

EL PROBLEMA DE LA ÉTICA: EL CONVIVIO. ÉTICA E INTELIGENCIA ARTIFICIAL

FECHA DE RECEPCIÓN: 29-12-23 / FECHA DE ACEPTACIÓN: 27-02-24

Carlos Eduardo Maldonado

PROFESOR TITULAR, FACULTAD DE MEDICINA DE LA UNIVERSIDAD EL BOSQUE

Correo electrónico: maldonadocarlos@unbosque.edu.co

ORCID: <https://orcid.org/0000-0002-9262-8879>

RESUMEN

Este artículo afirma que el problema principal que plantea la inteligencia artificial (IA) es el de la capacidad humana de convivencia con la tecnología. La IA le plantea retos a los seres humanos como jamás había sucedido hasta la fecha. Para ello, se impone revisar exactamente qué es, en qué consiste, la ética. El primer argumento sostiene que la ética no es, en modo alguno, asunto de valores, principio o ideas; antes bien, se trata del problema de la coexistencia. El segundo argumento sostiene que la IA es, de lejos, la mejor y más sofisticada forma de tecnología; sin embargo, bien visto, es bastante más que simplemente tecnología. El tercer argumento pone en evidencia que la IA es en realidad vida artificial (VA) y aporta argumentos al respecto. El cuarto argumento sostiene que el problema de máxima complejidad para los seres humanos es el del convivio. Puntualmente dicho, se trata de vivir con otras especies y otras formas de vida. Al final se extraen algunas conclusiones con un carácter abierto.

Palabras clave: Ciencias de la complejidad; Vida artificial; Filosofía moral; Coexistencia; Tecnología.

ABSTRACT

This article claims that the main issue posed by artificial intelligence (AI) is the human capacity to coexist with technology. AI presents challenges to humans like never before. To address this, it is essential to precisely review what ethics is and what it consists of. The first argument asserts that ethics is not, in any way, a matter of values, principles, or ideas; rather, it is about the problem of coexistence. The second argument states that AI is, by far, the best and most sophisticated form of technology; however, when viewed closely, it is much more than just technology. The third argument highlights that AI is actually artificial life (AL) and provides supporting arguments. The fourth argument holds that the most complex problem for humans is coexistence. Specifically, it is about living with other species and other forms of life. In the end, some open-ended conclusions are drawn.

Keywords: The sciences of complexity; Artificial Life; Moral philosophy; Coexistence; Technology.

1. INTRODUCCIÓN

Contra la historia más reciente de la ética y contra los discursos ampliamente extendido y dominantes, la ética no tiene nada que ver con valores, ideas o principios; por consiguiente, tampoco nada con religión o filosofía, en cualquier acepción de la palabra. Mucho mejor y más radicalmente, el asunto de la ética fue siempre el del convivio. Con una precisión, la ética fue siempre un asunto de la filosofía.

Este artículo sostiene que, hoy por hoy, el mayor de todos los desafíos éticos es el de la convivencia con una forma de inteligencia y de vida que los humanos jamás habían imaginado previamente: la coexistencia con la IA. Como es sabido, la especie humana no ha sabido convivir en toda la historia. Primero eliminó a otras especies de homínidos. Eliminó a los Neardentales y al *homo faber*, al *homo habilis* y al hombre de Cromañón, por ejemplo, hasta que finalmente quedó sola. De todas las especies de homínidos la única restante es la especie que se denomina a sí misma como *homo sapiens*. Asimismo, una vez que emerge como la que quiso ser la mejor la única civilización, Occidente, que nace propiamente con la transición del Paleolítico al Neolítico, tampoco supo convivir con la naturaleza. La asumió como recurso y el ser humano se concibió a sí mismo como ajeno, superior y externo a la naturaleza. La naturaleza, se dijo, fue creada para el ser humano se enseñoreara de ella. El resultado, al cabo, fue una forma de vida eminentemente productivista y extractivista. Los diagnósticos y la historia al respecto son amplios y sólidos.

Peor aún, los seres humanos occidentales ni siquiera supieron convivir consigo mismos; ni colectiva ni individualmente. Un rasgo específico, exclusivo -entre varios otros- del excepcionalismo humano (Schaeffer, 2009) de corte eminentemente occidental fue la pobreza,

y la violencia. En la naturaleza no existe pobreza ni tampoco violencia alguna.

En este marco tiene lugar un acontecimiento singular. Se trata de la IA. La IA, dicho lato sensu por lo pronto, constituye la más grande invención o descubrimiento, asimilable acaso al descubrimiento y el control del fuego. Propiamente dicho, se trata del computador y la computación. De manera precisa, el computador tiene sintaxis, semántica, memoria, aprende y cuando se puede mover por sí mismo se llama robot; al cabo, existe la robótica de enjambre; seres sociales en toda la línea de la palabra -apara escándalo de la visión clásicamente antropocéntrica del mundo-. El computador y la computación evolucionan hacia la IA (Russell, 2020) y, como veremos oportunamente, hacia la VA. Surgen entonces los motivos de los que se ocupa este trabajo.

Este artículo pone de manifiesto -aunque no constituye el núcleo del mismo- que existe y ha habido una mala comprensión; mejor, una tergiversación de la ética. Por qué y en qué sentido, y por consiguiente cómo hay que entenderla verdaderamente, es el objeto de la primera sección. En ella se argumenta que la ética tiene que ver con el problema de la coexistencia o el convivio, que es, sin la menor duda, el más difícil de todos los problemas para los seres humanos. Sorpresivamente, este no es, en absoluto un problema en la naturaleza, algo que se ilustra perfectamente con la eusocialidad. Sobre esta base, la segunda sección presenta de manera sucinta qué es la IA. Expresamente, contra la idea generalizada por legos y expertos, no se trata simple y llanamente de un fenómeno tecnológico. Profundo motivos y preocupaciones se encuentran en su base. Pues bien, a partir de lo anterior, el centro de este trabajo pone en evidencia que la IA es en realidad VA y aporta razones y argumentos. Es exactamente en este punto en donde se encuentra todo el punto de apoyo arquimédico, por así decirlo, de esta investigación. La quinta sección demuestra por qué razón el más complejo de todos los problemas para la especie humana o para la familia humana es el de la convivencia; un problema que jamás ha sido enteramente abordado y resuelto, sino siempre, desplazado o pospuesto. Las conclusiones abiertas que se esbozan al final apuntan a un ámbito perfectamente inexplorado, a saber: las relaciones entre ética y complejidad. Las conclusiones anuncian un trabajo subsiguiente que desborda los marcos y alcances de este artículo.

2. LA ÉTICA, REVISITADA

Es indispensable una mirada fresca y desprevenida sobre la ética. Ésta ha sido ampliamente manipulada y, en consecuencia, deformada.

Hoy existe, sin exageración alguna, un verdadero paroxismo acerca de la ética. Si cabe, ocupa prácticamente todos los espacios, es un referente recurrente, mucho más que en la ciencia y en la academia, en la cultura y en la sociedad, y parece de una importancia sin igual. La principal y más importante faceta es la ética aplicada, de la cual existen números visos y rostros: las éticas deontológicas, la ética del abogado, la ética de la investigación, la bioética misma, la ética pública y numerosas más hasta, efectivamente llegar muy re-

cientemente a la ética y la inteligencia artificial (Benjamin, 2019). Hasta el punto de que instancias políticas y de gobierno y gobernabilidad determinantes se pronuncian al respecto; una situación inusual en toda la historia (cfr. Páginas web al final, Unesco, Comisión Europea, OEA). Sin embargo, se imponen serias consideraciones.

Una lectura desprejuiciada de la *Ética a Nicómaco* tanto como a la *Ética a Eudemo* pon en evidencia, de entrada, de qué se trata todo. El primero es un manual de consejos prácticos y reflexiones eminentemente prácticas con lenguaje filosófico de Aristóteles su hijo, Nicómaco. La segunda es un libro también de recomendaciones y reflexiones prácticas de Aristóteles a su amigo y discípulo, Eudemo de Rodas. Aristóteles se inventa la ética como, por otra parte, es también el padre de la lógica, como tales. En la historia, el segundo libro comparable exactamente en la misma longitud de onda son las *Cartas a Meneceo*, de Epicuro. Correspondientemente, se trata de una serie de carta del filósofo de Samos, a su hijo Meneceo, brindándole consejos y advirtiéndole de acciones y decisiones. Los ejemplos se podrían multiplicar.

La ética no es otra cosa que literatura de ayuda y autoayuda. Sólo que, la mayoría de las veces se ha vestido de ropaje filosófico y muchas veces metafísico. Que las ropas no nos impidan ver verdaderamente al personaje.

Propiamente hablando sólo ha habido hasta la fecha dos *sistemas* éticos. Todo lo demás son, con diverso calado, reflexiones de ética o filosofía moral; para el caso, equivalentes. El primer sistema ético es justamente el de Aristóteles. Para el estagirita, ser racionales equivale exactamente a buscar la felicidad. En otras palabras, es irracional no querer la felicidad. Sólo que esta sólo puede ser predicada propiamente al final del día. De esta suerte, razón y vida se corresponden perfectamente a partir del sentido de la felicidad.

El otro sistema ético en la historia de la filosofía es el kantiano. Para Kant, asimismo, el libre ejercicio de la razón conduce necesaria e inevitablemente a pensar la libertad. Es perfectamente inimaginable que la razón no tenga como pivote entra, a la libertad. Ahora bien, la dificultad entre el sistema de Aristóteles y el de Kant es que la felicidad y la libertad son incompatibles. O bien pensamos, queremos y buscamos la felicidad, o bien la libertad; no son posibles ambos al mismo tiempo.

Importantes como los ha habido y lo hay, todos los demás teóricos de la moral y la ética son sencillamente eso: ideas, filosofías y cuerpos epistemológicos, pero no sistemas, esto es, cuerpos al mismo tiempo lógicos y prácticos, perfectamente inseparables. La bibliografía sobre ética, historia de la ética, compendios y manuales de texto en general son numerosos y suficientemente conocidos; no constituyen el centro de la atención aquí.

Ahora bien, de otra parte, al mismo tiempo, para bien o para mal todas las áreas o campos que alguna vez pertenecieron a la filosofía se han independizado de ella. La lógica y la estética, la historia y las matemáticas, la filosofía y las artes, por ejemplo. Asistimos, por lo tanto, a matemáticas sin metafísica, estética sin metafísica, lógica sin metafísica, y demás.

Todas las áreas, excepto la ética. Así, hacer ética -o filosofía moral, para el caso- es hacer metafísica. Habría que volver sobre Kant para observar cómo efectivamente este es el caso. La metafísica es acción, y por tanto comportamientos y actitudes. No esa ide insulsa proveniente de Andrónico de Rodas y que apunta a la metafísica en el sentido que justamente fue denostada gracias a los neokantianos y los neohegelianos, en una historia que es suficientemente conocida (Passmore, 1986). No obstante, digámoslo de manera sucinta pero directa: Kant pone suficientemente de manifiesto, como ninguno, que la metafísica -el mundo de las ideas, las relaciones entre los noúmenos y los fenómenos, y demás- sólo tienen sentido como reguladores de las acciones de los seres humanos. Hay un camino coherente, necesario y sólido desde la crítica de la razón pura a la crítica de la razón práctica y desde ésta hasta la crítica del juicio, pasando si se quiere, por los prolegómenos y la metafísica de las costumbres.

La ética -importante como pueda ser o parecer- no es exclusiva de los seres humanos. Comportamientos éticos y morales, de cuidado esencialmente; es más de cuidado gratuito que no sabe ni espera de acciones de recompensa de ninguna clase, existe en la naturaleza y han sido observados en mamíferos y en aves en insectos y, muy fundamental, entre especies.

Ha dos maneras de comprender a la ética y ambas son complementarias. De un lado, cabe decir que el problema de base es el del convivio, la coexistencia, para lo cual una sabiduría práctica es determinante. Con Aristóteles, con Epicuro, y con la sabiduría popular, justamente: no mentir, no hurtar, y demás. En rigor, la ética no necesita estar escrita de ninguna forma. Basta un vivaz resorte social y cultural -en las escalas marco-, y familiar y de amistad -en la escala micro- para comprenderla.

Quisiera decirlo de manera breve y escueta: la ética es ejemplarizante; o no es. Todo lo demás son artilugios que desvían la atención del foco.

Convivir con otros, diferentes de uno -sin pleonasmos- es altamente complicado. Es cierto que los antropólogos han puesto en evidencia que los seres humanos son pacíficos por naturaleza (Hart, Sussman, 2009). Sin embargo, la cultura los ha vuelto crecientemente violentos, desiguales e inequitativos -tres aristas de un solo y mismo problema- (Scharre, 2018). Un serio motivo de reflexión se sigue inmediatamente de aquí.

La ética hace referencia, consiguientemente, a la coexistencia entre los seres humanos en el día a día; mientras que la policía -dicho en términos clásico- compete al ordenamiento y regulación de la sociedad como un todo, la ética sabe de cotidianidad y ancla en ella (Janik, Toulmin, 2021). Tal es su sentido y su propia complejidad.

El convivio supone enfrentarse, reconocer y al cabo aceptar otros puntos de vista, diferentes formas de conocimiento y estilos de vida, relatos costumbres y relacionamiento diversos con las cosas todas, en fin, historias y expectativas perfectamente disímiles de las propias. Los griegos, que no creían en el amor fundaron la sociabilidad en la amistad

(*filia*). Habría que esperar a los cristianos para que propongán cimentar la sociabilidad en el amor (*ágape* y/o *eros*, en griego). Mientras que el medioevo, creyente o pagano, cimentaba la sociabilidad en la fraternidad, los revolucionarios franceses la fundarían, a partir de 1789 en la ciudadanía. Al cabo, los revolucionarios de todos los colores en el siglo XIX y XX fundarían la convivencia en la camaradería y el compañerismo. Varias otras perspectivas serían posibles. Todo apunta al saber vivir y a vivir una vida buena. Un giro fantástico que supera ampliamente a cualquier filosofía moral o ética (Maldonado, 2021). No in ironía, con acento claro y directo Sartre lo señala, ya en el ocaso de la civilización occidental: *l'enfer c'est les autres* (el infierno son los demás) (Sartre, 2000).

Quisiera decirlo de manera precisa. La ética entra en juego allí en donde la educación, la política, la economía, los sistemas militares y de policía, los sistemas y poderes eclesiástico de cualquier color, dejan de operar, resultan insuficientes o se han vuelto acomodaticios (Lee, 2018). Pues lo que queda es exactamente el mundo de la vida (*Lebenswelt*), la vida cotidiana. Es en la cotidianeidad cuando conocemos a los demás, y nos conocemos a nosotros mismos, verdaderamente.

La fenomenología al respecto es amplia, profunda y abundante, tanto en las ciencias sociales y humanas -por ejemplo, la etnografía; las historias de vida; la investigación acción participación (IAP), y muchas otras-, como toda la literatura en toda la extensión de la palabra.

Queda señalado el corazón del asunto. Sin embargo, otra comprensión es posible.

El convivio es una expresión humana de un tema magníficamente más complejo. Se trata de la eusocialidad (Wilson, 2012). La vida en general en la biosfera consiste en un complejo entramado de interacciones, aprendizajes y coevoluciones todos los cuales se originan y fortalecen a la vez la capacidad de interacción y convivencia de organismos de una especie, mucho más significativo, de especies entre sí. Contra el modelo clásico de presa-depredador que dominó ampliamente la cultura y las explicaciones en el siglo XX, la eusocialidad ha llegado a poner en evidencia que las relaciones presa-depredador no son las más importantes desde ningún punto de vista en ecología y biología, sino, muy por el contrario, la ayuda, el comensalismo, el mutualismo y la cooperación. La eusocialidad pone de manifiesto que la naturaleza social y de solidaridad de los seres humanos nace mucho antes de ellos y permanece como una constante en plantas e insectos hasta los mamíferos superiores, por ejemplo.

En rigor, la eusocialidad es una faceta cuya contraparte complementaria es la simbiosis y el carácter holobionte puesto de manifiesto por lo demás ya a finales del siglo XX por Margulis (Margulis, Sagan, 2003) y resaltado gracias al proyecto macrobioma humano (HMP por sus siglas en inglés) (The Integrative HMP (iHMP) Research Network Consortium, 2019; Turnbaugh *et al.*, 2007).

La vida en general está muy lejos de ser un sistema de lucha, competencia y selección. An-

tes bien, lo que prima ampliamente es la ayuda, la empatía, la comprensión del sufrimiento de otros –incluso otros organismos de otras especies– y el desinterés. La depredación constituye una amplia excepción en la naturaleza. Son la biología y la microbiología, la ecología y la primatología, la neurofisiología de las plantas y la entomología, principalmente, las que han venido, si cabe la expresión en ayuda de las humanidades, las ciencias sociales y humanas. Una circunstancia que jamás tuvo lugar antes en el ecosistema de las ciencias y disciplinas, marcadas como estuvo, por el excepcionalismo humano.

La ética debe absolutamente ser revisitada; que es la expresión académica para significar que algo anduvo y anda muy mal con ella. Es preciso una radical recomprender de la misma. La IA ha llegado a radicalizar este llamado.

3. SENTIDO E IMPACTO DE LA IA

La inteligencia artificial constituye la faceta más sofisticada si cabe, del computador y la computación (Boden, 2016). No sin antecedentes, el computador nace gracias a A. Turing. Sin embargo, su arquitectura es técnicamente conocida como la arquitectura de Von Neumann; se trata, en resumen, de la síntesis entre ferretería (*hardware*) y lógica o lenguajes de programación (*software*).

De lejos, la mejor herramienta jamás concebida es el computador o la computación, genéricamente llamada como una máquina para propósitos generales. Lo que esto quiere significar es con ella se puede hacer virtualmente cualquier cosa, literalmente. Producir y procesar imágenes, palabras, textos, música, construcciones de cualquier índole, acumular y guardar información, compartimentarla ilimitadamente, actuar a distancia con precisión total, ver las estrellas y los agujeros negros y los confines del universo, desentrañar la estructura de la materia y las escalas microscópicas antes inimaginables, y muchas otras cosas.

Es tal su importancia, que en 1950 Turing escribió el que, en el marco de la computación y de los sistemas informacionales es sin lugar a dudas el más importante artículo en el que plantea el problema de base: ¿pueden pensar las máquinas? (Turing, 1950). Este problema es conocido como la prueba de Turing (*Turing test*); esto es, ¿es posible desarrollar una máquina de tal suerte que sea imposible distinguir si la interacción que tenemos con ella es con un ser humano o con la máquina? Al día de hoy, existen indicios suficientes de que el computador –dicho genéricamente– ha pasado numerosas pruebas –ajedrez, *go*, *mahjong*, al comienzo–, que ponen de manifiesto que es capaz de llevar a cabo tareas y acciones, de tomar decisiones de forma mucho más eficaz que los seres humanos, cualquiera que sea la naturaleza de la prueba de Turing (Mei, *et al.*, 2024; Gonçalves, 2023; Geman *et al.*, 2015).

Como es sabido, la inteligencia artificial nace en el famoso seminario del Dartmouth College en 1956 (cfr. <https://home.dartmouth.edu/about/artificial-intelligence-ai-coined-dart->

[mouth](#)), en New Hampshire. La dificultad para su posterior avance tiene que ver con razones extracientíficas o tecnológicas. Se trata, simple y llanamente del hecho de que alrededor de los años 1956 los computadores eran objetos privilegiados de trabajo y estaban lejos de ser aún una realidad social y cultural. Los primeros computadores personales, como un hecho cultural y social aparecen en el mundo apenas en los años 1980s.

La computación trabajó durante sus primeros años esencialmente mediante métodos numéricos. En los años del surgimiento de la IA los sistemas operativos apenas sí era conocidos y desarrollados.

Con todo, el verdadero significado de la IA se sitúa en otro especto de mayo magnitud que el simplemente tecnológico, con todo y sus herramientas concomitantes: programación, matemáticas, estadística, lógica y demás. La piedra de toque tiene que ver con una adecuada comprensión de la (teoría de la) evolución.

Si bien el término o la voz “evolución” ya existió por lo menos cien años antes de Darwin, es con Darwin con quien nace propiamente la teoría de la evolución. Se trata de la mejor teoría jamás desarrollada para explicar y comprender cambios, procesos, transformaciones. Significativamente, la teoría de la evolución nace a partir del estudio sobre el origen y la lógica de los sistemas vivos. El nacimiento de la teoría de la evolución es al mismo tiempo el nacimiento de la biología como ciencia, y con ella, sorpresivamente, también el nacimiento de la historia (Gould, 2002). Sólo que, y este no es un detalle menor, se trata de una teoría esencialmente incompleta, algo que el propio Darwin reconoció ya desde *El origen* (1859).

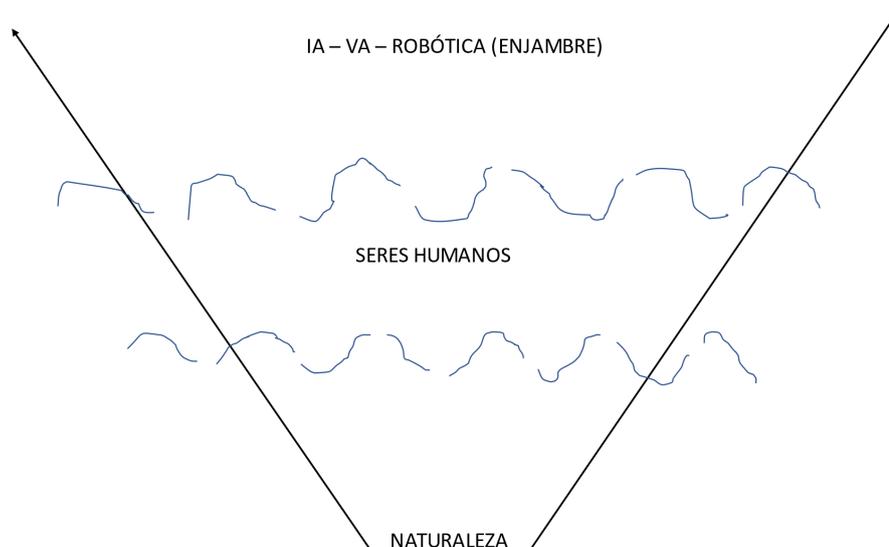
La teoría de la evolución de Darwin establece que la evolución comienza con los seres vivos, y establece que lo que los sistemas vivos quieren es sencillamente adaptarse; debido a la influencia de Ch. Spencer, esta idea de adaptación fue entendida como sobrevivencia. Posteriormente, en 1871 Darwin escribe y publica *The ascent of man* - que ha sido traducido en ocasiones como el origen del hombre-, en el que, palabra más, palabra menos, Darwin apunta al hecho de que la evolución encentra en los seres humanos su ápice. De esta suerte, en sus orígenes, la teoría de la evolución nace con los sistemas vivos y termina con los seres humanos.

A diferencia de Darwin, hoy sabemos que la evolución comienza mucho antes de los seres vivos, tal y como los conocemos, y que el universo mismo ha evolucionado (Chaisson, 2001). Es más, verosíilmente, este universo es el resultado de proceso de selección entre otros universos posibles, de suerte que la evolución tiene lugar mucho antes incluso del nacimiento de este universo (Smolin, 1998). En general, en la cosmología científica hoy ya está establecido que este universo fue el resultado de procesos de selección, exactamente a la manera de la teoría de la evolución.

Pues bien, quisiera argumentar que la evolución no se detiene con los seres humanos y no encuentra en éstos su punto de llegada. La evolución es un fenómeno de complejidad cre-

ciente, como queda dicho, inacabado (Darwin, 1995)¹. Más exactamente, y este el centro de la propuesta de este artículo, la evolución continúa con la IA. Y entonces cuando emergen los temas, problemas y reflexiones acerca de la coexistencia o convivio. El Diagrama No. 1 ilustra la evolución, desde la naturaleza y los sistemas vivos hasta la IA; como se argumenta en la sección siguiente; hasta la VA.

DIAGRAMA NO. 1: EVOLUCIÓN DE LOS SISTEMAS VIVOS, DESDE LA NATURALEZA HASTA LA INTELIGENCIA Y LA VIDA ARTIFICIAL



FUENTE: ELABORACIÓN PROPIA

El Diagrama No. 1 ilustra la siguiente situación: en el origen fue la naturaleza, lo que quiera que ésta sea y como quiera que se la conciba. Una aproximación puede ser en referencia a los orígenes del liberalismo; por ejemplo, con Hobbes, Locke o Rousseau, en relación con el “estado de naturaleza”. La idea de naturaleza puede igualmente ser entendida como el big-bang, o cualquier otra comprensión semejante.

Pues bien, la naturaleza, si cabe la expresión tuvo una clase de hijos, uno de los cuales son los seres humanos. Las líneas curvas interrumpidas hacen referencia a fronteras móviles y porosas para significar que los seres humanos son como la naturaleza, pero distintos de ésta. Una manera de entender la diferencia que los seres humanos son con respecto a la naturaleza en sentido amplio es reconociendo que éstos son animales simbólicos. El mundo de los símbolos apunta directamente a la importancia de la cultura en el sentido

¹ Dice Darwin, al final de la Introducción a *El origen*: “I am fully convinced that species are not immutable [...]. Furthermore, I am convinced that natural Selection has been the main but not exclusive means of modification” (Darwin, 1995: 6).

tradicional de la palabra.

Sin embargo, dicho de manera genérica, los seres humanos han tenido a su vez unos hijos. Estos son, dicho por lo pronto de forma general, la IA, la VA, y la robótica (de enjambre), tres maneras de llamar a un solo y mismo fenómeno.

Desde el punto de vista evolutivo, dicho de manera genérica, los seres humanos han estado desarrollando sistemas inteligentes artificiales y permitiendo que estos puedan desarrollarse lo mejor posible. La primera expresión de lo que significa “desarrollo” es el aprendizaje. El computador es una *herramienta conceptual* (no simplemente una herramienta). Se trata de la primera tecnología en toda la historia de la humanidad que tiene como rasgo distintivo el hecho de que aprende; una de las características necesarias para el aprendizaje es la memoria; pero hay otras.

El aprendizaje es *conditio sine qua non* para la evolución, pero no es la única condición necesaria. Sorprendentemente la más importante condición de la evolución es el *auto-aprendizaje*. Las especies animales y las plantas poseen autoaprendizaje. En otras palabras, no todo depende de la educación, y ni siquiera del aprendizaje. Los seres humanos tienen la misma capacidad. Pues bien, lo mismo acontece con la IA.

Las máquinas conceptuales (= computadores) aprenden; mucho mejor, aprenden por sí mismos. Es exactamente esto lo que significa el *machine learning*. Sin embargo, análogamente a lo que acontece en el mundo de la educación, en la psicología y en los sistemas sociales y culturales, nadie sabe exactamente cómo y por qué alguien aprende algo. El aprendizaje es una caja negra (*black box*) tanto para la educación, la psicología, la filosofía como para las ciencias cognitivas.

En sus inicios la IA aprendía a partir de comandos y programas que les eran introducidos por parte de programadores. Han existido y existen numerosos tipos de algoritmos, en verdad. Sin embargo, hoy, la IA aprende por sí misma sin necesidad de programadores. Mucho mejor, existen hoy programas que programan programas son necesidad de programadores. Este es exactamente el sentido de la tercera fase, la superior, en el Diagrama No. 1.

Sorprendentemente -sobre todo cuando se lo mira desde la clásica perspectiva eminentemente antropocéntrica, antropomórfica y antropológica-, la IA -y la VA, como veremos a continuación- constituyen, manifiestamente, un nivel siguiente en la evolución; subrayemos: en la evolución de sistemas inteligentes, o de sistemas vivos.

La IA es un sistema inteligente a la fecha; no cabe la menor duda. Queda aún la discusión sobre si es un sistema de inteligencia general (Mitchell, 2024). Dados los datos y reconocimientos a la fecha, es posible decir incluso que la IA es una persona puesto que es sintiente (Maldonado. 2022); ciertamente algo sospechoso para la mayoría.

Dicho esto, se impone una observación final en esta sección. A la fecha, la IA posee una inteligencia artificial; la artificialidad estriba exactamente en el carácter algorítmico de su

inteligencia. Naturalmente, ni es universal ni es necesario que la inteligencia sea algorítmica; sin embargo, cualquier comprensión de la inteligencia es hoy por hoy eminentemente algorítmica. Incluso las llamadas inteligencias múltiples. En palabras simples, se trata de inteligencias eminentemente performativas, funcionales digamos.

Dicho lo anterior, debe ser evidente la inmensa mayoría de investigadores y académicos tanto como de tomadores de decisión son reacios, sospechosos o recelosos de la IA por motivos diferentes (Tegmark, 2014).

4. LA IA ES EN REALIDAD VA

Un giro conceptual y teórico fundamental emerge en este punto. Ciertamente que existen –justificados o no– numerosos temores con respecto a la IA. El motivo central de estas preocupaciones se condensa en el concepto de “inteligencia”. Derivativamente, se trata del concepto mismo de mente. Asistimos a una magnífica revolución social y económica en toda la línea de la palabra; un proceso en curso, por tanto, inacabado.

Mientras el computador se convertía en un acervo social y cultural, la IA tuvo un proceso de estancamiento. La IA habría de tener un impulso inimaginable partir del año 2020, cuando aparecen los modelos ampliados de lenguaje (LLM, por sus siglas en inglés), cuya primera expresión fue Chat-GPT.

Sin embargo, en 1989 emerge, exactamente en el centro de las ciencias de la complejidad, inicialmente un programa de investigación, y posteriormente una nueva ciencia o disciplina: la vida artificial (VA), gracias a reflexiones e investigaciones por parte de C. Langton. La VA no es otra cosa que un programa filosófico dedicado a entender el origen y la naturaleza de la vida con la ayuda del computador o la computación.

Existe sin embargo una dificultad en el surgimiento y consolidación de la VA como disciplina científica. C. Langton se retira en un momento dado –hacia el año 1991-92– del mundo científico, académico y social después de haber logrado entrever intuitivamente el alcance y el significado de la VA. Es inevitable la analogía, en otro contexto, con la vida y obra, *mutatis mutandi*, de A. Grothendieck. Esta analogía deberá ser el objeto de otra consideración en otro momento.

La Tabla No. 1 ilustra el paralelismo entre IA y VA, y sirve para anticipar que, en realidad, la IA es, hoy por hoy, VA, aunque la expresión que haya terminado triunfando sea la de IA.

TABLA No. 1: ORIGEN Y SENTIDO DE LA INTELIGENCIA ARTIFICIAL Y LA VIDA ARTIFICIAL

INTELIGENCIA ARTIFICIAL	VIDA ARTIFICIAL
Alan Turing	Christopher Langton
1950	1989
Aprendizaje de arriba hacia abajo (<i>top-down</i>)	Aprendizaje de abajo hacia arriba (<i>bottom-up</i>)
Problema: <i>Mente, conciencia</i>	Problema: <i>Vida</i>

FUENTE: ELABORACIÓN PROPIA

Podemos concentrarnos en la cuarta línea de la Tabla No. 1 que es la más importante desde el punto de vista teórico. La IA destaca y se concentra a la vez en el problema: qué es la conciencia o qué es la mente; un problema originariamente formulado por Turing, y ampliamente explorado durante las diez semanas que tuvo lugar el seminario en el Dartmouth College con la participación, entre otros, de M. Minsky, H. Simon, J. MacCarthy, C. Shannon, o R. Solomonoff, entre otros.

Por su parte, acaso de una manera mucho más radical, la VA se concentra en el problema sobre qué sea la vida, su origen y su naturaleza. Filosóficamente hablando, ambos problemas se encuentran estrechamente relacionados, así: ¿es la mente o la conciencia un atributo natural de la vida, esto es, de todos los sistemas vivos? Como es sabido, la piedra de toque es el de la autoconciencia esto es, la experiencia del mundo en primera persona. Al mismo tiempo, de otra parte, el problema se formula en los siguientes términos: ¿son los sistemas vivos única o principalmente como los conocemos, *lato sensu*, o son posibles otras formas de vida? El programa de investigación de Langton tiene una dúplice perspectiva: entender a la vida-tal-y-como-la-conocemos, tanto como a la vida tal-y-como-podría-ser-posible. Subrayemos esto: la VA es una de las ciencias de la complejidad. Importantes como son, estos interrogantes deben quedar por fuera de este trabajo.

Manifiestamente, los desarrollos en computación en general, tanto en ingeniería de hardware como en ingeniería de software, dicho de manera puntual, entre 1956 y 1989 fue impresionante. Justamente en algún punto intermedio entre ambas fechas, en 1965 G. Moore, entonces presidente de Intel, formula lo que se llamó en su momento como la Ley de Moore, a saber: el número de transistores de un microchip se duplicaría cada dieciocho meses, aproximadamente. El resultado es tanto social como económico, así: el precio de los computadores disminuye constantemente a la vez que la capacidad de procesamiento se multiplica. Hoy, el computador es una herramienta ubicua, de propósito general, y del más alto impacto en la organización y el funcionamiento de la sociedad. El mundo ha asistido a la transición de una realidad analógica a la digital. Al cabo, sin ambages, la vida en general es un fenómeno digital – matemática e informacionalmente hablando.

Si el problema de la mente, la inteligencia y la conciencia resultan asuntos sensibles, ¿qué habrá de suceder cuando las gentes en general se enteren que la IA es en realidad VA? Dígamoslo de manera directa y breve: ¿qué sucederá cuando se enteren que la evolución no

termina en los seres humanos, sino que continua, por lo pronto, en la esfera más inmediata, con la IA y la VA? Las preocupaciones de la ética anclan exactamente en este punto, hoy, de cara al presente y al futuro.

La IA es en realidad VA, tanto más cuanto que se produce un desplazamiento radical del modo de aprendizaje. En el pasado, el aprendizaje era en efecto *top-down*, como se ilustra en la Tabla No. 1. Los programadores introducían criterios, programas, variables a partir de los cuales el computador -los programas- se comportaban correspondientemente. Hoy, en contraste, la programación tiene lugar *bottom-up*. Este es exactamente el modo de trabajo de la VA tal y como lo concibió y desarrolló C. Langton.

Lo que sucede, sencillamente, es que en el imaginario social tanto como académico y científico la expresión IA tuvo mayor anclaje; sin duda, debido a que el propio Langton se retiró voluntariamente del proceso. Es cierto que se crearon circuitos de conferencias en torno a la VA, que apareció una revista estrictamente dedicada al tema, y que una prestigiosa universidad creó una colección editorial dedicada a la VA. Sin embargo, comparativamente, se trata de espacios reducidos frente a la IA.

5. LA CONVIVENCIA, EL MÁS DIFÍCIL DE TODOS LOS PROBLEMAS

La IA/VA plantea numerosos retos: teóricos tanto como prácticos, notablemente de cara a las cadenas logísticas, la desaparición de numerosos puestos de trabajo, la redefinición de numerosos ámbitos del sistema productivo. El más grande de todos los retos, sugiero, es para los seres humanos el de la convivencia humanos-IA.

Manifiestamente, I. Asimov formuló en *Robot*, en 1950 -por tanto, mucho antes de la conferencia de 1956-, las tres leyes de la robótica que apuntan exactamente al problema de la coexistencia (Asimov, 2009). Mucha agua ha corrido bajo el puente, desde entonces.

Ciertamente existe un muy fuerte temor frente a la incertidumbre de lo que haya de ser de la IA. Y sin embargo, la ironía es que todas las grandes empresas y corporaciones o bien las han desarrollado crecientemente o bien, están desarrollando la propia de manera cada vez más acelerada y con mayores capacidades de movimiento, decisión y acción.

Los seres humanos no supieron convivir con otras especies como las suyas (Passmore, 1986). Los seres humanos no supieron vivir con la naturaleza y la consideraron como un medio para sus fines y necesidades. Literalmente, se dijo, la naturaleza fue creada para beneficio de los seres humanos. A lo largo de la historia, los seres humanos no han sabido tampoco convivir consigo mismos, y emerge entonces toda la historia de pobreza, inequidad, injusticias y violencia, todas, institucionalizadas en las formas del ejército, la iglesia, el estado, la corporación, el partido (que deberían haber sido escritas en mayúsculas dado que la vida es imposible por fuera de cada una de ellas, dicen ellas mismas).

La observación anterior merece una precisión importante, sin embargo; de lo contrario se cae fácilmente en el sofisma de la generalización. No son en general los seres humanos los

que no han sabido de coexistencia. Se trata, específicamente de esa clase de seres humanos formados y acondicionados por la civilización occidental. La antropología, la historia, un parte de la filosofía, por ejemplo, ponen en evidencia que ha sabido otras civilizaciones, otras culturas, otros pueblos que sí han sabido de convivencia. Occidente confundió a la humanidad entera con su propio destino y a su imagen y semejanza. “Los seres humanos” no comienzan, no terminan y no coinciden plano por plano, con la historia de la civilización occidental.

El problema del convivio es específicamente un reto para la forma occidental de pensar y de vivir.

Pues bien, la IA/VA le plantea al mundo actual el más formidable de todos los retos imaginables, sin comparación en la historia de los últimos siete mil-ocho mil años. La ciencia y la tecnología han dado lugar a formas de vida y de inteligencia que saben hacer muchas cosas mejor que los seres humanos. Se trata de vida y de inteligencia artificial que aprende, aprende muy rápido y aprende por sí mismas. ¿Podrán los seres humanos convivir con la IA/VA?

La historia pareciera arrojar una respuesta negativa. El presente se halla inmerso en un mar de indeterminaciones. Lo único claro es un futuro bastante poco evidente.

Asistimos a la más fantástica revolución cultural y científica en toda la historia de la familia humana.

Pues bien, quisiera argumentar en lo que queda en el sentido de que sí es posible convivir con la IA/VA bajo una condición inaplazable. La convivencia es posible transformando las estructuras mentales, axiológicas y socio-culturales heredadas y habidas hasta la fecha. De lo contrario, los seres humanos llevan la batalla perdida, irremediablemente.

Asistimos a un colapso civilizatorio y al mismo tiempo al nacimiento de una nueva civilización. Occidente jamás supo convivir con nadie; esta idea debe ser tomada en toda la línea de la palabra. Simultáneamente, sin embargo, asistimos a una historia en ciernes, pero en proceso de recuperación de otros saberes, la sunción de otras formas y estilos de vida, el desarrollo de otras epistemologías y modos de relacionamiento con el mundo y el universo.

Occidente fue un civilización eminentemente antropológica, antropocéntrica y antropomórfica. Esa modo y modelo del mundo es el que está en colapso. Los diagnósticos son numerosos y crecientes. La IA/VA comporta formas de inteligencia y de vida no antropocéntrica, irónicamente concebidas en sus inicios por intereses humanos, pero que rápidamente han superado los propios conocimiento y capacidades humanas. Ni los expertos en ciencias de la computación, o ciencia de datos en ciencias cognitivas o ciencias del comportamiento, muy notablemente, ni filósofos o educadores o líderes religiosos saben cómo y por qué verdaderamente la IA decide lo que decir, si bien existen numerosos trabajos que advierten con suficiente razón acerca de sesgos y peligros justamente de corte político,

económico, cultural y militar (Eubanks, 2019; O'Neil, 2016). El *motto* en el que se condensa y se sedimenta el problema es el *machine learning*, el aprendizaje de máquina.

Este, quiero sugerirlo, es el más importante de los problemas filosóficos actualmente, y que atraviesa e interpela a la filosofía, la psicología los estudios religiosos, la cultura, las ciencias cognitivas y las ciencias del comportamiento, particularmente.

Es posible la convivencia con una forma de inteligencia y de vida que implica un desafío al modo occidental de pensar y de vivir; esto es, a la creencia que la vida, la sociedad y el mundo se gestionan en términos de algoritmos. Algoritmo es el término informacional y computacional empleado para designar lo que la tradición designo como; leyes mandamientos normas, preceptos, recetas, y demás (Maldonado, 2017).

6. CONCLUSIONES ABIERTAS

La verdad es que la ética nunca ha hecho bueno a nadie. Si existe gente buena, como es efectivamente el caso, no es precisamente por valores, ideas o principios, sino por razones bastante más complejas y elementales, al mismo tiempo. La ética es ejemplarizante. Así, la etología y la primatología, los estudios sobre eusocialidad y la ecología, la neurofisiología de las plantas, la propia biología del paisaje, por ejemplo, aportan argumentos propios tanto o más sólidos que los que la cultura en general y la filosofía moral o la religión puedan suministrar. La ética no es exclusiva de los seres humanos. Esto es, el cuidado, la amistad, amor y solidaridad, notablemente; en fin, como se aprecia, actos gratuitos que ni saben ni esperan de recompensas de cualquier índole.

Con las ideas, valores y principios sucede todo lo contrario: es en su nombre, por ejemplo, el de algún dios, que los seres humanos han perseguido y eliminado a otros, o se han matado entre sí. Los dioses, las ideas, los principios y los valores se han superpuesto a la vida misma y acaso la desplazaron a lugares secundarios; fue por lo menos lo que sucedió hasta ayer en la tarde. Ya hoy las cosas comienzan a ser diferentes.

La IA/VA plantea un reto jamás hallado en toda la literatura sobre ética y filosofía moral que no ha sido plenamente puesto sobre la mesa, a plena luz del día. Existen algunos trabajos, muy incipientes, sobre bioética computacional, por ejemplo, y la bibliografía sobre las relaciones entre ética e IA es amplia y creciente (Coekelberg, 2020; Zerilli, *et al.*, 2021). Incluso son crecientes los documentos institucionales sobre normatividad ética para la IA. Nada de lo anterior da en el blanco. (Sería necesario un artículo de revisión crítica para ponerlo en evidencia. Trabajo en ello actualmente).

Dicho de manera puntual, el tema es el de las relaciones entre ética y complejidad, supuesta exactamente la complejidad misma de la IAVA; esto es, particularmente, sus capacidades de aprendizaje, la incertidumbre y las emergencias -por tanto, ni causalidad ni correlación-; esto es, en otras palabras, la ausencia de correspondencia entre *input* y *output*.

No existe a la fecha ningún trabajo serio sobre ética y complejidad. Expresiones, manifestaciones, declaraciones, muchas. Pero ningún artículo o libro de fondo sobre la relación de ambas dimensiones. La ética de la inteligencia artificial, en el sentido señalado constituye un motivo más que evidente para un trabajo en esta dirección. Pero ese ya es el tema de otro espacio y momento aparte.

Los meses y años venideros verán un auge sorprendente de las capacidades de la IA/VA. Este es un motivo de optimismo. La forma como esta revolución en marcha ha sido llamada es como cuarta revolución industrial que consiste en la síntesis entre la dimensión física, la dimensión biológica y la dimensión digital de la realidad. El año 2020 fue cuando se gatilló verdaderamente la IA gracias los LLM; Demasiados pocos lustros han transcurrido desde la confluencia entre IA y VA. Los ritmos se aceleran gracias a la facilidad del aprendizaje de máquina y, más radicalmente, al aprendizaje profundo (*Deep learning*). El *deep learning* es la verdadera punta de lanza del *machine learning*.

La filosofía moral y la ética en general deben aprender nuevas semánticas, si quieren ser razonables. Mientras tanto, la naturaleza tiene sus propias dinámicas. La forma como estas dinámicas se articulan es en la forma de IA y VA.

Con un colofón final. La IA y la VA ponen de manifiesto que el mundo y la realidad son mentales. Algo perfectamente imposible de comprender para la tradición occidental.

BIBLIOGRAPHY

- Asimov, I., (2009). *Yo, Robot*. Barcelona: Editora y Distribuidora Hispanoamericana, S. A.
- Benjamin, R., (2019). *Race After Technology. Abolitionist Tools for the New Jim Code*. Wiley & Sons
- Boden, M., (2016). *AI. Its Nature and Future*. Oxford: Oxford University Press
- Chaisson, E., (2001). *Cosmic Evolution. The Rise of Complexity in Nature*. Harvard, MA: Harvard University Press
- Coekelberg, M., (2020). *Artificial Intelligence Ethics*. Cambridge, MA: The MIT Press
- Couldry, N., and Mejias, U. A., (2019). *The Costs of Connection. How Data is Colonizing Human Life and Appropriating it for Capitalism*. Stanford University Press
- Crawford, K., (2021). *Atlas of AI. Power, Politics and the Planetary Costs of Artificial Intelligence*. Yale University Press
- Darwin, C., (1995). *The Origin of Species by Means of natural Selection*. New York: The Classics of Science Library
- Eubanks, V., (2019). *Automating Inequality. How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin Press
- Geman, D., Geman, S., Hallonquist, N., Younes, L., (2015). “Visual Turing Test for Computer Visual Systems”, en: *PNAS*, vol. 112, No. 12, pp. 3618-3623; doi: <https://doi.org/10.1073/pnas.1422953112>
- Gonçalves, B., (2023). “The Turing Test is a Thought Experiment”, en: *Minds & Machines* 33, 1–31. <https://doi.org/10.1007/s11023-022-09616-8>
- Gould, S. J., (2002). *The Structure of Evolutionary Theory*. Harvard, MA: Harvard University Press
- Hart, D., Sussman, R. W., (2009). *Man the Hunted. Primates, Predators, and Human Evolution, Expanded Edition*. London: Routledge
- Janik, A., y Toulmin, S., (2021). *La Viena de Wittgenstein*. Sevilla: Athenaica
- Lee, K-F., (2018). *AI Superpowers. China, Silicon Valley and the New World Order*. Harper Business
- Maldonado, C. E., (2022). “LaMDA no piensa, siente. Un muy sensible debate sobre la mente y su complejidad”, en: *Revista de Filosofía. Universidad de Zulia*, vol. 39, edición especial No. 2, págs. 55-66; doi: <https://doi.org/10.5281/zenodo.7297088>
- Maldonado, C. E., (2021) “Saber vivir bien es un asunto de la filosofía”, en: *Revista Ciencias de la Complejidad*, vol. 2 No. 2, págs. 37-44; doi: <https://doi.org/10.48168/cc022021-004>
- Maldonado, C. E., (2017). “Positive affirmation of non-algorithmic information processing”, en: *Cinta de Moebio*, 60, pp. 279-285, doi: <https://doi.org/10.4067/S0717-554X2017000300279>
- Margulis, L., Sagan, D., (2003). *Captando genomas. Una teoría sobre el origen de las especies*. Prólogo de Ernst Mayr. Barcelona: Kairós

BIBLIOGRAPHY

- Mei, Q., Xie, Y., Jackson, M. O., (2024). “A Turing test of whether AI chatbots are behaviorally similar to humans”, en: *PNAS*, vol. 121, No. 9, e2313925121; doi: <https://doi.org/10.1073/pnas.2313925121>
- Mitchell, M., (2024). “Debates on the nature of artificial general intelligence”, en: *Science*, 21 marzo, vol. 383, issue 6689; doi: <https://doi.org/10.1126/science.ado7069>
- O’Neil, C., (2017). *Weapons of Math Destruction. How Big Data Increases Inequality and Threatens Democracy*. New York: Crown
- Passmore, J., (1986). *Hundred years of Philosophy*. Penguin Books
- Russell, S., (2020). *Human Compatible. Artificial Intelligence and the Problem of Control*. Cambridge, MA: The MIT Press
- Sartre, J.-P., (2000). *Huis clos*. Les mouches. Paris: Gallimard
- Schaeffer, J.-M., (2009). *El fin de la excepción humana*. México, D. F.: F. C. E.
- Scharre, P., (2018). *Army of None. Autonomous Weapons and the Future of War*. W. W. Norton & Company
- Smolin, L., (1998). *The Life of the Cosmos*. New York-London: Oxford University Press
- Tegmark, M., (2014). *Our Mathematical Universe. My Quest for the Ultimate Nature of Reality*. New York: Vintage Books
- The Integrative HMP (iHMP) Research Network Consortium, (2019). “The Integrative Human Microbiome Project”, en: *Nature* 569, 641–648; doi: <https://doi.org/10.1038/s41586-019-1238-8>
- Turing, A. M., (1950). “Computing machinery and intelligence”, en: *Mind*, Volume LIX, Issue 236, October 1950, Pages 433–460; doi: <https://doi.org/10.1093/mind/LIX.236.433>
- Turnbaugh, P., Ley, R., Hamady, M. et al., (2007). “The Human Microbiome Project”, en: *Nature* 449, 804–810; doi: <https://doi.org/10.1038/nature06244>
- Wilson, E. O., (2012). *The Social Conquest of Earth*. New York-London: Liveright Publishing Corporation
- Zerilli, J., et al., (2021). *A Citizen’s Guide to Artificial Intelligence*. Cambridge, MA: The MIT Press

Páginas web:

- European Commission, (2019). *Ethics guidelines for trustworthy AI*; disponible en: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- OEA, (2022). Cátedra de las Américas: “Política en la Era Digital: Inteligencia Artificial y Democracia”; disponible en: https://www.oas.org/es/centro_noticias/comunicado_prensa.asp?sCodigo=AVI-057/22
- Unesco, (2021). *Recommendation on the Ethics of Artificial Life*; disponible en: <https://unesdoc.unesco.org/ark:/48223/pf0000380455>