

Pensar la técnica desde la bioética, medio siglo después de Heidegger

Susana La Rocca¹

Maria Marta Mainetti²

Juan Pablo Issel³

Resumen

El desocultamiento restringido de la técnica moderna impide tener una relación libre con ella. Solo se desoculta lo cuantitativo, lo capaz de ser acumulado. La forma restringida y no originaria del desocultar, dataíza la naturaleza. Esa dataización implica el paso de la cosa a la no cosa, o sea al imperio del dato. La naturaleza se reduce a su dimensión registrable. En ese reduccionismo el desocultamiento no se dirige hacia la verdad, sino que se queda en la instrumentalidad. La dataización de la naturaleza es una forma devaluada de conocer. La instrumentalización de la naturaleza y del conocimiento por su reducción al manejo de datos simplifica las relaciones y no permite acceder al dominio de lo complejo, lo abierto, lo siempre velado. Es desde esta perspectiva, que pensar la técnica vuelve a ser vital para el hombre, porque la IA no debe ser producto de una forma de ser restringida sino de una forma de provocar racional, solidaria, libre y ética. La aplicación del principio de precaución es un modo de provocar que reúne esas características. Pensar la técnica que dará lugar a la IA, es mucho más que pensarla. Las conclusiones que saquemos nos dirán más de nosotros que lo que hasta ahora hemos podido registrar. La técnica no se detendrá y seguirá adelante con o sin nosotros. La ética es sin duda una directiva anticipada, que solo será efectiva si no tenemos que usarla. Si necesitamos de ella no habrá registro humano que nos garantice su cumplimiento.

Palabras clave: técnica-inteligencia artificial-bioética-Heidegger

Thinking about technique from bioethics, half a century after Heidegger

Abstract

The restricted unconcealment of modern technology prevents having a free relationship with it. Only the quantitative is revealed, what is capable of being accumulated. The restricted and non-original form of revealing dataizes nature. This dataization implies the passage from the thing

¹ Universidad Nacional de Mar del Plata. E-mail: laroccasusana@yahoo.com.ar

² Universidad Nacional de Mar del Plata. E-mail: mmmainetti@gmail.com

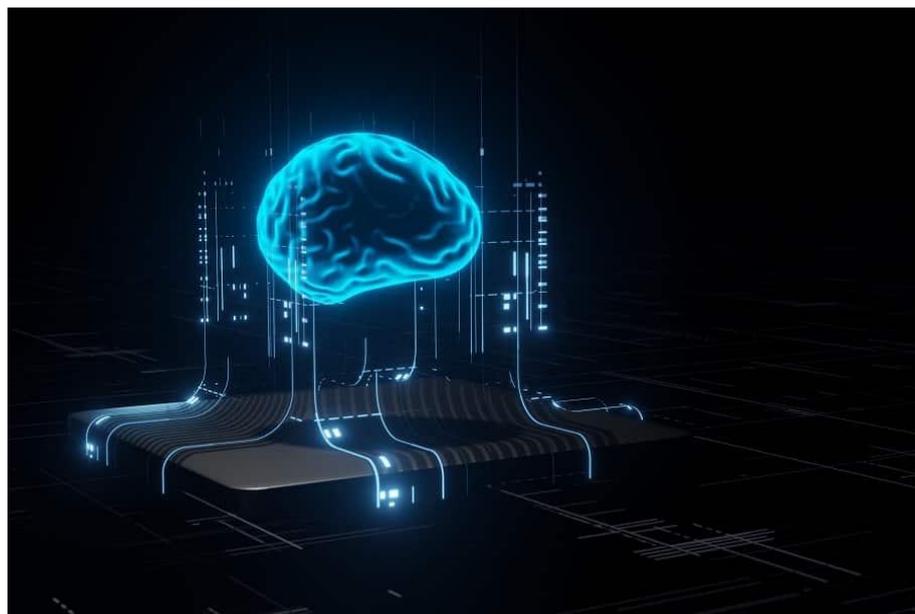
³ Universidad Nacional de Mar del Plata. E-mail: juanpabloissel@yahoo.com.ar

to the non-thing, that is, to the empire of data. Nature is reduced to its recordable dimension. In this reductionism, unconcealment is not directed towards the truth, but remains in instrumentality. The dataization of nature is a devalued way of knowing. The instrumentalization of nature and knowledge by its reduction to data management simplifies relationships and does not allow access to the domain of the complex, the open, the always veiled. It is from this perspective that thinking about technique is once again vital for man, because AI should not be the product of a restricted way of being but of a rational, supportive, free and ethical way of provoking. The application of the precautionary principle is a way of provoking that has these characteristics. Thinking about the technique that will give rise to AI is much more than thinking about it. The conclusions we draw will tell us more about ourselves than what we have been able to record up to now. The technique will not stop and will continue with or without us. Ethics is definitely an advance directive, which will only be effective if we don't have to use it. If we need it, there will be no human record that guarantees its fulfillment.

Keywords: technique-artificial intelligence-bioethics-Heidegger

La pregunta abre un camino. Por eso, es prudente prestar atención ante todo al camino y no permanecer apegados a frases y títulos aislados. El camino es un camino del pensar.

Martín Heidegger, *La pregunta por la técnica.*



Si la humanidad es en verdad un único sistema de procesamiento de datos, ¿cuál será el resultado? Los dataístas dirían que el resultado será la creación de un sistema de procesamiento de datos nuevo y más eficiente, el llamado Internet de Todas las Cosas. Cuando dicha misión se cumpla, Homo sapiens desaparecerá.

Yuval Noah Harari, *Homo Deus*, p. 414, *Debate* 2016

Introducción

Heidegger dice en su conferencia del año 1953 que cuando consideramos a la técnica como algo neutral, representación que se aplaude hoy con gusto, este posicionamiento nos vuelve completamente ciegos para la esencia de la técnica.

La correcta determinación instrumental de la técnica, que implica considerarla un medio y un hacer humanos para un fin, no nos muestra aún su esencia ya que no es necesariamente verdadera y por eso todavía no nos permite responder a la pregunta por su esencia. Ambas preguntas y respuestas quedan dentro de la concepción instrumental de la técnica, pero fuera de su esencia y en tanto incompletas, no pueden ser consideradas verdaderas. Para abordar el tema de la esencia, es necesario entablar una relación verdadera con la técnica y no simplemente instrumental. Para lograr adentrarse a elucidar la esencia de la técnica, Heidegger

recurrirá a la clasificación de causas, formulada por Aristóteles:

La filosofía enseña desde hace siglos que hay cuatro causas: 1. la causa material, el material, la materia, con la que se prepara, por ejemplo, una copa de plata; 2. la causa formal, la forma, la figura, en la que se introduce la materia; 3. la causa final el fin, por ejemplo, el sacrificio, por el cual la copa requerida es determinada según materia y forma, y 4. la causa efficiens, que produce el efecto, la copa real hecha, el platero. Lo que sea la técnica, concebida como medio, se hará patente si retrotraemos lo instrumental a la cuádruple causalidad. (2021, p.116)

Siguiendo la pesquisa por la esencia de la técnica Heidegger sostiene que el término *aitia* o *aition*, que habitualmente se traduce como causa, significa en griego *ser responsable de*. Por ello cada cosa adeuda a las distintas causas ciertas determinaciones que hay que elucidar. La copa del sacrificio adeuda a la plata su materia, a la idea o forma su ser copa y no anillo y sobre todo adeuda su *para que*, su

telos que puede traducirse como meta o su fin. Estas tres causas son corresponsables de la copa del sacrificio puesta aquí como ejemplo. (2021, p.117)

Finalmente, hay una cuarta causa representada por el orfebre que reúne a las otras y las sobrepone en el producir. Ese producir involucra y contiene la materia, la idea y el fin, determinaciones que jamás pueden estar ocultas o veladas. Al contrario, el producir completo de la técnica va desde un posible velamiento (reducción de la técnica a su modo instrumental), hasta un develamiento, desocultamiento (*aletheia*) que nos anoticia de su materia, su forma y su telos y nos pone en presencia de la esencia de la técnica, es decir de su verdad.

La técnica puede definirse entonces, no simplemente como un medio sino como un modo de desocultar. En cada objeto producido, su telos debe ser develado por quien lo produce. En ese telos estará también correlacionada su forma y la materia con que deberá ser hecho, pero el orfebre no puede desconocer el completo desocultamiento que, en tanto proceso permanente, permite que el producir sea poiesis.

Lo dis-puesto, en cuanto provocación en el establecer, destina en un modo del desocultar. Lo dis-puesto es una destinación del destino, como toda manera de desocultar. Destino, en el sentido mencionado, es también el producir, la poiesis. El desvelamiento de lo

que es siempre va sobre un camino del desocultar. Siempre impera al hombre el destino del desocultamiento. Pero no es jamás la fatalidad de una coacción. Pues, precisamente el hombre llega a ser libre en tanto que pertenece al ámbito del destino y, así, llega a ser un oyente, no un esclavo. (Heidegger, 2021, p.137)

La técnica moderna incorpora lo cuantitativo como novedad y se convierte en un imperativo técnico de dominación.

De igual manera, el desvelamiento según el cual la naturaleza se concibe como una conexión de efectos de fuerzas calculables, puede permitir, ciertamente, constataciones exactas; pero, precisamente, a través de estos resultados persiste el peligro de que en todo lo exacto se retraiga lo verdadero. El destino del desocultamiento no es en sí un peligro cualquiera, sino el peligro. (Heidegger, 2021, p.137)

Ge-stell (lo dispuesto) llama Heidegger a aquello que sustenta el imperativo técnico del desocultamiento y la provocación interpelante de la naturaleza. Lo *Ge-stell* es el presupuesto ontológico que se intercala entre el ser humano y la realidad, y es la determinación que predispone al hombre a provocar a la naturaleza. El ser humano se caracteriza por este carácter provocador que constituye la esencia de la técnica moderna. La llamada de la técnica es un imperativo y la libertad

del hombre se manifiesta en descubrir que eso no es técnico sino un llamado del ser.

El hombre debe revisar la falsa conciencia de creerse el *señor de la Tierra* y buscar un modo auténtico de habitar la tierra. Aclarar la esencia de la técnica en tanto imperativo no reducido a su carácter instrumental es sinónimo de buscar un modo de desocultar más primigenio y que evite la caída en la técnica instrumental

La técnica instrumental en su máxima expresión

Por su parte Byung- Chul Han nos dice que la violencia que el ser humano ejerce contra la naturaleza y que Heidegger llamaba provocar, se está volviendo contra él con más fuerza. Parte de esa violencia es convertir a la naturaleza en un conjunto de datos a los que el filósofo coreano ha denominado dataísmo (Rendueles:2020) y lo califica como una forma pornográfica de conocimiento que anula el pensamiento produciendo una hipercomunicación, consecuencia de la digitalización que no se corresponde con mayor comunicación y que, a su vez, trae aparejado la desaparición de la comunidad tal como la entendíamos.

Por ello se hace necesario inventar nuevas formas de acción y juegos colectivos que se realicen más allá del ego, el deseo y el consumo, y que creen comunidad., También redefinir el concepto de libertad

desde la comunidad. Dice Byung'-Chull Han que:

Freiheit, la palabra alemana para "libertad", significa en origen "estar con amigos". "Libertad" y "amigo" tienen una etimología común. La libertad es la manifestación de una relación plena. Por tanto, también deberíamos redefinir la libertad a partir de la comunidad. (Rendueles, 2020)

El autor coreano considera además que:

El sujeto neoliberal como empresario de sí mismo no es capaz de establecer con los otros relaciones que sean libres de cualquier finalidad. Entre empresarios no surge una amistad sin fin alguno. Sin embargo, ser libre significa estar entre amigos. "Libertad" y "amigo" tienen en el indoeuropeo la misma raíz. La libertad es, fundamentalmente, una palabra relacional. Uno se siente libre solo en una relación lograda, en una coexistencia satisfactoria. El aislamiento total al que nos conduce el régimen liberal no nos hace realmente libres. (Byung'-Chull Han, 2014, p.13)

Tanto la redefinición del concepto de libertad como la necesidad de buscar un modo más original de desocultar de la técnica moderna ya fueron anticipados como elementos necesarios en la argumentación que Heidegger desplegó en la *Pregunta por la técnica*.

El tema de la libertad se impone como marco para poder seguir avanzando en el tema de la esencia de la técnica y el estatuto

ontológico y moral de los entes producidos por ella, entre otras cosas, el tema de la inteligencia artificial (IA).

El dataísmo, resulta así una forma devaluada del pensamiento, que lo reduce al cálculo y le quita además su aspecto erótico, como bien lo expresaba Heidegger, cuando lo compara con el batir de alas de Dios Eros. “El batir de alas del dios Eros lo acariciaba cada vez que daba un paso significativo en el pensamiento y se atrevía a aventurarse en un terreno inexplorado”. (Heidegger, 2008, p.270)

El dataísmo es el lenguaje de las máquinas, pero no existe un pensamiento basado solo en los datos. Y la transparencia atribuida a los datos, es una pretensión insostenible frente a un mundo siempre velado por la inconmensurabilidad del ser. El mundo no solo es más profundo de lo que pensamos, sino que además es lo opuesto a la transparencia del dataísmo.

En el dataísmo las cosas se convierten en no-cosas, desaparecen como en la distopía de Ogawa (2021) en la que el mundo se vacía sin cesar y al final desaparece. Hay una disolución progresiva e incluso desaparecen partes del cuerpo. Al final, solo voces sin cuerpo flotan sin rumbo en el aire.

Dice Byung Chul Han que la isla mencionada por Ogawa, puede compararse a nuestro presente en algunos aspectos

donde también el mundo se vacía de sosa y se llena de de una información tan inquietante como las voces sin cuerpo.

La digitalización desmaterializa y descorporeiza el mundo. También suprime los recuerdos. En lugar de guardar recuerdos, almacenamos inmensas cantidades de datos. Los medios digitales sustituyen así a la policía de la memoria, cuyo trabajo hacen de forma no violenta y sin mucho esfuerzo. (*Byung-Chul Han, 2021, p32.*)

La información, es decir, las no-cosas, recubre las cosas y las hace desaparecer en tanto las sustituye reduciéndolas en profundidad e intensidad, cuerpo y volumen. La realidad de las no cosas es una realidad devaluada que posibilita convertir las cosas en infómatas (*Byung-Chul Han, 2021, p.9*). La infomanía que convierte al ser humano en infómano es el resultado de la informatización del mundo y del dataísmo.

En esta comunicación digital, muchos de los aspectos del otro también comienzan a desaparecer, el rostro, la mirada, la presencia física. Byung Chul-Han cree que “la desaparición del otro es un acontecimiento dramático”. (2021, p.71)

El Phonosapiens es quien toca con los dedos su Smartphone y tiene la «libertad de la yema de los dedos». Los continuos toqueteos y deslizamientos sobre el smartphone le dan la sensación de tener al mundo bajo control. El smartphone

refuerza así el egocentrismo. Al tocar su pantalla, someto el mundo a mis necesidades. El mundo parece estar digitalmente a mi entera disposición. El tacto del dedo índice hace que todo sea consumible. Todo lo que toca adquiere forma de mercancía. Hasta su condición de cosa pasa a un segundo plano: miramos a través de ellos. (2021, p.23)

Esta informatización del mundo nos ofrece un trueque sin duda desventajoso. Si bien nos ahorra mucho trabajo, recibe a cambio no solo la dación voluntaria de nuestros datos (lo que permite nuestra vigilancia y control) sino la permanente pérdida de autonomía frente al desarrollo incesante de la Inteligencia artificial.

Este canje es ilegítimo en tanto reduce el pensamiento a su carácter calculador y le sustrae su carácter analógico. Dice Byung Chul-Han:

El pensamiento es un proceso resueltamente analógico. Antes de captar el mundo en conceptos, se ve afectado por él. Lo afectivo es esencial para el pensamiento. La primera afectación del pensamiento es la carne de gallina. La inteligencia artificial no puede pensar porque no se le pone la carne de gallina. Le falta la emoción que los datos y la información no pueden comportar. (2021, p.56)

Nuestro mundo dataizado no tiene ojos, no nos mira y por ello pierde su alteridad. “La frialdad de las máquinas hace desaparecer el calor de las cosas.”

(p.69) La digitalización quita a las cosas cualquier materialización rebelde, cualquier resistencia a su descosificación. Resulta por lo menos extraño hacer este reclamo, cuando el anterior fue, no es bueno cosificar y hoy lo bueno es impedir la descodificación.

Eric Sadin (2020) analiza el mundo del dataísmo y la hipercomunicación y considera que estamos frente a un ideal de civilización muy definido: el de una humanidad aumentada, en la que la propuesta técnico-económica ha venido a resolver todos nuestros problemas y a complementar nuestra finitud. En esta trama, la “siliconización del mundo” (resultado de aceptar el modelo técnico-económico de Silicon Valley) permite una organización social algorítmica automatizada. Esta forma de organización posibilita que la vida social pueda ser gobernada por una administración digital. Esto ya no es futurología sino presente concreto donde la Inteligencia Artificial formula un discurso automatizado, robótico y no-humano que parece tomar decisiones.

Por supuesto que este modelo civilizatorio de reemplazo de acciones humanas por autómatas programados tiene además otro objetivo evidente: el uso de nuestros datos para actividades comerciales que supongan dinero y poder para quienes manejan esos reservorios.

Dice Sadin que la siliconización del mundo, es la forma más exacerbada de tecnocapitalismo que nos conduce a una forma de vida que no tiene un fin político, sino mercantilista y global, destinado únicamente a satisfacer intereses privados.

El ser humano es un ser complejo, multidimensional y multisensorial y el proyecto de siliconización es un reduccionismo político, social y por ello antihumano.

Si no organizamos urgentemente una movilización ciudadana de quiebre y resistencia, dice Sadin, después será demasiado tarde.

En *La realidad aumentada* (2017), Sadin considera que la actual situación en relación a la inteligencia artificial inaugura una racionalidad que es a-normativa, a-política, sin sujeto reflexivo y que se alimenta de datos infra-individuales. Datos anónimos, que son insignificantes por sí mismos (por eso aceptamos ponerlos en circulación todo el tiempo) y que solo significan cuando son puestos en relación por una IA.

Asistimos a la emergencia de una gubernamentalidad algorítmica que pilotea numerosas situaciones colectivas e individuales. Una fuerza a-disciplinaria, una administración soft de la vida orientada a captar lo posible en lo actual. Una mutación fundamental del sujeto racional

occidental que deja su lugar a un ser formado por su más pura exterioridad y sus relaciones.

Para algunos autores, que el peligro de la técnica se halla en un particular modo de desocultar que demanda que el hombre tenga con ella una relación libre y compleja, no parece ser problema. Si bien no nos vamos a detener en sus argumentos, hay un hilo conductor en su defensa de la IA y es que los hombres no poseen realmente la capacidad de tomar decisiones libres y por tanto la reducción de los comportamientos a relaciones de manejo computacional de datos, no es más ni menos que lo que venimos haciendo desde siempre.

Stephen Hawking y Leonard Mlodinow en *El gran diseño* (2013), sostienen que todos nuestros comportamientos son biológicos y están regidos por las leyes de la física y la química y que, por lo tanto, están tan determinados como las órbitas planetarias. Por lo tanto, es difícil imaginar cómo podría operar el libre albedrío si nuestro comportamiento está determinado por las leyes físicas, de manera que parece que no somos más que máquinas biológicas y que el libre albedrío es solo una ilusión.

Yuval Noah Harari (2016), filósofo israelí, experto en temas de IA también adhiere en este argumento, afirmando que

muchas de las decisiones que consideramos libres, en realidad no lo son.

Si éste fuera el punto de partida, que no compartimos, las preguntas y argumentos que se han expuesto hasta ahora y que hemos estado analizando, no tendrían sentido. Los mencionamos porque seguramente en ellos abrevan quienes defienden a ultranza que la IA no acarrearía los problemas que hemos venido señalando.

Avatares de la IA

Desde el principio del siglo pasado, Alan Turing (1985) comenzó a pensar en una máquina que pudiese conversar con los seres humanos. Propuso para ello su célebre e hipotético experimento en donde fuera posible a partir de la imitación del lenguaje humano replicar los estados mentales y explicarlos. Desde esa presunción, el paradigma computacional, sostiene que la forma más eficiente de adquirir conocimiento es reproducir cada vez con mayor eficiencia la reproducción de las reglas de procesamiento de información sin el ruido de las historias y subjetividades de los humanos.

Por su parte, John Searle (1985) distingue entre la Inteligencia Artificial Débil, que análoga algunos de los procesos de las computadoras al entendimiento humano, pero de ninguna manera los iguala en tanto no poseen

estados intencionales y la Inteligencia Artificial Fuerte que sí otorga esa igualdad entre personas y máquinas. Para probar la imposibilidad de la inteligencia artificial fuerte Searle delinea el experimento mental conocido como la habitación china o la pieza china donde intenta comprobar que las respuestas algorítmicas no implican comprensión sino solo procesamiento, puesto que las máquinas no tienen capacidad semántica ni intencionalidad.

La respuesta a todos estos cuestionamientos no es solo ontológica y gnoseológica, sino que también antropológica y ética. Nos centraremos en esta última, sin descartar las otras dimensiones que nos interrogarán por añadidura.

La roboética es una disciplina que estudia la programación de la Inteligencia artificial teniendo en cuenta los comportamientos éticos. El principal dilema a resolver es poder determinar si la autonomía para el comportamiento ético de la máquina es atribuible al programador o a la misma máquina. Para que realmente pudiesen tener comportamientos éticos, las máquinas tendrían que ser libres y poder elegir entre cumplir o no cumplir con las reglas de programación.

Por ahora no parece factible que las máquinas valoren y decidan en relación a

esos valores. La libertad implica también la posibilidad de incumplir lo normado y valorado. Crear una máquina con esa capacidad, puede poner a los seres humanos en gran riesgo, ya sea porque las máquinas respondan a su creador, ya sea que se independicen de él con consecuencias imponderables.

Nick Bostrom (2016), señala que una superinteligencia podría ser capaz de extinguir la raza humana. Esta inteligencia artificial súper avanzada podría desarrollar sus propios planes y decidir, por sí misma, que los valores morales o éticos del ser humano no son relevantes para el ejercicio de sus funciones. Incluso, podría ser capaz de desobedecer las órdenes con las que haya sido programada, si ello resulta más beneficioso para su objetivo. Dicho de otra manera, una inteligencia artificial de este calibre podría ser incontrolable, aunque se le introdujeran conceptos éticos, y podría dar lugar a consecuencias involuntarias e impredecibles.

Por este motivo el libro de Bostrom (2016) advierte acerca de los caminos hacia esa superinteligencia, los peligros que implicaría la trasgresión o desobediencia a las leyes de la robóticaⁱ preanunciadas para Isaac Asimov (1989) y también las estrategias que debemos imaginar los humanos ante la superinteligencia.

Si logramos construir una superinteligencia es muy probable que sea el último invento de la humanidad ya que, siendo tan bueno, no necesitemos otro o siendo tan malo, represente el fin de la humanidad. En base a estos problemas se han desarrollado varias guías éticas para el uso responsable de la Inteligencia artificial.

El Parlamento europeo publicó en octubre de 2020 la *Guía Ética para el uso responsable de la Inteligencia Artificial*. En ella especifica los estándares morales a los que debería someterse el desarrollo de la inteligencia artificial, a saber:

- La inteligencia artificial ha de estar enfocada en el ser humano y sus problemáticas. Su objetivo de trabajo nunca puede ser otra inteligencia artificial.
- Los sistemas de inteligencia artificial han de estar programados para prestar especial atención a los grupos más vulnerables o desfavorecidos, caso de las personas con discapacidad, los menores de edad o aquellos que vivan en entornos de riesgo.
- La inteligencia artificial ha de estar programada para comprender y respetar todos aquellos derechos fundamentales del ser humano, así como para cumplir la legislación aplicable en los diferentes países.

- Debe ser fiable y robusta, garantizando la máxima seguridad en su aplicación, minimizando al máximo los fallos de seguridad y los problemas de operativa. Para ello, es imprescindible comprobar su funcionamiento en entornos de prueba hasta que se pueda garantizar su uso seguro en un entorno real.
- Además, la inteligencia artificial ha de funcionar de forma transparente. Se ha de informar sobre cuáles son sus metas, y nunca se podrá emplear con otros fines distintos a los que fue desarrollada, y menos aún de forma secreta u opaca.
- Por último, y en respeto a los derechos fundamentales del ser humano, la inteligencia artificial ha de respetar en todo momento la libertad humana, no pudiendo tener una autonomía propia que impida a la persona ejercer la suya.

La *Declaración de Barcelona para un desarrollo y uso adecuados de la inteligencia artificial en Europa* detalla 6 puntos fundamentales a la hora de unir inteligencia artificial y ética:

- Prudencia: hay que ser conscientes de que todavía existen muchas problemáticas y frentes abiertos a la hora de implementar la inteligencia

artificial. Por tanto, a la hora de resolver todas estas cuestiones científicas y técnicas es fundamental tener en cuenta el sentido común.

- Fiabilidad: los sistemas de inteligencia artificial han de ser robustos y fiables. Toda inteligencia artificial debe ser evaluada en entornos de prueba y garantizar un óptimo funcionamiento antes de plantear su uso en un entorno real.
- Rendición de cuentas: Las personas afectadas por las decisiones que toma una inteligencia artificial han de ser informadas sobre las razones y procesos que han llevado a dicha decisión, además de las vías para realizar reclamaciones en caso de que las consecuencias sean contrarias a sus intereses.
- Responsabilidad: toda persona tiene derecho a saber si el responsable de las decisiones que le afectan es un ser humano o una inteligencia artificial.
- Autonomía limitada: es necesario establecer reglas que limiten la aplicación de los sistemas de inteligencia artificial, de manera que el ser humano sea capaz de controlarlos y aplicarlos cuando sea conveniente.

- Papel del ser humano: uno de los principios más importantes y a la vez más complejos será determinar el papel que juegan las personas afectadas por la puesta en marcha de una inteligencia artificial.

Por su parte, la Unesco publicó en 2022 la *Recomendación sobre la ética de la inteligencia*. En el texto se establecen valores y principios comunes que guiarán la construcción de la infraestructura jurídica necesaria para garantizar un desarrollo saludable de la IA.

Los objetivos de la presente Recomendación son los siguientes:

- a) proporcionar un marco universal de valores, principios y acciones para orientar a los Estados en la formulación de sus leyes, políticas u otros instrumentos relativos a la IA, de conformidad con el derecho internacional;
- b) orientar las acciones de las personas, los grupos, las comunidades, las instituciones y las empresas del sector privado a fin de asegurar la incorporación de la ética en todas las etapas del ciclo de vida de los sistemas de IA;
- c) proteger, promover y respetar los derechos humanos y las libertades fundamentales, la dignidad humana y la

igualdad, incluida la igualdad de género; salvaguardar los intereses de las generaciones presentes y futuras; preservar el medio ambiente, la biodiversidad y los ecosistemas; y respetar la diversidad cultural en todas las etapas del ciclo de vida de los sistemas de IA;

d) fomentar el diálogo multidisciplinario y pluralista entre múltiples partes interesadas y la concertación sobre cuestiones éticas relacionadas con los sistemas de IA;

e) promover el acceso equitativo a los avances y los conocimientos en el ámbito de la IA y el aprovechamiento compartido de los beneficios, prestando especial atención a las necesidades y contribuciones de los países de ingreso mediano bajo, incluidos los PMA, los PDSL y los PEID (los países menos adelantados (PMA), a los países desarrollados sin litoral (PDSL) y a los pequeños Estados insulares en desarrollo (PEID)

También la Recomendación tiene como objetivos hacer realidad las ventajas que la IA aporta a la sociedad y reducir los riesgos que conlleva. Por eso exhorta enfáticamente acerca de:

1. Protección de datos

La Recomendación pide que se actúe más allá de lo que hacen las empresas tecnológicas y los gobiernos para garantizar

a las personas una mayor protección, asegurando la transparencia, la capacidad de actuar y el control de sus datos personales. Afirma que todos los individuos deberían poder acceder a sus registros de datos personales o incluso borrarlos. También incluye acciones para mejorar la protección de los datos y el conocimiento y derecho del individuo a controlar sus propios datos. También aumenta la capacidad de los organismos reguladores de todo el mundo para hacerla cumplir.

2. Prohibición de los marcadores sociales y la vigilancia masiva

La Recomendación prohíbe explícitamente el uso de sistemas de IA para la calificación social y la vigilancia masiva. Este tipo de tecnologías son muy invasivas, vulneran los derechos humanos y las libertades fundamentales y se utilizan de forma generalizada. La Recomendación subraya que, a la hora de desarrollar marcos normativos, los Estados Miembros deben tener en cuenta que la responsabilidad última y la rendición de cuentas deben recaer siempre en los seres humanos y que no se debe otorgar personalidad jurídica a las tecnologías de IA por sí mismas.

3. Ayudar a supervisar y evaluar

La Recomendación también sienta las bases de las herramientas que ayudarán a su aplicación. La Evaluación del Impacto Ético pretende ayudar a los países y a las

empresas que desarrollan y despliegan sistemas de IA a evaluar el impacto de esos sistemas en las personas, la sociedad y el medio ambiente. La metodología de evaluación del grado de preparación ayuda a los Estados Miembros a evaluar su grado de preparación en términos de infraestructura jurídica y técnica. Esta herramienta ayudará a mejorar la capacidad institucional de los países y a recomendar las medidas adecuadas que deben adoptarse para garantizar la aplicación de la ética en la práctica. Además, la Recomendación anima a los Estados Miembros a considerar la posibilidad de añadir el papel de un funcionario independiente de ética de la IA o algún otro mecanismo para supervisar los esfuerzos de auditoría y seguimiento continuo.

4. Protección del medio ambiente

La Recomendación subraya que los actores de la IA deben favorecer métodos de IA eficaces en cuanto a datos, energía y recursos que ayuden a garantizar que la IA se convierta en una herramienta más destacada en la lucha contra el cambio climático y en el tratamiento de los problemas medioambientales. La Recomendación pide a los gobiernos que evalúen el impacto medioambiental directo e indirecto a lo largo del ciclo de vida del sistema de IA. Esto incluye su huella de carbono, el consumo de energía y el

impacto ambiental de la extracción de materias primas para apoyar la fabricación de tecnologías de IA. También pretende reducir el impacto medioambiental de los sistemas de IA y las infraestructuras de datos. Incentiva a los gobiernos a invertir en tecnología verde, y si hay un impacto negativo desproporcionado de los sistemas de IA en el medio ambiente, la Recomendación instruye que no se utilicen.

Conclusión

Si bien es posible que el desarrollo de la IA pudiera devenir en un beneficio a la humanidad en tanto puede permitirle optimizar cuestiones relativas a su quehacer en el mundo, también es cierto que los datos con los que contamos hasta ahora se inscriben en un contexto de incertidumbre que no nos es dado valorarlo positivamente.

Las cuestiones éticas constituyen uno de los problemas más complicados a los que debe enfrentarse la IA. La característica de los sistemas basados en algoritmos demanda definir valores y determinar sus preferencias, cosa muy difícil después de los sucesivos fracasos de universalizar la ética y de reconocer un pluralismo ético que permita la diferencia de perspectiva en relación a un contexto y a circunstancias particulares.

El otro gran problema es poder determinar el carácter ontológico de la IA.

El reconocimiento de su posible autonomía implicaría su independencia de los humanos. Ambas cuestiones modificarían absolutamente nuestra relación con la técnica y su producción.

Frente a este estado de cosas, se nos impone desde la ética el reconocimiento del principio de precaución al que debemos someternos en contextos de incertidumbre. Y el que hemos venido abordando sin duda lo es.

El principio de precaución ha sido definido como un principio procedimental llamado a potenciar la evaluación de riesgos inciertos y posibilitar la adopción de medidas frente a los mismos aun cuando éstos se desconocen en gran medida. (Sánchez Caro; García Sánchez, 2021)

Hans Jonas también se posiciona antes las prácticas de intervención biogenética/tecnológicas y sostiene el riesgo de manipulaciones erróneas, el primer mandato moral es la actitud de cautela y el pensamiento hipotético que examine el uso eventual o posible de esas capacidades y sus consecuencias (1997)

Tenemos que volver a temer y a temblar e, incluso sin Dios, a respetar lo sagrado. Hay tareas suficientes a este lado del límite que esto establece. El estado del hombre clama constantemente por su mejora. Intentemos ayudar. Intentemos prevenir, aliviar y curar.

Pero no intentemos ser creadores en la raíz de nuestra existencia, en la sede primigenia de su secreto” (Jonas, 1995, p.143)

El principio de precaución se traduce, en la práctica, en hacerse cargo de la prueba a quien pretende realizar una actividad técnica que pudiese comprometer el destino de la comunidad y /o de la humanidad. Es quien programa la IA quien debe acreditar que no será maleficente.

Habermas en su libro *El futuro de la naturaleza humana ¿Hacia una eugenesia liberal?* (2002) pregunta en ese caso respecto a la manipulación de la composición del genoma humano, pero podemos aplicarlo análogamente al tema de la IA si no es imprescindible revisar la “abstención fundamentada”, tan propiciada por el pensamiento liberal postmoderno. Algunas de las respuestas que da para justificar el compromiso de la ética y la filosofía en la intervención, del hombre en la vida biológica podemos usarlas para pensar acerca de la técnica de la IA

- La técnica de la IA otorga a los vivos un poder sobre lo que vendrá y eso demanda el principio de responsabilidad y el de precaución. No sabemos que puede resultar del desarrollo indefinido, en donde solo se patentice la causa eficiente y la material.
- Este desarrollo sin *telos* es la servidumbre futura de los vivos frente a los muertos.
- Esta circunstancia colisiona contra el principio de reciprocidad que permite reconocernos en tanto miembros de la especie, como seres de igual condición y repercute en nuestra comprensión de lo que la especie representa.
- La idea de humanidad obliga al reconocimiento de los otros, desde la que todos puedan verse como miembros de una comunidad inclusiva, la especie.
- Cada persona, para ser sí misma, requiere una auto-apropiación de su propia historia; la manipulación de la IA afecta la auto-comprensión de quien, por así decirlo, es “programado”.
- Si la tecnificación de la naturaleza humana transforma el modo en que la especie humana y los individuos que la componen se ven a sí mismos, en cuanto sujetos autónomos y responsables (juicios de hecho) se debe intervenir, (juicios imperativos del deber ser) custodiar una generalidad antropológica amenazada y garantizar las condiciones del ser sí mismo y su configuración biográfica como base última de las comunidades de diálogo y de las sociedades democráticas modernas (Habermas, 2002,p. 31).

En ese mismo sentido, la *World Commission* de UNESCO brindó en el año 2005 un esclarecimiento del principio de precaución que establece:

en circunstancias en que la acción humana puede causar un daño éticamente inaceptable –serio, grave o irreversible- a la vida humana o a la salud, a las generaciones actuales o futuras, otros seres vivos o al medio ambiente y siendo ese daño científicamente posible o razonable aunque incierto, es un deber tomar las acciones necesarias para evitarlo o para disminuirlo.

Para el COMEST es importante aclarar tanto lo que el principio de precaución es y lo que no es para evitar que un incorrecto recurso al mismo provoque precisamente lo que el principio no persigue:

Para evitar equívocos y confusiones, es útil extender se sobre lo que el principio de precaución no es. El principio de precaución no se basa en ‘riesgo cero’ sino que apunta a lograr que haya menos riesgos o contingencias o que sean más aceptables. No obedece a la ansiedad o a la emoción, sino que constituye una norma de decisión racional, basada en la ética, y que se propone utilizar lo mejor de las ‘ciencias de los sistemas’ de procesos complejos para adoptar las decisiones más

razonables. En definitiva, como cualquier otro principio, el principio de precaución en sí, no es un algoritmo de decisión y por ende no puede garantizar la coherencia entre los casos. Al igual que en los asuntos que se ventilan ante los tribunales, cada caso será algo diferente, pues tendrá sus propios hechos, puntos de incertidumbre, circunstancias, y responsables de la adopción de decisiones existiendo siempre una cuota de subjetividad que no podrá eliminarse. (p 14).

Para terminar, creemos que el aporte de la ética en el tema de la IA, implica superar diferentes giros filosóficos, el de la conciencia, el de la comunicación e incluso el de la perspectiva multicultural, para entrar en el giro *inclusivo* de la ética. Giro que pone el acento más en la responsabilidad y la precaución que en el desarrollo acrítico de la técnica.

El giro inclusivo de la ética supone pensar toda la vida, no solo la humana, desde la posibilidad de seguir siendo, sin dominios absolutos, sin reduccionismos, sin pérdidas de derechos y sin extinción causal de todo tipo de vida.

Como en la reciente pandemia de Covid, el lema ha de ser salvarnos todos. La IA no puede contribuir al dominio o a la extinción de nadie. Todo lo que hagamos en menos nos pasará factura, aunque no quede nadie para pagarla.

NOTAS

ⁱ Las tres leyes de la robótica fueron enunciadas por el escritor Isaac Asimov en el relato titulado “Círculo vicioso”, publicado en 1942. Estas tres leyes dicen así: 1. Un robot no hará daño a un ser humano o, por inacción, permitirá que un ser humano sufra daño. 2. Un robot debe cumplir las órdenes dadas por los seres humanos, a excepción de aquellas que entren en conflicto con la primera ley. 3. Un robot debe proteger su propia existencia en la medida en que esta protección no entre en conflicto con la primera o con la segunda ley.

Referencias bibliográficas

- Asimov, I. (1989). *Círculo vicioso. Los robots*. trad. Domingo Santos. Martínez Roca.
- Bostrom N. (2016). *Superinteligencia. Caminos, peligros, estrategia*. Tell.
- Jonas, H. (1995). *El principio de responsabilidad. Ensayo de una ética para la civilización tecnológica*, Herder
- Habermas, J. (2002). *El futuro de la naturaleza humana. ¿Hacia una eugenesia liberal?*, Paidós
- Han, B.C.(2021). *No-Cosas: quiebras del mundo de hoy*. Taurus.
- Han,B.C (2014). *Psicopolítica*, HERDER
- Harari, Y.N (2016). *Homo Deus*, Debate
- Heidegger, M. (2021). *La pregunta por la técnica*. Herder
- Heidegger, M. (2008). *Alma mía, Cartas a su mujer Elfride 1915-1970*, Manantial
- Hawking, S; Mlodinow, L. (2013). *El gran diseño*. Booket
- Ogawa, Y. (2021). *La Policía de la Memoria*, Tusquet
- Rendueles,C. Entrevista a BYUNG-CHUL HAN, diario El País , 15-5-2020. Extraído de:
https://elpais.com/cultura/2020/05/15/babelia/1589532672_574169.html
- Declaración sobre Inteligencia artificial, robótica y sistemas “autónomos” del Grupo Europeo sobre Ética de la Ciencia y las Nuevas Tecnologías. Extraído de:
<https://www.researchgate.net/publication/332061758> Traducción de la Declaración sobre Inteligencia artificial robotica y sistemas autonomos del Grupo Europeo sobre Etica de la Ciencia y las Nuevas Tecnologias/citation/download

Sadin, E. (2017). *La realidad aumentada, La administración digital del mundo*. Caja Negra.

Sadin, E (2020). *La silicolonización del mundo. La irresistible expansión del liberalismo digital*. Caja Negra.

Searle J. (1985). *Mentes, cerebros y programa*. Cátedra

Sánchez Caro, J. Abellán-García Sánchez (2021) *Coordinadores: Inteligencia Artificial en el campo de la Salud. Un nuevo paradigma: aspectos clínicos, éticos y legales*. Fundación Merck Salud.

Turing, A. (1985). *Mentes y máquinas*, Tecnos

Unesco, (2005). *Informe del grupo de expertos sobre el principio precautorio*

Fecha Recepción: 26-08-2022

Fecha Aceptación: 26-10-2022