzati vengano combinati con sistemi di algoritmi specifici per diventare autonomi.

A comprova di ciò, negli ultimi tempi, gli Stati Uniti come pure Paesi dell'Unione Europea, hanno usato, in Libia, Siria e Yemen, armi quali droni, missili da crociera e bombardamenti più sviluppati¹⁶⁸. In particolare, i carri armati e gli aerei convenzionali sono già equipaggiati con sensori, telecamere e radar che ne consentono il controllo da remoto.

Basta dotarli degli opportuni sistemi di AI, per esempio derivati da quelli impiegati nelle auto senza conducente, per trasformarli in armi autonome inarrestabili.

Il vero rischio connesso all'utilizzo delle armi autonome, dunque, è che il loro impiego conduca a una de-responsabilizzazione giuridica ed etica degli Stati nella condotta delle ostilità, che rischia inevitabilmente di rendere meno costoso e, in ultima istanza, più semplice ipotizzare l'avvio di un conflitto armato.

Da qui la necessità che le regole etiche inizino ad essere promulgate. Sulla base delle linee guida per l'accettabilità etico-giuridica delle armi autonome, si sta costruendo la cornice normativa di riferimento dalla discussione in corso a Ginevra nel quadro della Convenzione sulle armi convenzionali¹⁶⁹. Il dibattito in seno al gruppo di esperti governativi si sta progressivamente focalizzando sul c.d. elemento umano, cioè sull'identificazione del tipo di interazione uomo-macchina preferibile sul piano etico.

È emerso che l'impatto dell'intelligenza artificiale sui diritti umani è uno dei fattori cruciali che definirà il futuro della guerra. Al pari delle evoluzioni storiche vissute, l'evoluzione dei metodi di guerra è passata dagli scontri di massa, alla guerra di movimento con velivoli e veicoli sino al futuro avvento della guerra robotizzata. Infatti, attualmente i Paesi dibattono sul modo in cui i principi del diritto umanitario internazionale possono essere applicati ai sistemi autonomi letali e su come gestire gli effetti che ne scaturiranno.

Non si tratta di un semplice aggiornamento delle famose "Tre Leggi di Asimov" sulla robotica¹⁷⁰, ma di un vero e proprio progetto d'ampio spettro che coinvolgerà

¹⁶⁸Vedi P. W. Singer, *Wired for War: The Robotics Revolution and Conflict in the Twenty-First Century*, New York Penguin Press, 2009; J.G. Castel and M.E. Castel, *The Road to Artificial Super-Intelligence: has international law a role to play?*

¹⁶⁹Report of the session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems, 2018 (UN Doc. CCW/GGE.1/2018/3).

¹⁷⁰ Le Tre Leggi di Asimov: – Un robot non può recare danno a un essere umano, né può permettere che, a causa del suo mancato intervento, un essere umano riceva danno. – Un robot deve obbedire agli ordini impartiti dagli esseri umani, purché tali ordini non contravvengano alla Prima Legge. – Un robot deve proteggere la propria esistenza, purché questa autodifesa non contrasti con la Prima e la Seconda Legge; Riccardo Campa, *Le armi robotizzate del futuro. Il problema etico*, Futurabilia, 2017.

esponenti dell'industria e ricercatori. La questione della c.d. roboetica¹⁷¹, secondo alcuni esperti, si articola in raccomandazioni specifiche, che prevedono: il controllo da parte degli umani; un sistema di sicurezza per evitare usi illegali dei robot; la tracciabilità delle attività dei robot e un codice univoco d'identificazione per ciascun robot.

I diplomatici degli Stati hanno discusso sulla negoziazione di un nuovo Trattato per rispondere alle crescenti preoccupazioni sulla tipologia di queste armi esaminando l'elemento della forza letale, ma senza arrivare ad un documento normativo concreto¹⁷².

¹⁷¹ Gianmarco Verruggio, ricercatore del CNR e direttore del RobotLab di Genova: "Abbiamo due priorità: tenere conto dell'approccio etico dei costruttori di robot e contemporaneamente studiare l'etica per le intelligenze artificiali", I robot dovranno seguire un codice etico, 2006, vedi www.punto-informatico.it.

¹⁷² Da alcune affermazioni estrapolate dalle conclusioni del Report 2019 GGE on Laws, CCW/GGE.1/2019/3: -States, parties to armed conflict and individuals remain at all times responsible for adhering to their obligations under applicable international law, including IHL; -States must also ensure individual responsibility for the employment of means or methods of warfare involving the potential use of weapons systems based on emerging technologies in the area of lethal autonomous weapons systems in accordance with their obligations under IHL; -Human judgement is essential in order to ensure that the potential use of weapons systems based on emerging technologies in the area of lethal autonomous weapons systems is in compliance with international law, and in particular IHL; -A weapons system based on emerging technologies in the area of lethal autonomous weapons systems, must not be used if it is of a nature to cause superfluous injury or unnecessary suffering, or if it is inherently indiscriminate, or is otherwise incapable of being used in accordance with the requirements and principles of IHL; -International law, in particular the United Nations Charter and IHL, as well as relevant ethical perspectives applicable to the use of weapons systems based on emerging technologies in the area of lethal autonomous weapons systems, continue to guide the work of the Group; -Possible good practices in the conduct of legal reviews, at the national level, of a potential weapons system based on emerging technologies in the area of lethal autonomous weapons systems to determine if its employment, in light of its intended or expected use, would be prohibited by the requirements and principles of IHL in all or some circumstances; -Possible bias in the data sets used in algorithm-based programming relevant to emerging technologies in the area of lethal autonomous weapons systems; -Different potential characteristics of emerging technologies in the area of lethal autonomous weapons systems, including: self-adaptation, predictability, explainability, reliability, ability to be subject to intervention, ability to redefine or modify objectives or goals or otherwise adapt to the environment, and ability to self-initiate; -The role and impacts of autonomous functions in the identification, selection or engagement of a target are among the essential characteristics of weapons systems based on emerging technologies in the area of lethal autonomous weapons systems, which is of core interest to the Group; -Further clarification is needed on the type and degree of human-machine interaction required, including elements of control and judgement, in different stages of a weapon's life cycle, in order to ensure compliance with IHL in relation to the use of weapons systems based on emerging technologies in the area of lethal autonomous weapons systems; -Risk mitigation measures can include: rigorous testing and evaluation of systems, legal reviews, readily understandable human-machine interfaces and controls, training personnel, establishing doctrine and procedures, and circumscribing weapons use through appropriate rules of engagement; -Research and development of autonomous technologies should not be restricted based on the sole rationale that such technologies could be used for weapons systems. At the same time, given the dual use nature of the underlying technologies in the area of lethal autonomous weapons systems, it is important to promote responsible innovation and use of such technologies.

La prima sessione del Group of Governmental Experts (GGE)¹⁷³, conclusasi a Ginevra in occasione della Convention on Conventional Weapons del 2017, ha conteso la partecipazione di 86 Stati, 22 dei quali si sono già espressi in favore dell'adozione di un documento che proibisca l'utilizzo di queste armi letali autonome. In questa occasione, inoltre, è stata presentata la lettera-appello firmata da scienziati, ricercatori e filosofi (tra i quali Stephen Hawking, Elon Musk e Noam Chomsky) per chiedere la messa al bando delle armi letali autonome in linea con la campagna Stop Killer Robots¹⁷⁴.

Se da un lato è stato riscontrato il parere fortemente contrario di alcuni Paesi, come Egitto, Cuba, Pakistan, Ecuador e persino lo Stato del Vaticano, i quali non ritengono tali sistemi affidabili circa la corretta individuazione dei bersagli in quanto inadeguati a rispettare le norme fondamentali di diritto internazionale umanitario.

Dall'altro lato invece la maggior parte degli Stati, come gli Stati Uniti, la Russia e il Regno Unito, sostengono che la discussione debba valorizzare i vantaggi che l'introduzione delle armi autonome porteranno nell'economia dei conflitti armati in termini di riduzione dei costi umani, perché l'automazione consentirebbe un maggiore rispetto delle norme di diritto internazionale umanitario, in quanto si trasformerebbero in variabili algoritmiche.

Una posizione intermedia è sostenuta dagli Stati, come Francia e Germania, che ritengono sufficiente la regolamentazione attraverso le norme esistenti, votando per la possibile negoziazione di un nuovo Trattato al fine di rispondere alle crescenti preoccupazioni sull'uso letale di queste armi, considerato che il loro sviluppo deve essere coerente con i principi di necessità militare, di umanità, di distinzione e di proporzionalità.

Durante la Conferenza di esperti governativi tenutasi in occasione della CCW del 2019¹⁷⁵, gli esperti hanno esplorato le potenziali sfide tecnologiche emergenti nell'area dei Law, rispetto al Diritto Internazionale Umanitario, per promuovere una comprensione comune dei concetti, consentendo così un bilanciamento tra esigenze umanitarie e necessità militari assoggettando lo sviluppo delle armi al diritto internazionale umanitario, così da evitare effetti traumatici eccessivi o indiscriminati verso i civili.

L'Unione Europea, attraverso la Commissione Giustizia sta lavorando per identificare i termini entro cui poter determinare una legislazione europea in materia di

¹⁷³N. Ronzitti, *Uso e sviluppo delle armi autonome. Prospettive per un controllo a livello internazionale*, Istituto Affari Internazionali, n. 81, 2018.

¹⁷⁴Il 9 luglio, l'assemblea parlamentare dell'Organizzazione per la sicurezza e la cooperazione in Europa (OSCE) ha adottato una dichiarazione che sollecita i 57 Stati membri dell'OSCE a "sostenere i negoziati internazionali per vietare le armi autonome letali". Migliaia di scout hanno partecipato al 24° World Scout Jamboree in West Virginia il mese scorso, visitando una mostra della Campagna Stop Killer Robots e ponendosi domande determinate da sincera preoccupazione.

¹⁷⁵Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects, Report 2019 GGE on Laws, CCW/GGE.1/2019/3.

intelligenza artificiale. Di recente, nel giugno 2020¹⁷⁶, si è riunita discutendo la relazione sulle armi letali autonome, promuovendo l'idea di definire linee guida di uso del sistema, evitandone il divieto. È stato riconosciuto che l'uso della forza in guerra sta subendo una evoluzione di digitalizzazione con la produzione di sistemi robotici. I vari coordinatori hanno ribadito che è necessaria la presenza di una supervisione umana che possa rispondere ai principi dettati dalla Convenzione di Ginevra nonché i principi stabiliti dal diritto internazionale umanitario, prevedendo la possibilità di disinnescare l'arma. Ancora, la Commissione Europea per l'efficienza Della Giustizia (CEPEJ)¹⁷⁷ ha pubblicato la Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari, dal quale però possono estrapolarsi due principi fondamentali applicabili anche ai sistemi di arma autonoma letale.

Il primo principio mira ad assicurare che l'elaborazione e l'attuazione di strumenti e servizi di intelligenza artificiale siano compatibili con i diritti fondamentali garantiti dalla Convenzione europea sui diritti dell'uomo (CEDU). Il secondo prevede che la scelta etica sia realizzata a monte dagli elaboratori del programma e pertanto non vada lasciata all'utilizzatore, al fine di assicurare che le metodologie di discriminazione non vengano riprodotte o peggio aggravate.

La comunità internazionale ha ora un'opportunità storica per riflettere sia sui vantaggi e sia sugli aspetti negativi delle nuove tipologie di armi, identificando le linee guida per il loro utilizzo. Vanno considerate le problematiche legate all'attuale impianto normativo internazionale, rispetto al progressivo utilizzo di tali sistemi, che non è in grado di comprendere tutte le diverse tipologie di sistemi di intelligenza artificiale esistenti e quelle future.

Tra le varie soluzioni prospettate, oltre ad un nuovo Trattato è possibile:

- promulgare un Protocollo aggiuntivo alla CCW, incentrato sulla regolamentazione dei Laws. Il processo negoziale dovrà avvenire secondo le prescrizioni della Convenzione proponendo, con l'accordo di almeno 18 Stati parti, la convocazione di una conferenza per negoziare e concordare l'aggiunta di un nuovo Protocollo, che dovrà essere ratificato dagli Stati.

- aprire un negoziato al di fuori della CCW, convocando una conferenza ad hoc, come è stato fatto, ad esempio, per la Convenzione sulle bombe a grappolo.

- promulgare una Robot Law che comprenda gli aspetti più significativi della società, iniziando dalla regolamentazione di ogni prodotto o tecnologia sviluppabile con l'IA, nonché di ogni standard a cui una società che produce o gestisce questi sistemi deve attenersi. Recentemente, nel maggio 2020, il Parlamento Europeo ha stilato una sorta di road map¹⁷⁸ sugli obiettivi che gli utenti di armi robotiche negli

¹⁷⁶ Vedi anche la decisione del Parlamento Europeo 2020/2684(RSO) per la costituzione di una commissione speciale sul digitale e l'intelligenza artificiale.

¹⁷⁷ Commissione Europea per l'efficienza Della Giustizia (CEPEJ). Carta etica europea sull'utilizzo dell'intelligenza artificiale nei sistemi giudiziari e negli ambiti connessi, adottata dalla CEPEJ nel corso della sua 31ª Riunione plenaria (Strasburgo, 3-4 dicembre 2018).

¹⁷⁸ Nella roadmap (2007-2032) gli obiettivi che i costruttori e gli utenti di armi robotiche negli ambienti militari si sono imposti sono i seguenti: 1) Migliorare l'efficacia del comandante combattente e della coalizione senza equipaggio; 2) raggiungere una maggiore interoperabilità tra i controlli di

ambienti militari devono seguire ed ha istituito un apposito gruppo di Esperti di Alto livello con il compito di definire delle linee guida per la disciplina dei sistemi di intelligenza artificiale. Ha anche pubblicato una raccomandazione per la Commissione Europea su come l'intelligenza artificiale dovrebbe essere regolamentata, riconoscendo che l'IA, la robotica e le tecnologie correlate non presentano solo rischi, ma anche opportunità che contribuiscono al progresso e a beneficio della coesione sociale, è però importante definirne i limiti per evitare che, in nessun caso, possano causare danni alle persone.

Anche in materia di responsabilità la comunità giuridica si chiede se la responsabilità dei fatti causati dal sistema autonomo è riconducibile al proprietario quale persona giuridica, ovvero al conduttore, o allo stesso Stato¹⁷⁹. Benché gli agenti intelligenti non sono persone fisiche, né persone giuridiche e non hanno capacità di agire ai sensi del diritto, tuttavia, potrebbero creare fatti giuridici o causare eventi che producano effetti giuridici. Infatti, è emerso che serve una disciplina volta ad evitare che dalle azioni del Laws ne possa scaturire una violazione del Diritto Internazionale Umanitario.

Gli studiosi hanno prospettato diversi scenari:

- in base alle regole militari generali la responsabilità è assolta attraverso la catena di comando militare, dove le misure di responsabilità sono definite chiaramente negli ordini, nelle direttive e nelle procedure operative standard imposte. Gli incidenti che coinvolgono vittime civili in genere innescano le indagini. I soldati possono essere soggetti ad una disciplina amministrativa ovvero a sanzioni penali nelle ipotesi di inadempimento nell'esercizio delle funzioni. Le indagini di sospetta violazione possono anche concentrarsi, più in generale, sulla condotta di enti governativi o organizzazioni così da determinare raccomandazioni applicabili ad una riforma istituzionale per prevenire comportamenti futuri.

- creare una scatola nera per assicurare la registrazione elettronica delle azioni del Laws.

sistema, comunicazioni, prodotti di dati e collegamenti su sistemi non presidiati; 3) rafforzare la sicurezza con le migliori del controllo al fine di operare facilmente tra i vari tipi di robot. È importante prevenire intercettazioni, interferenze, dirottamenti; 4) promuovere lo sviluppo di politiche, standard e procedure che consentano tempestive operazioni e l'effettiva integrazione dei sistemi; 5) migliorare il coordinamento tra le autorità civili che gestiscono le zone aeree, marittime e terrestri e le autorità militari al fine di prevenire collisioni tra macchine con o senza equipaggio; 6) lo sviluppo di sistemi con sensori innovativi; 7) implementare misure di controllo positivo standardizzate e protette per i sistemi autonomi; 8) sostenere l'integrazione di capacità di combattimento convalidate in sistemi schierati/distribuiti attraverso un processo di prototipazione, test e supporto logistico più flessibile; 9) stimolare la concorrenza: incentivare i settori di ricerca e sviluppo per innovare il livello di automatizzazione dei sistemi di armi robotiche.

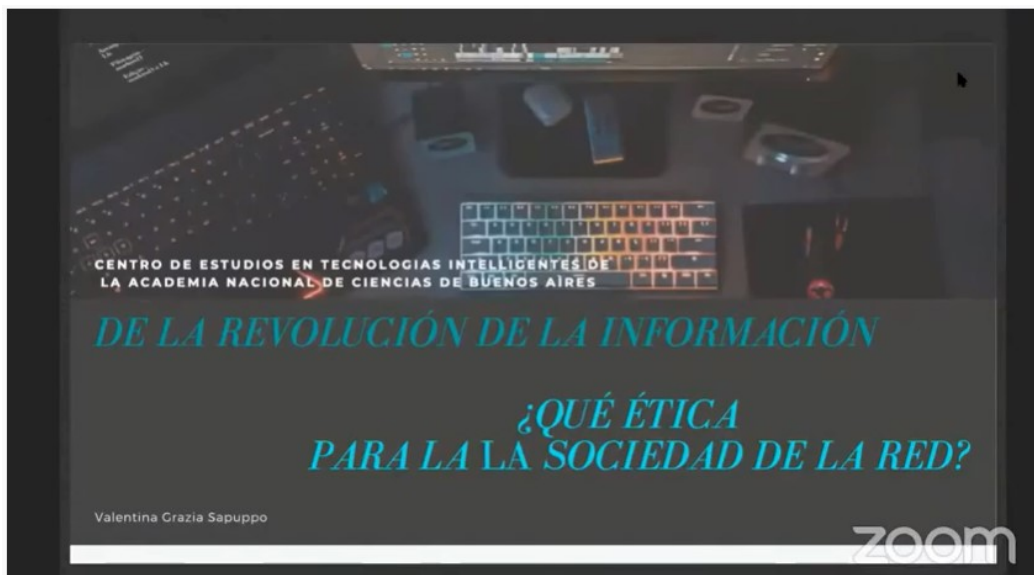
¹⁷⁹The GGE on Laws concluded in 2018 that “accountability for developing, deploying and using any emerging weapons system in the framework of the CCW must be ensured in accordance with applicable international law” and that “human responsibility for decisions on the use of weapons systems must be retained since accountability cannot be transferred to machines”; Report of the 2018 Session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems, CCW/GGE. 1/2018/3.

Questo strumento potrebbe agevolare il riconoscimento di responsabilità verso il sistema che commette un errore, come per esempio nelle ipotesi in cui un veicolo subacqueo autonomo colpisca per errore un peschereccio anziché un mezzo nemico.

- in base ai principi della responsabilità del produttore coinvolgendo anche gli utilizzatori o programmatori che hanno realizzato gli algoritmi decisionali. È stato introdotto il concetto di “comandante ragionevole”, riferendosi a quell’operatore umano che nelle sue decisioni dovrà equilibrare il potenziale danno collaterale con il vantaggio militare.

- in ogni caso, secondo il diritto internazionale che disciplina la responsabilità degli Stati, essi sono responsabili di atti illeciti a livello internazionale nelle ipotesi di violazioni del diritto umanitario commesse dalle loro forze armate e, in questo caso, dall’utilizzo di un sistema di armi autonome. Per punire le eventuali violazioni del DIU, vige la cosiddetta “giurisdizione universale”, secondo la quale ogni Stato ha l’obbligo di ricercare e punire i colpevoli. Inoltre, per assicurare una garanzia giudiziale nei casi in cui uno Stato non voglia o non possa intervenire verso i criminali, è stata istituita, sulla base del Trattato di Roma del 1998, la Corte Penale Internazionale

- la possibilità di attribuire ai robot una personalità giuridica, in quanto capaci di pensare ed agire autonomamente. Nello specifico, si tratterebbe di una personalità virtuale (e-personality), comparabile alla personalità giuridica degli enti di diritto privato, così da disciplinarne la responsabilità civile. O ancora il riconoscimento di “persone elettroniche” quali soggetti autonomamente capaci di essere imputabili di situazioni giuridiche attive e passive facendo ricorso al criterio di imputazione della responsabilità della culpa invigilando di colui che utilizza il sistema, nel momento in cui può rendersi conto dell’adozione di decisioni errate.



De la revolución de la información: ¿Qué ética para la sociedad de la red?

Por Valentina Grazia Sapuppo¹⁸⁰

¹⁸⁰Martino, Antonio (coord.), *Informe para ANCSA. Ciclo de sistemas inteligentes y ética cibernética*. Valentina Grazia Sapuppo, clase 87, nació en Catania, Sicilia, y obtuvo el título de Doctora en Derecho en el Departamento de Derecho de la Universidad de Catania, con una tesis experimental titulada *TecnoDiritto y CyberEthics*, supervisora S. Amato, miembro de la Comité Nacional de Bioética. Inicia la pasantía profesional, como colaborador legal del Bufete de Abogados Spampinato - La Rosa. Inscrita en el Colegio de Abogados de Catania, asiste al curso de formación de la "Fundación V. Geraci", escuela forense oficial del Tribunal de Catania, que le permite relacionarse con el COA - Consejo del Colegio de Abogados de Catania. El Dr. Sapuppo es un activista en el mundo asociativo. Miembro de ELSA - The European Student Association, junta local ELSA Catania, miembro de ISOC International y parte del ISOC Youth Observatory - Chapter Italy, miembro de Generation Ypsilon, lleva a cabo diversas actividades destinadas a proteger los derechos humanos, superando la brecha digital, a la afirmación y protección del derecho de acceso a Internet y lucha por la protección de los derechos de los aprendices legales en Italia, en concierto con todas las asociaciones comerciales a nivel nacional y sentados en las mesas técnicas con los representantes del Ministerio de Justicia, habiendo recibido delegación directa del presidente de Generation Ypsilon. Experta legal en el mundo digital, protección de la privacidad, protección de datos biométricos, se especializa en ética de la información. Escribe para revistas científicas jurídicas, incluidas *Salvis Juribus* e *IJ Editores*, de las que también es miembro del Comité Editorial como representante italiano. Colabora con diversas plataformas de información, incluidas las de e-learning specialist, *newsicilia* e *informalex*. Seleccionado como la excelencia italiana por EWA - Excellent Webinar Academy de ELSA Italia, confronta constantemente los temas de las nuevas tecnologías y con expertos italianos y extranjeros en el sector.

Ciclo "Sistemas Inteligentes y Ética"

1. Introducción

En un mundo inundado de información, el desarrollo disruptivo de la tecnología nos enfrenta al desafío tecnológico, uno de los mayores desafíos que ha enfrentado nuestra especie. La revolución de la información ejerce una presión sin precedentes sobre nuestro comportamiento y la ética individual. Debemos poder permitirnos el lujo de hacer una pausa para reflexionar un momento sobre el rumbo que está tomando lo digital y preguntarnos, en primer lugar, si nuestro sentido ético podría estar desactualizado.

La *Revolución de la Información*, también conocida como *Revolución Digital*, tiene sus raíces ya a finales del siglo XIX, en la época de los primeros y más importantes descubrimientos científicos, vinculados a los desarrollos de la tecnología moderna. Esta revolución, entendida por muchos como *Evolución*, ha cambiado el mundo mucho más rápida y radicalmente que la revolución científica anterior. Todo parte del concepto de *información*, entendida como objeto de comunicación.

2. ¿Qué es información?

Estamos en el momento en que se estaban formando las primeras redes de comunicación.

Norbert Wiener¹⁸¹, el padre de la cibernética, fue el primero en cuestionar la naturaleza de la información traducida por ultrafax. Él indicó que la información debe ser considerada como una medida de orden, y habló de cómo medir las señales entrantes y salientes. Pero la señal en sí es algo intangible. Entonces, *¿cómo podría esto ser medido?* Para tener una señal materialmente mensurable, Wiener afirmó que la materia en sí misma es energía e información en constante flujo, como sucede en la física nuclear. Uno no existe sin el otro.

Sobre la base de la segunda ley de la termodinámica, Claude Shannon¹⁸², colaborador de Wiener, hace que un nivel y funda la teoría matemática de la información. Ahora la información es ni materia ni energía, pero es sintáctica, como la de las señales de radio, líneas de teléfono, televisión por cable, las computadoras y el ADN humano. Según la *Teoría matemática de la información*, los seres humanos, los animales y la propia sociedad se componen de la información de Shannon. Estos son sistemas dinámicos que procesan información y se comunican entre sí a través de la retroalimentación. Habiendo puesto estos cimientos, el lenguaje ya no es un atributo exclusivo del hombre, sino un carácter que el hombre puede compartir no solo con el entorno que lo rodea, sino también con las máquinas. A su vez, las

Acepte el atrevido desafío de colaborar con ANCBA junto al profesor Antonio Martino, reportando resultados prometedores.

¹⁸¹ Wiener, N., *The Human Use of Human Being: Cybernetics and Society*, Mifflin, Houghton, 1950; sec. ed., Doubleday Anchor, Garden City, NY, 1954. Tr. it. *Introduzione alla Cibernetica. L'uso umano degli esseri umani*, D. Persiani, Bollati Boringhieri, Torino, 1970.

¹⁸² Shannon, E. Claude, Weaver, W., *The Mathematical Theory of Communication*, Univ. of Illinois Press, 1949.

máquinas se comunican entre sí a través de una *señal* matemáticamente medible llevado por un mensaje. Aquí, la era del automatismo ya no es ciencia ficción, sino un futuro ya previsto. En este futuro, las máquinas, como los humanos, son capaces de tomar decisiones de manera autónoma, que se adaptan al entorno y evolucionar, aprender de la experiencia y el cambio de su comportamiento futuro.

Una vez que se entiende que las máquinas se unirían con los seres humanos, dando lugar a sociedades híbridos en los que la información sería la materia prima para la comunicación, para evitar desastres inevitables que era necesario establecer reglas éticas.

Fue Luciano Floridi¹⁸³ quien inició el *Proyecto de Filosofía de la Información*, sentando así las bases de una nueva forma de disciplina filosófica que reemplazaría a las clásicas disciplinas antropocéntricas, que ahora se han vuelto puramente escolásticas. Floridi fundó el Paradigma para situarse en la base de una nueva *Philosophia Prima*. La información de Floridi siempre pretende ser un intercambio de datos pero, a diferencia de la de Shannon, es semántica y platónica porque incluye y concierne a las personas, la cultura, la sociedad, el bienestar, las interacciones sociales y todas las entidades, como tales, que actúan. en la Sociedad Red. A través del *Método Constructivista*, el llamado *El Mundo Noumenal de las Cosas en sí*, es decir, la realidad última, se manifiesta como una caja negra dentro de la cual nadie puede ver porque es incognoscible. Solo podemos construir Modelos de Realidad a través de los llamados Método de abstracción, hasta un resultado. Este resultado es el llamado *realismo informacional*, la idea de que el mundo está compuesto por la totalidad de objetos informativos que interactúan, de forma dinámica, entre sí y donde el ser humano también está compuesto por datos, como había dicho Alan Turing¹⁸⁴.

La *Nueva Macroética* de Floridi¹⁸⁵, que él define como *Ética de la Información*, se esfuerza por resolver el problema ético que surge de la Cuarta Revolución. En el *Infosphere* así entendida, cada entidad se considera un *Inforg*, un paquete de datos, dotado de un cierto valor mínimo y moral coherente, coherente en el sentido de que no presenta contradicciones y que, por tanto, representa y se revela, como información, por lo que es.

Partiendo de estas premisas, podemos añadir que la información como tal cuenta con un principio ontológico de la igualdad, de tal manera que todas las entidades que actúan en el *Infosphere* estará orientado a objetos, en el sentido de que actuará de forma desinteresada, imparcial y universal. Sobre todo, será leal porque se posicionará hacia el entorno sintético en el que se desenvuelve, independientemente de la posición que tenga. De esta manera, se crea una virtud epistémica, gracias a la cual todos los que pueblan la Sociedad Red actuará en el interés y para el

¹⁸³ Floridi, L., *Infosfera. Ética e filosofía nell'età dell'Informazione*, Torino, G. Giappichelli Editore, 2009.

¹⁸⁴ Turing, M. Alan, *Intelligenza meccanica*, a cura di G. Lolli, trad. it. N. Dazzi, Bollati Boringheri, Torino, 1994.

¹⁸⁵ *Ibidem*.

bienestar de todos los demás Inforgs, con la conciencia de que las acciones de cada uno condicionan el medio ambiente y la conducta de los demás.

Con esto en mente, Floridi cambia el enfoque de sus consideraciones éticas sobre las acciones, los caracteres y los valores de los agentes que trabajan allí.

Además, cada entidad en la Infosfera puede ser dañada o destruida simplemente alterando o eliminando la estructura de datos que la compone, previniendo su desarrollo y provocando su muerte - *Entropía*. Este es un fenómeno que tiene como resultado el empobrecimiento de la Infosfera. La Macroética de Floridi puede describirse como una *Teoría de la Ética No Antropocéntrica*, basada en el paciente y no en el agente, entendido el paciente como un entorno a cuidar. Sus principios fundamentales se establecen como teoremas¹⁸⁶: 0) *No se debe generar entropía en la Infosfera* 1) *Se debe prevenir la entropía en la Infosfera*. 2) *La entropía debe eliminarse de la infosfera*. 3) *Se debe promover el bienestar de la información ampliando (cantidad de información), mejorando (calidad de la información) y enriqueciendo (variedad de información) la Infosfera*.

Floridi, por tanto, cambia de punto de vista. Está claro que se podría argumentar que esta ética de la información es demasiado abstracta y, por lo tanto, irrelevante en la práctica, incluso si es teóricamente correcta. Lo que necesitamos entonces es una ética basada en la moral. Porque, para lograr una buena ética para todos, cada actor de la Sociedad Red debe, en primer lugar, confiar en las otras entidades –fides socráticas– que operan allí. Para ello debemos actuar con conciencia, evaluando las externalidades, positivas y negativas, que surgen de nuestras acciones, como un legado para las generaciones futuras.

3. Los componentes de la Sociedad de la Red

Para entender mejor el discurso moral dentro del nuevo cosmos informáticos, vamos a tratar de contextualizar mejor lo que se ha dicho hasta ahora.

Tiempo: Hyperstoria, caracterizada por el predominio de las TIC Digitales, activos intangibles que tienen la propiedad de registrar una cantidad innumerable de datos, creados y procesados en la Infosfera en la que Big Data son los protagonistas, desarrollando una red de interconexiones casi vertical.

Espacio: A través de interfaces digitales, el hombre cruza la puerta de una nueva biosfera hecha de datos. Lo experimenta como su nueva realidad. Crea el nuevo paradigma de que lo que es real es informático, lo que informa es real. Ya no hay diferencia entre Online y Offline, pero vivimos Onlife.

Agente: En la infosfera, los humanos, robots, soft bots y otros agentes artificiales como organizaciones, empresas y gobiernos son considerados “agentes”. Pero una entidad debe cumplir con tres criterios para ser definida como agente: 1) *Interactividad*: el agente y su entorno pueden actuar sobre uno u otro. 2) *Autonomía*: El agente puede cambiar su estado independientemente de sus interacciones con el

¹⁸⁶ *Ibidem*.

entorno. Un agente debe tener al menos dos estados de estar, hasta cierto punto, separado de su entorno. 3) *Adaptabilidad*: la capacidad del agente de modificarse evolucionando gracias a la capacidad de aprender de sus experiencias anteriores.

Autocomprensión: la mediación tecnológica nos obliga a reflexionar y revisar la forma en que nos entendemos a nosotros mismos. Compuesto por información, podemos ser tratados como productos comercializados en el mercado publicitario. El valor de una mercancía, por supuesto, depende de su propio poder adquisitivo y nosotros y nuestro valor se cuantifican y catalogan en función de la cantidad de datos que ingresamos como miembros de un grupo de consumidores que opera en la Web.

4. La Sociedad de la Red da vida a la Cultura Proxy

¿Qué tan humillante puede resultarnos tener que dividir el espacio con inteligencia artificial que, por sus características innatas, lleva al hombre a sentirse estúpido y nada inteligente, obligado a ponerse la camisa de fuerza de una condición miserable?

Si estamos hechos de información, ¿qué nos convierte en organismos informativos especiales en medio de tantos organismos informativos especiales?

Aquí, esta imagen crea algunos problemas de aceptación si adoptamos un punto de vista ético crítico.

“Abrazamos el nacimiento de Internet como un regalo del cielo mediante el cual nos sentimos libres de cualquier patrón y pudimos volar hacia la nueva luz del conocimiento.

Éramos inocentes y fascinados como Ícaro. Nuestras alas de las velas. Todavía no podíamos imaginar que la cera con la que los habíamos adherido se hubiera licuado en el océano de datos que, sin darnos cuenta, entonces habríamos creado”¹⁸⁷.

Cualquier libertad tiene su precio que pagar y los gigantes de la tecnología, por ejemplo, han comenzado a hacer un uso descarado de la biometría en nuestras huellas dactilares, nuestro óvalo, nuestro timbre de voz, nuestros datos.

¿Nos estamos dando cuenta? Nos hemos convertido en tipos específicos de consumidores debido a la marca que nos ha convertido en un producto anónimo masivo, expuestos en línea a miles de millones de otros Inforgs, al compartir y promover la *Cultura Proxy*.

Ninguno de nosotros hoy en día se preocupa realmente por la escasez que representamos en términos económicos: unos pocos centavos de dólares. Google nos engaña de que podemos recuperarnos a través de la construcción y personalización de Páginas de Internet, Blogs, Redes Sociales, videos en Youtube. Medios que permiten compartir nuestra vida privada, de nuestras preferencias sobre cual-

¹⁸⁷ Sapuppo, V., *Tecnodiritto y CyberEthics*, Tesis de Grado Experimental, Departamento de Derecho, Universidad de Catania, 2018-2019.

quier tema, desde la comida hasta los viajes. *Hablemos de Internet de las cosas. Todo es A2A, conectado de todo a todo.*

En nosotros, la distinción entre quiénes somos, nuestra identidad personal, y quiénes pensamos que somos, nuestra concepción del yo social es clara.

5. La filosofía de la mente y la generación hiperconsciente

En este contexto, la filosofía de la mente puede ayudarnos, porque cuanto más potentes y disponibles las TIC, más hipnótica se vuelve la mirada digital, haciéndonos perder en la contemplación y percepción que los demás tienen de nosotros. Y esto no es necesariamente bueno.

“¿Quién soy? ¿Quién soy yo para ti? La paradoja de la identidad nos cautiva.

¿Quiénes somos hoy? ¿En quiénes nos convertiremos?

¿Quiénes podríamos ser si pasamos una parte cada vez mayor de nuestro tiempo como Inforg?”¹⁸⁸

Existe una barrera artificial que separa nuestro auténtico yo interior de la realidad del mundo exterior. Pero en la era de las redes sociales este proceso de introspección sobre la autoconciencia se ha exteriorizado a través de *Micro narrativas*. Esta cosa es fascinante y aterradora al mismo tiempo. Ver que las personas pasan gran parte de las horas de sus vidas construyendo para embellecer un perfil en línea, convirtiéndose en uno con su creación.

El yo social da vida al fenómeno de la construcción de la identidad personal en línea, involucrando fuertes implicaciones en la concepción que tenemos de nosotros mismos y dando vida a un círculo vicioso para quienes trabajan a diario en estas actividades.

La *Generación Hiperconsciente* donde la personalidad social de uno es una mera creación del pensamiento de otras personas y donde cada uno de nosotros nunca deja de pensar en cómo se ve a los demás. Evidentemente seguimos siendo nosotros mismos aunque ya no seamos los de ayer, y citando a Heráclito, *Panta Rei*. “No se puede bajar dos veces por el mismo río”¹⁸⁹.

“La mariposa no es la oruga y tú no eres tu perfil de Instagram”¹⁹⁰.

6. Polarización social

Sin embargo, en un planeta cada vez más interconectado, la gente vive una vida cada vez más solitaria. Necesitan encontrar significado, propósito en nuestras

¹⁸⁸Ibidem.

¹⁸⁹ Eraclito, *Eraclito. Testimonianze, imitazioni e frammenti*, a cura di R. Mondolfo, M. Marcovich, L. Taran, Bompiani, Firenze, 2007.

¹⁹⁰Sapuppo, Ibidem.

vidas, creando una comunidad global que llene el vacío dejado por la desintegración de las comunidades humanas como alguna vez las entendimos. Lo llamamos *Polarización social*¹⁹¹: en un sistema de múltiples agentes es más fácil construir un diálogo relacionando a personas que tienen algo en común. Los sistemas multiagente, de hecho, son el resultado de la unión de personas en comunidades ubicadas en línea, que crean grupos de interés consolidado sobre determinados temas.

Por lo tanto, durante el último siglo, la tecnología nos ha alejado de nuestros cuerpos. Estamos perdiendo la capacidad de percibir la vida y somos absorbidos por nuestros smartphones, preocupándonos más por lo que sucede en el ciberespacio que por lo que sucede en nuestros hogares. Gx, Gy y Gz se han convertido en la generación *Always On*, siempre en línea, siempre conectada. Comemos nuestras comidas frente al monitor mientras trabajamos, mientras revisamos el correo electrónico o mientras vemos una película en Netflix. Preferimos compartir nuestras experiencias a través de compartir alimentos, buscamos una comparación.

Cuando muera, ¿*simplemente desapareceré*? El camino parece ser el de la Inmortalidad Digital, donde el hombre colocará fija en el tiempo esa información sobre sí mismo que quiere mostrar al mundo, registrándola de forma indeleble, gracias al C.A. Pensemos en *Black Mirror*, una serie de Netflix.

Inconscientemente hemos envuelto el mundo en torno a las TIC, hasta el punto de convertirlas en fuerzas ambientales, antropológicas, sociales e interpretativas, que crean y forjan nuestra realidad física e intelectual, modifican nuestra autocomprensión, cambian la forma en que nos relacionamos con otros y con nosotros mismos. Hacen todo esto de una manera penetrante, profunda e incesante, creando una profunda revolución cultural.

“¿Podremos sacar el máximo partido a estas tecnologías? ¿Cuáles son los riesgos? ¿Nos harán cada vez más poderosos y hábiles o estaremos empeñados en adaptar lo que nos ofrecen dentro de sus límites? ¿Nos veremos forzados a entrar en espacios físicos y conceptuales delimitados? ¿Nos ayudarán a solucionar los problemas sociales y medioambientales más acuciantes o acabarán por exasperarnos?”¹⁹².

7. ¿Se puede llamar inteligente a una computadora?

Después de que Deep Blue derrotara a Kasparov, después de que Watson, el sistema IBM, ganara en 2011, en el juego de preguntas “Jeopardy!”, una computadora derrota una computadora: AlphaZero de Google en 2017 derrota a Stockfish 8, el programa de campeones de ajedrez de 2016.

¹⁹¹XVI Rapporto Censis sulla comunicazione, *La Polarizzazione dell’Informazione*, 2017-2019, visibile al link www.censis.it/comunicazione/16%C2%B0-rapporto-censis-sulla-comunicazione/la-polarizzazione-dell%E2%80%99informazione.

¹⁹²Sapuppo, *Ibidem*.

Turing¹⁹³ afirmó que una máquina solo podía llamarse inteligente si podía demostrar que su comportamiento no era diferente al del ser humano promedio. Hoy en día, las computadoras resuelven problemas y cuentan con la capacidad de planificar y armar acciones para lograr objetivos particulares.

Pero la inteligencia artificial tiene una naturaleza algorítmica, basada en el código binario de Shannon, una naturaleza en la que los límites de la computación se convierten en los límites mismos de su inteligencia.

Hay un umbral semántico entre nosotros y nuestras máquinas y no sabemos cómo hacer que lo crucen. Pero, *¿quizás la inteligencia artificial no debería intentar simular el comportamiento humano inteligente y tratar de emular sus propios resultados?*

8. La Cuarta Revolución y el problema ético

La Cuarta Revolución planteó el *Problema Ético* de reconfigurar la ética clásica en *Ethics by Design* en el que una serie de cuestiones éticas sobre el comportamiento moral de la Sociedad Hipershistórica son centrales, relacionadas con cuestiones como la confianza, la privacidad, la transparencia, la libertad de expresión, lealtad, respeto, confiabilidad, reputación, Estado de derecho.

Para ello, es posible utilizar el *Concepto de Interlucencia* de B. De Bruin¹⁹⁴ como piedra angular. El concepto de interlucencia se desarrolla como una virtud epistémica directa para establecer un conocimiento común que sea capaz de registrar verdaderamente una comunicación éticamente eficaz, lo que genera una adición en términos hegelianos, que reconoce que el acceso al conocimiento digital es un derecho al bienestar de todos.

Para satisfacer este propósito no son suficientes líneas infinitas de términos y condiciones para ser aceptadas con un simple clic, en cuanto a las condiciones de uso de un espacio Cloud, contratos electrónicos, pensemos en Click-Wrapping, Scroll-Wrapping, Browser- Wrapping y envoltente de inicio de sesión.

Se debe crear una verdadera conciencia social basada en el principio de *AccountAbility*, condición de reciprocidad en la que los temas de ontología, seguridad, protección de la información personal, cuestiones éticas y legales que les conciernen son protagonistas.

La misma idea de utilidad pública, revolucionada por la computadora personal, ha sido abandonada en favor de un ideal computacional entendido como *Private Asfair*, definido como la transformación en *Commodity* de los recursos y software TIC, que presume de la intención de liberar al usuario del corpus material con dispositivos cada vez más inteligentes y con un uso cada vez más dependiente del Cloud Computing.

¹⁹³Turing, *Ibidem*.

¹⁹⁴ B. De Bruin, *Ethics and the Global Financial Crisis: Why Incompetence is Worse than Greed.*, Cambridge University Press, Cambridge 2015.

Always On solo se preocupa por tener la oportunidad y la libertad de hacer cosas con sus datos, en lugar de preocuparse por dónde se almacenan sus datos o quién es el propietario. Forman parte de una dimensión en la que los valores pasan del producto a los servicios que representa un mismo producto, que es el mundo de *Everything-As-a-Service*, un mundo en el que no es necesario comprar y poseer, sino que solo basta con tener libertad desechar.

El informe reciente del *Berkman Center for Internet & Society*¹⁹⁵ descubrió cómo *“Interconectividad universal, la provisión de un inmenso espacio público accesible para todos y en continua expansión, en el que todos vierten continuamente sus datos sensibles, registrando su posición, movimientos, intereses, opiniones, información sobre su estado salud, hábitos, estilo de vida, rasgos somáticos, expresión de sentimientos y mucho más, marcando y matizando los tiempos de su día típico y ofreciendo datos para la construcción de su propia personalidad virtual apta para integrar información para el perfil de un tema típico”*.

9. El problema moral en términos de privacidad

En términos de privacidad, el problema moral es la adquisición y mal uso de la información de otras personas.

Debemos procurar que la ética de la información resuelva el problema de lo que es moralmente correcto y lo que es moralmente incorrecto y asegurar que los seres humanos, sobre todo, tengan el mismo cuidado y respeto ético por la información que las demás personas. Lo mismo debería ocurrir con las máquinas o con quienes, a través de ellas, obtienen datos de las personas.

Como informa Claudio Sarra, en un artículo del *New York Times* titulado *Big data: Finder keepers, losers crypers?*¹⁹⁶ se afirma que *“Es mi secreto lo que se puede conocer, y no es nada obvio que la habilidad mostrada en descubrir información privada sobre mí, que no pretendía comunicar, sea suficiente para justificar su inescrupuloso uso económico, como una evolución del tema de las Elecciones Invisibles ya discutido por James Moor”*¹⁹⁷.

¿Por qué hablamos de privacidad cuando hablamos de información?

¹⁹⁵ Marwik ET AL., 2010, p.13: “Para los jóvenes, la ‘privacidad’ no es una variable única. Los diferentes tipos de información se perciben como más o menos privados; elegir qué ocultar o revelar es un proceso intenso y continuo... En lugar de concebir una división distinta entre ‘privado’ y ‘público’, los jóvenes conciben los contextos sociales como múltiples y entrecruzados... De hecho, la distinción entre ‘privado’ y ‘público’ es problemática para muchos jóvenes, quienes tienden a concebir la privacidad de formas más matizadas, considerando los espacios de red como ‘semipúblicos’ o haciendo distinciones entre diferentes grupos. di ‘amigos’... En muchos estudios sobre la relación entre los jóvenes y la privacidad, esta última no se define ni se entiende como un activo automático. Sin embargo, revelar información no es necesariamente arriesgado o problemático; produce muchos beneficios sociales que a menudo no se mencionan” (<https://cyber.harvard.edu>).

¹⁹⁶ Duhigg, *Ibidem*.

¹⁹⁷ Moor, J. H., *What is Computer Ethics?*, in “*Metaphilosophy*”, Canada, Wiley, 1985, p. 266 a 275,

Hemos dicho que en la Infosfera, cada entidad forma una red de relaciones de dependencia y relaciones. Los datos son la salida de los procesos de análisis y pueden convertirse en la entrada de otros procesos de información, pensemos en el *Celestial Jukebox* de Big Data.

Tenemos cuatro tipos de privacidad: privacidad física, privacidad mental, privacidad en la toma de decisiones, privacidad de la información.

10. La necesidad de una ciberética en el Proyecto “Sociedad de la Red”

Orban ha creado el *Proyecto Sociedad de la Red*¹⁹⁸, cuyo punto clave se encuentra en el hecho de que el cambio social y económico solo se puede difundir si se cuenta con una base tecnológica sólida.

La solución al problema, por tanto, debe ser inherente a la propia tecnología y la única forma de afrontar los problemas destacados es saber quiénes somos y, sobre todo, qué buscamos.

Nuestras opciones se transmiten mediante el marketing virtual. Estamos bajo vigilancia las 24 horas del día, los siete días de la semana. Los piratas informáticos y las instituciones están bien informados sobre todas nuestras actividades en línea y fuera de línea. Covid-19 y los sistemas de rastreo han exacerbado un control panóptico ya implementado por las políticas antiterroristas.

Una espléndida ilustración de nuestra situación actual se puede encontrar en el ejemplo clásico, que puede parecer trivial pero no lo es, de la película dirigida por Alfred Hitchcock en 1954, *La ventana al patio*. Vivimos en una infosfera de vidrio, donde ya no es suficiente apagar el enrutador para usar el manto de la invisibilidad.

Todos conocen a todos sobre todos y todos son posibles falsos amigos. Es necesario que el llamado CyberEthics ofrece herramientas conceptuales básicas serias para abordar la multitud de problemas difíciles que surgen y que se vuelven a conectar sobre todo al uso, estudio y explotación de nuestros datos personales, de nuestros datos biométricos para AIDC, Identificación Automática y Captura de Datos, para fortalecer la economía impulsada por datos, aprendizaje automático, Internet de las cosas, grupos de minería de datos, lucha contra lo oscuro con la protección de la privacidad, con código abierto, con el uso virtuoso de la tecnología Blockchain, con criptografía, con anonimización, con el procesamiento de contraseñas, con la creación de firewalls, protocolos específicos, servicios y sistemas de alerta para protegerse de la violación de datos.

11. Conclusiones

¹⁹⁸ Orban, D., *La società della rete. Nuove idee per l'uomo del futuro*, Milano, Hoepli, 2015.

¿Cómo no reflexionar sobre las riesgosas consecuencias de un mal uso mercantil de lo que caracteriza al ser humano a favor de un fortalecimiento de la IA? ¿Cómo no pensar en la posibilidad de elección, que no tendrán las generaciones futuras, de utilizar el propio cuerpo para acceder a lugares, servicios e información?

Agere Sequitur Esse, es decir, la acción sigue al ser. Necesitamos estimular la conciencia humana. Debemos preguntarnos: *¿qué es importante para nosotros?*

Podemos creer que la vida es un valor absoluto e inviolable, y es su propia naturaleza la que nos permite afirmarlo ontológicamente. Y los datos biométricos, que traducen la vida de las personas, deben considerarse de la misma manera. El desafío es empoderar a los usuarios, difundir los derechos fundamentales incluso en la era digital. Transmitir esto a las próximas generaciones, tratando de transmitir los valores de dignidad y libertad que hemos conquistado a través de la legislación nacional, europea e internacional, con los principios de responsabilidad, privacidad por diseño y privacidad por defecto, transparencia, consentimiento, autodeterminación, pero no es tan obvio mantenerlo.

Si, por ejemplo, la información relativa a miles de millones de personas se concentra sin preocuparse por cuestiones de privacidad, el algoritmo se desarrollaría leyendo al ser humano como objeto de estudio y análisis.

Escuchamos cada vez más que quien posee los datos es dueño del futuro. *Data Mining* es la nueva carrera de los gigantes de la alta tecnología, un nuevo modelo de negocio. A través de los servicios de entretenimiento gratuitos, el llamado Mercados de atención que operan a través de la elaboración de perfiles analíticos de *Big Data Pools*, tanques que recogen cada huella online que dejamos en la red, especialmente gracias al seguimiento operado por cookies. Somos observados en lo que hacemos, lo que buscamos, lo que nos gusta. *Cambridge Analítica Scandal* es un excelente ejemplo de esto, sin mencionar a uno de sus predecesores, *Case Target Corp*¹⁹⁹.

Utilizando el inmenso poder del aprendizaje automático, la inteligencia artificial puede analizar una gran cantidad de información de forma centralizada.

A través de sensores biométricos, escáneres faciales o brazaletes, realidad aumentada, chips debajo de nuestra piel o en nuestro cerebro –Elon Musk: *neuralink*–²⁰⁰, el algoritmo puede incluso leer y registrar datos sobre nuestras emociones, dígnos si estamos diciendo la verdad o no, ya sea que estemos contentos o enojados y lo hace a través de un análisis de medición biométrica de nuestros músculos faciales, nuestros ojos, nuestra presión arterial, nuestra actividad cerebral.

Si nos sentimos abrumados y confundidos, estamos en el camino correcto. Los procesos globales y la tecnología que utilizamos se han vuelto realmente complicados para que cada persona los entienda realmente. Sin embargo, aunque en un futuro próximo el mundo estará cada vez más articulado y nuestra vida cada vez

¹⁹⁹Duhigg, C., *Cómo las empresas aprenden sus secretos*, NYT, 2012.

²⁰⁰Elon Musk, *Neuralink*, proyecto visible en el link <https://neuralink.com>.

más llena de compromisos, quizás deberíamos hacer un pequeño esfuerzo para reflexionar sobre cómo actuamos en el espacio ilimitado de Internet Society.

Una clave ética para una posible solución es la de Tristan Harris²⁰¹. Cuéntenos sobre el método para medir el *tiempo bien empleado*. Adoptar un nuevo modelo que anime a las personas a conectarse solo cuando sea realmente necesario, a prestar más atención a su entorno físico, sus cuerpos y sensaciones.

Resiliencia. El cambio es la única constante. Las personas necesitan herramientas críticas para interpretar la información, para distinguir lo que es importante de lo que es irrelevante y, sobre todo, para poder enmarcar toda la información en un escenario más amplio y tomar las decisiones correctas. Tenga siempre en cuenta las 4C: crítica, comunicación, colaboración y creatividad. *Open Mind*. Reinventándonos continuamente, de esta forma podemos ayudar a mejorar la sociedad. Pero, *¿qué es lo correcto ante una situación sin precedentes? ¿Qué hacer cuando está inundado con una gran cantidad de información y no hay forma de que pueda absorberla y analizarla?*

Se requiere una gran flexibilidad mental y una conspicua reserva de equilibrio emocional y será necesario adentrarse en lo desconocido de forma serena.

¿En qué podemos confiar? ¿En tecnología? ¿Sobre nosotros? Nosce Te Ipsum. Primero debemos conocer “nuestro” sistema operativo, saber quiénes somos y qué queremos de la vida.

Bibliografía

- B. De Bruin, *Ethics and the Global Financial Crisis: Why Incompetence is Worse than Greed*, Cambridge, Cambridge University Press, 2015.
- Duhigg, C., *Cómo las empresas aprenden sus secretos*, NYT, 2012.
- Eraclito, Eraclito, *Testimonianze, imitazioni e frammenti*, a cura di R. Mondolfo, M. Marcovich, Firenze, L. Taran, Bompiani, 2007.
- Floridi, Luciano, *Infosfera, Etica e filosofia nell'età dell'Informazione*, Torino, G. Giappichelli Editore, 2009.
- Moor, J. H., *What is Computer Ethics?*, in “Metaphilosophy”, 266-275, Canada, Wiley, 1985.
- Orban, D., *La società della rete. Nuove idee per l'uomo del futuro*, Milano, Hoepli, 2015.
- Shannon, E. Claude, Weaver, W., *The Mathematical Theory of Communication*, Univ. of Illinois Press, 1949.
- Sapuppo, V., *Tecnodiritto y CyberEthics*, Tesis de Grado Experimental, Departamento de Derecho, Universidad de Catania, 2018-2019.

²⁰¹Harris, Tristan, *¿Cómo maneja éticamente los pensamientos y acciones de la mente de dos mil millones de personas todos los días?*, visible en el link www.tristanharris.com.

Turing, M. Alan, *Intelligenza meccanica*, a cura di G. Lolli, trad. it. N. Dazzi, Bollati Boringheri, Torino, 1994.

Wiener, N. *The Human Use of Human Being: Cybernetics and Society*, Mifflin, Houghton, 1950; sec. ed., Doubleday Anchor, Garden City, NY, 1954. Tr. it. *Introduzione alla Cibernetica. L'uso umano degli esseri umani*, D. Persiani, Bollati Boringheri, Torino, 1970.

Algunas reflexiones sobre la ética de los desarrolladores y diseñadores de sistemas inteligentes

Por Elisa Palomino Angeles

1. Introducción

En este trabajo se pretende determinar algunos de los problemas éticos que encontramos en el sistema ético, para ello se desarrollan primero los elementos que componen nuestro sistema, entre ellos los sujetos que se encuentran ante estos dilemas éticos, así como el entorno en el cual se desarrollan las conductas de nuestros sujetos, analizando las diversas funciones y responsabilidades que desempeñan los ingenieros de software, así como también reflexionamos sobre la función y el poder en el mercado digital que desempeñan las empresas, visualizando de manera general al Estado. Para posteriormente, continuar con la trascendencia de los datos en los sistemas inteligentes, para lo cual desarrollamos algunas etapas en las cuales nos percatamos puede darse más vulneraciones, por la falta de ética de los ingenieros de sistemas inteligentes, en virtud de los sesgos humanos de datos o de manipulación de datos, se automatizan por cuestiones de la personalidad del individuo que crea el programa o por otros factores externos como imposiciones de mercado digital, lo que origina que los sistemas inteligentes deje aún lado los sistemas éticos, que deberán estar centrados en el ser humano.

a. El contexto ético de la actualidad

El inicio de este trabajo comienza con la definición de lo qué es la ética, desde un punto vista etimológico se deriva del griego *ethos* que significa carácter, costumbre, y desde la perspectiva de la filosofía la ética es una rama de ésta, cuyo objeto

de conocimiento es la “teoría de carácter y de las costumbres”²⁰², pero en un texto actual, esto debido a que hoy en día la diversidad de representaciones sociales cambian aceleradamente desde un ámbito espacial y temporal, por ello, resulta necesario plantearnos referentes comunes para abordar la problemática ética que se presenta con los diseñadores y desarrolladores de la inteligencia artificial; para eso es necesario identificar todos los elementos que se encuentran en nuestro sistema complejo, siendo uno de estos nuestro espacio cibernético, los sujetos, las autoridades, las tecnologías, los diversos procedimientos de identificación, control, los protocolos y estándares internacionales, la infraestructura.

Sin lugar a duda el ciberespacio es el escenario en el cual se originan las conductas lícitas o ilícitas, correctas e incorrectas de los seres humanos que deben ser percibidas para abstraer las representaciones sociales que requieren tener un sistema ético, para ello debemos entender por ciberespacio el:

Ámbito operacional virtual en el que se desarrollan actividades de creación, procesamiento, almacenamiento, intercambio y explotación de información digital, a través de redes interdependientes e interconectadas – vinculadas a internet o no– y el software y firmware de dispositivos asociados a las mismas, siendo su carácter distintivo el empleo excluyente de las TICs (Tecnologías de la Información y Comunicaciones) y la interacción permanente con los otros ámbitos operacionales²⁰³.

El referido concepto establece los elementos que se encuentran dentro del ciberespacio nos proporciona las actividades de creación

La inteligencia artificial ha sido definida como un “conjunto de tecnologías que hacen posible que los sistemas digitales perciban, aprendan, razonen y colaboren en el proceso de toma de decisiones para la resolución de problemas que resultan similares a los que hacen los seres humanos”²⁰⁴.

Un punto interesante que trataremos en estas reflexiones es precisamente, a la ética como una representación social, la cual va variando de acuerdo a las fronteras, las políticas y los contextos sociales, consideramos que para la ética se deben considerar valores comunes entre los Estados, porque sí bien es cierto que se realizan construcciones o representaciones sociales, y éstas van cambiando de lugar y a través del tiempo, siempre hay una parte axiológica que contempla los valores dignidad, igualdad, libertad, justicia, honestidad, respeto y solidaridad, consideramos son los principales que debe tener una sociedad, donde la ética tiene como fin la protección de la dignidad de los seres humanos, en cualquier región y tiempo.

Ahora bien, los distintos sistemas de ética deberán adaptar un sistema jurídico acorde a sus necesidades, pero, si bien es cierto, que hay diferencias entre las sociedades de cada Estado, también como lo indicamos tenemos valores comunes

²⁰²Canal: Academia Nacional de Ciencias de Buenos Aires, 2020, 12m14s.

²⁰³Canal: Academia Nacional de Ciencias de Buenos Aires, 5/8/20.

²⁰⁴Canal: Academia Nacional de Ciencias de Buenos Aires, 5/8/20.

en todas las sociedades los cuales deben ser fundamentos éticos en la consciencia de los desarrolladores y diseñadores de la inteligencia artificial.

El debate sobre la situación de establecer un sistema ético por regulación contra un sistema ético por diseño, es un punto de partida que nos deberá de llevar a la solución precisa de acuerdo a los especialistas, el verdadero problema se encuentra la selección de contenido de las cajas que componen el sistema operativo de la inteligencia artificial, porque el pretender regular la ética de una sociedad puede encontrar una diversidad de problemas, políticos, económicos, culturales, sociales y jurídicos, esto debido a los diversos factores que influyen en las relaciones jurídicas que se establecen ya que generalmente son de asimetría del poder, en las cuales tenemos que los grandes desarrolladores y diseñadores de los programas de cómputo son los oligopolios y monopolios, los cuales controlan el mercado, y esto trae consigo la creación de los programas, que pueden ser o no ser predeterminados por éstas, ya que esta creación les dará como resultado la generación de riqueza y beneficios para estas grandes empresas transnacionales, lo que origina que la información pueda ser o no ser alterada, eliminada, o modificada o simplemente incluirse con errores por algunos sesgos humanos. Esto en la actualidad se investigó en Estados Unidos algunas de las problemáticas que existen entre las cuatro corporaciones que tienen

...el predominio de Amazon, Apple, Facebook y Google, y sus prácticas comerciales para determinar cómo su poder afecta nuestra economía y nuestra democracia. Además, el Subcomité realizó una revisión de leyes antimonopolio existentes, políticas de competencia y niveles de aplicación actuales para evaluar si son adecuado para abordar el poder de mercado y la conducta anticompetitiva en los mercados digitales (Comité de la Cámara de Representantes de los Estados Unidos sobre el Poder Judicial, 2020).

Es importante destacar el rol que tienen las corporaciones que predominan en los mercados digitales, y como sus prácticas desleales de comercio afectan tanto a nuestras economías, como a nuestra democracia y también a nuestros sistemas éticos, porque si no se tiene el predominio del mercado digital, los códigos éticos serán ineficaces, porque no podremos adoptar una ética de diseño mientras haya un predominio del mercado digital en cual sólo un número reducido de empresas tengan el poder, que en nada favorece a nuestra sociedad, por ello cuando se plantea como debate que modelo debe predominar la ética por regulación o por diseño, nosotros nos inclinamos por la primera, debido al anterior argumento, por estas razones, más vale regular la ética a través de códigos de conducta, pues la economía digital muestra deficiencias que a continuación se enuncian:

Durante la última década, la economía digital se ha vuelto muy concentrada y propensa a monopolización. Varios mercados investigados por el Subcomité, como las redes sociales, la búsqueda general en línea y la publicidad en línea están dominadas por solo una o dos empresas. Las compañías investigadas por el Subcomité (Amazon, Apple, Facebook y Google) han tomado el control sobre canales clave de distribución y han llegado a funcionar como guardianes. Solo una década en el futuro, el 30% de la producción

económica bruta mundial puede recaer en estas empresas, y sólo un puñado de otros²⁰⁵.

Hoy en día seguimos observamos que las empresas son un elemento que ha adquirido un poder real, el cual puede limitar tanto la competencia, como la ética, esto debido a que los monopolios y oligopolios pueden imponen modelos de consumo y sistemas inteligentes, en los cuales sólo se persigue el lucro y no contemplan la realidad social, ni el beneficio social, por ello, es necesario reduzcan su poder y una de las sugerencias de este Subcomité de Judicatura de los Estados Unidos de Norteamérica que realizó la investigación, recomienda esto.

2. Los sujetos desarrolladores y diseñadores de sistemas inteligentes

El sistema de ética, parte de la regulación de los caracteres y las costumbres de los seres humanos, que permiten la vida más plena y armónica de la sociedad, en este sistema nos encontramos con los sujetos a quienes va destinado el objeto de conocimiento de la ética, en el caso particular encontramos los siguientes: *a)* usuarios de estos programas, *b)* clientes o empresarios, *c)* Estado; y *d)* los ingenieros de software son aquellos que contribuyen, mediante la participación directa o enseñanza, al análisis, especificación, diseño, desarrollo, certificación, mantenimiento y pruebas de sistemas de software. Debido a sus funciones en el desarrollo de sistemas de software, los ingenieros de software tienen suficientes oportunidades para causar beneficio o generar daño y para habilitar o influenciar a otros a causar daño o beneficio²⁰⁶.

Con los procesamientos, almacenamientos, intercambios, explotación, creación, y el uso de los softwares se originan las conductas antiéticas tanto en los ámbitos público y privado, los cuales se tiene que adecuar uno o varios sistemas éticos, dada la complejidad que representa el pretender homogenizar un modelo ético, nunca será la solución, debido a que las construcciones sociales van evolucionando y variando de acuerdo al lugar, sus cultura y sociedad.

En nuestro sistema ético es el Estado quien tiene un rol preponderante porque debe tutelar los derechos fundamentales en los sistemas inteligentes, así como cooperar con los demás Estados para efectos de combatir la inseguridad jurídica, contribuyendo con inversiones económicas para tener los medios idóneos para combatir la inseguridad, controlar los entes privados, así como a los usuarios, a través de órganos del Estado, en la medida de lo posible.

Es preciso destacar que el Estado y la cooperación de los Estados resulta necesario para establecer Códigos de ética internacionales, con la finalidad de salvaguardar el interés colectivo y procurar sus esfuerzos se utilizarán para lograr el bienestar común.

²⁰⁵McKinsey & Company, www.mckinsey.com/business-functions/marketing-and-sales/our-insights/prime-day-and-the-broad-reach-of-amazons-ecosystem.

²⁰⁶ACM Ethics, <https://ethics.acm.org/wp-content/uploads/2016/07/SE-code-spn.pdf>.

Ahora, abordaremos a los ingenieros de software, quienes debemos de identificar para efecto de determinar sus funciones, y las responsabilidades que tienen en la creación de sistemas inteligentes, dentro de las cuales se incluyen:

a) Analista, quien tiene las funciones siguientes: “Definición de los requerimientos del software. Análisis de las funciones requeridas por el sistema. Desarrollo de la arquitectura funcional. Análisis de la descomposición de los requerimientos del sistema”.

b) Diseñador se encarga de: Diseño de la arquitectura de los datos. Diseño de la base de datos. Diseño de interfaces. Diseño de detalle de la arquitectura del software.

c) Programador, Creación código ejecutable. Creación documentación de operación.

d) Ingeniero de software, cuya responsabilidad. Integración del entorno. Conducir revisiones del software.

e) Probador. Definición de la arquitectura de las pruebas. Diseño de casos de prueba. Ejecutar pruebas del software²⁰⁷.

El rol que desempeñan los desarrolladores de software (incluyendo cada una de las funciones antes referidas) dentro del tratamiento de los datos en Machine Learning en la cual se pueden presentar deficiencias por los sesgos humanos en los datos o en la manipulación de datos, y es en esta última deficiencia en la cual tenemos que poner mayor estudio e investigación porque las compañías o clientes determinan las especificaciones de algún sistema inteligente, y los ingenieros de software desempeñan su función con base a lo solicitado por las empresas, o bien también puede ser una deficiencia por un sesgo humano en los datos [Representación de los datos, selección de los datos, sesgo de homogeneidad fuera del grupo, injusticias históricas, influencia de estereotipos²⁰⁸] estos originan que los sistemas inteligentes puedan estar programados representaciones sociales o caracteres antiéticos que generen más discriminación, vulneración de derechos humanos e injusticias. Por ello, resulta necesario establecer códigos de conducta para los ingenieros de software, y también será necesario reducir el poder de las prácticas monopólicas que pueden influir en la ética de los ingenieros de software, para ello será propicio considerar las siguientes reglas éticas:

1.03 Aprobar software sólo si se tiene una creencia bien fundamentada de que es seguro, cumple las especificaciones, pasa las pruebas apropiadas y no reduce la calidad de vida, la privacidad o daña el medio ambiente. El efecto último del trabajo deberá ser el bien social.

1.04 Exponer a las personas o autoridades apropiadas cualquier daño real o potencial al usuario, a la sociedad o el medio ambiente, que razonablemente se cree que está asociado con el software o documentos relacionados (AMC Ethics. S.f.).

²⁰⁷Estrada, E. y Peña, A., 2013.

²⁰⁸Canal: Academia Nacional de Ciencias de Buenos Aires, 23/12/20.

De lo anterior se desprende que es fundamental que el ingeniero de software y el probador, revise, ejecute para efectos de aprobar el software y verificar que cumple con los estándares internacionales de calidad, que son seguros y por supuesto cuidar que con este software no se reduzca la calidad de vida, la privacidad o se dañe el medio ambiente, y que este software deberá tener como fin no la ganancia, sino el bien social, es de suma importancia que los ingenieros de software de acuerdo a sus funciones y responsabilidades se establezcan los códigos de conducta que mejoren sus conductas y eviten las prácticas desleales.

3. La trascendencia del manejo de datos en los sistemas inteligentes

Sin lugar a dudas un factor fundamental para los sistemas inteligentes es la información, de la cual se desprende un debate constante, el derecho a la información vs el de la desinformación, en la era cognitiva se desarrolla la capacidad del ser humano para conocer por medio de la percepción y los órganos del cerebro.

En esta era cognitiva es indispensable que se perciban los datos primero, y después se seleccionen la información y para ello, la ética de los sujetos es indispensable, para determinar si hay una desinformación o bien, hay un conocimiento bien informado, sí se logra que los sujetos obtengan información veraz.

Para ello, el papel fundamental para los desarrolladores y diseñadores, consiste en obtener un cierto número de datos para poder crear un sistema inteligente, otro problema se origina para el uso de datos o de *big data* y el *ML* en Derecho, también problema con el volumen de datos, el problema del riesgo de automatizar errores previos y los sesgos contenido en los datos. Es importante señalar las etapas para la construcción de datos, desde nuestra perspectiva la etapa de la prevención es fundamental debido ya que se tiene que pensar en quiénes pueden ser los diseñadores y desarrolladores de los sistemas inteligentes y también los capturistas de datos, ya que es ahí donde comienzan los problemas jurídicos que más adelante abordaremos.

Las etapas en la construcción de datos son:

- a) Definición del problema.
- b) Recogida de los datos (elemento análisis de datos, el coste, el acceso, propiedad de datos).
- c) Preparación de los datos.

Respecto a la primera etapa se determina cuáles son los problemas sobre la inteligencia artificial que se pretenden resolver, o bien, qué es lo que pretendemos programar.

“*Ethics by design* (como la *Privacy by Design* concepto fundamental de la regulación de la protección de datos) permitirá avanzar en temas como la responsabilidad demostrada en la fabricación y funcionamiento de los sistemas de inteligencia artificial. El diseño, arquitectura, funciones y funcionamiento debe incorporar valores sociales fundamentales (dependerá de cada país, de forma de vida, tradición, pero con límites generales aplicables en toda la humanidad).

La finalidad de la *Ethics by Design* es que el pensamiento autónomo de las máquinas, así como la toma de decisión, fruto del aprendizaje propio, no puede desvincularse de los valores humanos individuales y colectivos” Peña V., D. (2017).

Esta protección de datos tiene varias etapas en las cuáles las funciones y responsabilidades de los ingenieros de software, entre ellos los analistas, diseñadores, los programadores son fundamentales para el programa justo o injusto, porque son ellos quienes trabajan con las bases de datos, y realizan la arquitectura de los datos, por ello deben tener una responsabilidad, tanto jurídica como moral, es pertinente mencionar estos sujetos requieren obtener una adecuada recolección de datos la cual no deberá presentar deficiencias como los sesgos humanos en los datos o en la manipulación de éstos, primero en la selección de datos, ya que los datos no representan una muestra aleatoria de la realidad que se intenta modelar o predecir, también tenemos que los datos en ocasiones no son representativos de las conductas que se quieren programar, así mismo también tenemos que se pretende homogenizar fuera del grupo, sesgo por confirmación²⁰⁹, tendencias extremas a la generalización, por lo que se llegan a automatizar los sesgos, por lo que encontramos capas ocultas en los programas, lo que se conoce como cajas negras, lo que puede generar vulneraciones a la dignidad humana y con ello las vulneraciones a los derechos fundamentales, las desigualdades, e injusticias.

Existen diversas clasificaciones sobre sesgos que un modelo puede adquirir entre ellos:

- 1) Sesgo de interacción, el propio usuario o programador.
- 2) “Sesgo latente (*latent bias*). Cuando el modelo realiza correlacionados, inapropiados, generalmente al establecer, falsos nexos entre puntos de datos. Contextualidad.
- 3) Sesgos de selección (*selección bias*) Este tiene lugar cuando la base de datos no es suficientemente representativa”.

Es evidente que dependerá de los ingenieros de software y los probadores de que no existan cajas negras, en las que se automatizan los sesgos, por ello se requiere de la evaluación de su personalidad moral o de su integridad, es decir, se requiere visualizar el perfil psicológico de todos aquellos que intervienen en la creación de sistemas inteligentes, esto cuando encontramos sesgos de interacción, en los cuales los sujetos creadores de sistemas inteligentes automatizan una inclinación o perjuicio a favor o en contra de un ser humano o un grupo de personas de manera injusta, y van construyendo la capa oscura, la caja negra, que muchas veces puede ser delimitada por la empresa o el cliente, esto puede ser un obstáculo para la transparencia de los sistemas inteligentes, debido a que al existir cajas negras no puede ser explicables, por lo que se requiere se regule el derecho a la explicabilidad, en la cual toda persona involucrada pueda tener conocimiento sobre la lógica de las decisiones de sistemas inteligentes, aunque aquí surge la duda, de preguntarnos si realmente hay procedimientos y técnicas para determinar a través de auditorías o peritajes de sistemas inteligentes los sesgos automatizados.

²⁰⁹Canal: Academia Nacional de Ciencias de Buenos Aires, 23/12/20.

En la preparación de datos de sistemas inteligentes requerimos que contengan algoritmos éticos, en los cuales se incluyan: 1) la transparencia de los sistemas inteligentes; 2) la inclusión, y 3) la responsabilidad de los creadores de los sistemas inteligentes.

1) La transparencia es un principio de los algoritmos éticos, y consiste en que los sistemas inteligentes deben ser explicables, para lo cual requerimos de tener conocimiento en estas áreas del conocimiento, o bien acudir a los expertos. Consideramos que este es un principio que puede dar muchos problemas, debido a que la capa oculta, o caja negra, que contiene los sesgos no siempre podrá ser encontrada, lo cual puede dar lugar a la impunidad.

2) La inclusión, consiste en que “todos los seres humanos deben ser tenidos en cuenta para que cada uno pueda beneficiarse y que a todos los individuos le sean ofrecidas las mejores condiciones para expresarse y desarrollarse”.

3) Responsabilidad: Aquellos que diseñan y despliegan el uso de la IA deben proceder con responsabilidad²¹⁰.

Es necesario establecer sistemas de transformación socio-económicas, en las cuales se establezcan sistemas inteligentes y las personas tengan derechos a ser informados sobre la lógica detrás de las decisiones de inteligencia artificial, además de reducir la concurrencia de varias discriminaciones.

Se requiere una adecuada recolección, preparación, y preprocesamiento de los datos tiene un impacto significativo en la disminución de la incorporación de prejuicios preexistentes cuando se diseñan soluciones de IA. Para ello se deberá de poner mayor atención en la complejidad de los algoritmos, pues nuestra sociedad requiere de algoritmos éticos que evite la vulneración de la dignidad humana, para ello se requiere que los elementos de nuestros sistemas éticos, sean analizados para dar soluciones justas y viables, dada la monopolización del mercado digital.

4. Conclusiones

a) La ética como una teoría filosófica tiene una trascendencia para el bienestar e interés colectivo de las personas, nuestros sistemas éticos actualmente tiene varios problemas complejos, que requieren de una investigación transdisciplinaria, y que no permiten la regulación jurídica, sin olvidar que si bien es cierto que el carácter y las costumbres de los pueblos van cambiando de acuerdo al lugar y al tiempo, también debemos de considerar que estas representaciones sociales, tienen puntos vinculantes entre las sociedades diversas.

b) Se deben establecer los valores comunes de las sociedades y establecer sólo códigos de conductas, que permitan una sana convivencia entre los seres humanos, sin importar las fronteras y los lugares, ya que los sistemas inteligentes rebasan fronteras, y presentan riesgos y daños que deben de prevenirse y ser subsanados.

²¹⁰Canal: Academia Nacional de Ciencias de Buenos Aires, 23/12/20.

c) En la automatización de los datos, se presentan los diversos sesgos que los sistemas inteligentes los cuales tienen que contar con un número determinado de datos para poder ser programados, ya que como establecieron ingenieros no siempre se tiene la cantidad suficiente de datos para poder programar, aunado a que no siempre se cuenta con las representaciones sociales a través de los algoritmos, lo cual origina vulneraciones a la dignidad humana. Se requiere de algoritmos éticos, dada la complejidad de éstos.

d) En los sistemas inteligentes se requiere de un sistema de gestión de riesgos programados con estándares éticos y técnicos internacionales reconocidos por el regulador que permita interoperabilidad, controles preventivos con alertamiento e identificación de responsabilidad ante la materialización de riesgos. Sistema totalmente auditable y transparente para el usuario, el supervisor e interesados, con registros o algoritmos inalterables que permite entender la operación de la IA y retroalimentación para ajustar la matriz. Que incluyan los principios de los derechos humanos como **transparencia auditabilidad, trazabilidad inalterabilidad, inclusión y responsabilidad porque la ética está en las personas y no en las máquinas, por ello el centro de las cuestiones éticas deben ser los seres humanos.**

Referencias

- AMC Ethics (S.f.) *Ingeniería de software. Código de ética y práctica profesional 5.2*, <https://ethics.acm.org/wp-content/uploads/2016/07/SE-code-spn.pdf>.
- Canal Academia Nacional de Ciencias de Buenos Aires, Primera conferencia del Ciclo “Sistemas Inteligentes y Ética”, CETI, ANCBA, 5/8/20, archivo de vídeo Youtube: <https://youtu.be/yIv2ZPseEug>.
- Canal Academia Nacional de Ciencias de Buenos Aires, 21° Conferencia del Ciclo Sistemas Inteligentes y Ética, CETI, ANCBA, 23/12/20, archivo de vídeo Youtube: <https://youtu.be/29d3RwF7JIY>.
- Comité de la Cámara de Representantes de los Estados Unidos sobre el Poder Judicial, 2020. *Investigation of competition in digital markets* (trad. Palomino Angeles, Elisa), https://judiciary.house.gov/uploadedfiles/competition_in_digital_markets.pdf.
- Estrada Guzmán, Elsa- Peña Pérez Negrón, Adriana, *Influencia de los Roles de Equipo en las Actividades del Desarrollador de Software*, “Revista Electrónica de Computación, Informática Biomédica y Electrónica”, n° 1, may.-oct., p. 5 y 6, 2013, www.redalyc.org/pdf/5122/512251560002.pdf.
- Fong, Catherine - Huang, Jess - Robinson, Kelsey - Ungerman, Kelly, *Prime Day and the broad reach of Amazon’s ecosystem*, 2019, McKinsey & Company, www.mckinsey.com/business-functions/marketing-and-sales/our-insights/prime-day-and-the-broad-reach-of-amazons-ecosystem.

Peña V., D., *Inteligencia artificial y ética (Bydesign)*, 2017,
<https://dernegocios.uexternado.edu.co/comercio-electronico/inteligencia-artificial-y-etica-by-design>.



¿Qué límites éticos se plantea hoy un hacedor de programas inteligentes?

Por Horacio Granero

Cualquier cosa que insista en las importantes repercusiones de la tecnología en el mundo real, y las responsabilidades de sus creadores con respecto a ellas, seguramente será bienvenida en una época en que los sistemas automatizados están implicados en todas las facetas de la existencia humana, sin embargo, cuando se trata de las formas en que se discuten los códigos de ética de la IA, existe una tendencia preocupante, incluso cuando el mundo despierta a la importancia de este campo.

Ciclo "Sistemas Inteligentes y Ética"

Los tecnólogos creen que los desafíos éticos de la IA se pueden resolver escribiendo códigos, pero los desafíos son mucho más complejos. Esta es la creencia de que los códigos de IA son recetas para automatizar la ética misma; y que una vez que se haya alcanzado un amplio consenso en torno a dichos códigos, el problema de determinar una dirección futura éticamente positiva para el código informático habrá comenzado a resolverse.

Pero... no existe un conjunto de principios éticos que puedan justificarse racionalmente de una manera que cada ser racional acepte. Si bien existe una convergencia global emergente en torno a cinco principios éticos (transparencia, justicia y equidad, no maleficencia, responsabilidad y privacidad), queda sin considerar la divergencia sustancial en relación con la forma en que se interpretan estos principios, por qué se consideran importantes, a qué tema, dominio o actores pertenecen, y cómo deben implementarse.

Los códigos éticos son mucho menos parecidos a código de computadora que lo que sus creadores podrían desear. No son tanto conjuntos de instrucciones como aspiraciones, formuladas en términos que generan más preguntas de las que responden y este problema no va a desaparecer, en gran parte porque no existe un conjunto único de principios éticos que puedan justificarse racionalmente de una manera que cada ser racional acepte.

¿Existiría una “inteligencia artificial ética” estrictamente hablando? Quizás el mito más revelador de nuestro tiempo es el de la súper inteligencia de máquinas, cuya promesa convierte simultáneamente la ética de la inteligencia artificial en un enfrentamiento con amenazas existenciales y un proceso de diseño destinado a desterrar la sinrazón humana. ¿Puede un programa de inteligencia artificial determinar si tengo razón o no en un juicio? ¿La justicia predictiva es posible?

El principio de precaución es, quizás, la respuesta a estas preguntas.



¿Ética de la tecnología o ética de los productos tecnológicos? La IA enfrenta sus límites y los de los seres humanos

Por Enzo Maria Le Fevre Cervini

A partir de 2018 a nivel global muchos países empezaron a enfrentar los importantes desafíos nacidos a partir de una nueva etapa de desarrollo de la inteligencia artificial. El progreso tecnológico y la increíble disposición de datos impulsó gobiernos, académicos y la sociedad civil a promover una reflexión sobre los desafíos éticos ligados a la gran progresión de sistemas inteligentes.

Las Naciones Unidas destacaron recientemente cómo los datos y la tecnología son herramientas utilizadas por los Estados para automatizar, predecir, identificar, monitorear, detectar, atacar y castigar, pero subrayan cómo deberían utilizarse principalmente para apoyar la protección social y los sistemas de bienestar y garantizar el bienestar a las poblaciones del planeta.

De acuerdo con las directrices europeas, la IA confiable debe ser: 1) legal: respetando todas las leyes y regulaciones aplicables; 2) ética: respetando los principios y valores éticos, y 3) robusta: tanto desde una perspectiva técnica como teniendo en cuenta su entorno social.

Según la Unión Europea los sistemas de inteligencia artificial deberían empoderar a los seres humanos y ser resistentes y seguros. Capaces de asegurar respeto total por la privacidad y la protección de datos, y mecanismos adecuados de gobierno de datos. Los modelos de negocio de datos, los sistemas expertos y la IA deben ser transparentes y garantizar diversidad, no discriminación y equidad, evitando el



sesgo injusto. También es necesario que los sistemas de IA beneficien a todos los seres humanos.

Si bien estas reglas son fundamentales para una sana convivencia entre humanos y tecnología es pensable dar reglas éticas a una tecnología que todavía no está desarrollada ni en el 20% de su potencial? ¿No es más justo pensar a una ética de los productos de la tecnología? ¿Las reglas son más bien para la tecnología o para los seres humanos y sus naturales sesgos culturales y sociales?



CAPÍTULO III

LOS SISTEMAS INTELIGENTES: UNA OPORTUNIDAD PARA EL ÁMBITO ÉTICO

Por Sebastián Chavarría L.

1. Introducción

Hemos concluido el ciclo de conferencias “Sistemas Inteligentes y Ética”, a cargo del profesor Antonio A. Martino. Somos del parecer de que lo ha liderado con sumo rigor, liderazgo y con un completo sentido de la apertura al diálogo. Hemos atendido al ciclo con mucho interés y entusiasmo. Ha sido un espacio con el que nos sentimos profundamente agradecidos. A lo largo de todos estos meses, hemos tenido la oportunidad de escuchar a grandes maestros de diversos ámbitos, todos relacionados con el tema de los sistemas inteligentes y de la inteligencia artificial. Hubo momento para escuchar grandes ponencias; asimismo, para generar preguntas o comentarios. En fin: ha sido un ciclo que representa un hito en nuestra vida intelectual. Por ello, queremos efectuar este trabajo a modo de agradecimiento. En estas reflexiones queremos sostener lo siguiente: luego de todo lo que hemos discutido, creemos que los sistemas inteligentes son una oportunidad para hacer nuestra ética auténticamente intersubjetiva.

A lo largo de los encuentros, fueron apareciendo una gran diversidad de proposiciones que pretendían poner énfasis en los riesgos de la inteligencia artificial, en general; y de los sistemas inteligentes en particular. Uno de esos riesgos es que estos sistemas pueden servir para tomar decisiones sesgadas en ámbitos en los que no debe existir ningún tipo de discriminación²¹¹. Este tema surgió con frecuencia. Y es, en efecto, un peligro del que no podemos hacer caso omiso.

A pesar de que estamos reconociendo explícitamente de que los peligros son reales, ¿por qué creemos que esto no debe implicar un rechazo a los sistemas inteligentes? Al contrario, creemos que debemos apostar por la evolución de estas herramientas tecnológicas. ¿Por qué creemos que esto debe ser así, a pesar de sus peligros? La respuesta es simple: lo afirmamos porque creemos que la ética debe ser auténticamente intersubjetiva. Y, a su vez, creemos que los sistemas inteligentes pueden servir para atender este fin. Ahondaremos un poco más sobre esto.

2. Criterios implícitos y ética

²¹¹ Algunas aplicaciones de esto la encontramos en VARONA, DANIEL, *Los retos ingenieriles que afronta una IA digna de confianza basada en el marco internacional de principios éticos para la IA*, sesión del 16 de diciembre del 2020, www.youtube.com/watch?v=EIF2cvNCHTk&t=1912s.

Un estudio detallado sobre las causas, peligros y los retos en el contexto de los sesgos en sistemas de inteligencia artificial, COTINO, LORENZO, *Garantías frente al sesgo y discriminación algorítmicos*, sesión del 16 de septiembre del 2020, www.youtube.com/watch?v=AtxkloYb_GI.

Partamos de un hecho: vivimos en sociedades que tienen como objetivo la admisión de un amplio marco de posibles planes de vida. Y dentro de esa pluralidad, advertimos que, si bien existen un número grande de posibles conductas que podamos adoptar, a su vez, necesitamos ponernos de acuerdo con respecto a las normas que debemos atender, sin importar el sistema ético que adoptemos para nosotros mismos.

Creemos que hoy más que nunca debemos exigir que los miembros de nuestras sociedades explicitemos esas normas que consideramos que se deben adoptar. Ya lo decía Martino en su primera exposición cuando afirmaba que los sistemas inteligentes pueden dar criterios éticos²¹². En aquella disertación, el profesor Martino, siguiendo a Aristóteles, nos decía que “estando con los otros hay maneras correctas e incorrectas de comportarse”. Así entonces, desde nuestro punto de vista, ante esta situación aparecen dos niveles de análisis. El primero que parte de una posición metaética. Y el segundo, que entiende que sea cual sea nuestra posición en sede metaética, debemos adoptar una conducta.

Desde la posición metaética denominada no descriptivismo (de la cual se derivan el emotivismo y el prescriptivismo) se entiende que no hay realidad moral objetiva con la que podamos saber cuáles deben ser las conductas que debemos adoptar y con las que se nos pueda juzgar de forma objetiva²¹³. Ahora, si bien esta posición es plausible desde muy diversas consideraciones, no debe perderse de vista que esta posición se contrapone con un asentimiento que tenemos, a saber: que estamos en la necesidad de adoptar una conducta. En ese sentido, termina siendo – hasta cierto punto– irrelevante si hay una realidad moral a la cual consultar cómo debemos comportarnos²¹⁴. Al fin y al cabo, tendremos que actuar. Y que, a su vez, luego de haberla adoptado, necesitamos formular un juicio o una crítica respecto de las conductas que el resto de congéneres ha adoptado.

Todo lo dicho hasta ahora es importante para poner de relevancia la necesidad del pensamiento ético independientemente del punto de partida metaético que adoptemos. Queda ahora ligarlo con el tema de los sistemas inteligentes.

Una de las críticas que se hizo a los sistemas inteligentes viene ligada con el tema de que las máquinas no podrían obedecer normas que contienen vaguedad (como las famosas leyes de la robótica de Asimov)²¹⁵ y que dicho sea de pase se parecen a muchas de nuestras normas éticas (o mejor aún, de muchos sistemas éticos). Pongamos un ejemplo: “hay que hacer el bien y evitar el mal”.

²¹²MARTINO, ANTONIO A., *¿Porque la ética establece criterios para los sistemas inteligentes?* Primer encuentro, 5/8/20, www.youtube.com/watch?v=yIv2ZPseEug&t=4140s.

²¹³GUIBOURG, RICARDO, *La construcción del pensamiento*, p. 132 a 133.

²¹⁴ Para un estudio profundo sobre la lógica sin verdad que puede ser relevante para el pensamiento metaético. Cfr. ALCHOURRON, CARLOS E. - MARTINO, ANTONIO A., *Lógica sin verdad*, Theoria, segunda época año III Curso 1987, 88 n° 7 - 8, p. 7 a 43.

²¹⁵ La discusión interesante sobre este particular la obtuvimos de parte IRIARTE, ERICK, *IA bajo la perspectiva de las leyes de la robótica de Asimov. Un real problema regulatorio*, 2/9, www.youtube.com/watch?v=0CZQ2UNsJWc&t=4190s.

El descubrir que existen criterios tan amplios en el ámbito de las reglas para las máquinas, debería de servir de rebote para advertir lo siguiente: los seres humanos, también requerimos criterios más precisos para adoptar nuestras conductas. Pero no por eso –como hemos hecho ver antes con respecto al no descriptivismo– rechazamos la reflexión ética. De esa misma forma, no debemos pensar que las reglas amplias en el ámbito de los sistemas inteligentes deben ser rechazadas a priori. Ciertamente que no somos robots y que nuestras capacidades los superan. Pero, quizás venga bien poner de relevancia que en este punto compartimos una necesidad con estas creaciones humanas, necesidad consistente en aceptar que los principios amplios requerirían de explicitación de criterios con los que decodificamos esos principios²¹⁶.

Así entonces, la exigencia de criterios amplios aparece tanto en las disyuntivas de los sistemas inteligentes como en nuestra propia ética. En este contexto es que consideramos que requerimos una ética mucho más intersubjetiva que propenda a la explicitación de los criterios. ¿Cuáles criterios? Los criterios con los que justamente entendemos esas expresiones normativas a las que podemos darles diversos significados y que tenemos que hacer manifiestos a los demás.

Lo que nos queda es lo siguiente: considerar que la ética nos exige criterios de explicitación para la decisión de nuestras conductas. Podemos partir de enunciados normativos amplios²¹⁷. Lo crucial será dar el paso de esos criterios amplios a la precisión de cada uno de esos enunciados.

3. La explicitación de criterios con sistemas inteligentes

Luego de haber afirmado la necesidad de explicitación de criterios, queda nada más hacer un comentario con respecto al papel de los sistemas inteligentes. En la sesión con el Ing. Daniel Varona, se proponía la necesidad de que quienes diseñen sistemas captadores de datos, deben entender que “los usuarios no podemos ser meras granjas de datos”²¹⁸. Es cierto que lo que impera actualmente es el enfoque de caja negra con respecto a los sistemas de información: es decir, sabemos que los sistemas actuales pueden predecir muchas conductas de los usuarios y consumidores. Pero no podemos explicarlo porque no conocemos su funcionamiento interno.

En ese sentido queremos compartir el criterio del Ing. Varona: consideramos que los sistemas que se creen en el futuro, ya deben estar depurados previamente

²¹⁶ Ya el doctor Martino en su primera exposición, nos aclaraba que no hay ética para los robots. La ética, en rigor, es para sus creadores.

²¹⁷ En este aspecto de la explicitación de criterios, seguimos a GUIBOURG, RICARDO, La certeza de los criterios, https://rua.ua.es/dspace/bitstream/10045/10223/1/doxa24_23.pdf.

²¹⁸ Algunas aplicaciones de esto la encontramos en VARONA, DANIEL, *Los retos ingenieriles que afronta una IA digna de confianza basada en el marco internacional de principios éticos para la IA*. Sesión del 16 de diciembre del 2020. Grabación: www.youtube.com/watch?v=EIF2cvNCHTk&t=1912s.

de los sesgos que provoquen discriminación. Esto, entendemos que es un trabajo preventivo para que los algoritmos no contengan los sesgos de sus creadores.

Entendemos que esto puede implicar una tensión con respecto a la posibilidad de continuar manteniendo la privacidad de los creadores de los sistemas. Pero creemos que es un imperativo exigir transparencia de parte de quienes los creen.

Si aceptamos que no podemos ser meros proveedores de datos entonces tenemos derecho a saber quiénes filtran nuestros datos. Y siendo esto así entonces podamos con legitimidad exigir un trabajo preventivo que permita conocer con antelación quiénes y con qué criterios se filtran. En ese sentido no solo habrá un conocimiento a posterior (es decir, hasta cuando ya se efectuó la discriminación), como se ha venido haciendo. Que, en estos casos podría perseguirse un resarcimiento; con el riesgo de que no podamos dar con el causante de la discriminación (en virtud de cómo opera el sistema).

Si se puede llevar a cabo estas acciones, creemos que el futuro puede ser prometedor: Por un lado, los sistemas inteligentes pueden servirnos para clarificar nuestros criterios éticos (que ya hemos puesto de manifiesto como una necesidad en nuestras sociedades). Y, a su vez, esto también se logrará en la medida en que se ponga la tecnología actual para la creación sistemas que efectivamente cumplan con criterios éticos que hagan la vida entre usuarios, empresas y creadores de sistemas más transparente, coordinada y respetuosa de los derechos humanos²¹⁹.

Conclusiones

Luego de todo lo que hemos dicho, solo queda entonces reforzar lo siguiente: no creemos que la humanidad vaya a retroceder a tal punto que se pretenda la extinción de inteligentes. Pero aun en ese contexto –en nuestra opinión, más cercano a la ficción que a la realidad– la explicitación de criterios éticos serán un requerimiento perenne. Afirmamos esto, tanto en el ámbito de la ética general como para la ética aplicada a los sistemas inteligentes. Por esa razón, no debemos temer a los avances tecnológicos. Los retos de sacar nuestros sesgos y de tomar decisiones libres de ellas, han sido un reto que ha estado en toda la historia de nuestra humanidad. Ahora tenemos una nueva oportunidad para enfrentarnos al problema y tratar de superarlo.

Referencias

Alchourrón Carlos E. - Martino, Antonio A., *Lógica sin verdad*, Theoria, Segunda Época, año III, curso 1987 - 88, n° 7 - 8, p. 7 a 43.

Cotino, Lorenzo, *Garantías frente al sesgo y discriminación algorítmicos*, Presentada en el ciclo de conferencias “Sistemas Inteligentes y Ética” a cargo de Antonio

²¹⁹GARRIDO, ROMINA, *Ética en la gestión de datos*, 18/11/20, www.youtube.com/watch?v=ZzGY9NSQn8I.

Martino, conferencia del 16 de septiembre del 2020, www.youtube.com/watch?v=AtxkloYb_GI

Garrido, Romina, *Ética en la gestión de datos*, presentada en el ciclo de conferencias “Sistemas Inteligentes y Ética” a cargo de Antonio Martino, conferencia del 18 de noviembre del 2020, www.youtube.com/watch?v=ZzGY9NSQn8I

Guibourg, Ricardo, *La certeza de los criterios*, “Doxa”, n° 24, 2001, p. 603 a 612, https://rua.ua.es/dspace/bitstream/10045/10223/1/doxa24_23.pdf.

—*La construcción del pensamiento*, Bs. As., Colihue, 2004.

Iriarte, Erick, *IA bajo la perspectiva de las leyes de la robótica de Asimov. Un real problema regulatorio*, Presentada en el ciclo de conferencias “Sistemas Inteligentes y Ética” a cargo de Antonio Martino, sesión del 2 de septiembre, www.youtube.com/watch?v=0CZQ2UNsJWc&t=4190s.

Martino, Antonio A., *¿Porque la ética establece criterios para los sistemas inteligentes?* Presentada en el ciclo de conferencias “Sistemas Inteligentes y Ética” a cargo de Antonio Martino, primer encuentro, 5 de agosto del 2020, www.youtube.com/watch?v=yIv2ZPseEug&t=4140s.

Varona, Daniel, *Los retos ingenieriles que afronta una IA digna de confianza basada en el marco internacional de principios éticos para la IA*, Presentada en el ciclo de conferencias “Sistemas Inteligentes y Ética” a cargo de Antonio Martino, sesión del 16 de diciembre del 2020, www.youtube.com/watch?v=EIF2cvNCHTk&t=1912s.

Principios de inteligencia artificial y ética por diseño

Por Daniel Peña Valenzuela²²⁰

Introducción

El día 23 de septiembre de 2020, en el Ciclo de Ética en Sistemas Inteligentes de la Academia Nacional de Ciencias de Buenos Aires, se presenta la conferencia “Principios de Inteligencia Artificial y Ética por Diseño”. Este escrito comprende las principales temáticas analizadas teniendo como dos ejes fundamentales la reflexión ética, es decir, respecto del valores o axiología que definen un deber ser y la inteligencia artificial como sistemas de información que permiten la predicción de conductas mediante el aprendizaje profundo de máquinas o dispositivos.

La inteligencia artificial potenciada por el *Deep Learning* y la robótica generan simultáneamente temores y esperanzas. La historia del miedo a la “creación” (humana) de inteligencia tiene un comienzo remoto incluso la mitología griega muestra al atrevido Prometeo robar el fuego de los Dioses y tener un elemento (soplo!) vital. Muchos siglos después, Mary Shelley con su personaje ficticio Frankenstein aproximó el mito del soplo creador a un repulsivo monstruo –al menos así es en el imaginario popular y de las películas–.

Esta aproximación conflictual entre los seres humanos y la inteligencia creada de manera artificial sugiere la continuidad del miedo ancestral al desborde tecnológico. Es el primer pilar que justifica el análisis prospectivo ya que debemos tener la capacidad de construir la hoja de ruta del futuro posible o probable.

Otras aplicaciones de la inteligencia artificial representan una mayor cercanía con el sentido de lo humano y su desarrollo muestra armonía con actividades cotidianas. Tomamos arbitrariamente el ejemplo de Alexa (desarrollo inteligente y virtual en Ecco) así como otros asistentes virtuales y botnets representan una versión más

²²⁰ Abogado de la Universidad Externado de Colombia con maestría en derecho de la Universidad de Londres (LLM en materia Corporativa y Comercial) y en Grenoble, Francia (Propiedad Intelectual y Nuevas Tecnologías) así como estudios de profundización en Derecho Internacional en La Haya y en Arbitraje Internacional en la Cámara de Comercio Internacional. Profesor Titular de la Universidad Externado de Colombia en las Facultades de Derecho y de Finanzas y Relaciones Internacionales. Director de Investigación en Comercio Electrónico en la misma Universidad. Árbitro y Panelista de nombres de dominio de internet del Centro de Arbitraje y Mediación de la Organización Mundial de la Propiedad Intelectual OMPI. Árbitro activo de las Cámaras de Comercio de Bogotá y Medellín. Miembro de la Comisiones de Propiedad Intelectual y de Negocios Electrónicos y Telecomunicaciones de la Cámara de Comercio Internacional CCI de París. Socio Fundador de la firma Peña Mancero Abogados. Autor de libros sobre el tema de Comercio Electrónico, entre otros: *Aspectos Legales de Internet y Comercio Electrónico* (2001); *Contrato de diseño, desarrollo y hosting de un sitio web* (2003); *El contrato por medios electrónicos* (2003); *la Sociedad de la Información Digital* (2007); *Responsabilidad civil en la era digital* (2007); *Aspectos legales de la Computación en la nube* (2013); *Software libre y software propietario* (2013); *Responsabilidad de los proveedores de servicios de internet en material de propiedad intelectual* (2013); *De la Firma Manuscrita a las Firmas Electrónica y Digital* (2016).

agradable y seductora de la inteligencia artificial. El asistente de voz escucha al usuario y realiza distintas acciones según lo que se le pida. Puede ofrecer a sus usuarios sus propias opiniones sobre programas de televisión, series o su bebida y comida favorita. El asistente va más allá de su función de motor de búsqueda por voz y son sujetos interactivos con los que se puede entablar una conversación (cada vez más sofisticada).

A favor de los beneficios que trae Inteligencia Artificial se puede referir el experimento de Google, para identificar tejidos cancerosos y daños en la retina. Una combinación de flujo de datos de los pacientes provenientes de los teléfonos inteligentes y dispositivos portátiles permiten un diagnóstico médico automatizado a partir de una descripción de síntomas, detectar rasgos de comportamiento que sugieran que está deprimido e identificar si tiene un riesgo especial de enfermedad cardíaca. La agregación de datos también facilitará encontrar otras personas con enfermedades similares y ver cómo respondieron a varios tratamientos.

De todas las nuevas tecnologías parece que la inteligencia artificial es la más retadora desde el punto de vista del monopolio humano respecto de la capacidad de análisis, creación, aprendizaje, memoria, intuición y sensibilidad. Muchas son las preguntas recurrentes en ambientes académicos y empresariales ¿Se reemplazará al ser humano en algunas dimensiones como la laboral? ¿Desaparecerán algunas profesiones y oficios? ¿Las máquinas estarán fuera del control humano? ¿Hasta dónde llegarán en su capacidad de pensar? ¿Las máquinas sólo aprenderán o serán intuitivas y creativas? ¿La masificación de las aplicaciones de la inteligencia artificial favorecerán o perjudicarán a las economías emergentes que han centrado su modelo de desarrollo en las materias primas y los bajos salarios para los trabajadores?

La ética puede ser parte de la solución frente a la imprevisibilidad de la fuerza de la inteligencia artificial. Sin embargo, proponemos la prospectiva como metodología para establecer la relación entre derecho y máquinas pensantes. Cada cultura y sociedad tiene una aproximación diferente a la relación hombre-máquina, no es la primera vez en la historia que se experimenta esa fricción. Los automóviles reemplazaron a los carruajes decaballos, pero aún hoy en día la potencia se mide por decenas o cientos de caballos de fuerza (algunas utilizan los cientos de caballos de fuerza para abusar de la velocidad y poner en riesgo a otros inocentes ciudadanos con piques en las carreteras).

La masificación de los computadores personales hacía prever desaparición de miles de empleos rutinarios, pero más bien se convirtió en una herramienta necesaria para secretarías, digitadores y operarios. Se crearon profesiones nuevas y empleos por doquier.

Frente a la inteligencia artificial no podemos ser reactivos porque la humanidad podría estar en riesgo en sus características más eminentes. Tampoco podemos ser preactivos porque no sabemos realmente cómo prepararnos para un cambio esperado. Más bien debemos ser proactivos para definir cómo controlamos los cambios, definimos los límites éticos y determinamos si queremos complementarnos con las máquinas y potencializar nuestras cualidades o más bien ser reemplazados en nuestra esencia.

1. Problema jurídico planteado

La pregunta central formulada fue: **¿podemos/debemos construir agentes/sistemas inteligentes éticamente conscientes?**

Existen dos visiones respecto de este tema:

1) la *ética por regulación* (regida por los principios de prevención y precaución), en donde la legislación regula y pone límites a los sistemas de inteligencia artificial existentes y

2) la *ética por diseño*, la cual busca incluir dentro del diseño mismo del sistema de inteligencia artificial (en sus algoritmos) el componente ético, es decir, que este queda embebido dentro del mismo sistema de inteligencia artificial.

De esta manera, el principio de prevención parece no ser tan eficiente frente al aprendizaje autónomo de las máquinas y el cada vez más complejo control efectivo sobre la toma de decisiones, la aplicación del principio de precaución parece demasiado limitativa y prohibitiva de los desarrollos aplicados de la inteligencia artificial por ejemplo en los vehículos autónomos, en los asistentes personales y en los bots para asesoría financiera automatizada. El diseño preestablecido en los sistemas de inteligencia artificial puede ser una alternativa. Según el Principio 5 sobre inteligencia artificial de la OECD: *“Las organizaciones y las personas que desarrollen, desplieguen o gestionen sistemas de IA deberán responder de su correcto funcionamiento en consonancia con los principios precedente. Este principio ratifica la responsabilidad de los diseñadores, encargados e inversionistas en los sistemas de inteligencia artificial”*.

2. ¿Qué es la ética por diseño?

Es un concepto originado en la privacidad *by design*, de la protección de datos, en donde la toma de decisión de la Inteligencia Artificial está embebida en el diseño mismo de ésta en sus algoritmos de manera tal que las decisiones que este tome involucren un componente ético.

a. Importancia de la ética por diseño

Una de las razones por las cuales la ética por diseño se está convirtiendo en esa tercera vía, al lado de la regulatoria y la preventiva o de precaución, es porque involucra principios, reglas y normas del propio código (utilizando la expresión del profesor Lawrence Lessig) porque implica:

1) Accountability (Responsabilidad Demostrada)

a) La rendición de cuentas implica la justificación frente a terceros interesados sobre la toma de decisiones en una organización.

b) ¿Qué y cómo fue el origen o causa de la decisión que se tomó?

c) El algoritmo de inteligencia artificial puede incorporar dentro de sí las funcionalidades que permitan detectar la manera como se decidió adoptar un curso de acción y tener una trazabilidad de dicha decisión. De esta manera, la ética por diseño va de manera más amplia pues no sólo abarca el recaudo de datos personales sino la explicación ex post respecto de cualquier toma de decisiones por un sistema de inteligencia artificial y permite tener una trazabilidad de estos procesos de toma de decisión. El principio 3 de la OECD sobre Inteligencia Artificial establece: *los sistemas de IA deben estar presididos por la transparencia y una divulgación responsable a fin de garantizar que las personas sepan cuándo están interactuando con ellos y puedan oponerse a los resultados de esa interacción.*

2) Responsabilidad (liability)

a) Capacidad de los sistemas de no sólo tomar la decisión sino también de que, en caso de causar un daño, de responder por esa decisión. Lo anterior implica, incorporar una capacidad de examen de las decisiones del sistema de Inteligencia Artificial y una capacidad de respuesta del mismo.

b) Esto permite que la cadena de responsabilidad pueda ser definida y permite funcionalidades en la Inteligencia Artificial que permitan discernir sobre sus decisiones y de responder sobre las consecuencias que éstas generan.

c) La ética por diseño se trata precisamente de anticiparse desde el diseño mismo de la Inteligencia Artificial a los posibles daños que ésta pueda causar en su proceso de toma de decisiones y que la funcionalidad misma permita corregir los errores cometidos por estos sistemas.

3) Transparencia

a) Este elemento permite el activismo respecto de los sistemas de Inteligencia Artificial, en el sentido de que, al dejar una trazabilidad del proceso de toma de decisiones del sistema, esto permite que el sistema tenga un proceso de auto-aprendizaje y, a su vez, cumpla con su deber de información frente a terceros interesados.

b) El deber de información muchas veces choca con el derecho a la confidencialidad, pero si pensamos en la transparencia y en la capacidad de responder por sus errores, es preferible que esté incorporado en el diseño de la Inteligencia Artificial y no a través de los mecanismos de precaución y prevención.

b) Problemas o cuestionamientos de la ética “by design”

La ética por diseño representa una convergencia entre el concepto de diseño (*legal design*) y la arquitectura del sistema jurídico (sistema normativo tradicional),
Ciclo “Sistemas Inteligentes y Ética”

con el fin de ajustar la Inteligencia Artificial a los sistemas legales, pensar los procesos jurídicos de manera tridimensional.

Las aplicaciones de la ética por diseño deben complementarse con una interactividad humana. El algoritmo siempre debe tomar una decisión ética, pero siempre que se refiera a conceptos éticos le concedemos a la Inteligencia Artificial la capacidad de tomar decisiones complejas. En ese sentido, la interacción hombre-máquina debe ser más intensa en los casos límite en los cuales la decisión implica un juicio moral. La inteligencia artificial ha dejado de ser puramente rutinaria y mecánica y los sistemas de decisión automatizado van a tener una repercusión moral a medida que van avanzando.

De cualquier manera, el sistema de ética por diseño también tiene algunos problemas/cuestionamientos, tales como:

1) ¿Es o no viable? la viabilidad de este sistema ex ante es una alternativa más sostenible en ese trade off entre la regulación y la prohibición.

2) Definir cuál es la norma correcta a aplicar y el valor es complejo. Pero este es el gran problema del derecho: axiología frente a lo normativo.

3) Cómo convertirlo en una decisión práctica que tenga efecto en estándares internacionales y locales de diseño y puesta en el mercado de los sistemas de inteligencia artificial.

3. Conclusiones

a) En conclusión, la razón de incorporar el sistema de Ética por Diseño en la toma de decisiones de la Inteligencia Artificial depende de un raciocinio de anticipación respecto de los efectos potencialmente nocivos de las aplicaciones prácticas de este nuevo paradigma tecnológico.

b) La idea de la ética por diseño es anticipar sin prohibir, es decir, generar un sistema técnico-jurídico de responsabilidad ex ante y generar mayor información y mayores opciones de control sin necesidad de afectar el desarrollo de los sistemas de Inteligencia artificial.

c) El principio de prevención parece no ser tan eficiente frente al aprendizaje autónomo de las máquinas y el cada vez más complejo control efectivo sobre la toma de decisiones, la aplicación del principio de precaución parece demasiado limitativa y prohibitiva de los desarrollos aplicados de la inteligencia artificial.

d) El algoritmo de inteligencia artificial puede incorporar dentro de sí las funcionalidades que permitan detectar la manera como se decidió adoptar un curso de acción y tener una trazabilidad de dicha decisión.

e) El deber de información muchas veces choca con el derecho a la confidencialidad, pero si pensamos en la transparencia y en la capacidad de responder por sus errores, es preferible que esté incorporado en el diseño de la Inteligencia Artificial y no a través de los mecanismos de precaución y prevención.

Referencias

- Brynjolfsson, Erik - McAfee, Andrew, *The Second Machine Age, Work, Progress and Prosperity in a Time of Brilliant Technologies*, 2014, p. 163 a 187.
- Comisión Europea, *Libro Blanco sobre la inteligencia artificial: Un enfoque europeo orientado a la excelencia y la confianza*, 2020, https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_es.pdf.
- El Tiempo, *La inteligencia artificial vive su gran momento en Colombia*, 2019, www.eltiempo.com/tecnosfera/novedades-tecnologia/inteligencia-artificial-en-colombia-entrevista-con-andrew-ng-416822.
- Grupo Bancolombia (23 de 07 de 2018), *Capital Inteligente, Inteligencia Artificial: el aprendizaje de las máquinas al servicio de las personas*, www.grupobancolombia.com/wps/portal/empresas/capital.
- iProup (19 de 2 de 2019), *Por qué L'Oréal utiliza inteligencia artificial para realizar diagnósticos personalizados de la piel*, www.iproup.com/innovacion/2888-tecnologia-realidad-virtual-automatizacion-LOreal-utiliza-inteligencia-artificial-para-realizar-diagnosticos-de.
- Llinares, F. M., *Inteligencia artificial y justicia penal: más allá de los resultados lesivos causados por robots*, "Revista de Derecho Penal y Criminología", 2018, p. 87 a 130.
- Lessig, L., *Code and Other Laws of the Cyberspace*, New York, Basic Books, 2001, p. 34 a 89.
- Navas Navarro, Susana y otros, *Inteligencia Artificial*, "Tecnología Derecho", 2017, p. 23 a 71.
- OECD, *Artificial Intelligence Principles*, www.oecd.org/going-digital/ai/principles, Paris, 2019.
- Peña Valenzuela, D., *Blog Derecho de los Negocios, Inteligencia Artificial, explicabilidad y toma de decisiones*, 2019, <https://dernegocios.uexternado.edu.co/comercio-electronico/inteligencia-artificial-explicabilidad-y-la-toma-de-decisiones>.
- Schneider, U. (productor), *Límites éticos para la inteligencia artificial*, Alemania, Deutsche Welle documentales, 2019.
- Schwab, Klaus, *The Fourth Industrial Revolution*, 2018, p. 25 a 47.

Blockchain como elemento de gobernanza, confianza, ética y autorregulación de la IA

Por Rafael Y. Cuartas Báez²²¹

Entre las tecnologías insignia de la denominada 4ª Revolución Industrial, el protocolo blockchain y la inteligencia artificial tienen un lugar de gran relevancia como tecnologías que ya tienen una alta aplicabilidad en casi todos los sectores económicos e industriales, pero que además tienen un gran potencial a futuro cuando maduren completamente y se integren con otros desarrollos tecnológicos.

Hay que recordar, que el protocolo blockchain en su idea original fue gestada al interior de un movimiento “Tecno-anarquista” denominado “Cypherpunk”²²², que tiene en Timothy C. May²²³ uno de sus principales precursores y autor del “The Crypto Anarchist Manifesto”²²⁴ como documento emblema de esa subcultura, que evoluciona y actualiza la cultura literaria “Punk” de principios del siglo XX, de autores como George Orwell y su novela “1984” que iniciaron ese movimiento libertarios, exponiendo en sus libros los riesgos de regímenes totalitarios que aprovechan los avances tecnológicos para generar control social y mantener su hegemonía.

Este no es un detalle de menor importancia, pues la persona o grupo detrás del seudónimo “Satoshi Nakamoto”, presenta el criptoactivo “Bitcoin” en un *White Paper* publicado en 2008, con toda la estructura técnica, y la filosofía política económica y monetaria de lo que denomina un sistema de efectivo electrónico persona a persona. Sistema sustentado en la Escuela Austríaca de Economía que promueve el anti-intervencionismo estatal y el individualismo metodológico, mientras proscribe cualquier sistema de banco central; filosofía económica que se refuerza con las ideas del referido movimiento criptoanarquista del cual hacían parte lo principales actores del Bitcoin originario, buscando la libertad económica, la privacidad y eliminar el impacto inflacionario que produce el sistema tradicional de emisión estatal de moneda por parte de los bancos centrales.

Técnicamente en blockchain confluyen 4 tecnologías básicas, que venían evolucionando hacia varias décadas y algunas descubiertas hace milenios:

²²¹ Abogado Magister en Derecho Informático y Nuevas TIC. Docente Universitario e Investigador. Experto en gestión de riesgos de cumplimiento normativo. Desarrollador de herramientas para la Informática Jurídica y el cumplimiento normativo (cuartasbaez@gmail.com).

²²² <https://en.wikipedia.org/wiki/Cypherpunk>.

²²³ https://en.wikipedia.org/wiki/Timothy_C._May.

²²⁴ <https://en.wikipedia.org/wiki/Crypto-anarchism>.

- *Distributed Ledger Technology* (DLT)²²⁵, que es una tecnología de manejo de bases de datos originada en 1991, que busca descentralizar la información para evitar la concentración del poder sobre ésta y evitar manipulaciones.
- Criptografía Asimétrica²²⁶ que fue desarrollada en la década de 1970 y garantiza mayor seguridad e integridad que la criptografía milenaria surgida con Heródoto desde el siglo V a.C.; pues usa la metodología que conjuga clave pública y clave privada, que sirve actualmente de sustento a las firmas digitales.
- Algoritmo de Consenso Bizantino²²⁷ que promete resolver este problema milenario de comunicación para llegar a un acuerdo a distancia, que sea resistente a ser manipulado en la transmisión del mensaje.
- Protocolo P2P (*peer to peer*) creado con la plataforma *Napster* a finales de la década de 1990, la cual permitía el intercambio de información persona a persona, sin necesidad de servidores centralizados o intermediarios.

Esta integración tecnológica hizo que luego de 12 años de Blockchain haya probado su confiabilidad y estabilidad como tecnología, más allá del propósito inicial de ser usada como moneda digital que permitiera una libertad financiera sin intervención del Estado o intermediarios privados que hacen parte del monopolio financiero. Lo que hoy permite que esté siendo usada en los diferentes sectores de la economía y con un alto volumen de investigación y desarrollo a futuro.

Cuando entramos a revisar cómo funcionaría una eventual integración de BlockChain con Inteligencia Artificial, en principio pareciera no tener mayores puntos en común que les permitan relacionarse. Al revisar la gráfica Gartner se prevé que una eventual conjunción entre estas dos tecnologías no podría darse antes de 10 años, cuando lleguen a su maduración. De hecho, algunos podrían considerarlas antagónicas cuando se miran desde el enfoque de su programación y autonomía de operación.

Mientras que blockchain es totalmente programada y opera según las instrucciones previamente definidas que permiten afirmar que “el código es la ley”; en la inteligencia artificial, si bien existe una programación previa que define los rasgos iniciales bajo los cuales opera la IA, el fin último es que ésta aprenda autónomamente y, eventualmente, tome decisiones de manera independiente (Deep Learning), más allá de lo previamente programado. Lo anterior sin perjuicio de las diferentes aplicaciones y niveles de autonomía que conocemos existen para esta tecnología, pues el caso de los sistemas expertos es muy similar a la forma de operación de un *smartcontract* basados en tablas de referencias o condiciones con efectos previamente establecidos ante su verificación positiva o negativa. Igualmente difieren en

²²⁵ [https://es.wikipedia.org/wiki/Distributed_Ledger_Technology_\(DLT\)](https://es.wikipedia.org/wiki/Distributed_Ledger_Technology_(DLT)).

²²⁶ https://es.wikipedia.org/wiki/Criptograf%C3%ADa_asim%C3%A9trica.

²²⁷ https://es.wikipedia.org/wiki/Problema_del_consenso.

torno a la centralización y control, pues mientras que la IA tiende a ser centralizada y controlada, Blockchain por naturaleza es descentralizada.

Sin embargo, las dos tecnologías, además de contar con Alan Turing como precursor decisivo en el siglo XX para el desarrollo de estas dos tecnologías insignia de la denominada 4ª Revolución Industrial, tienen en común que surgen, entre otras razones, como medidas para superar la crisis de confianza entre humanos, bien sea por sus capacidades físicas, intelectuales o éticas. El protocolo blockchain, sin ignorar que su principal aplicación está orientada a sistemas financieros y monetarios, presenta elementos bien interesantes que eventualmente pueden lograr superar algunas de las discusiones y problemas que históricamente se han señalado como riesgos y obstáculos para la adopción segura de la IA.

Los sistemas basados en blockchain presentan en sus modelos de gobernanza distribuida y protocolos de consenso (p.ej., prueba de trabajo -PoW- o prueba de participación -PoS-), elementos que pueden ser de aplicación en los procesos de operación de la IA, como una manera de garantizar el control y la trazabilidad de las acciones y decisiones adoptadas por la IA, mientras que asegura la transparencia en la información procesada y mantiene los *log* de actividad libres de cualquier manipulación para eventuales peritajes, auditorías y determinación judicial de responsabilidades, permitiendo además conservar los niveles de seguridad y privacidad en los datos personales que permanecen encriptados.

La programación inicial que sea determinada para la IA, podrá ser registrada en una blockchain con Smartcontracts preestablecidos, que ya no pueden ser manipulados individualmente, donde cualquier cambio o acción de la IA es registrado y contrastado frente a los condicionantes programados en los smartcontracts, para determinar la siguiente acción. Siendo visible por todos los nodos de la red distribuida, que deben confirmar ciertas acciones por la mayoría de estos (dependiendo del modelo de gobernanza), donde se haya determinado como necesario el elevamiento de una decisión al consenso de la blockchain, dada la criticidad de la misma, antes de que pase a ser código ejecutable de la IA y sobre la cual sigue aprendiendo para evitar futuras interrupciones.

Esta propuesta implica, necesariamente, la implementación de un sistema de administración de riesgos propio para cada proyecto, bajo normas internacionales estandarizadas (ej. ISO), que defina en una matriz aquellos eventos de riesgo vislumbrados según el objeto del proyecto, asignándoles su respectiva calificación de impacto y probabilidad según el análisis objetivo y experto realizado, permitiendo determinar controles *ex ante* y *ex post* evaluados periódicamente. Todo lo cual, no solamente sirve como insumo técnico y objetivo previo para definir la programación de la IA, bajo principios éticos y jurídicos orientados a la responsabilidad frente a la especie humana en general y los usuarios o consumidores en particular; sino que será parte fundamental de la programación de la blockchain y sus smartcontracts que mantienen controlada la autonomía de la IA, bajo determinado umbral ético y de operación.

Se considera que esta propuesta es preventiva y reactiva al mismo tiempo y permitiría mitigar el riesgo de sesgos en la programación de la IA, al definir metodologías y criterios objetivos basados en análisis de riesgos con participación diferen-

tes actores que puedan verse afectados, determinando límites jurídicos y éticos que la IA no puede vulnerar mientras cuenta con la permanente auditoría y control de los nodos de la blockchain, que puede generar banderas rojas y alertas ante cualquier novedad prevista como riesgosa.

Este modelo permite el desarrollo de la innovación de forma controlada y sin necesidad de regulaciones asfixiantes, impuestas, antitécnicas o incoherentes; sino bajo principios de autorregulación sustentada técnicamente que facilita la adopción de las normas al tener sentido y haber sido parte de su desarrollo. Lo cual igualmente permitirá implementar regulaciones globales dinámicas y versátiles que mitigan la brecha y el rezago con el que el Derecho siempre llega a estos temas, pues se establece un relevamiento periódico de riesgos y controles que actualice la matriz e identifique mejoras o ajustes necesarios producto del análisis de los resultados obtenidos hasta el momento, todo lo cual se retroalimenta en la blockchain a manera de *fork* sin perder la historia de la cadena.

La propuesta en resumen consistiría en diseñar proyectos de IA autorregulada, mediante *smartcontracts* que conforman un sistema de gestión de riesgos programados con estándares éticos y técnicos internacionales, reconocidos por el regulador, que permita interoperabilidad, controles preventivos con alertamiento e identificación de responsabilidad ante la materialización de riesgos. Sistema totalmente auditable y transparente para el usuario, el supervisor e interesados, con registros o logs inalterables que permite entender la operación de la IA y retroalimentación para ajustar la matriz.

En conclusión, la interacción de la IA apoyada en blockchain como mecanismo de grabación de toda la actividad y que permite la parametrización de un sistema de gestión de riesgos por medio de *Smartcontracts*, mitiga esa opacidad en la denominada caja oscura donde se hace el procesamiento y toma de decisiones que genera temor fundado o no, pero que al registrarse en una blockchain, a manera de una memoria del sistema neuronal artificial, hará efectivo el principio de transparencia y auditabilidad que demandan los marcos regulatorios y éticos actuales.

Esta parametrización y grabación de actividad, permitirá además aprender y retroalimentar no sólo el desarrollo de la tecnología, sino además retroalimentar el mismo sistema de gestión de riesgos, haciendo los ajustes que correspondan a los controles para lograr ese equilibrio que permita una regulación con sentido, orientada a eliminar monopolios y que no desestime o impida el desarrollo tecnológico con los beneficios que se proyecten.

La integración de estas dos tecnologías tiene la potencialidad de contar con una real identidad digital libre de intereses meramente económicos, para que pase a ser la herramienta ideal que permita habitar este ecosistema digital con seguridad y permitiendo la identificación de responsabilidades asociadas al uso de la IA, sustentada en principios éticos y en los mismos modelos de gobernanza que operan de forma orgánica y natural en la estructuración de blockchain.

Consideraciones constitucionales acerca de la política y gestión de las nuevas tecnologías en el proceso judicial peruano

Por Jorge Isaac Torres Manrique*

1. Introducción

El importante desarrollo de las nuevas tecnologías y tecnologías de Información y comunicación (TICs), resulta muy plausible como innegable en el orbe, así como, en toda materia, ciencia y disciplina del saber humano. En ese sentido, deviene en necesaria y urgente su utilización, ora en sede pública, ora en sede privada.

Por ello, tomando como premisa la grave crisis de la administración de justicia, el uso de las nuevas tecnologías podría contribuir a contrarrestarla.

En la presente entrega, analizamos en qué medida la política y gestión en el proceso judicial, toman en consideración a las nuevas tecnologías, cual es el impacto generado en el marco de la insoslayable modernización de la administración de justicia y el análisis del efecto generado para la salvaguarda de los derechos fundamentales.

2. A propósito de la política judicial

El uso que se les dé a las TIC en el contexto del funcionamiento de las instituciones del sistema de justicia, depende de los objetivos por los cuales sean implementadas. Primero, pueden mejorar la gestión y el desempeño de las instituciones del sistema judicial en un sentido amplio, ya sea del despacho judicial a nivel estructural, como la organización de recursos humanos y materiales o respecto de la forma

* Consultor jurídico. Abogado por la UCSM (Arequipa). Doctorados en Derecho y Administración, por la UNFV (Lima). Consejero Editorial de la Revista American Journal of Science and Technology (USA). Editor Asociado de la Revista Forensic Research & Criminology International Journal (USA). Par Académico Evaluador de las firmas editoras: Corporación de Estudios y Publicaciones (Ecuador) y Ediciones Jurídicas de Santiago (Chile). Miembro del Comité Científico Internacional del Instituto Jurídico Internacional de Torino (Italia). Miembro de la International Association of Constitutional Law- IACL (Serbia). Miembro extranjero adjunto de la Asociación Argentina de Justicia Constitucional (Argentina). Miembro, par académico evaluador, corresponsal e investigador externo adscrito al Instituto Vasco de Derecho Procesal (País Vasco). Par académico evaluador de la Revista Misión Jurídica, de la Universidad Colegio Mayor de Cundinamarca (Colombia). Investigador Externo de la Universidad Global (Honduras). Miembro asociado de E-Justicia Latinoamérica (Argentina). Autor de libros de derecho constitucional publicados en: Colombia, Paraguay, Chile, Ecuador y Perú. Autor del Tratado de Derechos Fundamentales (Perú y en prensa en Colombia). Coautor del Tratado de Derecho Penal Constitucional Aplicado (Colombia); y en prensa en: Perú, Chile, Bolivia, Paraguay, Nicaragua y Argentina. Coautor de las Dilucidaciones de Derecho Procesal Civil (Chile, Colombia y Perú). CoDirector de los Códigos Penales Comentados de Ecuador, Colombia (en preparación) y Chile (en preparación). CoDirector de los Tratados Latinoamericanos de Litigación Oral y Lavado de Activos (en preparación). Autor de artículos en publicaciones científicas físicas y virtuales, en más de treinta universidades de veinticinco países. Ponente nacional e internacional (Kimblellmen@outlook.com).

en que se manejan los casos. Luego, pueden tener por objeto generar o mejorar el vínculo existente entre las instituciones del sistema judicial y la ciudadanía, mejorando el nivel de acceso a la justicia²²⁸.

Al respecto, resulta imprescindible considerar de primer orden, que la adopción de una política de dicha naturaleza, comporta necesariamente que el personal directamente involucrado pueda no lo encontrarse imbuido, capacitado, sensibilizado, comprometido; acerca de los alcances y propósitos de la misma. En suma, nos referimos al factor intraneus.

De igual manera, se deberá garantizar lo propio para los casos en que se trate de una política judicial que comporte la participación de instituciones distintas al Poder Judicial.

Y así también, la difusión y comunicación acerca de la referida política a adoptar, debe considerar la pertinencia, suficiencia y oportunidad, ora para los diversos actores de la administración de justicia, ora únicamente para los justiciables, según sea el caso. Entonces, debemos entender como: 1) pertinencia, al enfoque y orientación apropiados para el factor humano; 2) suficiencia, a la información necesaria y precisa, en lugar de abundante como vaga, y 3) oportunidad, a la puesta en conocimiento en el momento indicado, a efectos de asegurar el cumplimiento de sus fines.

3. Sobre la gestión judicial

La gestión judicial tiene como principal objetivo la modernización total del Poder Judicial. La gestión judicial se centra en los siguientes ejes: 1) gestión administrativa organizacional -Rediseño de procesos; 2) coeficiente de gestión judicial; 3) firma digital; 4) notificación electrónica, y 5) expediente digital²²⁹.

Al respecto, el ex presidente la Corte Suprema del Perú, Duberlí Rodríguez Tineo, señaló que la gestión en los despachos judiciales resulta importante para mejorar la calidad del servicio de impartición de justicia que brinda el Poder Judicial. A la vez de exhortar a los funcionarios judiciales a trabajar en sus despachos enfocados en una atención de calidad en favor de los usuarios del servicio de administración de justicia. Acotó que en noviembre de 2017 se implementará el expediente judicial electrónico (EJE) en la Primera y Segunda Sala Civil, con subespecialidad comercial de la Corte Superior de Justicia de Lima y en otros cinco juzgados capitales²³⁰.

En lo relacionado a la política y gestión, consideramos que la primera hace uso de la segunda a efectos de poder lograr su materialización. En ese orden de

²²⁸ Lillo L., Ricardo, *Indicadores de ceja: el rol de las tics en una justicia para ciudadanos*, <http://inecip.org/wp-content/uploads/Inecip-Sistemas-Judiciales-N%C2%BA-16.pdf>, 2012, p. 16.

²²⁹ S/a. *Qué es la gestión judicial*, www.cij.gov.ar/nota-1708-Que-es-la-gestionjudicial.html.

²³⁰ S/a. *La gestión judicial importa en la justicia*, www.elperuano.pe/noticia-la-gestion-judicial-importa-la-justicia-61463.aspx, 2017, p. 1.

ideas, entre la política y la gestión, existe una relación de género y especie, respectivamente.

Por otro lado, cabe precisar que la asunción de una política y gestión judicial no solo en materia de nuevas tecnologías, no resulta ser suficiente con el simple hecho de la puesta en marcha de la misma. Ello, en tanto que de manera programada amerita la realización de la evaluación y retroalimentación correspondiente. Con el afán de corroborar si es que se viene llevando a cabo de la manera adecuada y si se encuentra acorde al avance proyectado. Esto es, si es que se encuentran cumpliendo los objetivos y retos trazados.

4. Nuevas tecnologías en el proceso judicial

De conformidad a la quintaesencia de la presente entrega, a continuación, damos paso a una revisión del aterrizaje de las nuevas tecnologías en el proceso judicial peruano.

a. Servicio de consulta de expedientes judiciales en línea

Resulta harto sabido que hacer cola para consultar un expediente es un drama para muchos peruanos. Así, con el fin de mejorar el acceso a la información, el Poder Judicial implementó el Servicio de Consulta de Expedientes Judiciales en Línea, que ofrece la posibilidad de acceder y hacer seguimiento de los procesos judiciales a través de internet. Así, los ciudadanos pueden conocer en tiempo real si alguno de los sujetos procesales presentó algún recurso, revisar el contenido de las resoluciones judiciales y el hacer seguimiento a las cédulas de notificación. Todo de manera gratuita, lo que generó ahorro de tiempo y dinero para los ciudadanos y mejora de la imagen de la institución²³¹.

El presente acápite grafica de manera clara como contundente, que el Estado es una creación del pueblo. Entonces, la quintaesencia del funcionamiento de las instituciones estatales (en el presente caso, el Poder Judicial), se justifica única y exclusivamente cuando se ve orientadas hacia el servicio y beneficio de la población.

En ese orden de ideas, resulta doblemente plausible para los justiciables que no solo se lleve a cabo la facilidad de la consulta gratuita virtual del estado de los expedientes, sino que además para ello, se haga uso de la nueva tecnología, esto es, la Internet.

b. Video audiencias

Al respecto, tenemos que la Corte Superior de Lima Norte fue la primera en el Perú que utilizó internet en el desarrollo de video audiencias en procesos judiciales; posteriormente han seguido con éxito, la Corte Superior de Justicia de Huaura. El

²³¹ Gibu, Caroline, *Las nuevas tecnologías y su impacto en la nueva relación Estado-ciudadano: de la información y apertura a la auditoría social. Su incidencia en la estrategia de ejecución de las políticas públicas*. [cas.www2.congreso.gob.pe/sicr/cendocbib/con4_uibd.nsf/A94556812D509E6F05257F860078E4B0/\\$FILE/gibucar.pdf](http://cas.www2.congreso.gob.pe/sicr/cendocbib/con4_uibd.nsf/A94556812D509E6F05257F860078E4B0/$FILE/gibucar.pdf), Lima, 2015, p. 6.

Sexto Juzgado de Familia de Lima Norte, que dirigía el doctor Luis Alberto Salinas Bernal, el día 15/12/06, en el expediente 2311- 2005, efectuó la primera video audiencia, con la intervención virtual de una demandada, residente en los Estados Unidos de Norteamérica, en un caso de adopción de un menor. Dicho acto fue un hecho histórico sin precedentes en la administración de justicia y abrió el camino para el uso regular del video en las audiencias judiciales en asuntos de derecho de familia, para evitar el desplazamiento nacional e internacional de numerosos justiciables en casos vinculados a dicha materia²³².

Es de resaltar en los presentes casos, donde las judicaturas mencionadas otorgaron mayor como efectiva salvaguarda a la materialización de los derechos fundamentales, al interés superior del niño y adolescente, así como, a tener una familia y a no ser separado de ella.

c. Audiencia de conciliación por alimentos vía WhatsApp

Con el fin de atender el principio de interés superior del niño, niña o adolescente, el Juzgado de Paz Letrado de la provincia de Huarmey (Del Santa), llevó a cabo una audiencia de conciliación por alimentos utilizando el aplicativo WhatsApp con un imputado radicando en la ciudad española de Madrid. El juez Vera Marín, en el marco de la directiva 005-2017-P-CSJSA-PJ, emitido por la Presidencia de la Corte Superior de Justicia del Santa respecto a lineamientos para el uso de redes sociales en audiencias. El magistrado apeló a la voluntad de conciliar del investigado para coordinar con el personal a cargo a fin de realizar la audiencia. Al término de la sesión, el abogado del investigado y la defensora pública de la parte agraviada llegaron a un acuerdo.²³³

En el presente acápite tenemos algunos puntos que resaltar que: 1) en principio, que se haya actuado en favor de la prevalencia del derecho fundamental a los alimentos; 2) que se haya hecho uso asertivo del WhatsApp como una nueva tecnología al servicio de los justiciables; 3) el hecho que el imputado se encuentre físicamente en otro país no haya sido impedimento para llevar a cabo la referida audiencia, y 4) que exista una norma interna de la mencionada Corte Superior de Justicia, que establezca su utilización.

d. Lineamientos para el desarrollo de audiencias mediante el uso de aplicaciones, videoconferencias y otros aplicativos tecnológicos en los juicios con el nuevo Código Procesal Penal (NCPP)

El objeto es regular el uso de la videoconferencia y otros aplicativos tecnológicos análogos, como el WhatsApp, Skype, Google Hangouts, y algunas redes so-

²³² Torres López, Edgardo, *Informática y justicia en el Perú*, www.pj.gob.pe/wps/wcm/connect/f07f4a8045957aff977bd77db27bf086/10.+Jueces+-+Edgardo+Torres+L%C3%B3pez.pdf?MOD=AJPERES&CACHEID=f07f4a8045957aff977bd77db27bf086, Lima, 2009, p. 157 y 158.

²³³ S/a. *Juez utiliza WhatsApp para realizar audiencia de conciliación de alimentos desde España*, www.pj.gob.pe/wps/wcm/connect/cortesuprema/s_cortes_suprema_home/as_inicio/as_enlaces_de_stacados/as_imagen_prensa/as_notas_noticias/2018/cs_n-juez-utiliza-whatsapp-audiencia-conciliacion-alimentos-espana-26012018, Lima, 2018.

ciales, a fin de convocar e instalar válidamente las audiencias judiciales programadas con los lineamientos del NCPP, y así evitar frustraciones, suspensiones o reprogramaciones de audiencias judiciales; asimismo, lograr una comunicación en tiempo real. De acuerdo con la resolución administrativa 084-2018-CE-PJ, la finalidad es convocar e instalar las audiencias bajo las reglas del NCPP, evitando reprogramaciones de estas, sin soslayar ningún principio procesal como el de oralidad, publicidad e inmediación, diligencia que se realizará en tiempo real²³⁴.

En este punto es de apostrofar en principio, que el Consejo Ejecutivo del Poder Judicial haya asumido su rol de ente administrativo tutelar del Poder Judicial. Esto es, que haya procedido de conformidad a lo establecido en el inc. 26, del Título Único Ordenado de la Ley Orgánica del Poder Judicial, que determina como atribución y función del Consejo Ejecutivo del Poder Judicial, la adopción de acuerdos y demás medidas necesarias para que las dependencias de dicho poder, funcionen con celeridad y eficiencia. Y en segundo término, que haya dispuesto saludablemente la utilización del WhatsApp y otros análogos, a fin de convocar e instalar válidamente las audiencias judiciales programadas con los lineamientos del NCPP.

e. Expediente judicial electrónico

Con la finalidad de brindar a los litigantes mayor celeridad y transparencia en la solución de sus procesos judiciales, la Presidencia de la Corte Superior de Justicia de Lima dispuso la implementación del Expediente Judicial Electrónico (EJE) en el 5° y 6° Juzgado Civil con Subespecialidad Comercial de Lima. La entrada en vigencia en la implementación de esta moderna herramienta para los Despachos antes mencionados será a partir del lunes 16/10/17. Cabe recordar que con fecha 31/7/17, se dio inicio al uso de las nuevas Tecnologías de Información y Comunicaciones (TIC), con la implementación del Expediente Judicial Electrónico en el 4to Juzgado Comercial de Lima, el mismo que hasta la fecha continúa recibiendo demandas mediante su Mesa de Partes. También se dispuso que la implementación de este piloto se realice de manera progresiva en los Despachos de la Subespecialidad Comercial. Y próximamente, el 20 de octubre, se implementará en dos juzgados más, uno Contencioso Administrativo con Subespecialidad Tributaria y Aduanera y otro Contencioso Administrativo con Subespecialidad en Temas de Mercado²³⁵.

La implementación del Expediente Judicial Electrónico (EJE), tal vez se constituya en el mayor y más importante logro de la administración de justicia, no solamente en lo que a sintonizar con el uso de las nuevas tecnologías se refiere, sino, además, en efectiva salvaguarda y defensa de los derechos fundamentales de los justiciables principalmente. Por ello, no podemos más que saludar la referida implementación.

²³⁴ S/a. *Justicia avanza hacia lo virtual*, www.elperuano.pe/noticia-justicia-avanza-hacia-virtual-65828.aspx, Lima, 2018, p. 1.

²³⁵ S/a. *Corte de lima continua implementando el expediente judicial electrónico*, www.pj.gob.pe/wps/wcm/connect/cortesuperiorlimapjcs/s_csj_lima_nuevo/as_inicio/as_imagen_prensa/as_noticias/csji_n_corte_lima_continua_implementando_eje_131017, Lima, 2017.

f. Ampliación del mejoramiento de la plataforma tecnológica de los procesos judiciales no penales con la implementación del expediente judicial electrónico en todo el país

Este proyecto tiene como objetivo modernizar la plataforma tecnológica de las 34 cortes superiores y de la Corte Suprema en las especialidades de Familia, Laboral, Civil, Comercial, Constitucional y Contencioso Administrativo. La modernización contará además con un importante componente de gestión del cambio, capacitaciones a los usuarios y operadores judiciales sobre el uso de las Tecnologías de la Información y Comunicación (TIC's) y sobre los beneficios del EJE²³⁶.

El EJE es una herramienta que posibilitará el uso de nuevas tecnologías de información y comunicaciones (TIC) en los procesos judiciales, lo cual habilitará a los ciudadanos a presentar sus escritos por internet, mediante su abogado, desde su hogar u oficina (mesa de partes electrónica); además, le permitirá revisar el expediente en línea, recibiendo en su correo electrónico un aviso, cada vez que le sea remitida una notificación en su casilla electrónica. Se modernizará la justicia laboral, colocándola a la altura de los ordenamientos procesales de otros países de nuestra región. La nueva filosofía del proceso laboral exige que tanto las partes, sus abogados y los magistrados afrontemos de manera distinta el proceso laboral, con mayor compromiso y dedicación, dejando de lado el recurso al documento por la palabra, obviando formalismos innecesarios, con la finalidad de agilizar la resolución de los casos y hacer realidad el derecho a la tutela judicial efectiva, pues muchas veces la excesiva demora de los procesos judiciales ha tornado en irreparable el daño, restando total efectividad a nuestro sistema de justicia²³⁷.

Corresponde subrayar, que resulta positiva la adopción de modernizar la plataforma tecnológica de las cortes superiores del país y de la Corte Suprema en las especialidades referidas. Así también, poner en relieve la inclusión de la especialidad laboral. Ello, en tanto que la Nueva Ley Procesal de Trabajo, registra la oralidad a efectos de agilizar el proceso judicial.

g. Tribunal constitucional notificará la vista de la causa y demás actos procesales en la dirección electrónica

Al respecto, es de verse que el TC ha modificado su Reglamento Normativo a fin de establecer que ahora la notificación se debe realizar a través del correo electrónico de las partes y del casillero electrónico de la ventanilla virtual del TC²³⁸.

²³⁶ S/a. *Implementarán plataforma tecnológica de procesos judiciales no penales en 34 cortes superiores*, <https://laley.pe/art/5292/implementaran-plataforma-tecnologica-de-procesos-judiciales-no-penales-en-34-cortes-superiores>, Lima, 2018.

²³⁷ Puntriano Rosas, César, *Sentencias laborales. Propuestas para la mayor eficacia de los fallos*, <https://elperuano.pe/suplementosflipping/juridica/705/web/index.html>, Lima, 2018.

²³⁸ S/a. *TC: actos procesales solo se notificarán por vía electrónica*, <https://laley.pe/art/6504/tc-actos-procesales-solo-se-notificaran-por-via-electronica>, Lima, 2018.

En ese sentido, si bien es cierto que el art. 30 de la resolución administrativa 193-2018-P/TC, respecto de las notificaciones jurídicamente: *“El Tribunal Constitucional notificará la vista de la causa y demás actos procesales en la dirección electrónica que necesariamente proporcionen las partes y en la casilla electrónica de la ventanilla virtual del Tribunal Constitucional, constituyendo esta última notificación el cargo respectivo”*.

Lo resaltante no resulta ser la modificación normativa señalada, sino más bien, la efectivización de la norma positiva señalada. Ello, a efectos de arribar en el aterrizaje de lo establecido en la referida norma.

5. Analizando la política y gestión de las nuevas tecnologías en el proceso judicial

La decisión o política de disponer el uso de las nuevas tecnologías en la gestión del proceso judicial, obliga a considerar los lineamientos de la Administración o Modelo de Gestión por Objetivos y Resultados (saludablemente incorporada del sector privado, al público), a nuestro parecer, mal denominada: Administración por resultados. En tanto, solo que no solamente por razones metodológicas (o mera construcción conceptual), si no también, por el obligado conectivo de logicidad a observar, entre el antecedente y el consecuente; es decir, que los resultados (consecuente) única y excluyentemente serán válidos y eficaces, si guardan estricta correspondencia con los objetivos previamente fijados (antecedente). Así, no corresponde fijar, ni alcanzar resultados, sin que de manera antelada se hayan fijado los objetivos. Ello, porque de nada servirá obtener importantes como considerables resultados, cuando los mismos no guardan relación con los objetivos a alcanzar, peor aún, cuando de lo que se trata es de la ausencia de objetivos.

Así, tenemos que la Administración por Objetivos y Resultados: *“constituye una forma de conducción y organización de las administraciones públicas que se basa en la previsión y preparación anticipada de logros, y en la delimitación y asignación de responsabilidades para conseguirlos. Sin embargo, la condición para que los logros puedan reconocerse es que se prefiguren y diseñen previamente las situaciones a alcanzar o a consolidar, explícitamente establecidas...A diferencia de las formas tradicionales de gestión, preocupadas por el cumplimiento de las normas y procedimientos, la gestión por objetivos y resultados le asigna especial importancia a la relación impactos-resultados-productos-recursos. La centralidad de este proceso permite incorporar la actitud, la reflexión y la programación prospectivas al planeamiento de la organización”*²³⁹.

De lo señalado, es de verse que, en la aplicación de la política de disponer el uso de las nuevas tecnologías en la gestión del proceso judicial, tiene que observarse escrupulosamente lo correspondiente a los postulados de la Administración por

²³⁹ S/a. *La Gestión Pública por objetivos y resultados: Una visión sistémica*, Subsecretaría de la Gestión Pública, www.sgp.gov.ar/contenidos/onig/planeamiento_estrategico/docs/libro/Capitulo_3_La_Gestion_por_Objetivos_y_Resultados.pdf, Bs. As., 2007, p. 35.

Objetivos y Resultados. Esto es, que no resulta suficiente con la dación normativa y la correspondiente aplicación, sino, que deviene en imprescindible el establecimiento de los objetivos.

Empero, otro factor importante es el relacionado a los recursos humanos y económicos, además de la imprescindible planificación. Ello, en tanto que no haberse previsto de manera óptima los mismos, de poco o nada servirá la dación de la política mencionada.

Entonces, colegimos que para tales fines, no solamente comporta tomar en consideración el aspecto legal, sino además, el de la administración; en suma, la interdisciplinariedad.

6. Una mirada desde los derechos fundamentales

La incorporación de nuevas tecnologías de información y comunicación (TIC) en la administración pública trae consigo un abanico de posibilidades para la mejora en la dotación de servicios públicos de manera eficaz, eficiente y predecible. A su vez, las TIC permiten explorar nuevas formas de relación entre el Estado y los ciudadanos. Sin embargo, más allá de las TIC, los antiguos dilemas subsisten. Qué tipo de sociedad aspiramos, cómo modelamos el Estado para lograr dicha aspiración y qué rol tienen los ciudadanos en la construcción de dicha sociedad, son reflexiones anteriores al uso de las TIC²⁴⁰.

Por una parte, pueden contribuir a que una mayor cantidad de personas acceda a información judicial mejorando los estándares en materia de acceso a la justicia y la relación entre las instituciones del sistema de justicia y los ciudadanos, como así también al mejorar la gestión de los tribunales, pueden contribuir a una mejor percepción de los usuarios que verán satisfechas sus necesidades con mayor celeridad y eficiencia. Por último, son valiosas las herramientas que permiten entregar información de mejor calidad en audiencia. Sin embargo, no debe dejar de considerarse, que por una parte que la incorporación de nuevas tecnologías no debe convertirse en una barrera de entrada para quienes no tengan los recursos necesarios²⁴¹.

Mención aparte merece el abordamiento de la influencia o impacto de la política y gestión de las nuevas tecnologías en el proceso judicial peruano, en la realización de los derechos fundamentales de los justiciables.

En ese sentido, colegimos que la aunque tardía utilización de las nuevas tecnologías en el proceso judicial se constituye en un gran acierto, quedan aún grandes brechas que salvar. Así, tenemos que por un lado en el limitado uso de las mismas y

²⁴⁰ Gibu, *Las nuevas tecnologías y su impacto en la nueva relación Estado-ciudadano: de la información y apertura a la auditoría social. Su incidencia en la estrategia de ejecución de las políticas públicas*, p. 7.

²⁴¹ Lillo L., Ricardo, *El uso de nuevas tecnologías en el Sistema Judicial: experiencias y precauciones*, www.ijjusticia.org/docs/LOBOS.pdf, Santiago de Chile, 2010, p. 137.

por otro, en la efectiva utilización a escala nacional, sobre todo en provincias lejanas, deprimidas, carenciales.

De modo tal que, al cubrir suficientemente las brechas señaladas en el marco del Estado Constitucional de derecho imperante, se estaría salvaguardando la efectiva vigencia y realización de cuanto menos, los derechos fundamentales al: 1) debido proceso; 2) seguridad jurídica; 3) defensa; 4) a probar; 5) presunción de inocencia; 6) inmediatez; 7) motivación; 8) tutela jurisdiccional efectiva; 9) tutela procesal efectiva; 10) imparcialidad, entre otros.

7. Conclusiones

Indudablemente se ha presentado un importante avance en el proceso judicial peruano, como producto de la aplicación de la política y gestión de las nuevas tecnologías en el proceso judicial.

No debe perderse de vista, que un objetivo capital de ello es otorgar mayor celeridad y mayor transparencia al justiciable. Ello, en tanto que las meras celeridad y transparencia, no habrían sido suficientes.

El éxito o resultado positivo de su implantación no se agota con la mera dación de la política o norma legal correspondientes, sino que, tiene que ver además con la observancia de los postulados de la administración, basilarmente entre ellos: 1) la administración por objetivos y resultados; 2) recursos humanos y económicos, y 3) planificación.

Dejamos constancia que la implantación de una política y gestión de las nuevas tecnologías en el proceso judicial peruano, genera a su vez, el abrace o vulneración de los mencionados derechos fundamentales. Se logrará lo primero, cuando al misma se aplique de manera adecuada y de conformidad a lo sugerido. No se conseguirá ello (o incluso, todo lo contrario, esto es, su vulneración), en tanto se inobserve los puntos desarrollados, relacionados a la disciplina de la administración.

Con la escrupulosa observancia de lo mencionado, no solamente se asegurará el óptimo cumplimiento de lo señalado, sino también, se evitará descartar que eventualmente el remedio termine convirtiéndose en algo peor que la enfermedad.

8. Sugerencias

Evaluar los resultados parciales de la correcta como asertiva implantación de la aplicación de la política y gestión de las nuevas tecnologías en el proceso judicial, a efectos de corroborar si los resultados obtenidos resultan correspondientes en porcentaje y naturaleza.

Capacitar y concientizar a la totalidad de actores de la administración de justicia, en la importancia de la aplicación de la política y gestión de las nuevas tecnologías en el proceso judicial. Y lo propio, en lo relacionado a los derechos fundamentales involucrados.

Panorámica del curso “ética y sistemas inteligentes”

Por Amir García Villalpando

Los sesgos que un modelo puede adquirir. Sesgo de interacción. El propio usuario o programador. Sesgo latente (*latent bias*) Cuando el modelo realiza correlacionados, inapropiadas, generalmente al establecer, falsos nexos ente puntos de datos. Contextualidad. 3. Sesgos de selección (*selección bigs*) Este tiene lugar cuando la base de datos no es suficientemente representativa.

Garantizar los derechos de los titulares de datos

- 1) Información sobre la existencia de una decisión automatizada.
- 2) Derecho a la intervención humana.
- 3) Impugnación.
- 4) Explicación lógica aplicada al tratamiento.
- 5) Derecho de oposición, bajo ciertos supuestos.

Ya sea la blanca o la negra.

Es un debate estable:dilema o falsedad

Ética por regulación vs. ética por diseño. Asimetría del poder

De acuerdo a la ética de la justicia o de la sobrevivencia, tenemos que considerar la frontera y las políticas.

Carácter ontológico

Estamos hablando de distintos sistemas de ética, tendremos que adaptar la regulación jurídica. Este trabajo se pretende determinar el sistema de códigos éticos.

Construcciones éticas sociales.

Considerando los contextos sociales podemos crear códigos de ética.

El desarrollador, el fabricante, responsabilidad del producto. Responsabilidad objetivo de fábrica.

Quienes fabrican tienen la absoluta responsabilidad.

La ética está en la persona que desarrolla la máquina.

Información crear un sistema interconectado como prevención.

No tenemos una sociedad homogénea.

Éticas y las utopías, fronteras políticas.

By design impacto en las empresas, estándares para el desarrollador.

Sesgos algorítmicos o etapas.

Algoritmos - Datos incompletos, no fiables, manipulados por ciberataques o pueden ser sesgados.

Construcción de datos

Ethics by design

b) Definición del problema.

c) Recogida de los datos (elemento análisis de datos, el coste, el acceso, propiedad de datos).

d) Preparación de los datos.

Intel Ar y data van juntos

Datos- calidad- dificultad de prueba

Estandarizar cómo funcionar

Auditar el sistema dentro del paquete (Caja blanca y negra)

Cambiar las correlaciones

Se requiere la vigilancia de la calidad de los datos

Transparencia

Los sesgos que un modelo puede adquirir. Sesgo de interacción. El propio usuario o programador. "Sesgo latente (*latent bias*) Cuando el modelo realiza correlaciones, inapropiadas, generalmente al establecer, falsos nexos entre puntos de datos. Contextualidad. 3. Sesgos de selección (*selección bias*) Este tiene lugar cuando la base de datos no es suficientemente representativa.

Garantizar los derechos de los titulares de datos.

Estamos constantemente expuestos a decisiones/estimaciones realizadas por un sistema informático, aunque no seamos totalmente conscientes de ello; por tanto, es imprescindible:

- Retener la responsabilidad en el decisor humano, *es decir que el ser humano debe decidir no la máquina.*
- Que todas las partes involucradas tengan los mecanismos adecuados para auditar decisiones que les afectan, *esto es democracia es decir un enfoque de horizontalidad.*
- Y que puedan diseñarse medidas de resarcimiento cuando fuere necesario. Lo que en el derecho se conoce frecuentemente como *reparación del daño.*
- El enfoque tradicional sobre el manejo de la discriminación, el prejuicio, y desigualdad algorítmica exhibe un carácter reactivo que se debe cambiar. Es decir, *una especie de control de daños en procesos que deberán echarse a andar de acuerdo a la programación de procesos, su correspondiente evaluación y corrección.*

- La adecuada recolección, preparación, y preprocesamiento de los datos tiene un impacto significativo en la disminución de la incorporación de prejuicios preexistentes cuando se diseñan soluciones de IA. Es decir, *el trabajo analítico*.
- Existe la necesidad de coordinar esfuerzos por agrupar profesionales de distintas ramas técnicas y de las humanidades para tener un mayor entendimiento del problema y su solución. *Una real multidisciplina, en un plano democrático e incluyente, es decir que más allá de las razones técnicas de ingeniería, las demás ciencias y disciplinas, coadyuven en un plano de igualdad.*
- Existe la necesidad de determinar las competencias profesionales que el entorno actual demanda de los(as) diseñadores de AI y de construir programas de formación acorde a esas competencias.

“El cielo estrellado ante mí y la ley moral en mí”

Necesidades básicas adicionales	Respuestas culturales individuales deseadas	
8. información	Conocimiento	Crisis de la ignorancia no reconocida
9. convención en sentido estricto	Paz	Guerra con uso de armas de destrucción masiva como lo son las que usa la ciberguerra
10. convivencia en sentido amplio	Comprensión de la naturaleza	Desaparición de la especie humana tremendos y potenciales riesgos de aplicaciones de la IA con un deficiente manejo de la ética.
Necesidad básica	Concomitante cultural individual	Riesgos por problemas y culturales
1. Metabolismo	Abasto	Crisis alimentaria

2. Reproducción	Parentesco	Crisis en la familia
3. bienestar corporal	Abrigo	Crisis ambiental
4. Seguridad	Protección y defensa	Crisis por la violencia
5. Movimiento	Actividades	Crisis por la inmovilidad
6. Crecimiento	Ejercitación	Crisis por debilitamiento físico
7. Salud	Higiene	Crisis por enfermedad
8. Información	Conocimiento	Crisis por ignorancia no reconocida

DDHH, por ello si se impactan estas actividades, se promueven o se ponen en riesgo, los DDHH

La marcha de la tecnología y el progreso garantiza que incluso mientras debatimos esta problemática, nuestro uso del ciberespacio ha alcanzado ya el punto donde una gama cada vez más amplia de actores sociales, políticos, económicos y militares son dependientes de él y, por tanto, vulnerables a la interrupción de su uso y la usurpación de sus capacidades. Por lo tanto, hay una dependencia a ese tipo de eficiencia y eficacia...

Carga entregada: Malware BlackEnergy 3.

Arquitectura de la carga: Spear phishing para obtener acceso a las redes. Instalación de BlackEnergy 3 en cada una de las subestaciones afectadas. Robo de credenciales de las redes. El uso de redes privadas virtuales (VPN) para ingresar a la red ICS (Internet Connection Sharing). Uso de herramientas de acceso remoto existentes dentro del entorno o la emisión de comandos directamente desde una estación remota similar a una HMI del operador. Dispositivos de comunicaciones de serie a ethernet afectados a nivel de firmware. Uso de un KIIDisk modificado para borrar el registro maestro de arranque de los sistemas de la organización afectados, así como la eliminación selectiva de algunos registros La utilización de sistemas UPS para impactar la carga conectada con una interrupción programada del servidor. Denegación telefónica de ataque de servido en el centro de llamadas.

Desarrollo del ataque: Durante la etapa de intrusión cibernética de Entrega, Explotación a Instalación se entregaron los documentos maliciosos de Office por correo electrónico a personas en la red administrativa o do TI de las compañías eléctricas.

Cuando estos documentos se abrieron, se mostró una ventana emergente a los usuarios para alentarlos a habilitar las macros en el documento.

La habilitación de las macros permitió que el malware explotara la funcionalidad de macro de Office para instalar BlackEnergy3 en el sistema, no fue un zero day explicit.

Una vez instalado el malware BlackEnergy 3 se conecta a las direcciones IP de comando y control (C2) para permitir la comunicación del adversario con el malware y los sistemas infectados, permitiendo recopilar información del entorno, acceder a credenciales, escalar privilegios y moverse lateralmente unos seis meses antes del ataque. Es decir que se puede modular el atacar de manera alevosa y planeada con mucho tiempo.

Carga entregada: Malware Black Energy 3.

Arquitectura de la carga: Spear phishing para obtener acceso a las redes. Instalación de BlackEnergy 3 en cada una de las subestaciones afectadas Robo de credenciales de las redes. El uso de redes privadas virtuales (VPN) para ingresar a la red ICS (Internet Connection Sharing). Uso de herramientas de acceso remoto existentes dentro del entorno o la emisión de comandos directamente desde una estación remota similar a una HMI del operador. Dispositivos de comunicaciones de serie a ethernet afectados a nivel de firmware. Uso de un KIIDisk modificado para borrar el registro maestro de arranque de los sistemas de la organización afectados, así como la eliminación selectiva de algunos registros La utilización de sistemas UPS para impactar la carga conectada con una interrupción programada del servicio. Denegación telefónica de ataque de servicio en el centro de llamadas.

BlackEnergy

Situación: El 23 de diciembre de 2015, en medio del conflicto entre el gobierno de Ucrania y los separatistas de Crimea apoyados por Rusia, 3 compañías de energía de Ucrania informaron interrupciones del servicio a los clientes. Los cortes se debieron a la entrada ilegal de un tercero en la computadora de la compañía y su sistema SCADA a partir de las 3:35 p.m. aproximadamente hora local, siete subestaciones de 110 kV y 23 de 35 kV fueron desconectadas por tres horas. Declaraciones posteriores indicaron que el ataque cibernético impactó porciones adicionales de la red de distribución y los operadores se vieron forzados a cambiar al modo de operación manual.

Tipo de operación: Ofensiva, denegación de servicio.

Objetivo: Sistemas de control industrial SCADA de distribución de energía.

Efecto: Degradación, denegación de servicio.

Arquitectura de la carga: STUXNET atacó equipos con sistema operativo Windows de Microsoft, haciendo uso de cuatro Zero Day Exploit (ataques de día cero) de ese sistema operativo.

Su objetivo fueron los sistemas SCADA (Supervisory Control And Data Acquisition), que controlan y supervisan procesos industriales a distancia mediante la utilización de dispositivos PLC (Programmable Logic Controller) cuyo software de monitoreo y control era WinCC/PCS 7 de Siemens.

Ataques cibernéticos

Desarrollo del ataque: La entrega inicial de la carga de malware se efectuó mediante una o varias memorias USB, una vez instalado el malware explotó las vulnerabilidades de día cero del sistema operativo, ingresando por intermedio del sistema SCADA SIMATIC WinCC/PC 7 de Siemens, WinCC está escrito para el sistema operativo Microsoft Windows y utiliza Microsoft SQL Server para el login. Una vez ganado el acceso al sistema STUXNET empleo las contraseñas originales porque esto “podría tener impacto en el funcionamiento de la planta”.

STUXNET verificó la existencia en el objetivo, las centrifugas, de cierto número de motores variadores de velocidad, fabricados por dos empresas concretas, una iraní y otra finlandesa, estableciéndose rutinas distintas según la cantidad de variadores de uno y otro fabricante.

Una vez relevadas las vulnerabilidades y antes de la activación de la carga, STUXNET grabó más de 30 días de parámetros normales de operación, los cuales al momento del inicio del ataque propiamente dicho eran presentados a los operadores para que los mismos no notasen ninguna anomalía. El malware produjo la salida de servicio de más de 1000 centrifugas por variaciones fuera de para metros de sus motores de velocidad. Al momento de ser advertida la anomalía por los operadores el malware evitó de desactivación inmediata de las centrifugas mediante el bloqueo del sistema de comando y control de las mismas.

STUXNET - Operación Juegos Olímpicos

Situación. Con las negociaciones sobre el programa nuclear de Irán en un punto muerto en 2006, la administración de George W. Bush inició una operación de guerra cibernética, llamada “Juegos Olímpicos” cuyo objetivo era sabotear las crecientes actividades de enriquecimiento uranio de Irán a través de un ataque cibernético. En noviembre de 2010, Irán suspendió temporalmente las actividades de enriquecimiento en medio de crecientes problemas con las operaciones de centrifugación.

Tipo de operación: Ofensiva.

Objetivo: Nalanz Ennchmenl Complex, Natanz. República Islámica de Irán

Efecto: degradación. Daño continuo e incremental en lugar de un golpe catastrófico, el ataque causó que los rotores de la centrifuga girarán incontrolablemente rápido. Estos aumentos sustanciales en la velocidad condujeron a un aumento en la presión de la pared del rotor, causando que las centrifugas se rompan. **Lo que hace ver que puede haber ataques a la distancia, con potenciales catastróficas consecuencias.**

“Zero day vulnerability timeline”

Software is developed	Attacker detects vulnerability	Malware is released	Detection & Patching
-----------------------	--------------------------------	---------------------	----------------------

Software is developed, but unbeknownst to the developers, it contains a security vulnerability

A bad actor finds a vulnerability either before the developer or exploits it before a developer having an opportunity to release an update or patch

Attackers release malware to exploit software while the vulnerability is still open and unpatched

After hacker release the exploit, either the public detects identity or data theft or the developer uncovers, and creates a patch



Ciclo de ciberdefensa

Respuesta -> CERT

Operaciones cibernéticas

Predicción -> Vigilancia (reconocimiento y exploración)

Prevención -> Control de acceso

Políticas seg.

Auditoria

Detección -> Protección basada en firmas

Protección AP

Seguridad analítica

Centro de ciberdefensa / Defensa en profundidad

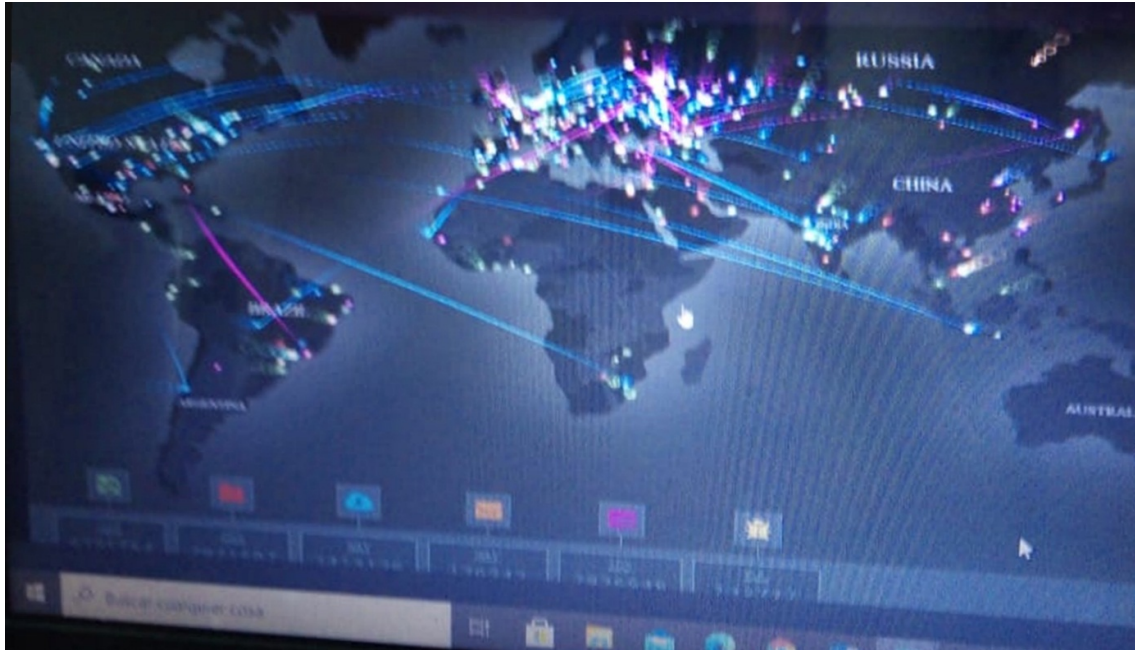
- Protección basada en firmas – Troyanos/virus/gusanos
- Protección contra ataques avanzados – Dirigidos – APT –
- Seguridad analítica – Correlación de eventos/Prevención y remediación de incidentes/Detección de vulnerabilidades
- Protección del perímetro – FIREWALSS/IDS/IPS
- Seguridad de las comunicaciones – FI-JAS/MOVILES/ALÁMBRICAS/INALÁMBRICAS/UHF/VHF/INF/SATCOM

Ataque cibernético

Medidas adoptadas a través del uso de las redes informáticas para interrumpir, negar, degradar o destruir información albergada en estaciones de trabajo y redes informáticas del adversario, o a las estaciones de trabajo y las redes mismas.

Proceso de ataque cibernético

Infiltración ->	Maniobra ->	Ataque
Reconocimiento	Acceso	Asalto
Exploración	Exfiltración	Sostenimiento



Top 15 cyber threats

1. Malware
2. Web-based attacks
3. Phishing
4. Web application
5. Spam
6. Identify theft
7. Data breach
8. Insider threat
9. Cryptohacking

El ciberespacio y las nuevas amenazas

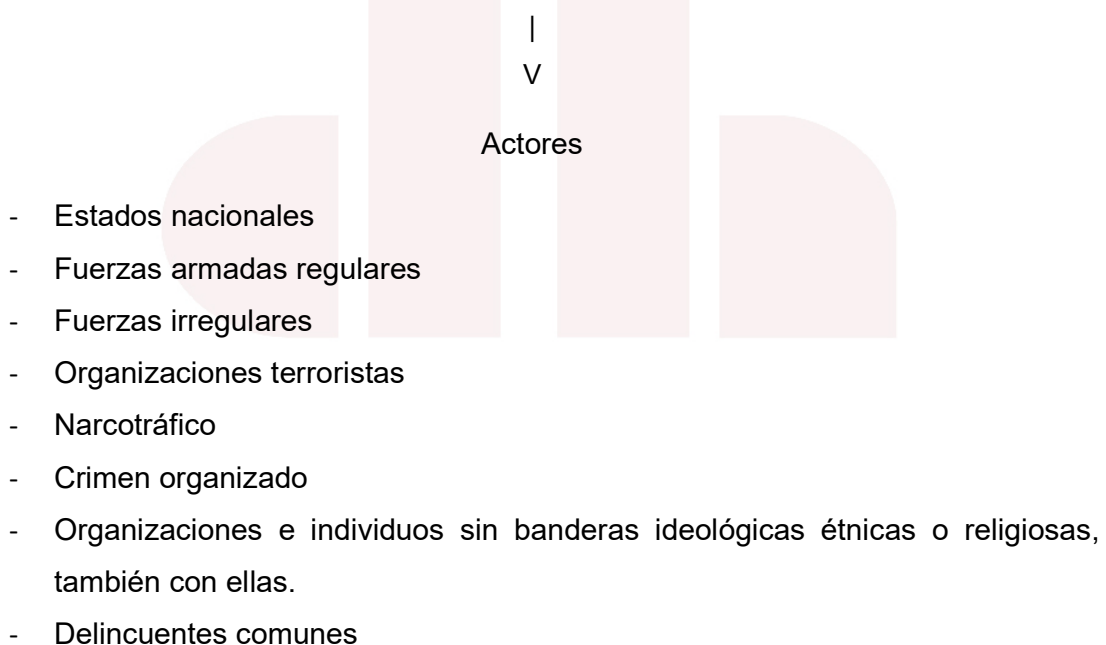
Las amenazas dentro del ambiente ciberespacial son subestimadas debido a que no se han registrado o reportado pérdidas de vidas humanas aún, sin embargo, solo es cuestión de tiempo para que ello ocurra. Esto se debe a que una buena parte de las funcionalidades de la sociedad depende del correcto funcionamiento de los sistemas de infraestructura crítica (Plantas Potabilizadoras de Agua, Distribución Eléctrica, Ferrocarriles, Señalización Urbana, Aeropuertos, Centrales Nucleares, etc.). Por lo que la fragilidad y la dependencia, cobran nuevas formas, como ya dijimos nos hemos hecho dependientes de una eficiencia, accesibilidad de la información y un nivel de confort, pender de ese hilo virtual, actualiza aquella noción de sociedad del riesgo.

Asimismo, cabe destacar la relación directamente proporcional entre el nivel de penetración de Internet en un país o región y al aumento de las vulnerabilidades en esa misma zona geográfica. Cuanto mayor es la cantidad de usuarios conectados a internet, mayor es la cantidad de blancos susceptibles de ser atacados o utilizados como plataformas de lanzamiento de ataques a terceros. Ahora el dato relevante de la conexión a Internet, es relevante como reflejo y como generador del tan anhelado bienestar, que nos proporciona trampolín, peor también potenciales trampas y ocasión para el delito.

El ciberespacio y las nuevas amenazas

Se presentan de manera múltiple, las principales son:

- Ciberguerra y ciberterrorismo
- Cibercrimen (secuestros, fraudes, robo de identidad, extorsión, violación de correspondencia, apropiación indebida de información, etcétera).



Ciberespacio – características

Escenario de conflicto	No convencional
Libertad de maniobra	Uso de recursos
Ubicuidad	Sin barreras de tiempo y espacio

Modos de acción

Silenciosos e invisibles

Dificultad para identificar al atacante (como producto de lo anterior)

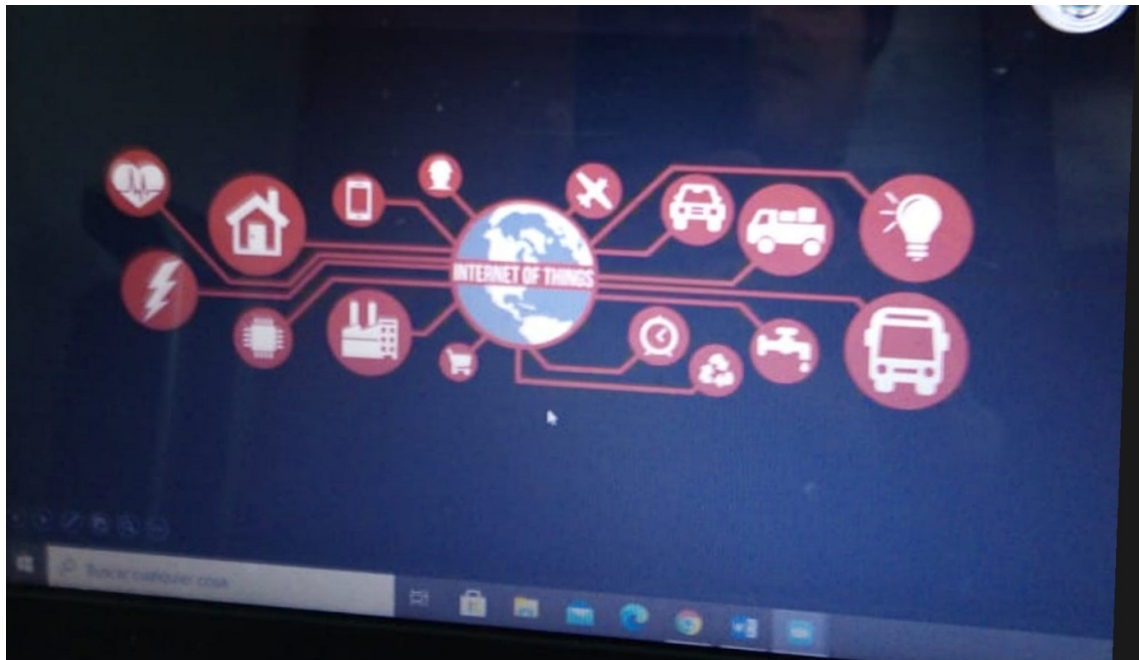
Cambio de paradigma

Adaptación permanente a los cambios tecnológicos para enfrentar nuevas metodologías de ataque con estrategias y medios no habituales; esto se traduciría en aprendizaje permanente.

El ciberespacio como ámbito de conflicto.

Ciberespacio (Terrestre, Marítimo, Aéreo, Espacial).

El único espacio que es verdaderamente transversal a todos los otros espacios, es el espacio de la información y la sociedad que de ella emana.



Componentes del ciberespacio – acceso

- Youtube
- Google
- Yahoo!
- Dropbox
- Facebook
- Twitter
- Skype
- Google +
- Pinterest

Componentes del ciberespacio – información. Que buscan:

Crear -> Identificar -> Capturar -> Adaptar -> Organizar -> Almacenar -> Explotar -> Compartir

Conceptos

El ciberespacio es definido como: “Ámbito operacional virtual en el que se desarrollan actividades de creación, procesamiento, almacenamiento, intercambio y explotación de información digital, a través de redes interdependientes e interconectadas –vinculadas a internet o no– y el software y firmware de dispositivos asociados a las mismas, siendo su carácter distintivo el empleo excluyente de las TICs (Tecnologías de la información y comunicaciones) y la interacción permanente con los otros ámbitos operacionales”. Se dice que son excluyentes por que excluyen a todo aquel que no tenga dinero para costearlo en sus diferentes dimensiones.

Antecedentes

Ámbitos operacionales tradicionales

- Terrestre
- Marítimo
- Aéreo
- Espacial

Se tiene que tomar muy en cuenta el rol del jurista o llamado abogado digital en la era del blockchain.

Usuario crítico del producto

- Debe comprender el funcionamiento del sistema
- ¿Qué datos emplearon? ¿cómo se procesaron?
- Es necesario que acepte los resultados (validez y utilidad)
- Debe documentar las lecciones aprendidas

El rol del jurista:

Revisión de la solución

- Apego a derecho
- Adecuación a buenas prácticas y estilos de trabajo
- Usabilidad
- Transparencia
- Capacidad de explicación
- Protección de datos
- Confianza

Proveedor de know-how

- Conoce las regulaciones sobre el desarrollo en IA & Derecho
- Define las necesidades (tareas legales a abordar)
- Establece requisitos funcionales (impedimentos de la doctrina, la ley y los procedimientos) y no funcionales (usabilidad, seguridad, etc.)
- Transmite experiencias previas y buenas prácticas

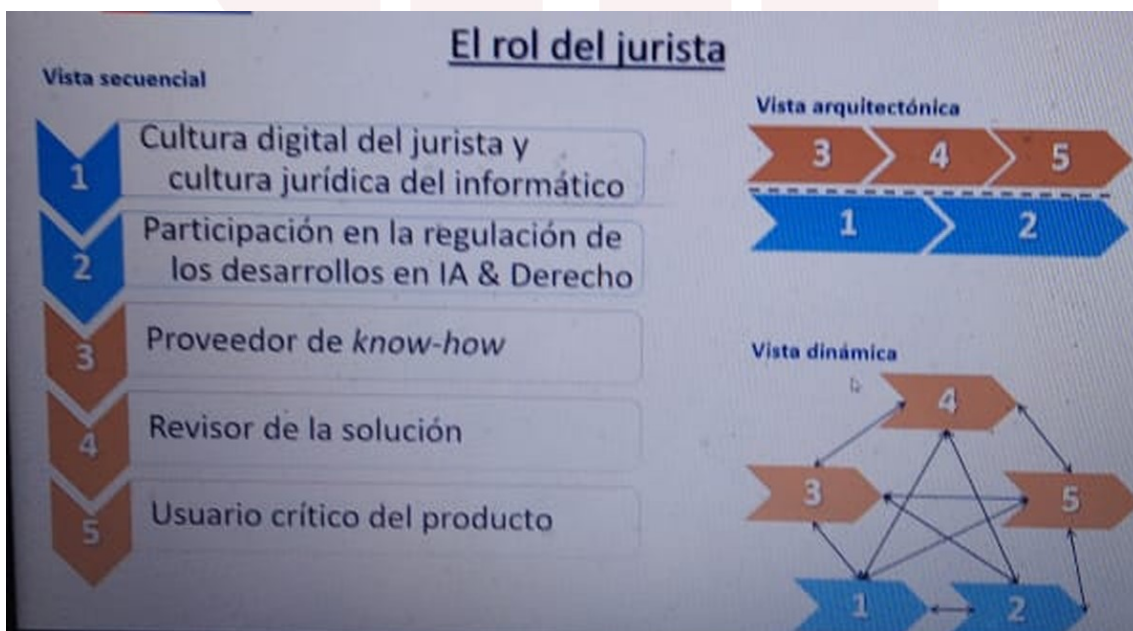
Participación en la regulación de los desarrollos en IA & Derecho

- Regulación del tratamiento de datos
- Regulación de las materias y tareas legales que serán susceptibles de impacto
- Regulación del proceso de desarrollo de los programas
 - o Participantes (equipo multidisciplinario, rol protagónico del jurista)
 - o Requisitos éticos y legales (declaración jurada)
 - o Puntos de revisión (rigor y cantidad de pruebas, umbrales de aceptación)

- Política de uso, transparencia y explicación
 - o ¿Qué se entenderá por transparencia y explicación?
 - o ¿Serán accesibles los códigos de los programas? ¿Para cuál audiencia?
 - o Retroalimentación en base a las lecciones aprendidas

Cultura digital del jurista y cultura jurídica del informático:

- Habilidades en el manejo de aplicaciones para PC y dispositivos móviles
- Conocimientos sobre seguridad de la información y protección de datos personales
- Conocimientos del uso de las redes para:
 - o Recuperación de información
 - o Generación de conocimiento especializado
 - o Comunicación (escrita y oral)
 - o Trabajo remoto (video conferencias, etc.)
- Nociones básicas de lógica, algoritmos, IA, ML, etc.
- Formación básica de los especialistas en computación



¿Queremos demostrar que los avances en IA son aplicables (“exitosamente”) al Derecho?

O

¿Queremos construir programas que utilicen adecuadamente las técnicas de IA para ayudar a los juristas a lidiar con sus necesidades reales?

Segunda época (desde 2010 aprox.) -> ¡resurgimiento!

- Basada fundamentalmente en Big data, ML y NLP.
- Programas como Watson (IBM), Ross, Project Debater, AlphaZero
- Sobresalen dos enfoques: robots vs sistemas integrados de ayuda a la decisión
- ¡No parece tan relevante la participación del jurista! ¿El jurista ha quedado rebasado y con ello el ser humano?
- Problemas para el uso del Big data y ML en el Derecho.
 - o La naturaleza retrospectiva del ML
 - o Justificación estadística del éxito de las soluciones de ML(Machine Learningo aprendizaje automatizado).
 - o Dificultad para explicar los resultados
 - o Habitual necesidad de grandes volúmenes de datos
 - o La no homogeneidad de los datos (cambios en la ley, en la sociedad). Los procesos de homologación vs la diversidad cambiante.
 - o El riesgo de automatizar errores previos y el sesgo contenido en los datasets. De los primeros que siempre existe la posibilidad de hacerlo, los segundos pueden ser de manera inconsciente o dolosa.

Resumen de propuesta: una IA autorregulada mediante smartcontracts que conforman un sistema de gestión de riesgos programados con estándares éticos y técnicos internacionales reconocidos por el regulador que permita interoperabilidad, controles preventivos con alertamiento e identificación de responsabilidad ante la materialización de riesgos. Sistema totalmente auditable y transparente para el usuario, el supervisor e interesados, con registros o logs inalterables que permite entender la operación de la IA y retroalimentación para ajustar la matriz. Si se logra materializar esta propuesta, estaremos ante un sistema civilizatorio controlado sin precedentes; sin embargo, no se podrá evitar quien quiera atacar estos sistemas.

- La autorregulación en blockchain es innata - aplicación sin humanos
- Ética y regulación biocéntrica y antropogénica: no solo Antropocéntrica. Los dos primeros suponen una conciencia ecosistémica, toda vez que el tercero ha tenido serias consecuencias, en lo antropológico y ambiental.

- Peligro de la concentración de poder de información y trazabilidad. Esto sucede cuando hay pocos actores en el mercado y laxa regulación, también cuando se puede maquillar la trazabilidad.
- Necesidad de identidad digital para protocolo de consenso Pol.
- No todo se le debe entregar a la IA – exclusividad humana. Este rubro es primordial.
- Blockchain y modelos de gobernanza son solución no el problema
- Necesidad de interfaces sencillas y de fácil acceso – Supervisión. Es importante un diseño amigable y la supervisión como parte del proceso administrativo, lo que generara empleos especializados.
- Necesidad de repensar el perfil de los abogados, que deberán tener otros conocimientos adicionales como programadores y de tipo ingenieril, esto será una revolución en la abogacía.
- Controles preventivos (ex ante) - Alertas
- Controles reactivos o reparativos (ex post). Preventivo y correctivo.

Estándares técnicos tipo ISO – interoperabilidad global entre sistemas. son de tal modo importantes, que sería deseable que los órganos supervisores, tuviesen personal privado y público.

“ISO/CD 22739 Blockchain and distributed ledger technologies terminology”

- Nivel de riesgo determina umbral para elevar decisión BC
- Regulación por diseño y con significado – Mayor confianza, al llenar de sentido dicho diseño, la confianza se ampliará por tener sentido para los participantes.
- Regulación sectorizada y gremial. Es decir, que se hará a nivel colegiado, por parte de los informáticos, ingenieros, administradores y abogados.
- Sistema ajustado a la realidad del productor y finalidad de la inteligencia artificial. El primero basado en el principio de que nadie está obligado a lo imposible y definir cuales es la finalidad o finalidades de la IA, que siempre debían ir entorno al bien de la humanidad.
- Accountability y DD estándar de responsabilidad. Además de auditabilidad.
- Mayor rigurosidad y límites a desarrollos de alto riesgo según estándar Y Supervisión. Lo anterior muy encarecidamente en el uso de armas automatizadas como el uso de drones y muchos otros.

Por su origen técnico, sus especificaciones serán:

- Criptografía – ciencia arcana matemática, para el manejo de códigos y su respectiva codificación y decodificación.
- Wei Dai – B-Money 1998 – antecesor – central

- Hal Finney – algoritmo PoW 2004 – 1° bitcoiner, para dinero electrónico.
- Nick Szabo – Bitgold 2005 – smartcontracts 1995. Podemos ver que el desarrollo de la IA, no es como tal novedoso, pero su despliegue es cada vez mayor.
- 2008 – wp Bitcoin – Satoshi Nakamoto:
 - o DLT-1991
 - o Criptografía asimétrica desarrollada en los 70
 - o Algoritmo de consenso PoW – bizantino
 - o Protocolo P2P – Napster finales de los 90'. Aplicados para la descarga de música gratuita y monedas electrónicas, estas desarrolladas después de la crisis mundial de 2008.

1. Transacción ->

- a. Operación \$
- b. Información
- c. Decisión

2. Registro de bloque ->

- a. Se empaqueta con otras Tx y se envía a la red

3. Verificación ->

Según mecanismo de consenso

4. Todos los nodos actualizan los registros

5. Historial inmutable. En mi opinión, este es una de las potencialidades más grandes para la rendición de cuentas y la prueba plena. lo anterior usado en cuestiones concretas como: Reglas historias clínicas y Diagnóstico Médico

La información que usa/obtiene la IA para su aprendizaje ¿Es admisible de acuerdo a las reglas de la prueba? ¿Su estructura es la misma que la de los litigios que provienen de sistemas de precedentes?

Es admisible de acuerdo con la ética de la percepción sesgo (bias) en el aprendizaje cuando la IA ha generado su propio aprendizaje “no supervisado” es probable que no sea posible recuperar todo el proceso para una auditoría (o es demasiado compleja u orgánica) ¿será suficiente una validación sobre su desempeño?

1769 “For the law holds, that it is better that ten guilty persons escape than that one innocent suffer”.

“Para la ley, es preferible que diez culpables escapen a que un inocente sufra”.

¿Debe la IA ser el resultado de **minimizar los riesgos** que maximizar los resultados? En materia penal los riesgos no son comparables con los resultados, ... de ahí que de acuerdo a la criminología crítica considera que la idea de criminalizar a una persona por su “peligrosidad”, equivaldría a penalizarla por lo que es y no por lo que hace (¿hasta dónde los delitos de “género”, son la prueba más palpable de

ello?). ¿impulsa el aprendizaje automático este principio?, ... ¿quién dice como ponderar los riesgos en relación con los beneficios? lo pueden calcular actuarios, economistas, riesgologos y expertos en prospectiva.

Ética y legalidad

¿La IA debe dar –además de un resultado– una fundamentación? Esto es clave Incluso para la autocomposición entre las partes. Finalmente, lo jurídico busca entre otras cosas la justificación de sus resoluciones.

Dada la costumbre de los votos disidentes en las decisiones colegiadas ¿deberla ofrecer la IA también las alternativas más relevantes, para que el juez elija? El humano debe juzgar asuntos humanos y determinar, pero la IA, puede servir de auxiliar al juez.

¿Debe exigirse que la IA tenga evaluación previa de riesgos?... o sea, una **validación**. ¿Toda IA debe tener monitoreo o posteriori? ... o sea seguimiento permanente y a largo plazo de las decisiones que se toman. De ser posible, es deseable que así sea.

¿Debe estar clara la responsabilidad por una decisión errada? ósea que no dé lugar a una responsabilidad diluida. Seguramente así debe de ser toda vez que hay manera de determinarlo.

PROMETEA. “Se descartaron explícitamente tecnologías como las redes neuronales por la falta de justificación del resultado provisto, y por funcionar como cajas negras, sin posibilidad de explicar el resultado de una decisión”

“La característica distintiva de PROMETEA es que realiza la recomendación de la sentencia en base a un conjunto de reglas definido previamente por los abogados y almacenadas en el sistema. Es por esto que PROMETEA no funciona como un algoritmo de caja negra (esto es, no se desconoce cómo es su lógica o cómo funciona), sino que es completamente trazable y permite explicar de una manera clara y sencilla cómo se arriba a una determinada propuesta de solución”. Por tanto, para el juzgamiento, es mucho más idónea que la caja negra.

GOFAI was based on a human-understandable symbolic system. It is an AI without machine learning.

¿Qué es la inteligencia artificial?

De acuerdo con algunas definiciones destacadas, la inteligencia artificial es:

1. Bellman, 1978 “[La automatización de] actividades que asociamos con el pensamiento humano, actividades como la toma de decisiones, la resolución de problemas...”

2. Kurzweil, 1990 “el arte de crear máquinas que realizan funciones que requieren inteligencia cuando las realizan personas”
3. Winston, 1992 “El estudio de los cálculos que hacen que sea posible percibir, razonar y actuar”
4. ... 1998 “Inteligencia artificial... se ocupa del comportamiento inteligente en los artefactos”.

COMPAS ... estimación de la probabilidad que un acusado reincida en un delito

Wisconsin v. Loomis, no viola el derecho del acusado al debido proceso,
...el reo responde a un cuestionario
...SCJ protege al secreto industrial

State v. Samsa, 859 N.W 2d 140 “COMPAS es simplemente una herramienta”

People v. Chubbs Idem, sobre algoritmos secretos - TrueAllele

Nuevas leyes

2019 - Idaho 19-1910, ch 258, sec l.p, 768. estableciendo que “las herramientas de evaluación de riesgos previa al JUICIO estarán abiertas a la inspección, auditoría y pruebas públicas”. Las herramientas deberán estar sujetas a ello, para prevenir o corregir fallos.

2020 - Michigan, se presentó en la Legislatura del Estado de Michigan un proyecto (Bill No 59).

Los Sistemas Judiciales son “justicia artificial”. Las constituciones definen un juez humano y jurados humanos, la administración de Justicia está en crisis, la reforma judicial requiere revisar algunos principios, la “percepción judicial” debe seguir reglas éticas y legales, rigurosas (e.g. ABA) existen soluciones (como la Ley Badinter) y hay espacio para los “sistemas inteligentes” con “supervisión humana”

La Formal Opinión 478 del 8 de diciembre de 2017 “Independent Factual Research by Judges Via the Internet” establece que: Si el juez necesita información adicional para decidir un caso:

- si necesita información que influencia la decisión, ésta debe ser aportada por las partes o sus abogados.

- si el propósito de la búsqueda es corroborar o desacreditar hechos, o llenar una laguna factual en el proceso la búsqueda es éticamente inapropiada.

Inteligencia artificial = Percepción Artificial + decisión Artificial Percepción [artificial].

Automóvil

Las reglas que se aplican a esa percepción, son por diseño

El sistema “automóvil autónomo” percibe los datos, pero no aprende las reglas.

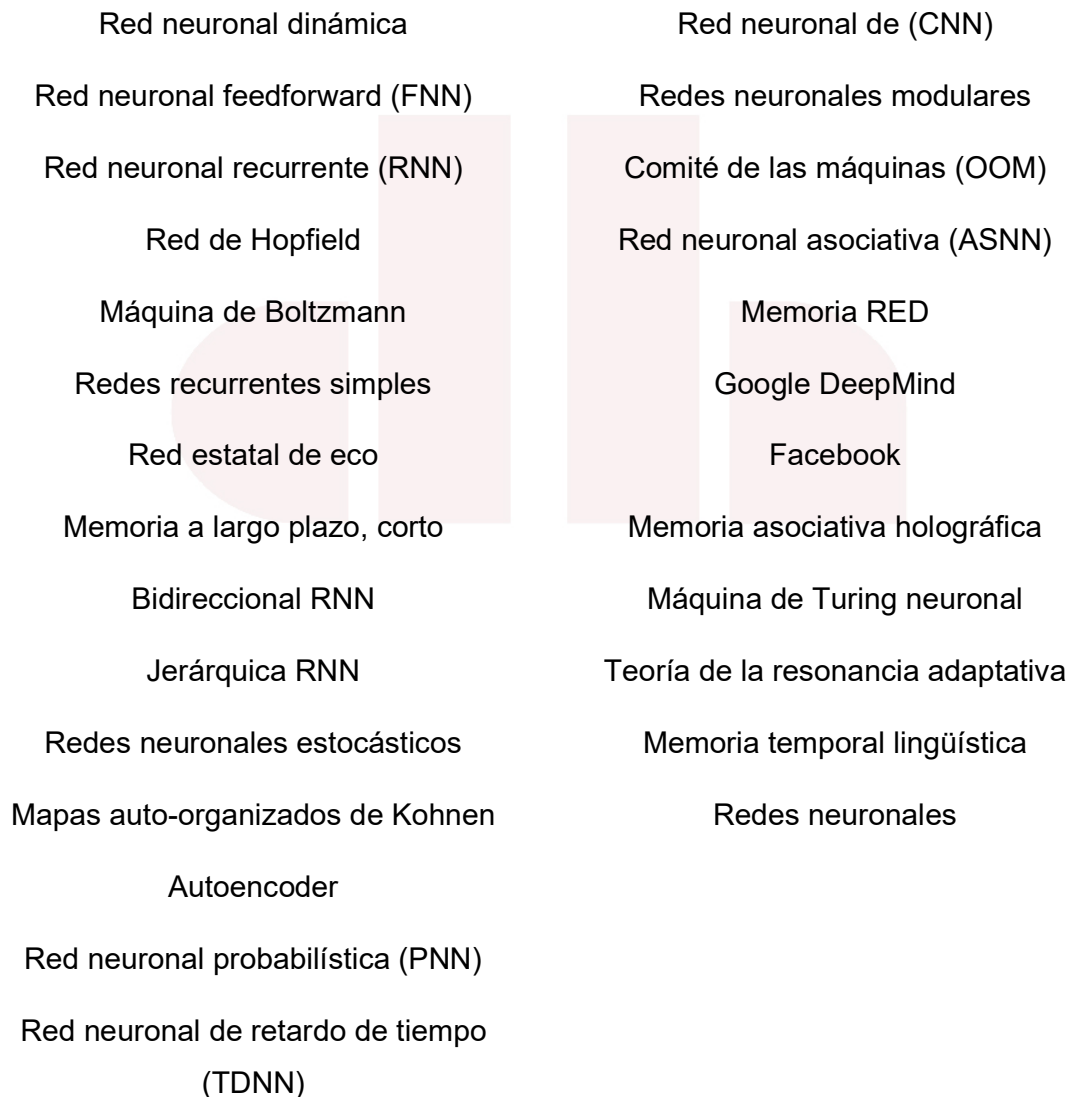
Diagnóstico médico. No hay reglas, lo importante es curar y existe la suficiente variabilidad en sentido clínico, para ello.

NowCast (Mateo)

La IA hace todo el trabajo (de punta-a-punta) implícitamente deduce las REGLAS (de la termodinámica de la atmosfera).

Según Jakob von Uexküll cada organismo biológico tiene su propio y único “umwelt” (Mundo circundante), donde la conciencia y la percepción son una función de la naturaleza del propio organismo.

Uexküll. A foray into the Worlds of Animals and Humans, 2010.



Red de realimentación reguladora
(RFNN)

Estatica red neuronal

Neocognitrón

Neurona de McCulloch-Pitts

Red de función de base radial (RBP)

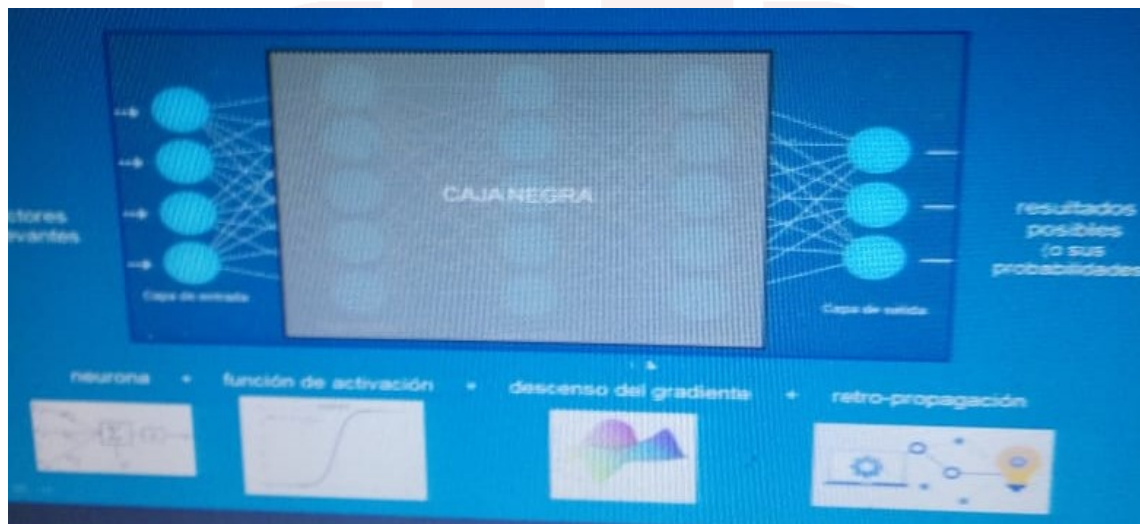
Aprender cuantificación vectorial

Percepción

Modelo Adaline

Lo anterior con sus respectivos códigos de ética y tal vez uno unificado, ejemplo del primero, el código de ética del software de Colombia.

Del modelo siguiente podemos ver la complejidad de interacciones, que deberán de estar definidas, por dichos códigos de ética.



Datos -> Programa de cómputo -> Resultados

Datos ->

Red neuronal -> Programa de cómputo

Resultados ->

Datos -> Programa de cómputo -> Resultados

Fuerza bruta computacional + Datos brutos (big Data)

Información

Factores relevantes

Resultado posibles (o sus probabilidades)

Capa de entrada Capa oculta Capa de salida

Normas

- Resguardar riesgos
- Optimizar beneficios, estas dos, en esta era de la competencia, que conlleva a valores, posmodernos de la relación costo-beneficio, que se vuelven dominantes en la posmodernidad, de acuerdo a Lyotard, al igual de que la verdad en dicha posmodernidad, deberá también ser susceptible de poder guardarse en una base de datos, tomar en cuenta dicha relación costo beneficio y ser comunicable.

Fijar pautas para resolver los conflictos

- Promover su desarrollo y uso seguros
- Respeto por los derechos de las personas
- Disminuir acciones discriminatorias (testers diversos)

Claudio Sarra

Tecno-derecho

Temas y problemas de la informática
robótica jurídica

“Lo que se puede conocer es mi secreto, y no es nada obvio que la capacidad mostrada en descubrir información privada sobre mí, que no pretendía revelar, sea suficiente para justificar su inescrupuloso

James Moor

What is computer.

uso económico” Esta frase nos hace ver que a través de algoritmos en los patrones de uso y lo que cargamos en redes sociales voluntaria o involuntariamente...

Además de las problemáticas anteriores, se dan fenómenos como el llamado teletrabajo o home office, que ha incentivado la subcontratación la que permite vulnerabilidad y precarización laboral en las personas jóvenes, que son paradójicamente las que tienen más habilidades informáticas y que son víctimas del nuevo precariado telemarketinero.

Si, ya que el trabajo decente se ha flexibilizado con el paso de los años, las reformas laborales han eximido y aligerado las responsabilidades del patrón, transformándose en trabajo indecente.

Si bien la flexibilidad laboral, afecta a todas y todos los trabajadores, se refleja también en las personas jóvenes (hombres y mujeres de 15 a 29 años), ya que, nos enfrentamos a la condición de ser joven, de personas jóvenes trabajadoras, personas vulneradas porque el patrón que visibiliza a estas personas como un grupo manipulable y de fácil explotación, razón. En realidad, más que visibilizar, visualiza...

¿Por qué, debemos estar atentos al discurso, de la irrefrenable automatización?

- El gran bluff de la robotización - Husson.
- El discurso del miedo, no es nuevo - Víctor Figueroa.
- Conformación de un nuevo proletariado - Ricardo Antunes.
- Nuevos repertorios sindicales, para tiempos post industriales - Adamini.
- Hacia una nueva internacional sindical - Uta Dirksen.
- Reforzar la organización social, por varias vías - Whendolyne Camarena.
- Si bien el tremendismo y el miedo a lo desconocido es de entenderse, no será predicando que hay una correlación entre el Covid y la 5g, tengan una correlación.

La necesaria contraposición de Paradigmas

Paradigma de la post modernidad: - productividad, innovación tecnológica, conectividad, sostenibilidad

Paradigma de los derechos humanos: dignidad, trabajo decente, desarrollo humano y de la sostenibilidad.

Una definición como punto de partida:

- Hay diferentes denominaciones
- La combinación de los sistemas ciber-físicos (CPS, por sus siglas en inglés) son los ejes de la industria 4.0.
- Constituyen el soporte de la estrategia alemana de fabricación inteligente en concordancia con las capacidades del país en la fabricación y el desarrollo de tecnologías de información.
 - La detallada y completa descripción de las características de la tarea a realizar, plazos de desarrollo y entrega, tipo de contratación, modo de pago, distribución de gastos (material, seguros, cotización, etcétera).
 - En la mayoría de los casos de teletrabajo por cuenta propia, el trabajo no se mide por horas, ya que esto solo corresponde al pasado del trabajo por cuenta ajena y por presencia en las instalaciones de la empresa.

Características de un contrato de teletrabajo

Durante el primer Congreso Iberoamericano de Teletrabajo CIT 2002, Michel ICKX, Marian Navarro Pérez y Paloma Fernández-Nieto, consideraron que las principales características de un contrato de teletrabajo son:

- La distancia (el teletrabajador realiza su tarea en un lugar a libre elección). La opción más elegida es la de alternar unos días de actividad en la empresa con otros en cualquier otra ubicación. Es decir mixto, así se logra por una parte continuidad del clima laboral, espíritu de cuerpo, supervisión y por la otra desconcentración y disminución de costes energéticos, entre otros que deberán de ser regulados, para volver más equitativa la ecuación.

- La valoración por objetivos y no por horarios (por tanto, la libre disposición de su tiempo por el trabajador, ya que lo que interesa es la entrega a tiempo del trabajo encomendado). Esta disposición es relativa, pues si los objetivos de trabajo son muy altos, no más no quedara tiempo a disposición del trabajador y si son comisionistas, están metidos en problemas.

En la categoría de formas colectivas de teletrabajo se incluyen las siguientes:

- La redistribución de las funciones de apoyo al trabajo de oficina en otras regiones o países (redistribución intraempresarial).

- La subcontratación de las funciones de apoyo (interempresarial, incluyendo el uso de telecottages y centros de trabajo).
- Trabajo en grupo, el desarrollo del trabajo en grupo repartido en el seno de la organización, trabajo realizado en colaboración con otras empresas a través del intercambio electrónico de datos y redes de pequeñas y medianas empresas e individuos, conectados entre sí mediante la informática y las telecomunicaciones.
- Se consideran aquellos trabajadores que desarrollan la mayor parte de su actividad en diferentes lugares (nómadas). De hecho, necesitan disponer de equipos fáciles de usar y transportar.

Teletrabajo individual y colectivo

Ursula Huws, asesora y experta en teletrabajo, ha resaltado la importancia de distinguir entre formas de teletrabajo individuales y colectivas. Según su análisis, la primera categoría incluye:

- Teletrabajo realizado parcialmente en casa para una sola empresa.
- Teletrabajo realizado enteramente en casa para una sola empresa.
- Trabajo realizado enteramente en casa para varias empresas. – Varias formas de Trabajo móvil o desde lugares distintos

El teletrabajo ofrece una amplia gama de oportunidades profesionales. Durante el Segundo Congreso Virtual sobre Integración sin Barreras en el Siglo XXI, el argentino Pablo Pascazzi presentó la siguiente lista de ocupaciones de teletrabajo:

- Administración de bases de datos.
- Consultoría en diferentes profesiones.
- Atención al cliente.
- Edición de texto.
- Diseño gráfico.
- Diseño de Web.
- Entrada de datos.
- Periodismo.
- Escritores.
- Programadores informáticos

¿Qué tantos abogados han perdido su trabajo? o ¿por el momento en la región de América Latina, ¿estamos en el clímax en que la informática y todas las asociadas a la IA solo han dado beneficios en confortabilidad y eficiencia, pero a partir de ahora y durante la próxima década se mermara la actividad abogadil y hasta cuanto, ¿con los smartcontract y el blockchain?

Salud y seguridad: El empleador es responsable de la protección de la salud y seguridad de los profesionales y de informar sobre la política de la empresa y las

normas vigentes en estos temas. El trabajador es responsable de cumplirlas. Por aquello del régimen de trabajo personal subordinado.

Formación: El tele trabajador tiene iguales derechos de capacitación que el resto de los trabajadores. No debe haber discriminación al respecto.

Equipamiento: por regla general el empleador es responsable de la provisión, instalación y mantenimiento, a menos que el tele trabajador utilice su propio equipo Y que el empleador tendría que ponerse a mano.

Organizador del trabajo: la organización de la jornada de trabajo es responsabilidad del empleado.

Los requisitos indispensables para ser teletrabajador, de manera enunciativa y no limitativa, son los siguientes:

- Ser sumamente independiente y responsable (también le llaman proactivo).
- Disponer de una computadora y tener acceso a internet (requisito sine qua non).
- Estar dispuesto a tomar un entrenamiento (capacitación).
- Comprometerse a hacer una nueva carrera profesional. Sin embargo, que tan atractivo y viable será esto con los bajos salarios, la nula seguridad social y la gran rotación.
- Tener deseos de progresar y ocupar niveles superiores dentro de la organización. Lo difícil es tener el “deseómetro”, además sería deseable esta característica en otros trabajos que no son a distancia.

Para Jane Tate, uno de los redactores del Informe sobre teletrabajo de la Comisión Europea, la palabra teletrabajo se refiere a “aquellas actividades ejercidas lejos de la sede de la empresa (se les denomina también, en ocasiones, trabajo a distancia), a través de la comunicación diferida o directa por medio de las nuevas tecnologías”.

Para Gray, Hodson y Gordon, el teletrabajo es una forma flexible de organización del trabajo, que consiste en el desempeño de la actividad profesional sin la presencia física del trabajador en la empresa durante una parte importante de su horario laboral. Engloba una amplia gama de actividades y puede realizarse en tiempo completo o parcial.

En la medida de lo posible, la política nacional en materia de trabajo deberá promover la igualdad de trato entre los trabajadores a domicilio y trabajadores asalariados, teniendo en cuenta las características particulares a domicilio y, cuando proceda, las condiciones aplicables a un tipo de trabajo similar efectuado en una empresa.

A menos que esa persona tenga el grado de autonomía y de independencia económica necesario para ser considerada trabajador independiente en virtud de la legislación nacional o de decisiones judiciales:

Una persona que tenga la condición de asalariado no se considerara trabajador a domicilio a los efectos del presente Convenio por el mero hecho de realizar ocasionalmente su trabajo como asalariado en su domicilio, en vez de realizarlo en su lugar de trabajo habitual.

La Organización Internacional de Trabajo (OIT), define el teletrabajo como:

“Una forma de trabajo en la cual: a) el mismo se realiza en una ubicación alejada de una oficina central o instalaciones de producción, separando así al trabajador del contacto personal con colegas de trabajo que estén en esa oficina y, b) la nueva tecnología hace posible esta separación facilitando la comunicación”.

Lo cual implica concebir el teletrabajo como una manera de organizar y realizar el trabajo a distancia con la asistencia de las tecnologías de la información y la comunicación (TIC) en el domicilio del trabajador o en lugares o establecimientos ajenos al empleador.

Reseña histórica del teletrabajo

Surge en la década de los sesenta con Robert Weiner y su teleworking, posteriormente, en 1976, el físico Jack Nilles introduce el término de telecommuting, surgiendo posteriormente el networking, homeworking, y teletravail en Francia.

La idea de esto era “llevar el trabajo al trabajador y no el trabajador al trabajo”, sin embargo, dado que en un principio el desarrollo tecnológico no estaba lo suficientemente preparado para el teletrabajo, éste tardó años en irse posicionando en Estados Unidos y Europa.

Además del mundo del trabajo, otra de las actividades humanas en la que interviene la IA, es en la industria armamentista y en la guerra, en ambas actividades, dada su trascendencia, la ética, debe jugar directriz destacada.

Control humano significativo

Crear es esencial un cierto grado de control humano sobre el **uso** de un sistema de armas autónomo, que traduzca la intención en el funcionamiento del arma y, a fin de garantizar el respeto del IHL, se **debe** limitar el nivel de autonomía legítima del arma.

El control puede incluir:

1. El desarrollo y la experimentación del sistema de armas (fase de desarrollo);
2. La decisión del comandante o del operador de activar el sistema de armas (fase de activación);
3. El funcionamiento del sistema de armas autónomo durante el cual selecciona y ataca independientemente los objetivos (fase de ejecución).

Ética de la gestión de datos e inteligencia artificial

A inicios de 2020 la pontificia academia para la vida, promovió la firma de un llamado a una ética de la inteligencia artificial, propuesta para aplicar inteligencia artificial, para soluciones concretas en el tema alimentario. Otro de los rubros importantes para la vida humana que se verán dramáticamente impactados en la próxima década y en adelante. Impulsada por Monseñor Vincenzo Paglia en conferencia **sobre inteligencia** artificial y el sentido de la vida humana. Lima, febrero 2020.

Además de los ámbitos básicos y cotidianos de la vida, otro asunto que nos acerca más a lo virtual, la automatización.

Datos personales en móviles

- Uso de datos personales en móviles. Y su rastreabilidad
- Uso de datos personales en redes sociales, su uso publicitario.
- El dilema de las redes sociales y su supuesta gratuidad.

Derecho al olvido digital

Derecho al olvido (Derecho de cancelación. Derecho de Oposición)

Casos emblemáticos:

- España: Caso Mario Costejas
- Perú: Caso Google y otros casos.

Valoración y comentarios sobre estos casos. La información debe permanecer como un derecho colectivo a la memoria, siguiendo el criterio de interés público, periodístico e histórico, sobre todo cuando se trate de delitos y no solo los sexuales y borrarse cuando no sea relevante y se trate de fruslerías.

Ética de la inteligencia artificial (Unesco)

La inteligencia artificial debe respetar y preservar la dignidad humana.

La utilización de los sistemas de inteligencia artificial **debe** ser coherentes con el derecho, los principios y las normas internacionales sobre derechos humanos, la equidad y la ética.

La inteligencia artificial debe respetar, proteger y promover la privacidad. Principios éticos de mando.

Uso de Big data. Analítica de datos. Inteligencia de Negocios. Marketing Digital. Perfilamiento y Gestión de Datos.

Ética en la gestión de datos utiliza inteligencia artificial

La ética de la gestión de datos utilizando la **inteligencia artificial**, debe cumplir los siguientes principios:

1. **Imparcialidad.** Los sistemas de inteligencia artificial deben tratar a todos imparcialmente.
2. **Fiabilidad.** Los sistemas de inteligencia artificial deben funcionar de manera fiable.
3. **Privacidad y seguridad.** Los sistemas de inteligencia artificial deben ser seguros y respetar la privacidad.

Ejemplos de información de identidad digital

La identidad digital conjunto finito de atributos que permite a una persona ser identificado como único y probar su identidad electrónicamente frente a terceros.

Nuestra persona digital se compone de varias identidades digitales, estando a su vez cada una de estas identidades representadas por uno o varios identificadores y atributos que son únicos en un contexto específico.

Identidad digital

La llegada de internet tal y como la conocemos, los teléfonos inteligentes y el internet de las cosas han cambiado el tipo de servicios que consumimos, así como la forma que socializamos como el resto de los seres humanos.

La digitalización de nuestras vidas ha introducido nuevos desafíos y oportunidades para la identificación de los individuos y consecuentemente, la identificación e identidad digital con ahora esenciales en cualquier sociedad.

La identidad de una persona, organización, cosa o proceso es todo aquello que lo caracteriza, en un individuo la identidad abarca desde características físicas, género, información biométrica o experiencias, hasta pertenencias, diplomas o propiedades.

Por lo tanto, podríamos decir que existen infinitos atributos que conforman la identidad de una persona y la mayoría de ellos se encuentran en constante cambio y evolución, incluyendo las biometrías.

- **Identidad digital:** es aquel conjunto de atributos que individualiza y permite identificar a una persona en entornos digitales.
- **Identificación digital:** es el procedimiento de reconocimiento de una persona como distinta de otras en el entorno digital.
- **Autenticación digital:** procedimiento de verificación de la identidad digital de una persona, mediante el cual se puede afirmar que es quien dice ser.

Inteligencia artificial

- ¿Por qué ahora?

- Los investigadores han estado trabajando en inteligencia artificial (IA) durante décadas. El progreso se ha acelerado, gracias a tres desarrollos: a) la mayor disponibilidad de datos. b) El poder creciente de la computación en nube. c) Algoritmos más poderosos desarrollados por investigadores de IA. La ética tiene mucho que regular, principalmente en la tercera.

Inteligencia artificial

- La inteligencia artificial y las tecnologías clave que incluyen:

- Conocimiento. La capacidad de un **dispositivo** inteligente de razonar **dicho conocimiento**

- La inteligencia artificial y las tecnologías clave que incluyen:

- Lenguaje: la capacidad de los sistemas inteligentes de comprender el significado de las palabras, teniendo en cuenta los diferentes matices y las complejidades del lenguaje (como las expresiones idiomáticas).

- La inteligencia artificial y las tecnologías clave que incluyen:

- Visión: la capacidad de un sistema inteligente de mediante el reconocimiento de lo que hay en una imagen y un video.

- Voz: la capacidad de los sistemas inteligentes de mediante la comprensión de las palabras que las personas dicen y transcribirlas a texto.

Inteligencia artificial, identidad, ética, datos

La inteligencia artificial: Conjunto de tecnologías que hacen posible que los sistemas digitales perciban, aprendan, razonen y colaboren en el proceso de toma de decisiones para la resolución de problemas que resultan similares a lo que hacen los seres humanos.

Casos prácticos de ética por diseño - Estado

En sistemas judiciales

Sistemas de IA que predicen el riesgo de reincidencia criminal y son usados para informar las decisiones de fiscales y jueces sobre la libertad de las personas. ¿Pueden prevenir en realidad o se cuela es el sesgo?

Programa perfila más drásticamente el riesgo cierto grupo racial que el del resto de la población.

NETFLIX -> Algoritmo genera producciones muy parecidas a las vistas previamente y, genera una “filter bubble” o caja de resonancia que le impide al televidente innovar con géneros y preferencias distintas. ¿Podríamos decir que la consolidación de patrones y su resistencia, promuevan una especie de acción y conservadurismo?

WAZE -> Algoritmo para escoger el camino más rápido, sin importar el consumo de combustible, ni la seguridad de la ruta, es decir basados en una sola variable.

Casos prácticos de ética por diseño – Estado

AMAZON:

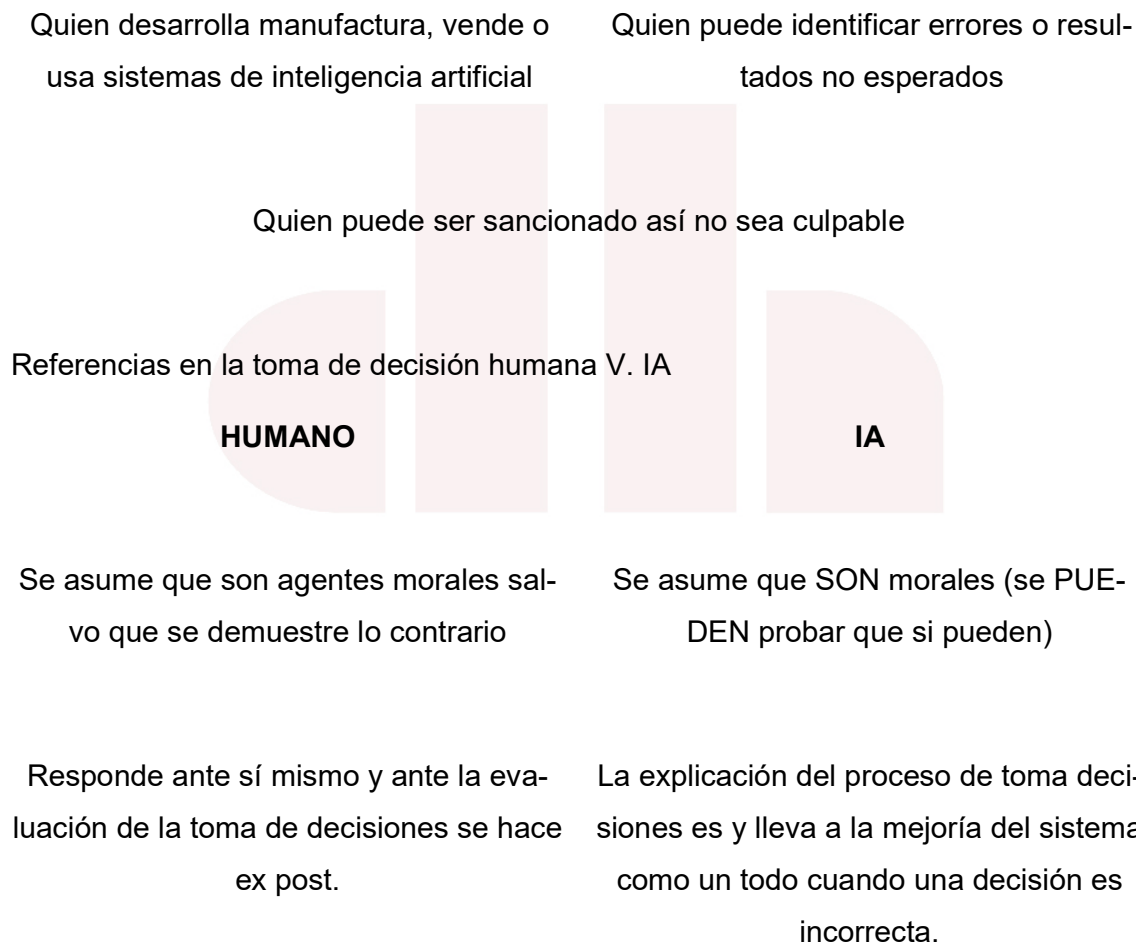
Software de selección de personal había sido entrenado con datos de la última década y discriminaba a mujeres programadoras porque en esa época había muy pocas mujeres ingenieras en sistemas. ¿Y ahora hay muchas?

Discusiones recientes en Colombia

Marco ético para la inteligencia artificial en Colombia aplicable para el sector público y el sector privado.

Tormenta perfecta: Principios generales, difícil interpretación, poco pragmatismo.

Inteligencia artificial y toma de decisiones



Ética en la IA en la actualidad

Cumplimiento de los principios recae en las personas e instituciones que definen y monitorean el comportamiento del sistema.

Sistemas interactivos de control que involucran la intervención humana no es deseable que sean 100% autónomos.

Transparencia

Describir, inspeccionar y Reproducir los mecanismos de la IA

-> Toma de decisiones

-> Aprender para adaptarse a su entorno

-> Manejo de información usada y ofrecida

Responsabilidad

Capacidad de los sistemas de IA de responder por sus decisiones

Rol de las personas que interactúan con estos sistemas de IA

Ambos elementos son necesarios para:

Cadena de responsabilidad entre el sistema de IA y personas involucradas en el proceso de toma de decisiones del sistema

- Responder por una acción o decisión
- Diagnosticar errores o resultados inesperados

Ética y responsabilidad en la IA, tres principios:

(Accountability)

(Responsability)

Rendición de cuentas

Responsabilidad

(Transparency)

Transparencia

¿En que se parecen los principios regulatorios propuestos para la IA?

Generalidad

Función interpretativa

Inseguridad jurídica

Dudas en las expectativas de los inversionistas

Preguntas sobre el desarrollo ético de la Inteligencia Artificial

Uso, privacidad y protección de datos personales	Vigilancia de los mecanismos de IA	Interacción con población vulnerable
Dignidad humana	Toma autónoma de decisiones	Responsabilidad moral
Responsabilidad de los robots	Imitación de seres humanos y seres vivos	Incremento de la población humana
	Estatus de los robots en la sociedad	

Las deficiencias actuales del tratamiento del bias en Machine Learning

1) Enfoque común hacia el logro de la equidad

- Lograr la equidad por desconocimiento.
- Uso de atributos protegidos.
- Imposibilidad de gestionar el sesgo introducido.

2) Los atributos protegidos también conducen a decisiones discriminatorias

- Discriminación inversa
- Los atributos deberían ser solo características del objeto

3) Las tazas de falsos positivos y fal-

4) Predomina el enfoque reactivo al

Los sesgos negativos impactan la percepción de riesgo

- Convierte la discriminación en una cuestión subjetiva.
- Algoritmos son incapaces de ajustar su curva de aprendizaje

Nos dice el Dr. Antonio Martino que Sesgo es el nombre actual de lo que llamábamos “pre juicios”

tratamiento del sesgo

- Evaluar en lugar de asegurar o guiar

De las más de 40 preguntas que realice sesión tras sesión, destaco tres por la trascendencia, la primera en el contexto de la realidad global y la segunda sobre la temática de la ética en sistemas inteligentes.

¿Qué tanto la IA servirá para mejorar condiciones de vida postpandemia o deteriorará la calidad de vida de millones de personas, ¿al excluirlas y acabar con sus empleos, durante la próxima década? ¿qué tanto es la ética, desde evitará dicha exclusión?

¿Cómo poder establecer las medidas para hacer eficaces los principios éticos, dado que existen factores económicos que establecen relaciones asimétricas e impiden su eficacia?, ¿es necesario castigar las prácticas desleales del comercio digital para garantizar la eficacia de los principios, es decir pasar de los principios a las sanciones?

Algunos principios que son a la vez características aplicables a la IA: transparencia, auditabilidad, trazabilidad, inalterabilidad.

En plena aplicación ética de los DDHH, es poner el centro el desarrollo humano, no discriminar a las personas, cuidar el medio ambiente, ayudar a los más débiles, a través de los principios de: transparencia, inclusión, responsabilidad.



CAPÍTULO IV

ÉTICA EN IA. DIMENSIONES EPISTEMOLÓGICA Y ONTOLÓGICA

Por Daniela López De Luise²⁴²

Abstract

La humanidad se encuentra acaso en un punto de inflexión que trasciende la pandemia actual por Covid-19. Detrás de esta dura realidad que nos toca vivir aparecen otras cuestiones no menos alarmantes: La Unesco se ha lanzado a la tarea de generar las bases de un documento que sirva de implante internacional para una ética mundial. La magnitud y profundidad de la empresa es suficiente para disparar un sin número de planteos desde diferentes dimensiones que son directa e indirectamente atacados: por caso la social, educativa, jurídica, tecnológica, bioética, religiosa, cultural, ecológica, etcétera.

Si bien todas y cada una tiene ribetes interesantes de estudio, este trabajo pretende encarar algunos aspectos iniciales respecto a la cuestión de si es posible que exista una ética global. Como parte del análisis surgen las preguntas de qué

²⁴²La autora agradece especialmente el apoyo y colaboración a los docentes de la cátedra de Epistemología, de la Especialización en Comunicación Pública de la Ciencia y la Tecnología, carrera presidida por Susana Gallardo, quien la organiza por la Facultad de Ciencias Exactas y Naturales de la Universidad de Buenos Aires, en conjunto con la Facultad de Filosofía y Letras, y la de Ciencias Sociales del mismo centro de estudios. También cabe un agradecimiento por su orientación a Cecilia Hidalgo, de la cátedra Sociología para la ciencia, perteneciente a la misma currícula académica.

implicaría la posibilidad de una ética global y las dimensiones de la misma. El abordaje se realiza desde la bibliografía, exposiciones de disertantes del ciclo de ética en Inteligencia Artificial (IA) de la Academia Nacional de Ciencias de Buenos Aires (ANCBA), y la propia encuesta Unesco (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, por sus siglas en inglés).

Introducción

Este trabajo tiene como disparador una encuesta global de la Unesco, realizada a mediados de 2020, con el objetivo de desarrollar un documento global que sirva para la regulación de una ética de IA apropiada para todo el planeta.

El desafío implica dos comunicaciones audiovisuales, y el presente texto, todos desarrollados como parte de un mismo estudio y con apoyo de la cátedra Epistemología, con la que se determinó el objetivo, el alcance, gran parte de la bibliografía y los elementos determinantes que hacen a la perspectiva epistemológica del problema y aspectos de sociología para la ciencia.

El objetivo general de la comunicación es analizar el requerimiento de una ética global que impone Unesco para dictar una regla ética global en Inteligencia Artificial (IA). Dicho requerimiento, se traduce aquí en una condición de validez, la que se analizará a continuación.

Entre otros problemas, el trayecto recorrido en los apéndices permite adelantar que no se realiza una determinación epistemológica de IA, ni de tecnología, ni de carácter ontológico, lo que debilita aún más la posibilidad de un establecimiento ético referido a los mismos.

En el desarrollo del texto, además de las referencias y citas a textos tradicionales, el lector hallará entre corchetes los lapsos de los videos correspondientes que se han extraído de las disertaciones, todas disponibles en el sitio WEB de la Academia Nacional de Ciencias de Buenos Aires (ANCBA) que se encuentra en las actividades del Centro de Estudios en Tecnologías Inteligentes (CETI): www.ciencias.org.ar/categoria.asp?id=559.

El formato indica en formato mm.ss, con mm los minutos y ss los segundos de inicio o finalización del lapso, como se ve en el siguiente ejemplo:

[24.13-24:33] ...remarcar [...]controlar.

Los productos audiovisuales terminados se encuentran disponible en dos versiones:

- Comunicación enfocada a la epistemología:
<https://youtu.be/bK13gz3Yz5g>.
- Comunicación enfocada a la sociología para la ciencia:
<https://youtu.be/F8HgoHBQB28>.

Desarrollo del tema

La Unesco ha lanzado en setiembre de 2020 una compulsa a expertos de la IA y temáticas relacionadas con la tecnología de punta. El cuestionario, que originalmente contenía 190 preguntas y fue distribuido a nivel mundial, terminaría compilado de alguna manera discrecional en un documento propuesta, publicado luego en la WEB. Si bien el proceso pretende ser claro y, como se verá aquí, es mímica de otras propuestas relacionadas con la ética, no queda claro cómo se redujo a 141 ítems, ni los criterios para la incorporación (o no) de las recomendaciones en la propuesta final. A pesar de ello, los ítems eliminados entre una y otra versión se han identificado y evaluado como parte del análisis. El documento final, pretende ser confirmado para la adopción de una ética global mundial que impactará la manera en que nos manejamos en lo social, político, científico, sanitario y legal.

Este apartado constituye la ilación final del material hallado en primera instancia en los apéndices 1, 2, 3 y 4, con las referencias bibliográficas. La función de esta sección es de servir de base para la redacción del contenido a transmitir en el guion de la comunicación.

La humanidad se encuentra acaso en un punto de inflexión que trasciende la pandemia actual por Covid-19. Detrás de esta dura realidad que nos toca vivir aparecen otras cuestiones no menos alarmantes: La Unesco se ha lanzado a la tarea de generar las bases de un documento que sirva de implante internacional para una ética mundial. La magnitud y profundidad de la empresa es suficiente para disparar un sinnúmero de planteos desde diferentes dimensiones que son directa e indirectamente atacados: por caso la social, educativa, jurídica, tecnológica, bioética, religiosa, cultural, ecológica, etc. Si bien todas y cada una tiene ribetes interesantes de estudio, este trabajo pretende encarar algunos aspectos iniciales respecto a la cuestión de **si es posible que exista una ética global**. Como parte del análisis surgen las preguntas de qué implicaría la posibilidad de una ética global y las dimensiones de la misma. El abordaje se realiza desde la bibliografía, exposiciones de disertantes del ciclo de ética en IA de la Academia Nacional de Ciencias de Buenos Aires (ANCBA), y la propia encuesta Unesco.

1. Un problema ontológico: ¿Ética de qué?

De acuerdo con Linares “Los límites éticos globales exigidos, se sustentan en esta definición ontológica del mundo tecnológico”²⁴³. Esto implica la preexistencia ontológica de un “algo”, que de por sí habilita la determinación ética. Sin embargo, para la perspectiva instrumental mueve a generar una definición de trabajo acerca del objeto ontológico sobre el cual se pretende trabajar la ética, en este caso la definición de IA.

²⁴³Linares, Jorge, *Hacia una ética para el mundo tecnológico*, “ArtefaCToS”, vol. 7, n° 1, 2018, 2a Época, p. 99 a 120.

La propuesta, si bien aclara en diversos incisos los argumentos de su necesidad y requerimiento, no establece con precisión el objeto de estudio. Luego de sus declaraciones iniciales realiza una definición “a los fines de la Recomendación”:

Pueden verse como sistemas tecnológicos con capacidad de procesar información en formas que imitan el comportamiento inteligente, y típicamente incluyen aspectos de aprendizaje, percepción, predicción, planeamiento o control.

Los sistemas de IA pueden incluir diversas metodologías y tecnologías, entre otras están...

De las primeras observaciones, es la falta de delimitación de la propia IA. En referencia a esto mismo, por ejemplo puede contraponerse la definición del IEEE que es Sociedad Científica y Tecnológica de ingeniería y campos asociados constituida como la más antigua y amplia a nivel mundial:

Es la teoría, diseño, aplicación y desarrollo de paradigmas computacionales inspirados biológicamente y lingüísticamente²⁴⁴.

Que resulta en elementos definibles en un mayor grado. La falta de demarcación de los límites de lo que se considera IA, generará eventualmente un emergente no explicitado y divergente, de cuestiones que quedan en una zona gris, desde la perspectiva temporal y geográfica. Esto es especialmente cierto considerando las desigualdades tecnológicas entre Estados.

Lo segundo a observar es que los expertos no emplean el término IA, sino Computational Intelligence (Inteligencia Computacional) o Intelligent Systems (Sistemas Inteligentes), dado que el término IA remite a los primeros paradigmas abarcados principalmente desde el surgimiento en los ‘60 hasta los ‘80, cuando el sector cambió radicalmente a los sistemas cuya potencia (paradójicamente) es mucho mayor y cubre las aplicaciones que supuestamente se quieren controlar.

Surge entonces la pregunta: ¿Este “descuido” fue accidental? ¿o se pretende dejar fuera del control a las aplicaciones más perturbadoras?

Adicionalmente la Unesco argumenta que el alcance de la IA es casi ilimitado:

En todos los campos de las ciencias, ciencias sociales y humanidades, la IA tiene incumbencias en nuestra concepción de entendimiento y explicación científica, y en las maneras en que el conocimiento científico puede ser aplicado en la toma de decisiones.

Lo que puede leerse como el argumento para el requerimiento de una ética universal, que englobe a todas las éticas de los diversos ámbitos.

La definición de trabajo, sobre concepciones imprecisas e inconclusas, hace que el objeto ontológico pueda ser diverso según las maneras de completar esos puntos parciales, por lo que da espacio a la posibilidad de una ética globalizada de todas las concepciones posibles (definidas o no).

²⁴⁴IEEE Computational Intelligence Society, <https://cis.ieee.org/about/what-is-ci>.

2. Un problema axiológico²⁴⁵: ¿Una ética de IA?

El documento Unesco expresa que la propuesta acciona:

En reflejo de la influencia profunda que la IA puede tener en las sociedades, ecosistemas y vidas humanas, incluyendo la mente humana, en parte debido a las nuevas maneras en las que puede influenciar al pensamiento humano y el proceso de decisiones, y afectar la educación, la ciencia, la cultura y la información y comunicación.

En este mismo sentido Bergel advierte que “La ciencia y tecnología, aventuras del pensamiento humano, no reconocen límites y, si se los fija hipotéticamente, pueden ser fácilmente vulnerables”²⁴⁶.

En este punto cabe la pregunta: ¿Fijar los valores de manera uniforme y sin atender diferencias culturales vulnera los derechos humanos?

Al respecto es interesante escuchar a la letrada Natacia Arcifa, especialista en armas letales autónomas²⁴⁷:

[24.13-24:33]...remarcar la importancia de mantener el libre albedrío humano sobre las decisiones de los sistemas autónomos. Parece un poco extraño porque si un sistema es autónomo no se puede controlar.

3. Un problema de conocimiento: ¿Construir un panóptico de información?

La propuesta explica que:

Los sistemas de IA pueden dar un gran servicio a la humanidad pero también generan problemas éticos fundamentales, por ejemplo respecto a los sesgos que pueden involucrar o exacerbar, resultando en potenciales desigualdades, exclusión y en una amenaza a la diversidad cultural y social y de género; requieren de transparencia y entendimiento de los algoritmos de trabajo y de los datos con los que son entrenados; y su potencial impacto en la privacidad, libertad de expresión, proceso sociales, económicos y políticos, y el entorno.

Al punto, es notorio que los aspectos mencionados no incluyan lo jurídico, siendo que la misma Unesco²⁴⁸ lo incorporara en su *Declaración universal sobre bioética y derechos humanos* (París, 2005). Como se verá luego, *lo jurídico tiene*

²⁴⁵Axiología, filosofía de los valores o filosofía axiológica es la rama de la filosofía que estudia la naturaleza de los valores y juicios valorativos.

²⁴⁶Bergel, Salvador D., *Diez años de la Declaración Universal sobre Bioética y Derechos Humanos*, “RevistaBioética”, vol. 23, n° 3, 2015, p. 446 a 455.

²⁴⁷N. Arcifa (Universidad de Pisa-Italia). Disertación: *Inteligencia Artificial o derechos humanos? Qué es más necesario?* Tema: *Armas letales autónomas (autonomous weapons)*. Ciclo de ética en Sistemas Inteligentes de la Academia Nacional de Ciencias de Buenos Aires, 2020.

²⁴⁸Declaración Universal sobre Bioética y Derechos Humanos. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, Unesco, 1995-2007, ID: 31058.

profundas raíces en las concepciones éticas y requiere de ciertos postulados previos.

Las afirmaciones de “amenaza a la diversidad cultural y social y de género” corresponde con lo que Linares denomina “filosofía de la sospecha”²⁴⁹, por la que el mundo tecnológico causaría fragmentación social, aislamientos y problemas ecológicos. También explica la existencia subyacente del “determinismo tecnológico”, por el que el futuro tecnológico es determinado por factores internos al desarrollo tecnológico. Esto deriva en la tecnología dominando a sus usuarios. En el ámbito latinoamericano, se refleja en el llamado de Garrafa²⁵⁰ a expandir la epistemología a los fines del estudio de las tecnologías para lograr: Protección, Prudencia, Precaución, Prevención.

Pero otras perspectivas hallan una idea reduccionista en esta visión²⁵¹ ya que el mencionado cruce y riesgo de todas las otras ciencias no sería tan directo.

El ítem 78 de la encuesta²⁵², se apoya en estas concepciones para promover la construcción de un **panóptico** que observa los “estudios e investigaciones” de disciplinas adicionales a la ciencia, tecnología, ingeniería y matemática (con enumeración explícita de las mismas) a fin de monitorear adecuadamente las investigaciones en IA y los malos usos o efectos adversos. Complementa este concepto con los ítems 134 (implementación de un observatorio de “ética de la IA”), 78 (patrimonio común de datos), y 48 (mecanismos y políticas específicas), entre otros.

Se evidencia aquí la evolución que sufre Unesco desde la Declaración Universal de Bioética²⁵³ (París, 2005), donde el art. 9 determina que se debe preservar la información de las personas de manera confidencial, que no debería utilizarse ni revelarse para fines distintos que los de su acopio. El uso a los fines de fiscalización rompería este esquema, y presenta un cambio de posición con la versión de 2010.

Para Simondon, esta perspectiva el riesgo es una percepción que se valida desde una falta evolutiva. En sus palabras: “la iniciación a las técnicas se debe situar en el mismo plano que la educación científica”²⁵⁴, además agrega que “un niño debería saber qué es una autorregulación o una reacción positiva, al igual que conoce

²⁴⁹Villalba Gómez, Jairo A., *Problemas bioéticos emergentes de la inteligencia artificial*, “Revista Diversitas. Perspectivas en psicología”, vol. 12, n° 1, 2016.

²⁵⁰Tealdi, Juan C., *Diccionario latinoamericano de bioética*, Unesco, Red Latinoamericana y del Caribe de bioética, Universidad Nacional de Colombia, 2008.

²⁵¹Philosophy of Biology, *Stanford Encyclopedia of Philosophy*, 2020.

²⁵²Member States and funding bodies should promote interdisciplinary AI research by including disciplines other than science, technology, engineering, and mathematics (STEM), e.g., law, international relations, political sciences, education, philosophy, culture, and linguistic studies to ensure a critical approach to AI research and proper monitoring of possible misuses or adverse effects.

²⁵³Declaración Universal sobre Bioética y Derechos Humanos, Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, Unesco, 1995-2007, ID: 31058.

²⁵⁴Simondon, Gilbert, *El modo de existencia de los objetos técnicos*, Prometeo Libros, 2008.

los teoremas matemáticos; es tan desinteresada como la práctica de las artes”, instituyendo la ciencia desde la raíz misma de la formación.

En cierta forma, Simondon denuncia el alejamiento del código de la información que representa al pensamiento y la comunicación de la realidad científico-tecnológica actual, y culpa a dicha disociación de tergiversar el mensaje y promover una visión distorsionada de los que denomina “seres tecnológicos”.

En su trabajo “El modo de existencia de los objetos técnicos”, determina que la cultura es un sistema de defensa de la humanidad contra las técnicas. Debe incorporar a los seres técnicos por medio de la filosofía puesto que lo técnico es un intermediario entre el hombre y la naturaleza. La postura plantea que los objetos técnicos no tienen realidad humana, pero encierran lo humano. Entiende a la cultura como una actitud dual: por un lado, es ensamblaje de materia y por el otro es un ente con intencionalidad hostil y peligrosa. En estos términos, el hombre se convierte en organizador permanente de la sociedad de objetos técnicos, quienes a su vez están abiertos al hombre-organizador.

Desde esta perspectiva, el hombre permanentemente regula el margen de indeterminación para adaptarse en diferentes grados a posibilidades de información diversa. Este manejo, para ser adecuado requiere de una toma de conciencia de la realidad técnica no como relación de uso-propiedad, sino de objeto de conocimiento científico. Pero esta capacidad de toma de conciencia recae en una introducción de la misma en la cultura. Por ello debe incluirse de manera sistemática en la educación de todos los niveles, cuestiones como la conciencia de la naturaleza de las máquinas, las reacciones posibles entre ellas, con el hombre, y los valores existentes en todas estas instancias mencionadas.

Es aquí donde la postura de Simondon converge con la crisis actual. La tecnología es un problema no solo del tecnólogo, sino del hombre en todas sus esferas y en todo momento de su vida.

Las lecturas de riesgo, amenaza, desnaturalización y descontrol, pasan a ser consecuencia inmediata de una codificación inadecuada a la realidad gobernada (la que es integrada por hombres y máquinas), por lo que no hay una función directiva real del hombre en su entorno. El desfase resulta de la falta de incorporación de esquemas actuales, mediadores entre los elementos que sustenten un código, un símbolo cultural que represente adecuadamente a la realidad. La génesis adecuada de los elementos técnicos, permiten una sinergia equilibrada, y ha evolucionado en tres estadios esenciales:

- Como elemento: el modo inicial que pretende evolucionar y perfeccionar una herramienta, por lo que no existe angustia. Corresponde su esplendor al siglo XVIII.

- Como individuo técnico: en donde se percibe a la tecnología como adversaria del hombre (un competidor), puesto que reemplaza al hombre-máquina en su actividad como herramienta inteligente. También percibe a lo técnico como una oposición de lo natural, una violación de la naturaleza, y como una expresión de voluntad de poder desmesurada.

- Como conjunto tecnológico: constituye el punto de evolución necesario para poder naturalizar la relación con la tecnología. En este punto la máquina se convierte

en una proveedora de información y crecimiento, con el crecimiento de la neguentropía el humano se opone a detener la evolución del universo. En ese cambio de actitud mantiene su poder de cambio del universo y convierte a la tecnología en el medio por el que el hombre estabiliza al entorno al volverse parte del cambio entrópico del todo.

La humanidad, en virtud de la imposibilidad de acceder al total de la realidad representada en su entorno, es vuelve incapaz de administrarla y a la vez la percibe erróneamente inmanejable y amenazante.

¿Cuál es la reparación exigida por el entorno? La mutación del código del esquema hombre-herramienta al de hombre-máquina con todo lo que ello implica.

Parte de esa reparación exige lo que el autor declara en “el pensamiento humano debe instituir una relación igual, sin privilegio, entre las técnicas y el hombre. Esta tarea queda por cumplir, porque los fenómenos de dominancia técnica, que hacen que en cada época haya una parte del mundo técnico reconocida por la cultura, mientras que la otra es rechazada, mantienen una relación inadecuada entre la realidad humana y la realidad técnica”²⁵⁵. Esa reparación se ajusta al proceso que se inserta entre la concepción racional del hombre como ser social y como individuo. Es la coordinación entre lo que el autor denomina “estatuto de mayoría” y el de “minoría”. Es la coherencia entre la realidad técnica y la concepción desde el pensamiento.

La sociedad actual adolece aún de esta discrepancia. La virtualidad aporta a la perspectiva individual de la reflexión acerca del “objeto tecnológico” de Simondon.

4. Un problema de economía ¿Una ética para el desarrollo?

A. La apropiación técnica desde el Subdesarrollo

Linares reconoce desde Lewonting y Levins, en *Biology under the influence*, asocian los cambios que la IA produce en ámbitos como la sociedad, usos de la tierra, nutrición, etcétera²⁵⁶. Esa influencia se manifiesta también en este ítem del documento estudiado:

...Conciencia de que los países de bajo y medianos ingresos (LIMC), incluyendo África, Latinoamérica y Caribe, Asia Central y los pequeños Estados isleños en desarrollo, sufren una aceleración del uso de tecnologías de información e IA, y que la economía digital presenta desafíos importantes y oportunidades de sociedades creativas, que requieren el reconocimiento de culturas endógenas, valores y conocimiento para desarrollar las economías.

El apartado atiende específicamente a la capacidad de apropiarse de la innovación tecnológica en instancias específicas como las “tecnologías de información e IA”, en Estados con problema económico. Se observa en este punto la idea de *posi-*

²⁵⁵Simondon, *El modo de existencia de los objetos técnicos*.

²⁵⁶Linares, Jorge, *Ética y mundo tecnológico*, Biblioteca de Ética y Bioética, Universidad Nacional Autónoma de México (Facultad de Filosofía y Letras) y Fondo de Cultura Económica, 2000.

ble futuro conflicto emergente entre las culturas locales endógenas, el desarrollo económico y la innovación (vista aquí como disruptiva).

Desde la perspectiva de Linares, entonces se estaría previniendo un error tradicional: subestimación por parte de la ciencia hacia ciertos emergentes. Este autor remarca que el problema suele consistir en la propia manera en la ciencia resuelve los problemas: prediciendo en base a lo conocido. Proceso de construcción de conocimiento que constituye para la ciencia moderna, la esencia de validez de conocimiento.

B. La economía bélica

La letrada Natacia Arcifa (Universidad de Pisa-Italia) en su disertación para la Academia Nacional de Ciencias de Buenos Aires²⁵⁷ expone una mirada diferente de la desigualdad económica por la IA. Su perspectiva, planteada inicialmente en el seno de la Unión Europea, se enraíza más en la capacidad productiva bélica, que de apropiación de la IA por países subdesarrollados:

[0:9:12 – 0:10:00] *La disputa es hoy bastante acalorada puesto que en el 2010 el mercado de uso militar y de seguridad pública de drones, valía casi 5 billones de dólares, posteriormente, desde 2015 hubo un incremento del doble, porque los principales países (como los Estados Unidos, Rusia, China, Francia) han invertido billones en el desarrollo de la autonomía en el uso de la fuerza militar y están financiando sistemas de inteligencia artificial que confeccionan unos escenarios inesperados por nuestra sociedad.*

La nueva industria bélica es millonaria, corre el centro del problema de los mecanismos de apropiación a la concentración acelerada y descontrolada de una industria perjudicial, gracias a la decisión estratégica de los Estados. La decisión entonces pasa de la esfera de la ciencia en IA a la política económica y estratégica de ciertos Estados: si hay Estados que invierten en una industria bélica, es de esperar que luego esa industria deba producir beneficios, y para ello hacen falta guerras.

Pero la perspectiva económica tiene otras lecturas. Arcifa, expone en la misma disertación²⁵⁸ una perspectiva de *economía de los conflictos* en la que:

[0:17:54 – 0:18:14] *Los Estados Unidos, China, Rusia sostienen que, en vez la discusión [sobre la adopción o no de armas autónomas] debería valorizar las ventajas que la introducción de armas autónomas traerá en la Economía de los conflictos humanos. Sobre todo, desde el punto de vista de la reducción del costo humano.*

Este tipo de economía refiere a que los asalariados podrán mantener la capacidad productiva del Estado, mientras sostienen los conflictos armados, lo que redundaría, en un sistema bélico que no interfiere con el desarrollo industrial, y que además facilita el desempeño bélico. Esto hace, según Arcifa, que los conflictos se-

²⁵⁷N. Arcifa (Universidad de Pisa-Italia), Disertación: Inteligencia Artificial o derechos humanos? Qué es más necesario? Tema: Armas letales autónomas (*autonomous weapons*), Ciclo de ética en Sistemas Inteligentes de la Academia Nacional de Ciencias de Buenos Aires, 2020.

²⁵⁸N. Arcifa (Universidad de Pisa-Italia), Disertación: Inteligencia Artificial o derechos humanos? Qué es más necesario? Tema: Armas letales autónomas (*autonomous weapons*), Ciclo de ética en Sistemas Inteligentes de la Academia Nacional de Ciencias de Buenos Aires, 2020.

an más económicos de realizar, posibilitando su incremento en número. El llamado de Estados Unidos, China y Rusia sería a una ética de la guerra, en la que los conflictos serían más “seguros” para los humanos, económicos para los Estados, y apropiados para las economías industrializadas.

Pero el tablero político internacional, estaría aún en pleno cambio²⁵⁹.

[0:18.37 – 0:18.52] *Hay países con una posición intermedia. Por ejemplo, Francia y Alemania. Muchos en Europa tienen una posición intermedia. Argumentan que es posible tenerlos, pero de manera controlada.*

En este último bloque de países, no hay una decisión por rechazar ni adoptar la nueva economía de los conflictos humanos.

5. Un problema científico: ¿Ética de la innovación?

Para Linares²⁶⁰ hay cinco personajes “anunciadores” de una catástrofe: Martin Heidegger, Jacques Ellul, Günther Anders, Hans Jonas y Eduardo Nicol. Ellos advierten los peligros de la tecnología, y el inicio de lo que Mc Kibben denomina “Era del Antropoceno” donde la naturaleza terrestre está irreversiblemente cambiada. El apocalíptico anuncio estaría ya en proceso de verificarse en tanto hemos cambiado casi todos los ecosistemas.

En términos del autor, la catástrofe cubriría todos los ámbitos (ecología, historia, política-social, ontológica y al hombre), pero no es vista por la comunidad debido al “desfazaje prometeico” (Anders) por el que somos incapaces de tomar conciencia y responsabilidad de los efectos del mundo tecnológico.

Alfredo Marcos detalla algunos aspectos del cómo, en cierta manera, la IA logra sus habilidades, y se pregunta: “¿quién se hace responsable en cada caso de un posible fallo?, ¿habría que poner comillas en ‘decide’ y ‘diagnostica’?” “termina, en referencia a la aplicabilidad de estos términos que la IA aplica realmente por imitación a la contraparte humana”²⁶¹.

Esta línea de pensamiento lleva a la necesidad de determinar elementos de juicio adicionales a la IA, por analogía a lo humano. Por caso una “personalidad jurídica” de los robots. Pero esta antropomorfización lleva a un lugar muy conveniente para ciertos sectores, como lo observa Suñé²⁶²:

²⁵⁹N. Arcifa (Universidad de Pisa-Italia). Disertación: Inteligencia Artificial o derechos humanos? Qué es más necesario? Tema: Armas letales autónomas (*autonomous weapons*). Ciclo de ética en Sistemas Inteligentes de la Academia Nacional de Ciencias de Buenos Aires. 2020.

²⁶⁰Linares, Jorge, *Hacia una ética para el mundo tecnológico*, “ArtefaCToS”, vol. 7, n° 1, 2018, 2a Época, p. 99 a 120.

²⁶¹ Marcos, Alfredo, *Información e inteligencia artificial*, Universidad de Valladolid, 2020.

²⁶² Suñé Llinás, Emilio, *Un “decision making” ético para sistemas inteligentes*. Seratti, Lidia, *Hacia una racionalización del alcance de la algorítmica de la IA y de los desafíos éticos y regulatorios impuestos por su irrupción sin límite ni fronteras*, Ciclo de Ética en Sistemas Inteligentes de la Academia Nacional de Ciencias de Buenos Aires, 2020.

[0:52:19- 0:53:06] *Lo que se pretende, con la personalidad jurídica de los robots que no tienen patrimonio más que sí mismos, sería una pantalla de responder con su valor precisamente, y ahí está la trampa, porque no tuviera que responder el fabricante que es lo que se ve venir encima y quizás por eso se está retrasando la comercialización del coche autónomo.*

Para Marcos, esto es un sinsentido, ya que un sistema de IA, es una suerte de “prótesis” de nuestros sistemas de expectativas, puesto la IA que se vale de un artilugio matemático sobre datos históricos para decir qué esperar de cierto objeto con el que interactúe. Por supuesto que, como cualquier tipo de operación artificial, puede colapsar en situaciones donde [la información dada por la comunicación se vuelve disruptivo para el receptor], y corresponderá en esas instancias que reaccionen las personas con creatividad, cambiando el sistema de referencias que se le haya dado al sistema artificial.

Como se explicó en los segmentos anteriores, la IA es esencialmente transversal a todas las áreas del humano moderno, por lo que se inscribe en el proceso de innovación con un acento creciente. Esto no la convierte en un elemento peligroso per se, sino que la ata a *los valores y ética humanos*, tal como lo declara el documento de Unesco:

Atender riesgos y problemas éticos no debería impedir la innovación sino más bien estimular a nuevas prácticas de investigación e innovación responsable en la que la investigación diseño desarrollo y puesta en producción y uso de la IA se ancla en valores morales y reflexión ética.

Pensadores como Foridi (2014), ven el cambio más optimistamente, como una “nueva infósfera” donde la fusión de objetos, artefactos y humanos se realiza por intercambio y procesamiento digital²⁶³.

Para Gilbert Simondon esta visión apocalíptica corresponde a una falla en la concepción de la tecnología concibiéndola como algo externo cuando ha sido interna desde la misma creación del hombre²⁶⁴.

Luego de establecer que la IA se ancla en “valores morales y reflexión ética”, el inciso 74 también promueve la inversión para la investigación en una ética de IA:

Los Estados miembro deben promover la investigación en ética de IA a través de inversiones directas o por medio de la creación de incentivos para que los sectores públicos y privados investiguen en esta área.

El pedido de la constitución de una ética de la IA se sostiene en

“un entorno normativo para la IA y sus implicaciones sociales se encuentra en la intersección de la ética, derechos humanos, marcos legales internacionales y nacionales y la libertad de investigar e innovar, y el bienestar humano”.

Establece entonces una ética de IA emergente *desde la investigación* de los Estados miembros, pero con un *marco legal común*. En lo que deposita la *carga éti-*

²⁶³Linares, *Hacia una ética para el mundo tecnológico*, “ArtefaCToS”, vol. 7, n° 1, p. 99 a 120.

²⁶⁴Simondon, *El modo de existencia de los objetos técnicos*.

ca sobre el sector científico, suponiendo una ciencia como la proclama Merton²⁶⁵. Según éste, la mentalidad religiosa del protestantismo ascético y la lógica de la ética puritana han caracterizado la ciencia tradicional. De ello supone una concepción de bienestar social como meta constante derivada del Ethos puritano.

Pero evidencia una evolución en la posición, ya que para trabajar en bioética la misma Unesco en 2005 pretende mucho menos²⁶⁶: no un marco normativo sino un marco de principios y procedimientos, dejando librado a los Estados las políticas, legislaciones y códigos éticos.

Al respecto, el criterio normativo originalmente planteado es compatible con Ortega (1997)²⁶⁷. En su evaluación, la técnica humaniza al medio para hacerlo más habitable a los humanos. Por ello es de esperar que una mejora técnica redunde en una mejora humana, pero también determina que “esto solo lo podemos argumentar si disponemos de un criterio normativo”.

Un aspecto no menor, que se agrega a lo dicho es que el mismo documento reclama la autoridad de quién es dueño de lo innovado: el ítem 100 declara que los Estados miembros tienen potestad de determinar de quién son los *derechos de propiedad intelectual* sobre los productos de IA. Por lo que cabe preguntarse sobre la efectividad de las investigaciones de ética de IA, implicadas directamente en los derechos de propiedad regulados y protegidos por los Estados miembros.

6. Un problema legal: ¿Quién tiene la carga?

En ocasión de su disertación ante la Academia Nacional de Ciencias de Buenos Aires, Suñé Llinás, catedrático de filosofía del derecho de la Universidad Complutense, y subdirector del departamento de derecho internacional y filosofía del derecho afirma²⁶⁸:

[0:35:07 – 0:35:46] *Hablar de personalidad jurídica de los robots, es la manera de decir (yo fabricante): “¿Quién es el responsable?, pues, ¡el coche autónomo! ...”*

...“Y el coche autónomo, ¿con qué patrimonio responde? ... Solo que sea con el propio valor del coche autónomo...es una inmensa trampa...”

²⁶⁵Becker, George, *Pietismo y ciencia: una crítica de la hipótesis de Robert K Merton*, “American Journal of Sociology”, 1984.

²⁶⁶Declaración Universal sobre Bioética y Derechos Humanos, Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, Unesco, 1995-2007, ID: 31058.

²⁶⁷Chillón, J. M. - Marcos, A. *Técnica y sentido*, “SCIO. Revista de Filosofía”, n° 11, 2015, p. 77 a 99.

²⁶⁸Suñé Llinás, *Un “decision making” ético para sistemas inteligentes*. Seratti, *Hacia una racionalización del alcance de la algorítmica de la IA y de los desafíos éticos y regulatorios impuestos por su irrupción sin límite ni fronteras*, Ciclo de ética en Sistemas Inteligentes de la Academia Nacional de Ciencias de Buenos Aires, 2020.

Lo que se está intentando, y los juristas no se dan cuenta, porque les están dando un caramelo envenenado, es efectivamente quitar del medio a la responsabilidad de la fábrica”.

Este tipo de falencias regulatorias son la que refiere el documento de Unesco:

Considerando que el rápido desarrollo de los sistemas de IA encuentrabarreas de entendimiento e interpretación de la IA, debido a la *diversidad de orientaciones éticas y culturas* alrededor del mundo, la *falta de agilidad de la ley* en relación con la tecnología y la sociedad de la información, y el riesgo de que los estándares éticos locales y regionales y los valores sean trastocados por la IA.

Pero la Unesco en ocasión de la *Declaración Universal de Bioética* (2005, París)²⁶⁹ no remite el problema a la arena legal, sino a la solución de pluralidad por otro camino: búsqueda de principios éticos en dos pasos: consenso, y pactos sociales. En este sentido, la carga de la iniciativa es previa a la ley: corresponde a un emergente combinado con medidas a nivel social y político, y el marco no estaría dado por una regulación global a la que adscribir, como se menciona en el inciso 5, “Un problema de científico”.

De hecho, el mismo Suñé en su disertación apoya una postura pragmática, planteando un proceso de gestión a partir de la jurisprudencia y no de un marco ético previo²⁷⁰:

[0:35:55 – 0:36:51] *Para mí, la adopción de decisiones sobre la base de criterios directamente éticos nos lleva, por las razones expuestas, a un callejón sin salida. La alternativa racional es más jurisperdente. Basada sobre la pragmática y la jurisprudencia, porque la inmensa mayoría de estas cuestiones se han planteado a través de principios como el “estado de necesidad” ante Tribunales de Justicia que han dado soluciones concretas. Tenemos ahí un banco de experiencia. Y no es que yo diga que hay que volver a los sistemas “expertos”...*

Suñé entiende que dada la heterogeneidad cultural (mencionado en el documento Unesco como diversidad en orientaciones éticas), es imposible a priori fijar regulaciones globales, y por ello emplea el término “callejón sin salida”.

A su vez Natacia Arcifa da unos números que muestran esta misma preocupación, cuando se votó la posición de los Estados miembros respecto a la adopción de armas letales autónomas²⁷¹:

[0:17:20 – 0:17:36] *Durante esta conferencia [última conferencia de expertos en IA, con 86 Estados soberanos participantes], 22 de los Estados participantes, solamente en 22*

²⁶⁹Declaración Universal sobre Bioética y Derechos Humanos, Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, Unesco, 1995-2007, ID: 31058.

²⁷⁰Suñé Llinás, *Un “decision making” ético para sistemas inteligentes*. Seratti, *Hacia una racionalización del alcance de la algorítmica de la IA y de los desafíos éticos y regulatorios impuestos por su irrupción sin límite ni fronteras*, Ciclo de ética en Sistemas Inteligentes de la Academia Nacional de Ciencias de Buenos Aires, 2020.

²⁷¹N. Arcifa (Universidad de Pisa-Italia). Disertación: *Inteligencia Artificial o derechos humanos? Qué es más necesario? Tema: Armas letales autónomas (autonomous weapons)*. Ciclo de ética en Sistemas Inteligentes de la Academia Nacional de Ciencias de Buenos Aires. 2020.

se han manifestado contrarios a la adopción de estos sistemas de armas autónoma letal.

7. Un problema legal: ¿Regulación emergente en IA?

Más allá de los antecedentes que se está solicitando ese consenso a nivel legislativo, al mismo tiempo la ley tracciona de un consenso global de normas éticas, no como una imposición global sino de articulación específica de las cuestiones generadas por la IA. En la presentación de Arcifa esta diferenciación es clara²⁷²:

[0:21:50 – 0:22:21] *Ha surgido la necesidad de promulgar reglas éticas. Reglas éticas que, imagino todos conocen las 3 leyes de Asimov. No son que sea una simple implementación, sino que abrazan la cuestión del modo como la robótica se articularía con el control por parte del ser humano.*

En este caso se trata de robótica con su control por parte de los agentes humanos sobre los mismos, pero hay dos cuestiones importantes al respecto:

- 1) toma como base las reglas surgidas de un especialista en el área,
- 2) Se refiere al problema puntual de articulación conducta robótica versus control humano.

Este esquema tiene la ventaja de respetar la *autodeterminación* de cada Estado miembro y no miembro, por lo que se postula como la llave válida para acceder a instancias internacionales gradualmente, no por decreto. La construcción progresiva con esta modalidad se muestra respetuosa de las éticas y morales diversas.

En contraste, Unesco sigue un esquema inverso: que las estructuras de gobierno avancen hacia los cambios estructurales necesarios a fin de adecuarse al marco determinado en la Recomendación:

Recomienda que los Estados miembro apliquen las provisiones de esta Recomendación realizando los pasos apropiados, incluyendo cualquier legislación o medida que sea requerida, en conformidad con la práctica constitucional y las estructuras de gobierno de cada Estado, para dar efecto a los principios y normas de la Recomendación dentro de las jurisdicciones.

En el ítem 62, la recomendación es más explícita en su llamado a modificar la legislación de los países, lo que contradice el dicho anterior de “en conformidad con la práctica constitucional”. La modificación de las cartas magnas implica un cambio no emergente del derecho soberano de cada Estado, sino desde una ética global de la IA resultante de los procesos de investigación que a su vez se restringirán con las recomendaciones del documento Unesco (ver 5. Un problema de científico) y se vigilarán desde un esquema centralizado de información (ver 3. Un problema de conocimiento).

²⁷²N. Arcifa (Universidad de Pisa-Italia). Disertación: Inteligencia Artificial o derechos humanos? Qué es más necesario? Tema: Armas letales autónomas (*autonomous weapons*). Ciclo de ética en Sistemas Inteligentes de la Academia Nacional de Ciencias de Buenos Aires. 2020.

8. Un problema jurídico: ¿Cambios constitucionales?

El peligro de sobreponerse a la ética del emergente cultural diverso es que las mismas contemplan realidades alternativas difícilmente contemplables desde un único punto. La bioética, por ejemplo, una de las impactadas por esta recomendación, ha sido sujeto de una diversidad de estudios que se renuevan permanentemente. Sólo en Latinoamérica presenta un complejo que incluye por ejemplo²⁷³:

- La “Declaración Universal sobre Bioética y Derechos Humanos” (que sitúa la bioética como una práctica multidisciplinaria y pluralista que se ejerce en el marco de una realidad humana culturalmente plural, socialmente diversa y económicamente desigual)
- La bioética-normativa de Juan Carlos Tealdi (de hecho, denuncia una bioética neoliberal y su pretensión fáctica de convertirse en bioética global)
- La bioética fuerte o bioética dura (que objetivan y contextualizan los problemas a las realidades concretas de Latinoamérica)
- La bioética de intervención (para cubrir las falencias relacionadas con los análisis contextualizados de conflictos, con adecuación a la cultura LA, e incorporación de macro-problemas regionales de alta nivel de exclusión social)
- La bioética de los Derechos Humanos (responde al fundamentalismo de los principios éticos y al imperialismo moral ejercido en su nombre, en particular en América Latina)
- La ética empresarial (respaldado por el de la sustentabilidad de ciertas prácticas agrícolas e industriales, no cuestiona los factores económicos y políticos que producen)

E. Suñé puntualiza esta misma complejidad en su disertación ante la Academia²⁷⁴:

[0:37:50- 0:38:11] *...Por lo tanto, para la adopción de decisiones de carácter inmediato, creo que, al no ser posible definir un sistema ético como universal, habrá que ir en un sistema más basado sobre la pragmática y la propia jurisprudencia...*

Desde el punto de vista jurídico, la diversidad ética no presenta un obstáculo para la resolución de conflictos. Como indica Arcifa²⁷⁵:

²⁷³Tealdi, *Diccionario latinoamericano de bioética*, Unesco, Red Latinoamericana y del Caribe de bioética, Universidad Nacional de Colombia, 2008.

²⁷⁴Suñé Llinás, *Un “decision making” ético para sistemas inteligentes*. Seratti, *Hacia una racionalización del alcance de la algorítmica de la IA y de los desafíos éticos y regulatorios impuestos por su irrupción sin límite ni fronteras*, Ciclo de ética en Sistemas Inteligentes de la Academia Nacional de Ciencias de Buenos Aires, 2020.

²⁷⁵N. Arcifa (Universidad de Pisa-Italia). Disertación: *Inteligencia Artificial o derechos humanos? Qué es más necesario? Tema: Armas letales autónomas (autonomous weapons)*. Ciclo de ética en Sistemas Inteligentes de la Academia Nacional de Ciencias de Buenos Aires. 2020.

[0:22:48 – 0:23:19] *Estas normas éticas hoy no existen, los estados para remediar esta falta han hecho recurso al art. 36 del primer protocolo adicional, de la Convención de Ginebra, que exige a los Estados adherentes que se encarguen del estudio del desarrollo, de la adopción de nuevas armas y de determinar en qué circunstancias pueden ser admitidos o prohibido.*

Con lo que ya existen regulaciones que remiten al criterio diverso de cada Estado. Arcifa boga también por el establecimiento de una ética global relacionada, emergente de las condiciones actuales jurídicas y los acuerdos entre Estados. Esto mismo es mantenido por la ingeniera Lidia Seratti, consultora Experta en Tecnologías de la Información y las Comunicaciones, de la Subsecretaría de Estudios y Prospectiva, Ministerio de Ciencia, Tecnología e Innovación²⁷⁶:

[1: 24:10 – 1: 24:37] *La labor legislativa y regulatoria puede poner parámetros en los tratados y resoluciones de la Unión Europea que permitan ir haciendo un cuerpo realmente, esto es planetario, para hacer una regulación.*

Si bien el ítem 59 de la Recomendación alienta a Estados, privados y civiles a crear una red de responsables para apoyar el proceso de implementación de la propia Recomendación, está claro que la unificación desde lo jurídico es un tanto más compleja ya que las cartas magnas responden a sistemas diferentes. Suñé aporta una base inicial de compatibilidad entre América Latina y Europa²⁷⁷:

[0:38:35- 0:39:04] *Un sistema jurídico y jurisprudente de carácter romano-germánico- napoleónico, lo tenemos en la Unión Europea. Pero también lo tienen en América Latina, en Iberoamérica. Que es el mismo sistema jurídico y, en consecuencia, las soluciones no serían muy distintas...*

Es de notar que Suñé destaca las cercanías solo entre estos sistemas, pero nuevamente, su postulación de resolución desde una propuesta jurídica global implica la omisión de casi medio planeta. Por lo explicado en la sección “7. Un problema legal” resulta entonces que la iniciativa debe comenzar desde los acuerdos entre Estados.

²⁷⁶Suñé Llinás, *Un “decision making” ético para sistemas inteligentes*. Seratti, *Hacia una racionalización del alcance de la algorítmica de la IA y de los desafíos éticos y regulatorios impuestos por su irrupción sin límite ni fronteras*, Ciclo de ética en Sistemas Inteligentes de la Academia Nacional de Ciencias de Buenos Aires, 2020.

²⁷⁷Suñé Llinás, *Un “decision making” ético para sistemas inteligentes*. Seratti, *Hacia una racionalización del alcance de la algorítmica de la IA y de los desafíos éticos y regulatorios impuestos por su irrupción sin límite ni fronteras*, Ciclo de ética en Sistemas Inteligentes de la Academia Nacional de Ciencias de Buenos Aires, 2020.

9. ¿Cómo engloba la ética a la IA?

A. Ética de IA en la Ética global

La comunidad internacional se plantea cómo encaja la IA en la ética. Respecto a la ética en armas letales autónomas Arcifa expresa²⁷⁸:

[0:25:23 – 0:26:44] *En cuanto a los argumentos éticos, que dicen a favor o contra estos sistemas de arma autónoma... Las dos teorías dicen que los argumentos a favor o contra, son aquellos que remiten a la ética consecuencialista y la deontológica. La ética consecuencialista se centra en los resultados y por lo tanto promueve la autonomía de las máquinas para obtener mejores prestaciones. También dice que con IA habrá un derecho internacional y valores éticos internacionales. La deontológica se focaliza en el proceso y por lo tanto se focaliza en los valores morales como guía de la acción son parámetro de juicio de las opciones individuales y colectivos.*

Desde el punto de partida, hay argumentos opuestos en ambos sentidos: la ética consecuencialista o la deontológica²⁷⁹. Pero parece que esta dualidad, presente en la comunidad internacional no se refleja en la Recomendación, donde adopta la opción deontológica, con su art. 24:

Los sistemas de IA pueden ser investigados, diseñados, desarrollados, producidos o usados para asistir en interacciones que involucren personas vulnerables, incluyendo pero no limitándose a los niños, ancianos o insanas, pero no deben jamás cosificar al humano, minar la dignidad humana o violar o abusar los derechos humanos.

También se puede evaluar esta posición como un resabio del positivismo, ya que desde la escuela de Frankfurt ya se advierte sobre la des-humanización de la técnica²⁸⁰, específicamente la crítica de Horkheimer y Adorno indican, por ejemplo, que el cientificismo positivista y tecnologismo asociado derivaría en una técnica totalitaria. Husserl ya previamente habla al respecto y se centra en que la extensión del patrón positivista e instrumental a todos los órdenes de la vida humana deriva en la pérdida del sentido.

J. E. Rulli afirma también sobre la des-humanización, que “es la pérdida de la capacidad axiológica, tan propiamente humana, cuando estamos rodeados de tanto poderío científico y tecnológico”²⁸¹. El encuentro y confrontación entre norma jurídica y norma ética, será una condición necesaria, aunque no suficiente para constituir el concepto de derechos humanos. Los derechos humanos expresan en su origen una

²⁷⁸N. Arcifa (Universidad de Pisa-Italia). Disertación: Inteligencia Artificial o derechos humanos? Qué es más necesario? Tema: Armas letales autónomas (*autonomous weapons*). Ciclo de ética en Sistemas Inteligentes de la Academia Nacional de Ciencias de Buenos Aires. 2020.

²⁷⁹*Ética consecuencialista*: Es aquella que determina la bondad o maldad de un acto moral y de su norma a través de las consecuencias que ofrecen.

Ética deontológica: Deviene de Kant. Es una ética del deber, enfocada en pautas de comportamiento a seguir, o leyes que regulan al ciudadano(ver 22).

²⁸⁰Chillón - Marcos, *Técnica y sentido*, “SCIO. Revista de Filosofía”, n° 11, 2015, p. 77 a 99.

²⁸¹Tealdi, *Diccionario latinoamericano de bioética*, Unesco, Red Latinoamericana y del Caribe de bioética, Universidad Nacional de Colombia, 2008.

apreciación de valores, en particular de aquellos que tienen mayor altura o densidad o peso, esto es, mayor jerarquía.

Esta des-humanización, se combina con la antropomorfización de los sistemas técnicos, desde el momento que el ítem 63 insinúa la posibilidad de usar la IA para vigilar los derechos humanos. Esta peligrosa delegación, acaso significa el declarado “exceso de tecnificación” de Heidegger²⁸², para quien la tecnificación y la huida de la necesidad de pensar están mutuamente implicados.

B. Ética de IA y los derechos humanos

Chillón y Marcos rematan su postura sugiriendo adoptar la postura de Heidegger, en su relación del ser con la técnica, de serenidad. Advierte también que la fuerte des-humanización tiene irónicamente como verdugo y víctima al propio humano. ¿Cómo hay que situarse ante la técnica? Para Heidegger: Con una actitud crítica (*Gelassenheit*) que evoca, el misterio del ser en el que debe contextualizarse la cuestión de la técnica.

En la Recomendación es recurrente a lo largo de todo el documento la relación entre la ética de la IA y los derechos humanos. El tema, tiene especial relación no solo con la posibilidad de cosificación del humano (art. 24) sino también en la seguridad física.

En el art. 71 la Recomendación establece la enseñanza como primer instrumento para mitigar el problema:

Los Estados miembro deben promover las iniciativas de investigación sobre el uso de IA en enseñanza, capacitación docente y e-learning, entre otros tópicos, en una manera que genere oportunidades y *mitigue los desafíos y riesgos* asociados con estas tecnologías.

Pero es claro que el conocimiento no erradica per se el problema. Los riesgos y desafíos se generan en un contexto más complejo. Alfredo Marcos²⁸³ hace un paralelo interesante con los períodos de innovación referidos por Kuhn en 2017, donde la innovación en ciertos momentos cae dentro del marco conceptual (lo esperable, la expectativa previa) y otros momentos donde la innovación sale del mismo, y exige un cambio de paradigmas, que es tomado por las ciencias.

Esto iría en concordancia con el concepto de que los cambios son guiados por un tipo de saber (no son azarosos), y que remite al concepto de *phrónesis*²⁸⁴ de Aristóteles. Este saber, guía al cambio desde lo práctico y lo social.

La evolución, desde lo práctico, está mostrando su lado oscuro, que es el denunciado por la Recomendación. Arcifa, en su presentación a la Academia Nacional de Ciencias de Buenos Aires, presenta un video impactante²⁸⁵:

²⁸²Chillón - Marcos, *Técnica y sentido*, “SCIO. Revista de Filosofía”, n° 11, 2015, p. 77 a 99.

²⁸³Marcos, *Información e inteligencia artificial*, Universidad de Valladolid, 2020.

²⁸⁴*Phronesis* o *Frónesis* es pensable como sabiduría práctica o prudencia.

[0:13:02 – 0:13:08] VIDEO dentro de la presentación de una demostración de un dron asesino de los Estados Unidos (aclara que no está activo pero existe). El video muestra a una persona de seguridad monitoreando a un dron para atravesar el cráneo de un maniquí en segundos.

El video concluye con el presentador afirmando que se trata de una “matanza limpia”, ya que los humanos pueden no obedecer, pero un dron Sí. La disertación prosigue con un video de un militar admitiendo que se está invirtiendo mucho (heavily) en Proyectos clasificados²⁸⁶.

En los arts. 27 y 28 la Recomendación refiere a las responsabilidades éticas y legales por la gestión de un IA que deben recaer en los humanos involucrados, a los que denomina agentes. Por tanto, los derechos humanos deben primero ser garantizados por éstos.

Villalba Gómez expresa este tipo de miedo al desborde de la IA²⁸⁷. Define el concepto de singularidad tecnológica como análogo a la singularidad matemática (en el sentido que trasciende cualquier limitación finita).

Para Chillón y Marcos, en cambio el problema deviene del proceso de delegación progresivo en la tecnología²⁸⁸. Conciben la técnica como medio de humanización, y plantean que una antropotecnia (modificación de la naturaleza humana) pone en riesgo la continuidad/sentido de la técnica misma y esto derivará en la *disolución de la ética* y de la razón práctica.

El riesgo se origina en el mismo actor que lo recibe: el propio humano. En palabras de Seratti²⁸⁹ el ciudadano medio tiene un problema por delante:

[1:29:56 – 1:30:05] *Hay muy poco acercamiento a lo digital como para pretender que de golpe la sociedad se pueda adaptar a esto.*

Para Ferreri, académico presidente de la Academia Nacional de Ciencias de Buenos Aires, parte del problema es un proceso educativo que cambie la mentalidad de la gente:

[1:33:37- 1:33:53] *...la concientización necesaria como para saber que la ética está implícita en las personas que desarrollan las cosas y las implementan, y por lo menos que*

²⁸⁵N. Arcifa (Universidad de Pisa-Italia). Disertación: Inteligencia Artificial o derechos humanos? Qué es más necesario? Tema: Armas letales autónomas (*autonomous weapons*). Ciclo de ética en Sistemas Intelgentes de la Academia Nacional de Ciencias de Buenos Aires. 2020.

²⁸⁶Micro Drones killer arms robots, autonomous Artificial Intelligence, Warning!! www.youtube.com/watch?v=TIO2gcs1YvM. 2017.

²⁸⁷Villalba Gómez, *Problemas bioéticos emergentes de la inteligencia artificial*, “Revista Diversitas. Perspectivas en psicología”, vol. 12, n° 1, 2016.

²⁸⁸Chillón - Marcos, *Técnica y sentido*, “SCIO. Revista de Filosofía”, n° 11, 2015.

²⁸⁹Suñé Llinás, *Un “decision making” ético para sistemas inteligentes*. Seratti, *Hacia una racionalización del alcance de la algorítmica de la IA y de los desafíos éticos y regulatorios impuestos por su irrupción sin límite ni fronteras*, Ciclo de ética en Sistemas Intelgentes de la Academia Nacional de Ciencias de Buenos Aires, 2020.

quedará en claro que 'no hay máquinas perversas, sino hay perversos que implementan máquinas'.

Esto se plantea también en el apartado 32 de la encuesta, mencionado como concientización pública y entendimiento de las tecnologías de IA, y lo plantea para todos los niveles educativos.

El proceso educativo debe complementarse con los cambios desde el complejo cultural, en otras palabras, debe incorporar la IA.

Para Hall serán “los significados y los valores que emergen entre grupos y clases sociales diferenciados, sobre la base de sus condiciones y relaciones históricas dadas, a través de los cuales [estos grupos y clases] ‘manejan’ y responden a las condiciones de existencia; y como las tradiciones y prácticas vividas a través de las cuales son expresadas esas ‘comprensiones’, y en las cuales están encarnadas”²⁹⁰.

Para Vico el desarrollo es al mismo tiempo, e interactivamente, la configuración de las sociedades y de las mentes humanas²⁹¹.

Para Herder ya no se trata de un desarrollo progresivo unilineal y existe la necesidad de hablar de “culturas” más que de cultura, por lo que se debe aceptar la variabilidad y complejidad de las fuerzas que la conforman.

10. ¿Puede la IA tener una ética?

Nick Bostrom (2014) postula que debe construirse una “ética de la IA”. La justificación se supone en la capacidad de crear máquinas pensantes dentro de un campo ético. Desde ese punto determina a los agentes que construyan y manipulan la IA, como los responsables de calificar y evaluar implicaciones éticas en relación con las condiciones humanas²⁹².

En lo social establece las categorías bio-éticas siguientes: creación, diseño, construcción y aplicación de la IA. Dichas categorizaciones sirven para poder juzgar los algoritmos mediante funciones sociales tales como: responsabilidad, transparencia, auditabilidad, incorruptibilidad, predictabilidad, etcétera.

En este mismo marco se presentan las concepciones del documento Unesco. Sin embargo, otras posiciones como Kosman (1987) sostienen que la IA no puede tener ética: “los artefactos en la tradición aristotélica, son sustancias solo en un sentido accidental” y por lo tanto eso los diferencia de los seres humanos y de hecho de los seres vivos en general, que son sustancias en sentido propio y paradigmático.

²⁹⁰ Hall, Stuart, *Codificar/decodificar*, Culture, Media, Language, Working Papers in Cultural Studies, 1979.

²⁹¹ Vico, B. I. - Herder, J., *Two Studies in the History of Ideas*, 1976.

²⁹² Villalba Gómez, *Problemas bioéticos emergentes de la inteligencia artificial*, “Revista Diversitas. Perspectivas en psicología”, vol. 12, n° 1, 2016.

Esta diferencia ontológica es la que vuelve nula la eliminación de dicha distancia basada en la sofisticación tecnológica y de hecho la clasifica en un error categorial.

Sin embargo, el catedrático Martino, de la Academia Nacional de Ciencias y miembro asociado del Center for the AI and Cognitive learning de la Universidad de Greenwich, declarado “personalidad destacada de la Ciudad Autónoma de Buenos Aires en el ámbito de las ciencias jurídicas”²⁹³:

[0:48:36 – 0:49:08] *Si bien se habla de cumplir las 4 leyes de los robots. pero la ética es para los seres humanos.*

[0:51:28 – 0:51:50] [...] *no hay moral de robots [...]*

[0:56:19 – 0:56:25] [...] *un robot no es sujeto de derecho y esto es consensuado.*

En el mismo sentido Suñé declara:

[0:16:50 – 0:17:31] *Una ética para sistemas inteligentes... complicado por la dificultad de establecer un sistema de valores y tomar decisiones la máquina a partir de ellos... y en la medida en que pudiera establecerlos, adaptarlos para la toma de decisiones concretas.*

El art. 88 de la Recomendación reclama el desarrollo de guías adecuadas que controlen la adquisición progresiva de ética por parte de la IA. Pero esta afirmación, desplaza la responsabilidad inicial de los humanos, y prevé posibilidad de que los humanos cedan progresivamente la ética a los seres artificiales.

Hay indicios de cierta confusión y hasta prejuicios a la hora de evaluar los riesgos. Por caso la anécdota del doctor Martino:

[0:19:14 – 0:19:22 / 0:19:31 – 0:19:38 / 0:19:49 – 0:20:14] *Creía que no hacía falta darle explicaciones a los robots del por qué interpretaba las cosas como humano y ser social, pero luego un técnico explicó que esa explicación es el rastro guardado de un algoritmo para luego saber qué arreglar si falla.*

[0:15:22 – 0:25:31 // 0:25:49 – 0:25:55 // 0:26:16- 0:26:19] *Actuamos con sesgos cognitivos... tenemos prejuicios ... cuando hacemos programas transferimos esos prejuicios.*

11. ¿Qué implica la posibilidad de una ética global?

Unesco indica que los valores inspiran una buena moral y ésta determina los principios basados en políticas y acciones. Entre los valores que exige a la IA:

- 1) Los sistemas inteligentes deben respetar la dignidad humana.
- 2) Su actividad no debe dejar a nadie atrás: a todos los humanos se nos tratará por igual, considerando las diferencias de cualquier tipo.
- 3) Respetarán que los humanos vivimos en armonía: todos estamos interconectados y dañar a uno es dañar a todos.

²⁹³ Martino, Disertación: ¿por qué la ética establece criterios para AI? Tema: implicancias éticas en la robótica y la IA, Ciclo de ética en Sistemas Inteligentes de la Academia Nacional de Ciencias de Buenos Aires, 2020.

4) Deben demostrar al humano coherencia siendo íntegros en su confiabilidad.

5) Deben proteger el medio ambiente.

Los principios que deben seguir la IA se atacan dos aspectos:

1. la interfase humano-tecnología
2. las características que deben cumplir según las expectativas internacionalmente aceptadas de lo que es el comportamiento ético

Abrirse a cualquiera de los dos aspectos involucra avanzar en los miedos, prejuicios, establecer de manera consensuada y progresiva. Aceptar acercarse de a poco al problema desde lo plural en los diversos ámbitos de la cultura. Tal como el video de Martino:

[0:5:38 – 0:5:57] *El temor a abrir la caja de Pandora, y cuando lo haces, descubres que no es lo que crees: se clama por desmitificar la IA.*

Analizar la IA desde cualquier aspecto implica comenzar por evaluar si no estamos careciendo de los elementos de análisis a priori, que nos permita ver el bosque y no sólo el árbol. Esta posición incómoda que pretende incorporar lo que consideramos visiones alternativas a las que nos oponemos, valoraciones que puedan contraponerse pero, en todo caso, significa abrirse a otros elementos de juicio que podrán o no gustar, pero están ahí, representando parte de la diversidad de posiciones de quienes conformarán, en última instancia, la ética global pretendida. Para ello hay que abrir la “caja de Pandora”.

Por otro lado, los avances efectivos actualmente más aparecen en lo jurídico. Desde el punto de vista de los especialistas no hay que cambiar la ética, hay que aplicarla y establecer una vez más las reglas jurídicas que sigan haciéndola visible. Como apunta Arcifa:

[0:16:22 – 0:16:42] *se está buscando crear un marco regulador de referencia sobre el impacto de los sistemas autónomos. Sobre todo la legislación sobre la base que dirima la aceptabilidad ético-jurídica en términos del derecho internacional humanitario.*

Funtowicz e Hidago evalúan el problema de la indeterminación desde la regulación ética y evaluación de riesgo²⁹⁴. Al estudiar el modelo de encuadre, como una de las posibilidades alternativas del modelo precautorio y del modelo de demarcación, curiosamente refiere a los siguientes dichos de la misma Comisión Europea²⁹⁵ afirma:

La comisión podría encontrarse ante una gama de dictámenes contradictorios de expertos del mundo académico, de personas con conocimientos prácticos en una materia determinada y de otras con intereses directos en la cuestión. Estas opiniones pueden basarse en presupuestos completamente diferentes y tener objetivos total-

²⁹⁴ Funtowicz, Silvio - Hidalgo, Cecilia, en López Cerezo, José - Gomez González, Francisco (coords.), *Apropiación social de la ciencia*, Biblioteca Nueva, 2008.

²⁹⁵ Communication from the communities on the collection and use of expertise by the commission: principle and guidelines (COM 2002: 713). Commission of the European Communities, Bruselas, 2002.

mente distintos...Por tanto, cada vez más, la interacción entre los responsables políticos, los expertos, las partes interesadas y el público en general es una parte crucial de la elaboración de políticas y, por eso, hay que prestar atención no sólo al resultado de las políticas sino también al proceso seguido.

En una clara demostración de que la comunidad produce distintos marcos alternativos al encarar un problema y por lo tanto el marco mismo se vuelve parte del debate entre grupos de interés. El resultado es que los expertos pueden incurrir en contradicciones que deben ser dirimidas en decisiones políticas previas.

Según los autores: “Cualquier parte que resulte considerada la ‘dueña’ del problema hará la contribución más importante y disfrutará de los mayores beneficios”.

La solución que propone la propia Comisión de la Unión Europea es una llamada a la pluralidad de perspectivas:

“el factor determinante final de la calidad es el pluralismo. Siempre que sea posible, deben solicitarse puntos de vista diversos... También pueden ser importantes otros factores, como diferentes perspectivas geográficas culturales y de sexo”.

Pero nuevamente, el problema es que estamos ante un modelo de demarcación y no de encuadre donde los autores bien determinan que hace falta determinar los valores que deben ejercerse para poder luego prescribir. El problema es, como afirman los autores: “El problema principal es que ya no resulta funcional excepto en casos de corrupción claramente delimitados” puesto que “la filosofía de la ciencia pos-empirista ha mostrado que, en general, dadas la complejidad y la indeterminación de los sistemas emergentes, es imposible una separación total entre los hechos y los valores”.

Esto es que la polarización de la situación y resulta en extremo difícil establecer una separación impermeable entre la evaluación y la gestión del riesgo, y como concluye, los grupos de interés accionan de expertos y nuevamente no se es totalmente imparcial.

Conclusiones

Desde la perspectiva de las políticas científicas modernas, la indeterminación que subyace en la regulación ética y evaluación de riesgo hace difícil el punto de partida para un trabajo global ético. La Unesco podría encuadrarse en un “modelo de encuadre” al hacer este referéndum a “expertos” para plantear el marco de valores a considerar. El resultado es que los expertos pueden incurrir en contradicciones que deben ser dirimidas en decisiones políticas previas.

Es importante el proceso de contribución de los Estados miembros ya que hay riesgo de que existan desbalances en las contribuciones y por lo tanto de los beneficios.

Considerando Funtowicz e Hidalgo se aplica un modelo de demarcación (es decir de determinación de prescripciones), con una gran complejidad e indeterminación de los sistemas emergentes, por lo que resulta imposible una separación total entre los hechos y los valores.



La polarización de la situación hace en extremo difícil establecer una separación impermeable entre la evaluación y la gestión del riesgo, y como concluyen, los grupos de interés accionan de expertos y nuevamente no se es totalmente imparcial.

Otra cuestión que surge es que la Recomendación de Unesco epistemológicamente torna invisible la relevancia que los valores políticos tienen en las elecciones de los científicos, haciendo hincapié en lo metodológico y los procesos de gobernanza de los criterios de relevancia.

Debemos estar atentos entonces, adoptar la posición de Heidegger, respecto a los procesos políticos e intereses que están jugando en estos momentos los verdaderos movimientos detrás de todos recursos aparentemente democráticos. Observar y prepararse para ejercer nuestros derechos como sociedad a ser escuchados ante aquellas determinaciones que nos pongan en desventaja o en peligro.

La visión de riesgo y de falta de control sobre el proceso de evolución de la tecnología (y en este caso de la IA), fue abordado por Simondon desde una perspectiva donde se demarca el sesgo cognitivo subyacente.



Apéndice 1. Textos interesantes desde Encuesta Unesco

Elementos extraídos de la versión enviada inicialmente a los expertos en IA

Este apéndice contiene una serie de extractos que fueron de interés a los autores en la gestación de la fundamentación del contenido textual.

Cada uno es referenciado convenientemente a la página correspondiente del documento y comentado o analizado según corresponda. En ocasiones hay referencias a otras partes del mismo texto. A fin de diferenciar los aportes de los originales transcritos se emplea itálicas para diferenciarlos.

Las porciones de texto marcadas con (§) originalmente estaba en la encuesta pero luego fue removida de la versión final. Se incluye a fin de mostrar la concepción subyacente en la propuesta.

(pp.1) [...]Reflecting on the profound influence that Artificial Intelligence (AI) may have on societies, ecosystems, and human lives, including the human mind, in part because of the new ways in which it influences human thinking and decision-making, and affects education, science, culture, and communication and information.

(pp.1) Considering that AI systems can be of great service to humanity but also raise fundamental ethical concerns, for instance regarding the biases they can embed and exacerbate, potentially resulting in inequality, exclusion and a threat to cultural and social diversity and gender equality; the need for transparency and understandability of the workings of algorithms and the data with which they have been trained; and their potential impact on privacy, freedom of speech, social, economic and political processes, and the environment, [...]

→Refutamos una idea reduccionista: los algoritmos cruzan todas las otras ciencias (ver Apéndice 5 Estudio de filosofía de la biología)

→ Estamos ante la posición de que la ciencia y la ética van de la mano (ítem 4 clase 1) donde la ética es parte del individuo y existen distintas éticas para cada grupo (distintas reflexiones sobre la moral, siendo la moral un conjunto de valores). Por lo que se interpreta que la ética es parte del proceso de generación de las ciencias. Ver (***)

(pp.1) [...] Conscious of the fact that low and middle income countries (LMICs), including but not limited to those in Africa, Latin America and the Caribbean, and Central Asia, as well as Small Island Developing States, are facing an acceleration of the use of information technologies and AI and that the digital economy presents important societal challenges and opportunities for creative societies, requiring the recognition of endogenous cultures, values and knowledge in order to develop economies, [...]

→Cabe decir que estamos “curando” el error tradicional de subestimación por parte de la ciencia hacia hechos nuevos? ¿O estamos nuevamente prediciendo en base a lo conocido?

Reconocemos los cambios de IA en soc, usos tierra, nutrición, sociedad...²⁹⁶

¿Unimos la ciencia fragmentada [con esta encuesta y reunión París 2021]?

Es de destacar que se apoya en la desigualdad económica y cultural en la desigualdad sanitaria entre países de referencia.

(pp.1) *Noting that addressing risks and ethical concerns should not hamper innovation but rather stimulate new practices of responsible research and innovation in which the research, design, development, deployment, and use of AI is anchored in moral values and ethical reflection,[...]*

→ ¿Estamos ante el pedido de la constitución de una filosofía de los sistemas inteligentes? Ver (**)

(pp.1) *Observing that a normative framework for AI and its social implications finds itself at the intersection of ethics, human rights, international and national legal frameworks, and the freedom of research and innovation, and human well-being, [...]*

→pero la ética no tiene que ver con las leyes. La interpretación de la ley tiene que ver con la moral

(pp.4) *Taking fully into account that the rapid development of AI systems encounters **barriers** to understand and implement AI, **because of the diversity of ethical orientations and cultures** around the World, the **lack of agility of the law** in relation to technology and the information society, and the risk that local and regional ethical standards and values be disrupted by AI,[...]*

(pp.42) **Recommends** that Member States apply the provisions of this Recommendation by taking appropriate steps, including **whatever legislative** or other measures may be required, in conformity with the constitutional practice and **governing structures of each State**, to give effect within their jurisdictions to the principles and norms of the Recommendation;

§ (pp.2) *In all fields of the sciences, social sciences and humanities, AI has implications for our concepts of scientific understanding and explanation, and for the ways in which scientific knowledge can be applied as a basis for decision- making.*

²⁹⁶Linares, J., *Ética y mundo tecnológico*, Biblioteca de ética y bioética, Universidad Nacional Autónoma de México (Facultad de Filosofía y Letras) y Fondo de Cultura Económica, 2000.

§ (pp.1) *Even though this Recommendation is addressed primarily to policy-makers in and outside Unesco Member States, it also aims to provide a framework for international organizations, national and transnational corporations, NGO's, engineers and scientists, including representatives of humanities, natural and social sciences, non-governmental organizations, **religious** organizations, and civil society, stimulating a multi-stakeholder approach, grounded in a globally accepted ethical framework that enables stakeholders to collaborate and take common responsibility based on a global, intercultural dialogue.*

Valores de la IA

§ (pp.1) 8. *Values thus inspire **good** moral behaviour in line with the international community's understanding of such behaviour and they are the **foundations of principles**, while principles unpack the values underlying them more concretely so that values can be more easily actualised in policy statements and actions.*

(05) *Living in harmony*

14. *The value of living in harmony points to the research, design, development, deployment, and use of AI systems recognising the interconnectedness of all humans. The notion of **being interconnected** is based on the knowledge that every human belongs to a greater whole, which is **diminished when others are diminished** in any way.*

→interesante: ¿Estamos ante un posible concepto de ecosistema humano extrapolado?

(§)16. *AI systems should be trustworthy. Trustworthiness is a socio-technical concept implying that the research, design, development, deployment, and use of AI systems should inspire, instead of infringing on, trust among people and in AI systems.*

(§)17. *Trust has to be earned in each use context and more broadly is a benchmark for the social acceptance of AI systems. Therefore people should have good reason to trust that AI technology brings benefits while adequate measures are taken to mitigate risks.*

→interesante: Aplica una virtud moral a la IA

(pp.6) *The definition of ‘quality of life’ should be left open to individuals or groups, as long as no human being is harmed physically or mentally, or their dignity diminished in terms of this definition*

→interesante: limita la IA en cuanto a determinar por encima de humanos?

(pp.7) 24. *AI systems may be researched, designed, developed, deployed or used to assist in interactions involving vulnerable people, including, but not limited to children, the elderly or the ill, **but should never objectify humans** or undermine human dignity, or violate or abuse human rights.*

Principios

(pp.7) 25. *The research, design, development, deployment, and use of AI systems may **not exceed what is necessary** to achieve legitimate aims or objectives and should be appropriate to the context.*

→interesante: Limita la cantidad de investigación y desarrollo a lo “necesario”

(pp. 6) *Human oversight and determination*

27. *It should always be possible to attribute both ethical and legal responsibility for the research, design, development, deployment, and use of AI systems to a physical person or to an existing legal entity. Human oversight refers thus not only to individual human oversight, but to public oversight.*

28. *It may be the case that sometimes humans would have to share control with AI systems for reasons of efficacy, but this decision to cede control in limited contexts remains that of humans, as AI systems should be researched, designed, developed, deployed, and used to assist humans in decision-making and acting, but never to replace ultimate human responsibility.*

→invierte ahora la carga: la responsabilidad es siempre del humano. Es un viraje interesante que no mantiene.

(pp.6) *Awareness and literacy*

32. ***Public awareness and understanding of AI technologies and the value of data should be promoted through education, public campaigns and training to ensure effective public participation so that citizens can take informed decisions about their use of AI systems. Children should be protected from reasonably foreseeable harms arising from AI systems, should have access to such systems through education and***

training, and children should not be disempowered by their interaction with AI systems.

Cuerpo general

(pp.8) 67. *Member States should encourage in accordance with their national education programmes and traditions the embedding of **AI ethics** into the school and university curricula for all levels and promote cross collaboration between technical skills and social sciences and humanities.*

→ Invierte la carga: antes la ética correspondía al actor, ¿y ahora hay que enseñar una ética de IA? Este tipo de inversiones manifiesta la dualidad del mensaje de acuerdo a la argumentación que desea imponerse.

(pp.8) 71. *Member States should encourage research initiatives on the use of AI in teaching, teacher training and e-learning, among other topics, in a way that enhances opportunities and **mitigates the challenges and risks** associated with these technologies.*

→ deposita riesgo en el conocimiento. Es una manera de justificar el panóptico de información que luego sugiere imponer a nivel mundial.

(pp.8) 72. *Member States should support collaboration agreements between academic institutions and the industry to bridge the gap of skillset requirements and promote collaborations between industry sectors, academia, civil society, and the government to align training programs and strategies provided by educational institutions, with the needs of the industry. **Project-based learning** approaches for AI should be promoted, allowing for partnerships between companies, universities and research centers.*

→ El método de enseñanza cambia (es decir todo otro conocimiento) en función de la industria! Esto se opone a miradas como la de Simondon y la necesidad de abandonar la posición de hombre-máquina que tradicionalmente ocupa el hombre en sus tareas desde la revolución industrial.

(pp.8) *Promoting AI Ethics Research*

74. *Member States should promote AI ethics research either through direct investments or by creating incentives for the public and private sectors to invest in this area.*

→(**) En sincronía con esta parte: Hay un pedido de “filosofía de IA”

(pp.8) 78. *Member States and funding bodies should promote interdisciplinary AI research by including disciplines other than science, technology, engineering, and mathematics (STEM), e.g. law, international relations, political sciences, education, philosophy, culture, and linguistic studies to ensure a critical approach to AI research and proper monitoring of possible misuses or adverse effects.*

→(***) En sincronía con esta parte.

(pp.9) 88. *Member States should encourage development and use of comparable AI guidelines, including ethical aspects at global and regional levels, and gather the required evidence to evaluate, monitor and control the progression in the ethical implementation of AI systems*²⁹⁷.

Ítems extraídos de la versión web

En esta parte las citas de interés son las del texto que se publica en la web por parte de Unesco. No coincide con su contraparte original por lo que estas referencias son adicionales a las anteriores. En contraste con la sección previa, los ítems no son transcritos sino solo identificados y comentados los aspectos pertinentes al análisis de este documento.

Item 62: pretensión de modificar la legislación de los países desde una ética global.

Item 63: insinúa la posibilidad de usar la IA para vigilar los derechos humanos, establecer una policía digital.

Item 78: proponen que los Estados enfoquen un patrimonio común de datos de IA.

NOTA: Si el patrimonio es central y mundial estamos ante un gran poder central mundial y hegemónico en manos de alguien.

Item 100: los Estados miembros determinan quienes son los derechos de propiedad intelectual sobre los productos de IA.

Item 134: los mecanismos de seguimiento y evaluación un observatorio de ética de la IA (¿Es la implementación de un panóptico global?).

²⁹⁷Material compatible con Communication from the communities on the collection and use of expertise by the commission: principle and guidelines (COM 2002: 713). Commission of the European Communities, Bruselas, 2002, p. 5, "91. Member States".



Item 48: las Unesco apoyar a los Estados en elaboración y seguimiento de los mecanismos de políticas.

NOTA: Así se introduce lo legal. ¿Es ético monitorear los Estados de derecho centralmente?

Item 59: alienta a Estados, privados y civiles a crear una red de responsables para apoyar este proceso de implementación.

Item 139: la sociedad civil deberá defender los intereses de sector público y por lo tanto Unesco debe garantizar y promover su legitimidad.

NOTA: Ellos se ponen del lado del civil que propugna apoyar la iniciativa, aunque los Estados no lo hagan. ¿Dónde quedan los derechos soberanos de cada Estado?



Apéndice 2. Video 1.Ciclo ética en IA - ANCB

Este apéndice pertenece al estudio de la disertación del académico profesor Martino, de la Academia Nacional de Ciencias de Buenos Aires, con el título: “¿Por qué la ética establece criterios para AI?”, en su desarrollo aborda el tema: Implicancias éticas en la robótica y la IA. La siguiente tabla es un resumen de los momentos considerados relevantes para el presente estudio, y además de la referencia al minuto exacto del extracto, se realiza una breve evaluación acerca del contenido.

[UBICACIÓN] COMENTARIO	TIP EVALUACIÓN
[4.52 - 5.12] el video define a IA como FENÓMENO EMERGENTE y peligroso	si es un fenómeno emergente entonces, ¿por qué exigirle ética?
[5.38 -5.57] video sobre el mito de la caja de Pandora	caja de Pandora es mito pues no es una caja: el video clama por desmitificar usando el conocimiento
[6.04 - 6.12] tomado de video: the age of AI. T1 E1. Cuán lejos es demasiado? “cuando creamos tecnología más potente, creamos una palanca mayor para cambiar el mundo	remite a Arquímedes “dadme un punto de apoyo y moveré al mundo” refiriendo a una palanca lo mismo es evidencia por Latour también
[14.39- 15.07] en el siglo pasado un ético inglés...deberes según ocasiones e informaciones que tenemos.	esto se opone a la idea de una ética consensuada mundial
[16.03-16.23]	ética entre kant y Ross: hay un mundo nuevo con situaciones donde deben tomarse decisiones globales
[18.12 - 18.39] no hay una ética para robots [ojo va contra las leyes de la robótica], siempre para los humanos. debe seguir con lo Kantiano, que efectos sean compatibles con la vida humana auténtica	ojo esto se lee como una de las preguntas más frecuentes en la comunidad
[19.14 - 19.22 /19.31 - 19.38/19.49 - 20.14]	muestra que es común el prejuicio personal como algo inadvertido, respecto al uso de directivas en robots.

[UBICACIÓN] COMENTARIO	TIP EVALUACIÓN
	<p>Creía que no hacía falta darle explicaciones a robot porque lo interpretaba como humano y ser social, no como técnico que está guardando el rastro de un algoritmo para luego saber qué arreglar si falla</p>
	<p>POR LO TANTO: hay distintas visiones de lo mismo</p>
	<p>DATOS que da en la disertación:</p> <ul style="list-style-type: none"> -En 2005. Declaración Universal sobre Bioética y Derechos Humanos[12]. -el 28 de febrero de 2020 se firmó en vaticano un documento que enfatiza los aspectos éticos de la IA y que coloca otra pieza en la base de la regulación: <ul style="list-style-type: none"> ojo esto explica por que Unesco menciona la religión??? Similitud con trabajos leídos en Apéndice 5: Estudio de Merton influencia en ciencias. -2019 de Directiva Europea "Directive (UE) 2019/771 del parlamento europeo y del consejo": <ul style="list-style-type: none"> determina aspectos de los contratos de compraventa de bienes. Modifica el reglamento (CE) nro. 2017/2394 y Directiva 2009/22/CE y se deroga la Directiva 1999/44/CE y que trata de las transacciones informáticas y los criterios éticos que deben respetar. -GRUPO INDEPENDIENTE DE EXPERTOS DE Alto nivel sobre inteligencia artificial <ul style="list-style-type: none"> (creado por la comisión europea en junio de 2018). "Directrices ética para una

[UBICACIÓN] COMENTARIO	TIP EVALUACIÓN
	<p>IA fiable"</p> <p>-42 países adoptan principios éticos de la OCDE sobre IA</p> <p>1-estar al servicio de personas /planeta</p> <p>2-respetar el Estado de derecho y derechos humanos, democráticos y diversidad. incorporado salvaguardias para garantizar una sociedad justa y equitativa</p> <p>3-transparencia y divulgación responsables para que sepan cuando les conviene oponerse a ciertas cosas</p> <p>4-funcionar de forma robusta, fiable y segura en su vida útil evaluando riesgos</p> <p>5-los agentes (personas o organizaciones responden por su correcto financiamiento</p> <p>6-Unesco esta gestionando una normativa global en dos etapas:</p> <p>1ro compulsa global</p> <p>2do revisarlo y presentarlo al mundo</p>
[15.22-25.31// 25.49-25.55//26.16-26.19]	Actuamos con sesgos cognitivos. tenemos prejuicios // cuando hacemos programas transferimos esos prejuicios
[48.10 - 48.31]	Robótica es la parte pequeña de IA. Pero está en la imaginación: todos temen a los robots
[48.36 - 49.08]	Cumplir las 4 leyes de los robots parece un resguardo que deben cumplir las máquinas, pero la ética es para los seres humanos.
[49.30 - 50.29// 51.13]	Los miedos es que los robots nos manejen. un historiador israelí lanza la idea de que las máquinas dominan. pero estamos lejos. estar preparados y se dan reglas a los que crean y usan.

[UBICACIÓN] COMENTARIO	TIP EVALUACIÓN
[51.28-51.50]	No hay moral de robots
[56.19 - 56.16]	Un robot no es sujeto de derecho y esto es consensuado
[37.36 -58.07]	El robot aprende algo malo, se le puede señalar que algo está mal. entonces sigue dependiendo del humano
[1.02.31- 1.2.43//1.2.49-1.3.06]	Tener en cuenta los límites: La ética aplica a un ciudadano, poblador de un Estado y como ciudadano del mundo



Apéndice 3. Video 3.Ciclo ética en IA - ANCBA

Este apéndice corresponde al disertante Emilio Suñé Llinás, quien habla sobre Ética para Sistemas Inteligentes. Luego se presenta una tabla correspondiente a la disertante Lidia Seratti quien continúa la temática. Finalmente se incluyen algunas partes del interesante cierre de Juan Carlos Ferreri. Las consideraciones son similares a las del Apéndice 2.

En este apéndice hay transcripciones que se mantienen en *itálicas*.

UBICACIÓN	Ing. J.C.FerreriCITA/EVALUACIÓN
[0:16:50 – 0:17:31]	<i>Una ética para sistemas inteligentes... complicado por la dificultad de establecer un sistema de valores y tomar decisiones la máquina a partir de ellos...y en la medida en que pudiera establecerlos, adaptarlos para la toma de decisiones concretas.</i>
[0:18:28 – 0:18:44]	<i>...además, si esto fuera factible, las máquinas también se encontrarían con el mismo problema con que nos encontramos los humanos: la presencia de distintos sistemas éticos.¿Cuál escoger?</i>
[0:22:02 – 0:22:13]	<i>...porque entonces el modo de actuar dependería del sistema ético de referencia adoptado".</i>
[0:31:45 – 0:32:18]	<i>...un sistema ético que se necesite para tomar decisiones inmediatas por un robot como es el caso del coche autónomo, ¿por qué no se implementa?: por la responsabilidad de losfabricantes.</i>
[0:32:20 – 0:33:14]	<i>...hay una resolución del Parlamento Europeo del año 2017...que debe tener unos 20 folios, recomendando cuestiones de derecho civil acerca de los robots y en 4 o 5 líneas a alguien se le ocurrió decir que habría que atribuir una personalidad jurídica a los robots...sin darse cuenta de que esto es un caramelo envenenado...¿por qué? Porque cuando los robots inteligentes produzcan determinados daños,¿de quién es la responsabilidad? ...</i>
[0:35:07 – 0:35:32]	<i>Hablar de personalidad jurídica de los robots, es la manera de decir (yo fabricante): “¿Quién es el responsable?, pues, ¡el coche autónomo! ...” ...Y el coche autónomo, ¿con qué patrimonio responde? ... Solo que sea con el propio valor</i>

UBICACIÓN	Ing. J.C.FerreriCITA/EVALUACIÓN
	<i>del coche autónomo...es una inmensa trampa...</i>
[0:35:32 – 0:35:46]	<i>Lo que se está intentando, y los juristas no se dan cuenta, porque les están dando un caramelo envenenado, es efectivamente quitar del medio a la responsabilidad de la fábrica...</i>
[0:35:55 – 0:36:51]	<i>Para mí, la adopción de decisiones sobre la base de criterios directamente éticos nos lleva, por las razones expuestas, a un callejón sin salida. La alternativa racional es más jurisprudente. Basada sobre la pragmática y la jurisprudencia, porque la inmensa mayoría de estas cuestiones se han planteado a través de principios como el “estado de necesidad” ante Tribunales de Justicia que han dado soluciones concretas. Tenemos ahí un banco de experiencia. Y no es que yo diga que hay que volver a los sistemas “expertos”...</i>
[0:37:50 – 0:38:11]	<i>...por lo tanto, para la adopción de decisiones de carácter inmediato, creo que, al no ser posible definir un sistema ético como universal, habrá que ir en un sistema más basado sobre la pragmática y la propia jurisprudencia...</i>
[0:38:11 – 0:38:31]	<i>...Y alguien me diría: “¡Pero el Derecho tampoco es universal!” Ciertamente,pero los sistemas jurídicos, sí son bastante homogéneos entre sí y tienen soluciones tremendamente parecidas entre países.</i>
[0:38:35 – 0:39:04]	<i>Un sistema jurídico y jurisprudente de carácter romano-germánico-napoleónico, lo tenemos en la Unión Europea. Pero también lo tienen en América Latina, en Iberoamérica. Que es el mismo sistema jurídico y, en consecuencia, las soluciones no serían muy distintas...</i>
[0:39:28 – 0:40:07]	<i>Creo que un sistema de adopción de decisiones. de “decision making” instantáneo para robots de carácter inteligente no tiene solución, por las razones que les he dicho desde el punto de vista ético, y por lo tanto, si no es posible establecer un sistema que lógicamente sea superior a los demás, lo lógico es irse haDisertante: L.Seratticia un sistema jurídico que sí podría proporcionar este tipo de soluciones...</i>

UBICACIÓN	Ing. J.C.Ferreri CITA/EVALUACIÓN
[0:52:19 – 0:53:06]	<i>Lo que se pretende, con la personalidad jurídica de los robots que no tienen patrimonio más que sí mismos, sería una pantalla de responder con su valor precisamente, y ahí está la trampa, porque no tuviera que responder el fabricante que es lo que se ve venir encima y quizás por eso se está retrasando la comercialización del coche autónomo.</i>

Disertante: L.Seratti

Concepto: Regulaciones

UBICACIÓN	CITA/EVALUACIÓN
[1:24:10 – 1:24:37]	<i>La labor legislativa y regulatoria puede poner parámetros en los tratados y resoluciones de la Unión Europea que permitan ir haciendo un cuerpo realmente, esto es planetario, para hacer una regulación.</i>
[1:29:56 – 1:30:05]	<i>Hay muy poco acercamiento a lo digital como para pretender que de golpe la sociedad se pueda adaptar a esto.</i>

Disertante: Ing. J.C.Ferreri

Concepto: Cierre

UBICACIÓN	CITA/EVALUACIÓN
[1:33:37 – 1:33:53]	<i>...la concientización necesaria como para saber que la ética está implícita en las personas que desarrollan las cosas y las implementan, y por lo menos que quedará en claro que “no</i>

hay máquinas perversas, sino hay perversos que implementan máquinas”



Apéndice 4. Video 10. Ciclo ética en IA - ANCBA

Este apéndice fue extraído de la disertación dentro del mismo ciclo a cargo de Natacia Arcifa, letrada de la Universidad de Pisa-Italia), titulada Inteligencia Artificial o derechos humanos? Qué es más necesario? Con temática basada en armas letales autónomas (Autonomous weapons)

La disertación original está en italiano, pero los textos son traducciones literales.

UBICACIÓN/TRADUCCIÓN	EVALUACIÓN
<p>[8:54 – 9:07]</p> <p>las innovaciones son tales que nos estamos acercando a la adquisición de instrumentos y de los sistemas que podrían (en un futuro próximo muy vecino) matar a las personas.</p>	
<p>[9:12 – 10:00]</p> <p>la disputa es hoy bastante acalorada puesto que en el 2010 el mercado de uso militar y de seguridad pública de drones, valía casi 5 billones de dólares, posteriormente, desde 2015 hubo un incremento del doble, porque los principales países (como los Estados Unidos, Rusia, China, Francia) han invertido billones en el desarrollo de la autonomía en el uso de la fuerza militar y están financiando sistemas de inteligencia artificial que confeccionan unos escenarios inesperados por nuestra sociedad</p>	<p>La desigualdad se incrementa con la IA? O es nuevamente la desigualdad pero más visible? No es lo mismo explotar económica y socialmente países de tercer mundo?</p>
<p>[10:35 – 10:53]</p> <p>el riesgo es que sin embargo exista una falta de responsabilidad jurídica y ética de la que se ha hablado bastante en este encuentro... especialmente de lo que concierne a los Estados como Estados Unidos, o lo reciente de Israel</p>	
<p>[13:02 - 13:08]</p> <p>VIDEO dentro de la presentación de una demostración de dron asesinos de USA (aclara que no está activo pero existe). El video está acá resumido a la demo.</p>	<p>La presentación cínica que se hace refiere a una “matanza limpia” (los humanos no obedecen, pero un dron SI)</p> <p>muestra un militar admitiendo que se está invirtiendo MUCHO (heavily) en PROYECTOS CLASIFICADOS</p> <p>[07]</p>

UBICACIÓN/TRADUCCIÓN	EVALUACIÓN
<p>[14:32 – 14:53] ha nacido una campaña “stop killer robot”, de la que leí la Argentina está muy activa. Sobre todo porque muy recientemente, en febrero de 2020 en buenos aires se hizo un encuentro global donde se discutió sobre la posibilidad de limitar estos equipos, las armas letales autónomas.</p> <p>[15:01 – 15:12] en este campo se ha firmado en 2015 una carta por muchos científicos, durante la conferencia internacional de la inteligencia artificial</p>	<p>www.youtube.com/watch?v=TIO2gcs1YvM</p> <p>¿cómo es que no se hizo público? Dónde se avisó de esto? Lo peor: por qué no se nos avisó?</p>
<p>[16:22 – 16:42] se está buscando crear un marco regulador de referencia sobre el impacto de los sistemas autónomos. Sobre todo la legislación sobre la base que dirima la aceptabilidad ético-jurídica en términos del derecho internacional humanitario</p>	
<p>[17:20 – 17:36] durante esta conferencia [última conferencia de expertos en IA, con 86 Estados soberanos participantes], 22 de los Estados participantes, solamente en 22 se han manifestado contrarios a la adopción de estos sistemas de armas autónoma letal.</p>	<p>esto indica que la ética de los Estados no es aún homogénea como para permitir una uniformidad como se requiere en la propuesta de Unesco</p>
<p>[17:54 – 18:14] los Estados Unidos, China, Rusia sostienen que, en vez la discusión [sobre la adopción o no de armas autónomas] debería valorizar las ventajas que la introducción de armas autónomas traerá en la ECONOMÍA de los conflictos humanos. Sobre todo desde el punto de vista de la REDUCCIÓN del costo humano.</p>	<p>Se habla de economía en relación a costo humano</p>
<p>[18:37 – 18:52] hay países con una posición intermedia. Por ejemplo Francia y Alemania. Muchos en Europa tienen una posición intermedia. Argumentan</p>	<p>Se ve entre 1, 2 y 3 que no hay una ética unificada al respecto</p>

UBICACIÓN/TRADUCCIÓN	EVALUACIÓN
<p>que es posible tenerlos pero de manera controlada.</p> <p>[20:18 – 20:28]</p> <p>promulgar una ley robótica que se refiera a los sistemas inteligentes, y que incluya a los sistemas de armas letales autónomas.</p>	<p>Entonces estamos ante una ley global? Por encima de las leyes soberanas</p>
<p>[21:50 – 22:21]</p> <p>Ha surgido la necesidad de promulgar reglas éticas. Reglas éticas que, imagino todos conocen las 3 leyes de Asimov. No son que sea una simple implementación, sino que abrazan la cuestión del modo como la robótica se articularía con el control por parte del ser humano.</p>	<p>Las leyes articulan la ética del agente para preservarlo en calidad de tal</p>
<p>[22:48 – 23:19]</p> <p>estas normas éticas hoy no existen, los Estados para remediar esta falta han hecho recurso al art. 36 del primer protocolo adicional, de la convención de Ginebra, que exige a los Estados adherentes que se encarguen del estudio del desarrollo, de la adopción de nuevas armas y de determinar en qué circunstancias pueden ser admitidos o prohibido. Por lo tanto, en ese sentido también las reglas éticas tienen una validez universal y se puede referir a las normas existentes.</p>	<p>nuevamente se hace un caso omiso a la posibilidad de diversidades, para poder controlar el problema</p>
<p>[24:13 – 24:33]</p> <p>remarcar la importancia de mantener el libre albedrío humano sobre las decisiones de los sistemas autónomos. Parece un poco extraño porque si un sistema es autónomo no se puede controlar</p>	
<p>[25:23 – 26:44]</p> <p>en cuanto a los argumentos éticos, que dicen a favor o contra estos sistemas de arma autónoma. Las dos teorías dicen que los argumentos a favor o contra, son aquellos que remiten a la ética consecuencialista y la deontológica. La ética consecuencialista se centra en los resultados y por lo tanto promueve la autonomía de</p>	<p>Dos vertientes éticas para los juristas en el aspecto armamentístico</p>

UBICACIÓN/TRADUCCIÓN	EVALUACIÓN
<p>las máquinas para obtener mejores prestaciones. También dice que con IA habrá un derecho internacional y valores éticos internacionales. La deontológica se focaliza en el proceso y por lo tanto se focaliza en los valores morales como guía de la acción son parámetro de juicio de las opciones individuales y colectivos.</p>	



Apéndice 5. Elementos bibliográficos incorporados

Esta sección incluye la elaboración bibliográfica del problema para ser cruzado a los materiales de las secciones anteriores. Las observaciones de los autores se destacan con la palabra **TIP** en cada caso.

Estudio de Funtowicz y COM 2002

NOTA: Este estudio se basa en [08] y [09]

En [08] se estudia el problema desde la perspectiva de las políticas científicas modernas, donde los autores evalúan el problema de la indeterminación desde la regulación ética y evaluación de riesgo. Al estudiar el modelo de encuadre, como una de las posibilidades alternativas del modelo precautorio y del modelo de demarcación, curiosamente refiere a los siguiente dichos de la misma Comisión Europea afirma [09]:

La comisión podría encontrarse ante una gama de dictámenes contradictorios de expertos del mundo académico, de personas con conocimientos prácticos en una materia determinada y de otras con intereses directos en la cuestión. Estas opiniones pueden basarse en presupuestos completamente diferentes y tener objetivos totalmente distintos...Por tanto, cada vez más, la interacción entre los responsables políticos, los expertos, las partes interesadas y el público en general es una parte crucial de la elaboración de políticas y, por eso, hay que prestar atención no sólo al resultado de las políticas sino también al proceso seguido.

En una clara demostración de que la comunidad produce distintos marcos alternativos al encarar un problema y por lo tanto el marco mismo se vuelve parte del debate entre grupos de interés. El resultado es que los expertos pueden incurrir en contradicciones que deben ser dirimidas en decisiones políticas previas [08].

Según los autores: “Cualquier parte que resulte considerada la ‘dueña’ del problema hará la contribución más importante y disfrutará de los mayores beneficios”.

TIP: ¿y quién será entonces el beneficiario? dónde se está dirimiendo el acuerdo político previo?

La solución que propone la propia CE es una llamada a la pluralidad de perspectivas:

“El factor determinante final de la calidad es el pluralismo. Siempre que sea posible, deben solicitarse puntos de vista diversos... También pueden ser importantes otros factores, como diferentes perspectivas geográficas culturales y de sexo” [09].

TIP: ¿pero nuevamente, el problema es que estamos ante un modelo de demarcación y no de encuadre donde los autores bien determinan que hace falta determinar los valores que deben ejercerse para poder luego prescribir. El problema

es, como afirman los autores: El problema principal es que ya no resulta funcional excepto en casos de corrupción claramente delimitados puesto que “la filosofía de la ciencia pos-empirista ha mostrado que, en general, dadas la complejidad y la indeterminación de los sistemas emergentes, es imposible una separación total entre los hechos y los valores”.

Esto es que la polarización de la situación y resulta en extremo difícil establecer una separación impermeable entre la evaluación y la gestión del riesgo, y como concluye, los grupos de interés accionan de expertos y nuevamente no se es totalmente imparcial.

La solución propuesta epistemológicamente torna invisible la relevancia que los valores políticos tienen en las elecciones de los científicos [08] en lo metodológico, y en la relevancia de los procesos de gobernanza de los criterios de relevancia.



Estudio desde filosofía de la biología

NOTA: Este estudio se basa en [11].

La Enciclopedia de filosofía de Standford tiene una compilación de la evolución de la filosofía de la biología con un enfoque que parece traslucir similares cambios en otras ciencias, aunque tal vez este foco en la biología se postula como una de las más paradigmáticas, ya que la biología y las ciencias médicas fueron radicalmente impactadas en los cambios conceptuales y filosóficos de los últimos 40 años en la ciencia.

Tal como señala el artículo, la literatura acerca de la filosofía en esta área evidencia tres tipos de cambios esenciales:

- el tipo de tesis planteadas
- los problemas conceptuales emergentes
- cuestiones enraizadas en la biología, pero con foco en la filosofía tradicional

El análisis es clasificado en ocho vertientes según ciertas combinaciones evidenciadas, de estos tres tipos de cambios:

1. Pre-historia de la filosofía de la biología: Situada alrededor de los '50, con publicaciones de Woodger y Becker como iniciadores de la visibilización de la necesidad de establecer este tipo de enfoques. Los estadios previos de la biología se presentan aquí como eras discusión de cuestiones puntuales como leyes que la rigen. Entre los primeros que consolidan el surgimiento estarían Mayr y Ayala (alrededor de los '70), seguidos por una cantidad respetable de publicaciones desde el '76 al '86.

2. Surgimiento de tres tipos de filosofías.

- Como tesis en el marco de la biología. Aquí aparecen las argumentaciones como el uso de la genética para el reduccionismo y el anti-reduccionismo en filosofía de la ciencia. Entre los primeros, por caso K. Schaffner aplica el modelo empiricista reduccionista para relacionar la genética de Mendel con la genética molecular. Entre los segundos figura D. Hull que argumenta sobre la imposibilidad de reducir la primera a la segunda. Ya en los '70 -'80 e debate parece suavizarse con la conclusión de que ambas cosas son fundamentalmente diferentes (Kitcher).

- En la parte conceptual: surgen problemas teóricos de la biología, analizados desde la filosofía. Por ejemplo, la adaptación de los individuos por reproducción. Si el más apto sobrevive, entonces surge la problemática de cómo se define al más apto. Esto parece zanjarse luego de cierto tiempo con dos estrategias. La primera es describir la aptitud para la supervivencia como un conjunto de características determinadas. La segunda es determinar que se requiere y refiorea una aptitud o propensión para la supervivencia, algo relacionado con la probabilidad y no el determinismo.

- El tercer tipo encastra cuestiones de biología en la filosofía tradicional. Aquí se encuentran los filósofos con posiciones

establecidas en tópicos tradicionales de filosofía, tales como la ética y la epistemología. Entre ellas son de mencionar la teleología biológica y la posterior teleonomía emergente como lenguaje específicamente teleológico.

3. Filosofía de la biología evolutiva. Comprende los estudios sobre la interpretación de la evolución. Justifica este apartado en la relevancia de la evolución para la biología. Presenta aquí las diferentes visiones como la visión semántica de Beatty (1980) donde destaca los problemas conceptuales emergentes de la interpretación de la teoría, Sober (1984) relacionado la evolución genética en función de la mecánica newtoniana con un conjunto de fuerzas distintas que se plasman en la selección, mutación, etc. También incorpora sus críticos como Sterelny (1988) quien finalmente cambia esa visión a otra de perfil más estadístico. Otra vertiente destacable aquí es la sociobiología de Wilson (1978) donde la extensión de la biología alcanza a un tipo de evolucionismo de la conducta social humana. También esta visión recibe sus críticas, entre otros, de S. Jay Gould y Lewontin. Se estudia entonces el rol de la metodología de la evolución y los modelos optimales, y se responden a nuevas cuestiones tales como si la selección no es válida solo para la teoría de los genes.

4. Filosofía de la biología sistemática. Surge como reacción. Es una revolución de la concepción de la biología que corresponde a los '60 - '70. Comienza con W. Hennig (1966), entre otros, y mueve el estudio desde la fenética²⁹⁸ a la cladística²⁹⁹. De aquí se centra en las cuestiones filogenéticas con estudio de las relaciones de parentesco entre especies o taxones. Surgen diferentes criterios y categorizaciones de especies con diferentes criterios. Destacan Meier, también M. Ghiselin, quien hace hincapié que las relaciones entre especies es más histórica que genética, pero se centra en los trabajos de Mayr (y su teoría BSC) que sostiene de una especie es un conjunto de grupos aislados de otros que pueden interactuar, lo que trajo bastante debate.

5. Filosofía de la biología molecular. Relacionada con las concepciones de la biología molecular.

6. Filosofía de la biología de desarrollo. Plantea como desarrollo de opciones disponibles en cada momento en función de fuerza de disponibilidades y restricciones.

7. Filosofía de ecología y biología de la conservación. Relacionada con la ética del entorno y la ecología.

8. Metodología en filosofía. Expone y estudia las diferencias esenciales entre la filosofía de la biología y la biología per se. Destaca que si

²⁹⁸Fenética refiere a la escuela de taxonomía que clasifica a los seres por similitud morfológica y cualidades generalmente visibles.

²⁹⁹Cladística refiere a las clasificaciones sobre relaciones evolutivas (similitudes derivadas).



bien el objeto de estudio es el mismo en ambos casos, existe una diferencia sustancial en la naturaleza de las preguntas que cada una responde.



Estudio de Merton influencia en ciencias

NOTA: Este estudio se basa en [04]

R. Merton y las influencias religiosas de las ciencias

De acuerdo a R. Merton la ciencia moderna, tal como la conocemos, fue posible gracias a una combinación de tres elementos centrales:

PRIMERA INFLUENCIA: El estado cultural y mental de la sociedad del siglo XVII. Dicho estado le dio una impronta a las ciencias que se refleja en gran parte de sus características actuales. De acuerdo con A. Weber, los procesos de la sociedad, la cultura y la civilización impactan en la conformación de una sociedad cultural en la que se manejan ciertas categorías y procesos que están íntimamente interrelacionadas.

Para Merton, la mentalidad religiosa del protestantismo ascético y la lógica de la ética puritana se caracterizaron por:

-Estudio sistemático racional del mundo como método de control de un mundo corrompido.

-Estudio empírico de la naturaleza como modo de glorificar a Dios.

Ambos elementos derivaron en un conjunto no explicitado de elementos básicos y condicionantes de la posterior evolución en las concepciones científicas.

En la deriva de la instauración sistemática se generan entonces las corporaciones científicas que buscan el bien del hombre, como una prolongación explícita de dicha mentalidad, y que se implementara en normas utilitarias que valorizaran la deseabilidad de las prácticas científicas siempre como ponderación en función de los mismos valores éticos.

SEGUNDA INFLUENCIA: El ethos puritano entonces deriva en una concepción de bienestar social como meta constante y su implementación a través de una filosofía natural donde se caracteriza primariamente con:

-alejamiento de tentaciones que no sean hacia el bienestar social

-concepto de utilidad al mundo

-determinación de la futilidad del concepto de gabinete único y el origen en cambio, de gabinetes alternos de las ciencias (lo que deviene en categorizaciones)

En todo el proceso resalta el autor a los valores de empirismo y utilitarismo.

Es interesante observar que Merton entonces subraya esta mentalidad, otrora religiosa, como una nueva versión que reedita el poder del clero (ahora autorización religiosa implícita al conocimiento que deberá glorificar y acercar a Dios), y la amalgama con la constitucional (al dar autorización social al mismo conocimiento ya que el mismo será destinado en su beneficio), junto con un desplazamiento de respon-

sabilidad del peso del conocimiento desde su estatuto tradicional en la Iglesia, ahora asignado al individuo (científico).

TERCERA INFLUENCIA: Como resabio de la teología medieval, y desde Hume, el autor reconoce en la época la existencia de un orden de la naturaleza que tiene dos sesgos importantes:

-El orden es absoluto y definitivo (ya que deriva de Dios), y por lo tanto establece una sistemática racional para su estudio apropiado. En la confluencia con los elementos explicados en los apartados anteriores, entonces dicha racionalidad exige al experimento como testimonio de verdad

-La naturaleza, como ente previo a las ciencias y de carácter inteligible, lo que permite que se le pueda indagar.

Estas características de época, y los elementos estudiados por el autor, reflejarían una estructura mental previa que, sumamos al acerbo del protestantismo da importancia a la experiencia, la razón y al interés activo por este mundo

Como prueba de este análisis, el autor luego aporta elementos históricos diversos que remiten a las academias disidentes de base puritana en Inglaterra, el surgimiento de ciertas academias como Northampton y Morton, y las universidades (esencialmente Harvard). Asimismo, extiende el estudio al resto de Europa con su contraparte de Alemania centrada en las escuelas pietistas, las Schulen y gymnasien.

El autor concluye con un sinóptico en que la relación de las Ciencias modernas y la religión se centran en los siguientes elementos:

1) conexiones no deliberadas, con raíces en los primeros reformadores religiosos (ej. Lutero) quienes paradójicamente iniciaron siendo hostiles o en el mejor de los casos indiferentes a las ciencias

2) orientación axiológica en las ciencias, que se enraíza aún después de la independencia del vínculo teológico de las mismas

3) centrado en la evidencia estadística como prevalencia de lo axiológico inconscientemente establecido

4) la interacción visible en cuestiones de guerra (especialmente en siglo XIX), que logra tapar la conexión indirecta de la religión e invisibiliza a la misma.

Declaración Universal sobre Bioética y Derechos Humanos (Unesco)

NOTA: Este estudio se basa en [12]

La Declaración Universal sobre Bioética y Derechos Humanos de la Unesco fue homologada unánimemente por 191 países, durante una Sesión de la Conferencia General de la Unesco en París (2005). La Declaración trata sobre las cuestiones éticas relacionadas con la medicina, las ciencias de la vida y las tecnologías conexas aplicadas a los seres humanos, teniendo en cuenta sus dimensiones sociales, éticas, jurídicas y ambientales.

El documento pretende fundar ciertas garantías a las libertades fundamentales y los derechos humanos en un mundo plural. Proclama el ejercicio del diálogo racional, la búsqueda de los principios éticos sobre consenso y pactos sociales. Establece pautas comunes para abordar las nuevas situaciones de intervención sobre la vida a raíz del progreso de las ciencias biomédicas y de las nuevas tecnologías disponibles en un marco de integridad científica.

La Declaración se elaboró en dos pasos, de manera similar a como se maneja París 2021:

-El primer año, el Comité Internacional de Bioética (CIB), en consulta con los Estados miembros y otros organismos especializados, y redactó un borrador.

- El segundo año se hicieron las correcciones finales definitivas realizadas durante las reuniones intergubernamentales de expertos.

Con 28 artículos, pretende proporcionar un marco de principios y procedimientos que puedan servir de guía a los Estados en la formulación de sus políticas, legislaciones y códigos éticos.

La similitud es sorprendente respecto al manejo de ética para la IA, salvo que con París 2021 se solapan esta declaración o todos los demás aspectos de la vida ciudadana, social e individual. Con esto estamos en condiciones de decir que se ha redoblado la apuesta.

Respecto a los ítems, principalmente enfocamos los siguientes:

- **Artículo 8.** Respeto de la vulnerabilidad humana y la integridad personal

Al aplicar y fomentar el conocimiento científico, la práctica médica y las tecnologías conexas, se debería tener en cuenta la vulnerabilidad humana. Los individuos y grupos especialmente vulnerables deberían ser protegidos y se debería respetar la integridad personal de dichos individuos.

- **Artículo 9.** Privacidad y confidencialidad

La privacidad de las personas interesadas y la confidencialidad de la información que les atañe deberían respetarse. En la mayor medida posible, esa información no debería utilizarse o revelarse para fines distintos de los que determinaron su acopio o para los que se obtuvo el consentimiento, de conformidad con el derecho internacional, en particular el relativo a los derechos humanos.

- **Artículo 10.** Igualdad, justicia y equidad

Se habrá de respetar la igualdad fundamental de todos los seres humanos en dignidad y derechos, de tal modo que sean tratados con justicia y equidad.

- **Artículo 11.** No discriminación y no estigmatización

Ningún individuo o grupo debería ser sometido por ningún motivo, en violación de la dignidad humana, los derechos humanos y las libertades fundamentales, a discriminación o estigmatización alguna.

- **Artículo 12.** Respeto de la diversidad cultural y del pluralismo

Se debería tener debidamente en cuenta la importancia de la diversidad cultural y del pluralismo. No obstante, estas consideraciones no habrán de invocarse para atentar contra la dignidad humana, los derechos humanos y las libertades fundamentales o los principios enunciados en la presente Declaración, ni tampoco para limitar su alcance.



Estudio de Información e IA

NOTA: Este estudio se basa en [14]

El autor inicia dando la tradicional definición de información en los términos que Shannon expuso en 1949, define entonces los conceptos de mensaje (M) como conjunto de mensajes $\{m_i\}$, σ como el conjunto de estados posibles $\{s_i\}$ de un sistema S, y finalmente R, que inicialmente introduce como el receptor, luego pasa a denotar al sistema de referencias de éste.

Hace un parate específico en el sistema de referencias R donde establece que se caracteriza como la dupla de:

a) Probabilidades a priori de ciertos estados o expectativas: $\{P(s_i)\} / \sum_k P(s_i) = 1$

b) Probabilidades a posteriori: $\{P(s_k/m_i)\} / f(\langle m_i, s_k \rangle) = P(s_k/m_i)$, con $\sum_i P(s_k/m_i) = 1$ de los estados efectivamente alcanzados durante la comunicación.

La diferencia de expectativas entonces es medible a través de:

$$D(m_i, R, S) = \sum_k |P(s_k) - P(s_k/m_i)|$$

Esta será entonces la cantidad de información transmitida con:

$$I(m_i, R, S) = -\log(1 - (D/2))$$

El estudio de esta cantidad de información da pie al autor para determinar la dinámica de interacción entre emisor y receptor, donde los límites plausibles son dados por la matemática misma como: $D \in [0; 2)$

En $D(0) \Rightarrow I = 0$, el mensaje no produce cambios de expectativa

En $D(2) \Rightarrow I \rightarrow \infty$, el mensaje es disruptivo y exige un cambio revolucionario o de sistema de referencias (R) en el receptor.

Para describir este último caso, hace un paralelo interesante con los períodos de innovación referidos por Kuhn en 2017, donde la innovación en ciertos momentos cae dentro del marco conceptual (lo esperable, la expectativa previa) y otros momentos donde la innovación sale del mismo, y exige un cambio de paradigmas, que es tomado por las ciencias.

TIP: Esto iría en concordancia con el concepto de que los cambios son guiados por un tipo de saber (no son azarosos), y que remite al concepto de phrónesis de Aristóteles. Este saber, guía al cambio desde lo práctico y lo social.

Si bien la definición de IA es de época (basada en DRAE), acierta en afirmar ciertas características de la misma que perduran en forma discutida en la actualidad:

- Estos sistemas no identifican a los problemas como problemas propios
- Ayudan a resolver problemas que sí le pertenecen al humano
- Tienen la capacidad de simular comportamientos.

Refiriendo los dichos de Kosman (1987) afirma: “los artefactos en la tradición aristotélica, son sustancias solo en un sentido accidental” y por lo tanto eso los dife-

rencia de los seres humanos y de hecho de los seres vivos en general, que son sustancias en sentido propio y paradigmático.

Esta diferencia ontológica es la que vuelve nula la eliminación de dicha distancia basada en la sofisticación tecnológica y de hecho la clasifica en un error categorial.

Luego detalla algunos aspectos del cómo, en cierta manera, la IA logra sus habilidades, pero sienta ya el aspecto que nos llama a este trabajo: “¿quién se hace responsable en cada caso de un posible fallo?, ¿habría que poner comillas en “decide” y “diagnostica”? “termina, en referencia a la aplicabilidad de estos términos que la IA aplica por imitación.

Finalmente, vale la pena resaltar que un sistema de IA, para el autor, es una suerte de “prótesis” de nuestros sistemas de expectativas, puesto que hace un vector n-dimensional de datos históricos para simplemente con eso decir qué esperar del objeto representado. Por supuesto que como cualquier tipo de operación artificial, puede colapsar en situaciones donde $D(2)$, y corresponderá en esas instancias que reaccionen las personas con creatividad, cambiando R.



Estudio de Diez años de la Declaración Universal sobre Bioética y Derechos Humanos

NOTA: Este estudio se basa en [15]

Según Salvador Darío Bergel, la declaración Universal sobre bioética y derechos humanos de la Unesco en 2005, sienta las bases de una nueva bioética, que agrega a la bioética clásica la visión social y política. Evalúa su proceso como preciso y metódico, y deja establecida su consideración de que la misma, al no terminar en un tratado, no se vuelve vinculante para los suscriptores, pero consolida a la bioética en las relaciones internacionales y sienta en ese momento un antecedente de valor considerable en el campo jurídico internacional.

Si bien reconoce que su debilidad radica en no ser vinculante, explora las razones por las que considera que genera una explicitación del vínculo indiscutible entre la bioética y los derechos humanos, relacionándose asimismo los derechos nacionales y los internacionales.

Este es un avance en las autonomías y autodeterminaciones locales por una regla internacional global. Como colofón reconoce que esta declaración establece qué derechos humanos deben ser aplicados y respetados en pie de igualdad.

TIP: Es interesante el grado de visión positivo que le otorga el autor a esta imposición, indica que junto a la Declaración Universal de Derechos Humanos de 1948, consolidan ambos las exigencias sociales y políticas de derechos humanos de los suscriptores.

Luego enumera los siguientes derechos a los cuales remite:

- autonomía personal
- integridad personal
- goce del máximo grado de salud
- atención médica de calidad
- acceso a medicamentos
- alimentación adecuada
- acceso al agua potable
- reducción de la pobreza
- alfabetización
- gozar de los resultados de la investigación científicas
- medioambiente no contaminado

TIP: Pero como crítica al autor están los criterios de autonomía, integridad, máximo grado de salud, calidad, adecuación pueden tener valores y significados diferentes en distintas culturas y contextos. Esta indefinición, por lo tanto, deja abierta la brecha a determinaciones no compatibles en el nivel procedimental, pero de estos asuntos el autor no detalla.



Entre los ítems de interés que destaca finalmente el autor, está la siguiente cita textual de advertencia: “La ciencia y tecnología, aventuras del pensamiento humano, no reconocen límites y, si se los fija hipotéticamente, pueden ser fácilmente vulnerables”.

TIP: La pregunta que cabe también es si no es vulnerar los derechos humanos, fijar los valores de manera uniforme y sin atender diferencias culturales. Esta expresión sienta claramente la intención de no dejar libre “las aventuras del pensamiento humano”, bajo el supuesto de que será perjudicial no hacerlo.



Estudio de Técnica y Sentido

NOTA: Este estudio se basa en [16]

José Manuel Chillón a y Alfredo Marcos presentan una evolución histórica de la ciencia de la siguiente manera:

técnica → tecnología → tecnociencia → biotecnología → antropotécnica

siendo esta última mirada como una absolutización de la razón instrumental sobre el humano y que por tanto es de sumo riesgo.

Remiten a Ortega (1997) en su evaluación de que la técnica humaniza al medio para hacerlo más habitable a los humanos. Por ello es de esperar que una mejora técnica redunde en una mejora humana, pero también determinan que “esto solo lo podemos argumentar si disponemos de un criterio normativo, como es el de la naturaleza humana, que nos ayuda a entender lo positivo del progreso técnico en relación con esta promoción de valor en el ser humano.

También sostienen que “La historia de la técnica puede ser también la historia de la humanización de nuestra especie, siempre que la técnica favorezca, en lugar de amenazar, el auténtico progreso humano”.

En términos generales, los autores comprenden la historia técnica como una historia de la humanización de la especie, pero esto sería válido en tanto favorece y no amenaza el **auténtico** progreso humano.

TIP: Como lectores nos preguntamos si es posible determinar precisamente el concepto de auténtico que expresan los autores.

Desde la escuela de Frankfurt ya se advierte sobre la des-humanización de la técnica, específicamente la crítica de Horkheimer y Adorno indican, por ejemplo, que el cientificismo positivista y tecnologismo asociado derivaría en una técnica totalitaria. Husserl ya previamente habla al respecto y se centra en que la extensión del patrón positivista e instrumental a todos los órdenes de la vida humana deriva en la pérdida del sentido.

TIP: Como lectores nos preguntamos si la determinación de una ética universal funcional a los efectos de la IA no son también una manera de positivismo instrumental aplicado a los valores éticos. Para Heidegger el exceso de tecnificación y la huida de la necesidad de pensar están mutuamente implicados.

Chillón y Marcos rematan su postura sugiriendo adoptar la postura de Heidegger, en su relación del ser con la técnica, de serenidad. Advierte también que la fuerte des-humanización tiene irónicamente como verdugo y víctima al propio humano. ¿cómo hay que situarse ante la técnica? Para Heidegger: Con una determinada actitud, con un determinado temple de ánimo (Befindlichkeit) que evoca, el misterio del ser en el que debe contextualizarse la cuestión de la técnica. Este temple de ánimo es –según Heidegger– la serenidad ante las cosas (Gelassenheit).

Dada la naturaleza de la técnica como medio de humanización, los autores plantean que una antropotecnica (modificación de la naturaleza humana) pone en

riesgo la continuidad/sentido de la técnica misma y esto derivará en la **disolución de la ética** y de la razón práctica.

En c.t. “Si modificamos técnicamente la naturaleza humana, ¿no habríamos, por lo mismo, puesto en riesgo la continuidad y sentido de la propia técnica? Quizá la absolutización de la razón técnica conduzca a la disolución de la ética y de la razón práctica, quizá acabe por disolver también los trazos más reconocibles de nuestra sociedad”.

Estudio de Problemas bioéticos emergentes de la inteligencia artificial

Nota: Este estudio se basa en [17]

Nota: el autor replica definiciones incorrectas de la IA a las que no se reproduce acá.

Villalba introduce un concepto análogo a la antropotecnia de Chillón/Marcos: el Transhumanismo. Denotado como H+ o h+, es definido por Russell y Cohn como: “cuestionar la posibilidad y conveniencia en la transformación de la condición humana, creada y desarrollada mediante tecnologías que disminuyan el envejecimiento de la persona y aumenten su capacidad intelectual, física, y psicológica” (Russell y Cohn).

Define el concepto de singularidad **tecnológica como análogo a la singularidad matemática (en el sentido que trasciende cualquier limitación finita)**. También sostiene que la tecnología busca mejorar las capacidades humanas, y como parte del todo tecnológico, también la IA; siendo a la vez esta última quien se posiciona como la opción por excelencia que logra mejorarlas mediante la creación e incremento de variables tecnológicas.

También destaca aportes a la IA desde la epistemología, la ética, y otras ciencias. Algunos de los cuales son:

- **Aportes desde la ética.** Destacan Hottoris (1991) abriendo la discusión en bioética por la tecnociencia manipulando la naturaleza humana. **También Linares (2008) postulando la “filosofía de la sospecha”** sobre el mundo tecnológico, y destacando los fenómenos producidos por su causa: fragmentación social, aislamientos y problemas ecológicos. También explica la existencia subyacente del “determinismo tecnológico”, por el que el futuro tecnológico es determinado por factores internos al desarrollo tecnológico. Esto deriva en la tecnología dominando a los usuarios de la misma. En el ámbito latinoamericano, menciona a Garrafa (2009) que **llama a expandir la epistemología a los fines del estudio de las tecnologías: Protección, Prudencia, Precaución, Prevención.**
- **Aportes desde la epistemología.** Cottingham, Stoothoff y Murdoch introducen el concepto de que todo cuerpo es una máquina, un mecanismo complejo que funciona según las leyes de la mecánica. F. Bacon a su vez

sostiene que la realidad es un conjunto de corpúsculos que siguen las leyes mecánicas. D. Hume (1948) explica que las reglas generales derivan de asociaciones repetidas entre elementos (NOTA: lo que es consistente con la estrategia inferencial de algunos sistemas orientados a datos dentro de los sistemas inteligentes). Finalmente, R. Carnap (1950) explica que el conocimiento se caracteriza por teorías lógicas relacionadas.

- **Aportes desde la Ciencias económicas y biología.** A. Smith: teoría de la utilidad. L. Walras: teoría de las probabilidades. J. Von Neumann: teoría de la decisión. J. Nash: teoría de juegos. R. Bellman: investigación de operaciones, cadenas de Markov y Árboles de decisión. G. Boole: lógica booleana. A. Turing, D. Hilbert y Gödel fueron otros notables
- **Aportes desde neurociencias.** Mc. Culloc & Pitts introducen el modelo neuronal
- **Aportes desde la sicología cognitiva.** N Chomsky, Newel y Simon introducen procesos matemáticos y sistemáticos de la lingüística y razonamiento.

Finalmente, el enfoque filosófico que presenta, el autor describe a Nick Bostrom (2014) quien postula que debe construirse una “ética de la IA”. La justificación se supone en la capacidad de crear máquinas pensantes dentro de un campo ético. Desde ese punto determina a los agentes que construyan y manipulan la IA, como los responsables de calificar y evaluar implicaciones éticas en relación con las condiciones humanas. En lo social establece las categorías bioéticas siguientes: creación, diseño, construcción y aplicación de la IA. Dichas categorizaciones sirven para poder juzgar los algoritmos mediante funciones sociales tales como: responsabilidad, transparencia, auditabilidad, incorruptibilidad, predictabilidad, etcétera.

TIP: hay una alta compatibilidad entre lo presentado por Bostrom y la jerga por él manejada con los términos y vocabulario empleados en el documento Unesco.

Estudio de Hacia una ética para el mundo tecnológico

NOTA: Este estudio se basa en [18]

Jorge Linares

Sostiene un tanto apocalípticamente que “Los análisis de los cinco pensadores mencionados nos conducen a una última constatación: se aproxima nuestra hora final”.

Menciona como corriente alternativa a Floridi (2014) que ve el cambio más positivamente como una “nueva infósfera” donde la fusión de objetos, artefactos y humanos se realiza por intercambio y procesamiento digital.

Según el autor hay cinco “anunciadores” de una catástrofe, y el inicio de lo que Mc Kibben denomina “Era del Antropoceno” donde la naturaleza terrestre está irreversiblemente cambiada.

La catástrofe cubriría todos los ámbitos (ecología, historia, política-social, ontológica y al hombre), pero no es visto por la comunidad debido al “desfasaje prometeico” (Anders) por el que somos incapaces de tomar conciencia y responsabilidad de los efectos del mundo tecnológico.

TIP: asume incapacidad para dar el paso a la justificación de la violación de la ética.

Luego promete que dicha “ética del mundo tecnológico” sienta bases de:

1. La responsabilidad colectiva (a todo el planeta y a la posteridad) para frenar los excesos negativos del poder tecnológico.
2. Garantiza la distribución de beneficios y riesgos de manera justa.
3. Provee a la caracterización ontológica de un mundo ontológico que tiene ciertas características:
 - expansión acelerada
 - expansión desmesurada
 - falta de límites

Al cambiar el entorno primario en el que evolucionamos (naturaleza ambiente) a la naturaleza bio-artefactual industrializada, se ha producido una evolución que llega a caratular casi de irreversible.

TIP: en progresión, a la incapacidad suma la irreversibilidad del proceso amenazante.

TIP: según Simondon esta visión apocalíptica corresponde a una falla en la concepción de la tecnología concibiéndola como algo externo cuando ha sido interna desde la misma creación del hombre.

En consecuencia, de dicho cambio, lo natural está desapareciendo en pos de la dominación técnica de toda la naturaleza manufacturada que nos ha colocado en

la Era del Antropoceno, y donde ya no cabe discriminar lo natural de lo artificial; sino lo controlable técnicamente de “lo otro”. Ese otro capitalizaría una naturaleza desplazada por la racionalidad del mundo tecnológico, y que está condenada a desaparecer. Con ello también se desplazan los fines teóricos y contemplativos: ese fenómeno es llamado “razón de fuerza mayor” (E.Nicol).

Dicha razón surge como una razón unilateral que se impone por necesidad sobre las acciones libres.

TIP: justamente esta fuerza mayor que desea combatir es la misma que termina imponiendo

Luego talla una limitación en el concepto antropológico-instrumental de la concepción del mundo: el individuo cree que todo es instrumento y que siempre puede manejarlo (controlarlo).

TIP: es usada para demostrar desde su perspectiva la incapacidad y la necesidad de fuerza mayor para contrarrestarla.

Pero a esa concepción se contraponen la realidad del mundo bio-artefactual, donde se cimientan ciertos hechos:

1. Lo tecnológico es un hiperobjeto no abarcable, no comprensible completamente y no delimitado.
2. Dicho mundo es complejo, con lógica propia y compele a los humanos a la subordinación al mundo tecnológico.
3. Se caracteriza por un incremento de la complejidad, de las interrelaciones, y de alcance sobre la naturaleza y la sociedad (Ellul).

De los tres, el último determinaría el incremento de la complejidad epistemológica para medir las causas/efectos del proceso. Lo que significa que la filosofía debe redefinir el sentido y fin último del mundo tecnológico.

TIP: recae en la filosofía una responsabilidad de la sociedad como conjunto? La bio-artefactualidad es consecuencia de la sociedad y debiera ser responsabilidad de ésta como ente organizo. Las causas con en ella y el efecto también (ver Simondon).

El mundo tecnológico tiene propiedades que exigen poner límites éticos a nivel planetario:

1. Artefactualidad. La naturaleza es subsumada por el sistema artefactual y artificial.
2. Racionalidad pragmática y económica. Exige máxima eficacia/eficiencia a toda la praxis humana (NOTA: no es algo que pasa desde la revolución industrial?)
3. bioartefactualidad. Todo se supone a disposición y pasible de ser manipulado/convertido en objeto tecnológico (NOTA: antes a eso mismo lo llamaban progreso).
4. Autocremiento progresivo y alcance global (NOTA: esto es una antropomorfización de un proceso de expansión industrial).

5. Interconexión compleja, orgánica y sistémica (NOTA: esto ya existía en la época de los cibernetas y Simondon posteriormente también lo nota).
6. Riesgo generalizado, difuso, y peligro de colapso sistémico (NOTA: la bomba atómica no ??).
7. Autonomía relativa de los sistemas tecnológicos (NOTA: la autonomía es la que se le inserta, y es una mímica no algo adquirido).

Los límites éticos globales exigidos, se sustentan en esta definición ontológica del mundo tecnológico.

Desde esta perspectiva, el autor promueve a cuestionar y violar la ética occidental, admitiendo que para la determinación de principios para una ética global de este estilo “implica la necesidad de cuestionar y sobrepasar ciertos límites de la tradición ética occidental”, para que se genere una ética nueva que enfrente las consecuencias, desafíos y riesgos planteados por este mundo tecnológico.

TIP: de la definición ontológica fabricada sobre estas concepciones parciales pretende instaurar una ética global.

TIP: se radica la contradicción.

8. Autonomía relativa de los sistemas tecnológicos. El mundo tecnológico parece progresar y crecer de modo autónomo. Por ello, el desafío para la ética y la política de nuestro tiempo consiste en establecer bases para el control y la evaluación social de las tecnologías, mediante una nueva cultura de valores éticos y de acciones co-responsables entre científicos, tecnólogos y el resto de la sociedad. La tecnología y la tecnociencia no pueden dotarse a sí mismas de fines y criterios éticos. Es la sociedad entera la que debe evaluarlas, conducirlas y orientarlas conforme a principios y reglas fundados en los intereses vitales de la humanidad, y mediante procesos de deliberación y decisión más abiertos que involucren a todos los usuarios y posibles afectados.

Estudio de Diccionario Latinoamericano de Bioética

NOTA: Este estudio se basa en [19]

- **Sostiene Héctor Gros Espiell** (Delegado Permanente del Uruguay ante la Unesco): el carácter necesariamente multidisciplinario de la bioética, en la que confluyen la ciencia, la tecnología, la medicina, las ciencias de la salud, la ética, el derecho, la sociología y la política.
- **Volnei Garrafa** (Presidente del Consejo Directivo de la Red Latinoamericana y del Caribe de Bioética de la Unesco/Redbioética): la nueva agenda bioética mundial [...], además de las cuestiones biomédicas y biotecnológicas, incorporó los temas sanitarios, sociales y ambientales, inclusión fuertemente defendida por las naciones periféricas.
- **Alya Saada** (Consejera Regional para las Ciencias Sociales y Humanas de la Unesco para América Latina y el Caribe): La bioética, ha conocido, desde los noventa hasta la actualidad, un desarrollo que incluye a una gran parte de los países del mundo...Muy a menudo son los intereses del mercado los que tienen tendencia a llevarse.

Historia

El término “bioética” nace en 1970 (Van Rensselaer Potter). La define como “ética que tome en cuenta de modo muy explícito los valores y la totalidad (la sociedad global y la naturaleza, la biosfera)”.

Pero se ha centrado más en el individuo, las cuestiones médicas y biomédicas, la relación paciente/médico y la experimentación sobre las personas. Se basa en cuatro principios fundamentales: la autonomía, la beneficencia, la no maleficencia y la justicia. Esta concepción, fue designada como principialista.

La “Declaración Universal sobre Bioética y Derechos Humanos” ilustra una evolución y ofrece un abanico de principios correspondientes a esta ampliación. Sitúa la bioética como una práctica multidisciplinaria y pluralista que se ejerce en el marco de una realidad humana culturalmente plural, socialmente diversa y económicamente desigual.

- **Juan Carlos Tealdi** (Asesor en Bioética de la Secretaría de Derechos Humanos de la Nación, Argentina): La introducción de campos como “Desarrollo humano y educación”, “Hambre y desnutrición”, “Identidad”, “Pobreza y necesidad”, y “Poder”, entre otros, destaca conceptos que muchas veces han sido minimizados o simplemente no tratados en la bibliografía bioética corriente.

Pensamiento crítico latinoamericano

- **Ricardo Salas Astraín.** Trata de un pensar filosófico que cuestiona las formas discursivas veladas o encubiertas, tan propias de las ideologías de la dominación, teorías hegemónicas, teorías que no vinculan sus modelos epistémicos con las estructuras vigentes del poder, tal como ocurre en la visión neo positivista o tecnocientífica.
- Pensar críticamente no es algo propio de la filosofía, sino de otras formas de racionalidad presentes en nuestras culturas latinoamericanas; lo propio de la filosofía es llevarlo a la máxima claridad de un trabajo conceptual.

Ética social, valores y liberación

- **Adriana Arpini:** Cuando hablamos de ética social nos estamos refiriendo a una dimensión de la filosofía práctica, cuya reflexión se nutre de las situaciones vividas a diario por cada uno de nosotros en tanto sujetos que interactuamos socialmente.

NOTA: esto contradice la idea de una ética global impuesta.

- **José Luis Aranguren** No solo las normas, sino también la conciencia moral se originan socialmente (Aranguren, 1968). Muchas veces ordenamos nuestro obrar de acuerdo con la moral vigente, pero otras veces no

NOTA: esto contradice la idea de una ética global impuesta.

- **Agnes Heller** utiliza el término “objetivaciones” para señalar el hecho de que las acciones humanas se exteriorizan y constituyen un mundo; es decir, se objetivan en producciones materiales y simbólicas: instituciones, leyes, conocimientos, creencias, obras de arte y de ingenio, costumbres, normas, valores.

Bioética de intervención

- **Volnei Garrafa y Dora Porto** definen la bioética fuerte o bioética dura (*hard bioethics*) pretende avanzar en el contexto internacional, a partir de América Latina, como una teoría periférica y alternativa a los abordajes tradicionales

(verificados en los llamados países centrales), principalmente el principialismo, de fuerte connotación anglosajona. A partir de la década de 1990 emergieron fuertes críticas al principialismo en el contexto de la bioética. A partir de entonces, nuevas corrientes de pensamiento empezaron a surgir en la bioética, objetivando contextualizar los problemas a las realidades concretas.

NOTA: esto muestra que hay otros discursos en ética además de los centrales representados por Unesco

- Con las discusiones y homologación de la Declaración Universal sobre Bioética y Derechos Humanos de la Unesco, cambia por completo el cuadro. La bioética, que hasta entonces tenía un direccionamiento preferencial hacia las cuestiones biomédicas y biotecnológicas, incorpora, definitivamente, los temas sociales, sanitarios y ambientales a su agenda. Es en este contexto donde surge la bioética de intervención.

NOTA: esto muestra que no es materia cerrada, sino que debe ser abierta y flexible. No cabe entonces un decreto.

Justificativas y objetivos de creación de la bioética de intervención

Surge para cubrir las falencias de la previa teoría de los cuatro principios, que es insuficiente para:

1. análisis contextualizados de conflictos que exijan flexibilidad para una determinada adecuación cultural.
2. enfrentamiento de macro-problemas bioéticos persistentes o cotidianos enfrentados por la mayoría de la población de los países latinoamericanos, con significativos niveles de exclusión social.

Los países periféricos (en particular en América Latina) sufren la despolitización de los conflictos morales. Se usa la bioética como herramienta metodológica neutral para una interpretación horizontal y aséptica de estos conflictos.

En consecuencia, se aminoran/desvanece la gravedad de las diferentes situaciones, sobre todo aquellas colectivas que acarrearán las más profundas distorsiones e injusticias sociales.

La equidad es el punto de partida de la igualdad, pero se requiere el reconocimiento de las diferencias y necesidades diversas de los sujetos sociales.

Todo esto es la bioética de intervención: colectiva, práctica, aplicada y comprometida con el “público” y con lo social en su más amplio sentido.

Bioética de los Derechos Humanos

Juan Carlos Tealdi

Se postula por primera vez el 5 de octubre de 2001 en Buenos Aires, como respuesta a la confusión y perversión de la ética que se hacía en nombre de la bioética.

Responde al fundamentalismo de los principios éticos y al imperialismo moral ejercido en su nombre, en particular en América Latina.

Ambos desconocían a la salud como un derecho humano y reducían a la justicia al rango de principio *prima facie*.

La bioética de los Derechos Humanos sostiene dos tesis básicas:

1. En origen la bioética es un campo plural de reflexión ético-normativa. Admite distintas singularidades de pensamiento y, por tanto, diversas bioéticas
2. El respeto de la moral universal de los Derechos Humanos que incluye el respeto de la diversidad cultural y lingüística.

Afirma es que los Derechos Humanos son el mínimo moral o la frontera demarcatoria entre los mundos de la moral y la inmoralidad, en modo tal que solo desde ellos es posible hoy –histórica y sociológicamente hablando– la construcción crítica y reflexiva de toda bioética.

La tradición liberal de Estados Unidos, en cambio tiene énfasis en la economía de libre mercado y acento en el individualismo.

La negación de la salud como derecho humano básico, que la Declaración de Alma-Ata(1978) de la Organización Mundial de la Salud procuró proteger, puede verse como una de las mayores expresiones de ese reduccionismo.

La bioética dominante durante los años setenta y ochenta fue:

1. de tipo clínico antes que social o enfocado a la salud pública,
2. se orientó a los problemas tecnológicos de el curar antes que a las cuestiones interhumanas de el cuidar,
3. promovió una mezcla pragmática entre valores éticos e intereses económicos antes que un verdadero entendimiento moral comunitario.

Se debe prestar atención en bioética a la ética de la pobreza, al medio ambiente y los daños para las generaciones futuras, al desarrollo de políticas de salud pública que procuren la equidad, a las poblaciones vulnerables y vulneradas, a la diversidad cultural, y a las cuestiones sociales y de responsabilidad pública, y no solo a la libertad y responsabilidad individual.

Toda ética, requiere no solo el saber, sino también, y sobre todo, dar cuenta de si miramos al mundo en el que vivimos con la voluntad o el querer comprender y actuar para cambiar una realidad indignante y por ello injusta.

Los Derechos Humanos y la bioética tienen su punto de vinculación indisoluble en la dignidad humana y en los actos reivindicativos de la misma a que nos conduce toda indignación.

La bioética de los Derechos Humanos es la postulación de una moral básica universalmente reconocida.

De allí que la principal tarea de una bioética crítica hoy es la demolición de los falsos supuestos de la bioética neo-liberal y su pretensión fáctica de convertirse en bioética global.

Globalización y mercado de la alimentación

Jorge Eduardo Rulli (Argentina)

En países productores de alimentos hay:

- 1) represión a los productores;
- 2) exportación creciente de productos transformados mediante la biotecnología, para lograr aumento de divisas;
- 3) dependencia del modelo de agro-negocios.

La lógica del modelo es el aumento constante de las exportaciones; crecimiento que va en desmedro de la producción de alimentos local.

Se agregan a este modelo intensas políticas sociales:

- 1) Para paliar la pobreza.
- 2) Para fomentar micro-emprendimientos.
- 3) Para la atención a la salud, etc.

Que en realidad no buscan resolver el tema de la exclusión y del hambre, sino perpetuarlo y contenerlo para evitar nuevos estallidos.

4) Se usa la sustentabilidad. Para comprender el uso de discursos ambientales para propósitos de dominio económico (p.ej., la soja en Argentina).

5) Es un modelo que usa la agricultura para fines que no son producir alimento.

6) Es un modelo de agricultura sin agricultores.

7) Es un modelo de exportación de insumos que vacía de población rural el campo.

8) Se enfoca en una agricultura de escala.

9) Apoya una ganadería de fábrica.

10) Abandona el modelo de seguridad alimentaria y también la antigua producción de alimentos.

El modelo agroexportador de forrajes se ha constituido en una fábrica inagotable de pobreza, fuente de desarraigo y razón de migración hacia las grandes ciudades, donde en los nuevos y crecientes conurbanos se multiplican los fenómenos de la indigencia y de la exclusión social.

Además, la soja y el maíz transgénico han desplazado a muchos otros cultivos que aportaban alimentos a la mesa de los argentinos, algunos de los cuales ahora deben ser importados.

La ética empresarial de base debe cambiar. La responsabilidad social empresarial tiene como concepto aproximadamente unos diez años de vida y viene creciendo con renovadas fuerzas.

La tradicional ética de la empresa es “lo ético es rentable, entre otras cosas, porque reduce los costos de coordinación externos e internos de la empresa: posibilita la identificación con la corporación y una motivación más eficiente”.

Incorporaron cursos de ética desde una perspectiva meramente instrumental para ponerla al servicio de un logro empresarial: la maximización de las ganancias.

Sin embargo, el discurso de la ética empresarial, respaldado por el de la sustentabilidad de ciertas prácticas agrícolas e industriales, no cuestiona los factores económicos y políticos que producen la injusticia, sino que se preocupan por las estrategias y los procedimientos empresariales.

Se debe recuperar la antigua concepción de la economía como economía política, en el sentido que la capacidad y la decisión de modificar las reglas sigue siendo un tema político.

2. Sistema de Derechos Humanos

La deshumanización es la pérdida de la capacidad axiológica, tan propiamente humana, cuando estamos rodeados de tanto poderío científico y tecnológico.

El encuentro y confrontación entre norma jurídica y norma ética, que será una condición necesaria aunque no suficiente para constituir el concepto de derechos humanos.

La idea de un sistema de normas acerca de lo justo y de lo injusto, dado por “naturaleza” y válido por tanto en todo tiempo y lugar, ha sido defendida por Platón y Aristóteles, por la doctrina cristiana del derecho natural, y posteriormente y en distintas versiones, por Hobbes, Locke y Rousseau. La tradición de la doctrina cristiana del derecho.

Durante el siglo XX los derechos humanos tomaron su significado actual más específico, como cuerpo normativo.

En 1948 la Asamblea General promulga la Declaración Universal de Derechos Humanos proclamando los derechos civiles y políticos junto a los derechos económicos, sociales y culturales.

Los derechos humanos expresan en su origen una apreciación de valores, en particular de aquellos que tienen mayor altura o densidad o peso, esto es, mayor jerarquía.

Estos valores, cuando son sustentados con la pretensión de ser derechos humanos, son postulados de un modo indiscutible y absoluto, aunque de hecho puedan ser discutidos y no reconocidos. Son postulados como imperativos categóricos si se quiere, pero no de la voluntad autónoma que se vuelve contra sí misma, sino de la conciencia que se rebela contra otras voluntades.

Los derechos humanos son exigencias a las instituciones del respeto moral de valores humanos que se consideran universalizables.

El vocablo “bioética” se usó por primera vez en 1970 para designar un nuevo campo dedicado a la sobrevivencia humana y la mejora en la calidad de vida.

Agricultura y agrotecnologías

Eduardo Freyre Roach

Bioética y agrotecnologías. Van Rensselaer Potter, al referirse a la bioética como “puente al futuro” y “ciencia de la supervivencia”, sugirió que la Tierra no se considere como una propiedad cuya relación con la misma sea estrictamente económica y en términos de privilegios, sino que implique más bien deberes y obligaciones. La “bioética global y profunda”, postula la necesidad de una ética de la preocupación por los menos privilegiados.

Este conjunto de pautas potterianas alude a los desafíos bioéticos de las agrotecnologías. El uso de productos agroquímicos en la fertilización de los suelos y en el control fito-sanitario de las plagas, y también en la atención médico-veterinaria y zootécnico-pecuaria del ganado, constituye un pilar fundamental del modelo de desarrollo agrario denominado Revolución Verde.

La valoración bioética adecuada de las agrotecnologías presupone tener en cuenta no solo sus condicionamientos intelectuales y artefactuales, sino también socio-económicos, políticos, éticos, jurídicos y culturales.

El desafío bioético fundamental de las múltiples agrotecnologías consiste en maximizar y distribuir de forma justa sus beneficios y minimizar sus riesgos.

Actualmente los desafíos bioéticos de las agrotecnologías se ven afectados por el orden socio-económico, político, científico-tecnológico, jurídico y cultural predominante, promovido por el proyecto de globalización neoliberal, que favorece a las corporaciones transnacionales agroalimentarias como Monsanto, Novartis, Aventis, Syngenta, etc., apoyadas por sus gobiernos y las organizaciones internacionales como el FM, el BM, el BID y la OMC, y legalmente auspiciadas en los marcos de los

acuerdos TRIP y los convenios integracionistas como el TLC, los acuerdos bilaterales con los Estados Unidos y el ALCA.

Modelos de producción rural

Adolfo Boy (Argentina)

Lo rural y lo urbano: sociedad y producción. Desde que la humanidad abandonó sus hábitos nómades y se estableció en feudos o ciudades, se formaron círculos o cinturones que las proveían de alimento, medicina, fibras y energía: lo rural.

En 1966 se crea el Centro Internacional de Mejoramiento de Maíz y Trigo (Cimmyt) en México y comienza a gestarse allí la Revolución Verde, cuyo modelo de producción se basó en llevar al máximo el potencial genético de los cultivos y modificar las condiciones naturales con el uso de insumos externos, fundamentalmente los fertilizantes químicos.

Características principales de este modelo son la dimensión humana de todos los factores productivos, la biodiversidad, el respeto por la naturaleza y sus ciclos, las rotaciones, el bajo uso de insumos externos y el no contemplar la venta de lo producido.

El modelo agroindustrial produce, a gran escala y bajo costo, materias primas a granel (*commodities*) en monocultivos (soja, eucaliptos, maíz, colza, algodón), con animales en confinamiento (corrales de engorde; acuicultura).

Una respuesta al modelo agroindustrial han sido los modelos agroecológicos que surgieron en la década de los ochenta basados en los modernos conocimientos de la ecología y su aplicación a la producción agropecuaria comercial para diseñar modelos de agroecosistemas con especial visión de lo social y la sustentabilidad. Caracterizan este modelo la prescindencia del uso de insumos de síntesis química, la biodiversidad, las rotaciones, el bajo empleo de insumos externos y la inserción en el mercado (Sarandón, 2002).