

“Inteligencia artificial”. Ontología y ética

“Artificial Intelligence”. Ontology and Ethics

Jorge Martínez Barrera

Universidad Gabriela Mistral (Chile)

jorge.martinezbarrera6@gmail.com

 <https://orcid.org/0000-0001-6137-6237>

Resumen: En el presente artículo expondré dos aspectos de la hasta ahora llamada Inteligencia Artificial. El primero de ellos es el *ontológico*; intentaré una caracterización de lo que ella es y con este fin me serviré de algunas distinciones propuestas por Aristóteles para distinguir un ente natural de uno artificial. El segundo aspecto es el *ético*. En este sentido, y sin renunciar a Aristóteles, recojo algunas propuestas contemporáneas, tomando como punto de partida las críticas de Luc Ferry a Hans Jonas, especialmente en lo referido a la “heurística del temor”.

Palabras clave: Inteligencia artificial, instrumentalidad, Aristóteles, Hans Jonas, Luc Ferry.

Abstract: In the present article I will present two aspects of the so far called Artificial Intelligence. The first of them is *ontological*; I will try to characterize what it is and for this purpose I will use some distinctions proposed by Aristotle to distinguish a natural entity from an artificial one. The second aspect is *ethical*. In this sense, and without renouncing Aristotle, I take up some contemporary proposals, taking as a starting point Luc Ferry's criticisms of Hans Jonas, especially with regard to the "heuristics of fear".

Keywords: Artificial Intelligence, instrumentality, Aristotle, Hans Jonas, Luc Ferry.

1. Breve introducción

No se puede negar que el interés por la Inteligencia Artificial (IA) ha surgido de manera explosiva en los últimos años. La casi invasión temática de la IA sobre todos los aspectos relacionados con las decisiones humanas ha generado las más diversas actitudes y posiciones respecto de su alcance e implicancias en la vida personal, social y política. No obstante, no podemos negar que nos hallamos frente a cierto estado de perplejidad respecto de algunos aspectos insoslayables relativos a su estatuto de nuevo ingrediente de nuestra realidad cotidiana. Sugeriré entonces algunas propuestas desde la filosofía clásica, especialmente aristotélica, que, espero, puedan constituir un paliativo para este desconcierto.

Estas reflexiones tienen dos momentos: uno ontológico, o sea, lo que la IA es; y otro ético, es decir, como *debiéramos* situarnos frente a ella. Vamos al primero, el *ontológico*.

2. ¿Qué es la IA? El artificio y el ente natural

Aristóteles, como es obvio, no emplea la expresión IA. Y de haberla conocido probablemente hubiera rechazado semejante nomenclatura porque, como bien señala el Prof. Alfredo Marcos de la Universidad de Valladolid, lo que la inteligencia artificial tiene de inteligencia no es artificial, y lo que tiene de artificial no es inteligencia.¹ Dicho sea de paso, el Prof. Marcos prefiere hablar de "sistemas de control delegado" (CoDe)².

¹ La mayor parte de las referencias a Alfredo Marcos me han sido sugeridas por el Prof. Luca Valera, a quien agradezco por ello.

² Sara Lumbreras, citada por Marcos (ver Alfredo Marcos, "La Inteligencia artificial y el efecto Toy Story", http://www.fyl.uva.es/~wfilosof/webMarcos/textos/textos2022/Scio_IA.pdf, consultado el 20 de junio de 2022), sugiere "herramientas de apoyo a la decisión":

Volvamos al filósofo griego. Para él existe una diferencia inconmensurable entre un *artefacto* y un *ente natural*. Si bien no contamos con una *filosofía del artefacto* en el *Corpus* aristotélico, sus referencias a la naturaleza permiten no obstante establecer una clara distinción entre los entes naturales y los artificiales. El mismo Aristóteles no se priva de recurrir a la comparación con los artefactos para afinar su definición de un ente natural y de la naturaleza misma. Para decirlo en términos más actuales, digamos que esa distinción entre lo *natural* y lo *artificial* depende de la presencia de la *iniciativa* en los entes naturales, a diferencia de los artefactos. Ellos, los naturales, poseen un principio de acción o desarrollo que les es intrínseco y esencial. Sin temor a distorsionar la idea aristotélica, digamos que la *naturaleza*, como principio de acción en el ente donde reside a título esencial y no accidental, es la misma esencia de ese ente, pero considerada desde un punto de vista dinámico.³ Los entes naturales son semovientes, o "automóviles", por decirlo de una manera provocativa. Lo que nosotros llamamos un "automóvil", pues, no lo es, aristotélicamente hablando. Para que la "esencia" de un vehículo se ponga en movimiento necesita de, por lo menos, un agente externo: alguien que lo ponga en marcha y lo conduzca (o al menos que le proporcione ciertas instrucciones, como en el caso de los vehículos "inteligentes").⁴ Y, lo más importante, para llegar a ser lo que son, los artefactos necesitan de un principio eficiente extrínseco, sin el cual ni siquiera existirían: el hombre. Una bacteria, en cambio, no necesita de ningún agente externo para llegar al ser, reproducirse, alimentarse y crecer, por ejemplo. Basta con que cuente con el entorno apropiado y ella sola se encarga del resto. El lenguaje, que no es ingenuo, lo expresa claramente: un ente natural se engendra o nace; un artefacto se fabrica o construye. Un ente natural muere. Un artefacto se rompe o deja de funcionar, pero no muere, aun cuando en el lenguaje

³ Es importante esa aclaración de Aristóteles en cuanto al carácter de "esencial" de la naturaleza, pues podría darse el caso de un médico que se cura a sí mismo. *Física*, II, 1, 192b 21-24. El médico no cura al hombre en cuanto médico, sino a la naturaleza del hombre, en este caso, su parte visible.

⁴ *Física*, II, 1, 192b 27-29.

coloquial solemos emplear esa expresión.⁵ Solemos decir: “mi teléfono está muerto”, pero quien lo fabrica no dice: “nació un teléfono”, o “he dado a luz un teléfono celular”.

Un artefacto, estrictamente hablando, no posee “naturaleza”, y, por lo tanto, jamás tomará la *iniciativa*.⁶ El ChatGPT, por ejemplo, es un artefacto o artificio que jamás preguntará nada porque preguntar algo es una *iniciativa*.⁷ ChatGPT responde, no pregunta. Y sus respuestas, cuando se trata de asuntos comprometidos, están tan ingeniosamente elaboradas algorítmicamente, que difícilmente, por no decir nunca, proporcionan elementos de juicio esclarecedores cuando se trata de obtener respuestas a asuntos éticamente complejos. La *corrección política* parece ser el mejor criterio que guía el diseño de los algoritmos de respuestas. ChatGPT, así como cualquier otro artefacto tecnológico, ya se trate de un programa computacional (*softwares* y aplicaciones) o de un dispositivo (*hardwares* y máquinas), son entes incapaces de hacer algo por sí mismos. No solo hay que proporcionarles el entorno apropiado, sino incluso la misma dinámica de funcionamiento, o sea, los famosos algoritmos en el caso de los dispositivos “inteligentes”; hay que darles todo para que *funcionen*. Un ente natural no “funciona”. Sé muy bien que por analogía podemos atribuir “funcionamiento” al ente natural o a sus partes, pero *stricto sensu*, no, como ya señalé más arriba.

Por otra parte, el algoritmo, por más desarrollado que esté, es producido por una mente humana creadora de determinados

⁵ Notemos esto: un ser viviente, por ejemplo, el hombre solo puede “producir” un solo ser viviente de su misma especie, pero puede “producir” un incontable número de artefactos. Dios, como causa eficiente del ser viviente, puede “producir” un incontable número de seres vivientes, mientras que no produce un solo artefacto. Esto, ciertamente, no es una limitación de su naturaleza, sino una mostración de su omnipotencia, la cual no precisa de máquinas para el gobierno del universo.

⁶ La instrumentalidad del artefacto queda clara cuando examinamos el modo en que lo definimos: su “esencia” solo puede expresarse mediante la identificación de su utilidad, cosa que no sucede con los entes naturales.

⁷ Me estoy refiriendo a ChatGPT como un ejemplo que está a punto de volverse obsoleto, aun cuando todavía su uso no se ha masificado e incluso se estudien actualizaciones. Digo esto porque al momento de escribir estas líneas está irrumpiendo con fuerza un nuevo modelo de IA desarrollado por Google: Gemini, cualitativamente diferente de ChatGPT, ChatGPT Plus y BingChat.

procesos. Pero esos procesos vienen marcados por las mismas limitaciones de la inteligencia humana: son incapaces de prever respuestas a todas las situaciones imprevistas, como no sea cambiando el algoritmo... que a su vez será incapaz de prever nuevas situaciones imprevistas. A eso se reduce el "*machine learning*". El ser vivo en cambio, dispone de una asombrosa dinámica de adaptación a entornos desfavorables, aun cuando no pueda preverlos. En el caso del viviente humano, su capacidad de respuesta frente a imprevistos es inagotable.

Primera conclusión: la iniciativa es de los entes naturales, no de los artefactos o artificios.

3. Segundo asunto, el ético

Ha sido Hans Jonas, en su monumental obra *El principio de responsabilidad. Ensayo de una ética para la civilización tecnológica*, publicada en 1979, uno de los primeros en abordar sistemáticamente el descomunal reto que supone llenar lo que él llama un vacío moral respecto de la tecnología. Jonas afirma que las éticas tradicionales no están en condiciones de dar una respuesta apropiada a las perplejidades éticas surgidas en ambientes de alta densidad tecnológica. Para él, la ética tradicional sufre una seria limitación: su campo de aplicación jamás puede superar la duración de una vida humana. En cambio, ahora, gracias a nuestras posibilidades de manipular la naturaleza en su más íntima constitución, podemos afectar la vida de las generaciones futuras. Suponga Ud. lector, que vive en el siglo III a.C. y desea vengarse de su enemigo dándole a beber una pócima tóxica. Esta acción no afectará a los posibles hijos, nietos y bisnietos de su enemigo. Pero si un científico inescrupuloso cree que es bueno manipular el microecosistema genético para que las personas tengan la posibilidad de elegir el biotipo de su descendencia, sí que hijos, nietos, bisnietos, tataranietos y poblaciones enteras podrían verse afectados. Frente a esto la ética clásica, según Jonas, no tendría respuesta. El razonamiento es muy seductor y en principio inatacable. Y es aquí donde Jonas introduce su "principio de responsabilidad", cuyo primer imperativo es que todas nuestras acciones deben ser

compatibles con la vida de otros seres humanos, ahora y en el futuro.

Sin embargo, este “principio responsabilidad” implica, para Jonas, un compromiso con una *heurística del temor*: temor de lo que podemos hacer. Esto parece muy sabio, pero también es algo que indigna al francés Luc Ferry, quien no concibe que Jonas haya escrito lo siguiente: “la filosofía moral tiene que consultar antes a nuestros temores que a nuestros deseos, para averiguar qué es lo que realmente apreciamos”.⁸ Y poco más abajo: “(...) aquí resulta imprescindible la guía del temor. Este es el caso de la ética orientada al futuro que andamos buscando”.⁹

Respuesta de Ferry en su libro *La innovación destructiva* (2014): “Tenemos miedo de todo: del sexo, del alcohol, del tabaco, del exceso de velocidad, de los OGM, del agujero de ozono, de las nanotecnologías, de las microondas, de Turquía, de la globalización, del calentamiento global, de las antenas de los celulares y de mil cosas más abominables aún.” Y prosigue:

Hoy en día, bajo el efecto de las ideologías ecologistas y pacifistas provenientes de la Europa del Norte (...) el miedo ha cambiado de estatuto. Ya no es más considerado como un signo de puerilidad, sino todo lo contrario, como el primer paso hacia la sabiduría entendida en el sentido de ese maldito principio de precaución. En el libro de Hans Jonas, *El principio responsabilidad*, obra que sirve de Biblia a los ecologistas alemanes, encontramos un capítulo titulado de manera significativa “Heurística del miedo” (...).

Para un filósofo griego, esta idea le habría parecido grotesca y de un indefendible absurdo, pues el ser humano dominado por la angustia era lo contrario del sabio. El miedo nos vuelve brutales y malvados, incapaces de pensar libremente y de abrirnos a los demás. Basta ver en qué nos convertimos cuando una pequeña fobia se apodera de nosotros, por ejemplo, la fobia a los insectos, a los reptiles, a los ascensores que se rompen, a los ratones (...); cuando esto pasa, toda reflexión sensata se desvanece para dar lugar a la estupidez y al repliegue sobre sí mismo. El sabio es exactamente lo contrario. Es aquel que, como Ulises, habiendo superado sus miedos, se

⁸ Hans Jonas, *El principio de responsabilidad. Ensayo de una ética para la civilización tecnológica* (Barcelona: Herder, 1995), p. 66.

⁹ *Ibid.*

encuentra en situación de pensar y amar libremente, capaz de ser inteligente y de abrirse a los demás.¹⁰

Los ejemplos siguen, pero la indignación de Ferry es evidente ante esta claudicación de la valentía, uno de cuyos teóricos sistemáticos es, según el autor francés, justamente, Hans Jonas.

Así entonces, una vida es plenamente humana y aparece en toda su grandeza cuando tiene que vérselas con los imprevistos y estos imprevistos poco importan que se den en el siglo de Aristóteles o en el de Jonas. *Nihil novum sub sole*: la única constante en una vida auténticamente humana es la novedad. Los dilemas planteados por la IA no difieren en gravedad del dilema de la Antígona de Sófocles. No tiene sentido entonces, según Luc Ferry, una ética cuyo punto de partida es el miedo.

En los textos evangélicos puede verse también esa exhortación a no temer. Uno de los pasajes más importantes, si no el más importante, Lucas, 1:30, resume este llamado: "No temas María, pues tú gozas del favor de Dios". Y señalo de paso que en el Evangelio abundan los pasajes donde Jesús comienza sus enseñanzas con esta invitación: "no temáis". El miedo es mal consejero, y hasta tal punto lo es que ha usurpado el nombre de "prudencia". Para un griego, la prudencia es lo más alejado del temor.

Si adoptamos la postura de Ferry, inferiremos que el temor es una actitud errónea frente a las interpelaciones de la IA; por eso es preciso hacer lugar a la actitud opuesta, es decir, el coraje o valentía. Una vez más, Aristóteles tiene algo que decir respecto de esto y nos permitirá complementar la *pars destruens* de Ferry con una *pars construens*.

Cuando el Estagirita se refiere a la virtud del coraje o valentía (*andréia*) en la *Ética Nicomáquea* intenta restringir un poco el vasto ámbito de aplicaciones dado por Platón en el *Laques*, 191 d-e, refiriéndola principalmente a la disposición habitual virtuosa del guerrero. Es de provecho recordar que el miedo es definido

¹⁰ Luc Ferry, *L'innovation destructrice* (Paris: Plon, 2014), pp. 9-12. (Traducción propia).

como “la espera de un mal”.¹¹ El valiente en cambio, en su acepción principal, lo es frente al más terrible de los males, que curiosamente no es el de la muerte, sino el más noble de todos: el de la guerra.¹² Sin embargo, y se ha reparado poco en esto, la acepción derivada más importante del coraje o valentía es la valentía política (*andréia politikê*).¹³ Esta valentía es la del ciudadano dispuesto a soportar los riesgos y peligros de sus intervenciones en la vida pública cuando no teme decir la verdad. Decirlo todo, sin temor al escarnio, es lo que los antiguos llamaban la *parresía*, una actitud, si no virtuosa en sí misma, al menos no carente de virtud.¹⁴

Para el caso que nos ocupa entonces, y puesto que la IA tiene indudablemente sus proyecciones en la configuración de nuestro modo de convivir, no parece que una actitud de temor sea la más apropiada para abordarla. Así pues, es pertinente sugerir esta actitud de espíritu valiente como marco general para afrontar nuestras vacilaciones éticas respecto de la IA y no temerle, lo cual implica también no rechazarla de plano y no esperar solamente un mal de ella. Y si esto es así, cabe reflexionar también acerca de su especial instrumentalidad.

No se me escapa que otorgar a la IA el estatuto de “instrumento” podría ser una jibarización de sus alcances. Un algoritmo no es un instrumento a la manera en que un martillo lo es: con Jonas nos planteamos una ética para la civilización tecnológica, mientras que en la era pretecnológica no disponemos de una ética de los artefactos que vaya más allá del uso que hagamos de ellos. Sin embargo, aun así, el carácter de artefacto o artificio de la IA la condena sin remedio a no ser otra cosa que eso, aunque Stanley Kubrick haya imaginado hace casi 60 años una

¹¹ *Ética Nicomáquea*, 1115a 9.

¹² *Ibid.* 1115a 30 ss.

¹³ *Ibid.* 1115b 17.

¹⁴ Es probablemente Michel Foucault uno de los primeros, si no el primero en el siglo XX, que recoge esta vieja noción y su relación con la valentía en algunos de sus trabajos, especialmente en tres de los cursos dictados en el Collège de France: 1981-1982, 1982-1983 y 1984. Ver Claudia Mársico, “La noción de *parrhesía* en M. Foucault a la luz de los estudios sobre las filosofías del círculo socrático. *Aufklärung*”. *Revista de Filosofía* 4, n° 3 (2017): 47-60.

supercomputadora, Hal 9000, capaz de tomar la iniciativa. No obstante, lo de Hal fue debido a un error en la programación, obviamente de origen humano. ¿Habrà que pensar en una ética para creadores de algoritmos? De hecho, esa es la propuesta de Marta Muñiz, quien acuñó el término "Algor-ética".

Su propuesta puede resultar algo ingenua al retrotraernos a un modelo ético ya prácticamente en desuso, la ética principialista. En realidad, la "ética principialista", surgida en el ámbito de la *bioética*, puede ser criticada en tanto los principios que propone pertenecen a tradiciones filosóficas muy diferentes y acaso incompatibles.¹⁵ Lo que puede haber de errado en este principialismo bioético no anula el hecho de que la gramática elemental de los sistemas de la moralidad es *deóntica*. La forma primaria de toda ética es el *deber*, y no hay por qué suponer que esto sea incompatible con una ética de virtudes.

En todo caso, y para ir finalizando, me permito citar lo que Muñiz considera los principios subyacentes en la Algor-ética:

(...) <los principios de la Algor-ética> son: la transparencia (los sistemas de inteligencia artificial deben ser comprensibles para todos); la inclusión (estos sistemas no deben discriminar a nadie, ya que todo ser humano tiene la misma dignidad); la responsabilidad (siempre debe haber alguien que se responsabilice de lo que hace una máquina); la imparcialidad (los sistemas de IA no deben seguir ni crear prejuicios); la fiabilidad (la inteligencia artificial debe ser fiable) y la seguridad y la privacidad (estos sistemas deben ser seguros y respetar la privacidad de los usuarios).¹⁶

Parece un buen comienzo, pero es solo eso, un comienzo.¹⁷

¹⁵ La primera versión de la ética "principialista" puede verse en Tom Beauchamp y James Childress. *Principles of Biomedical Ethics* (Oxford: OUP, 1979).

¹⁶ Marta Muñiz, 2022: <https://ethic.es/2022/11/la-algor-etica-o-los-principios-eticos-de-la-inteligencia-artificial/>

¹⁷ Una muy interesante propuesta ética es la de Alfredo Marcos, "Silencio tecnológico". *Scio. Revista de Filosofía*. 15: 2018: 157-176.

4. Conclusión abierta, con unas pocas propuestas

En primer lugar, he tratado de mostrar que las enseñanzas de Aristóteles siguen y seguirán estando vigentes, y que ellas pueden proporcionarnos criterios ontológicos y éticos para tratar con la IA. Si hay alguna diferencia es que hoy, más que nunca, urge la Filosofía, urge la Ética, urgen, en suma, las Humanidades.¹⁸ Es necesario vencer la inercia de una pereza de pensamiento, cuyo efecto más nocivo es el reemplazo de las decisiones prudentiales por los procedimientos automatizados.

En segundo lugar, necesitamos reemplazar la ética del miedo por una de la valentía en la que lo primero, *a contrario sensu* de la propuesta de Jonas, el temor no lleve la voz cantante. Esto va de la mano con la recuperación de la *parresía*.

Esta valentía, indisolublemente unida a la *parresía*, revitalizará el amor por la verdad, que es lo propio de la Filosofía.¹⁹

Este cambio de actitud, este cultivo de la inteligencia (natural) y de la virtud, nos hará merecedores, tal como lo escribe Aristóteles, del amor de los dioses, que aman en nosotros lo que más nos asemeja a ellos: la inteligencia, sí, pero también la sabiduría, que nunca serán artificiales.²⁰

Referencias

Aristóteles. *Física*. Libro I-II. Traducción, introducción y comentario por Marcelo Boeri. Buenos Aires: Biblos, 1995.

Aristóteles. *Ética Nicomáquea*. Introducción por Emilio Lledó Íñigo. Traducción y notas por Julio Pallí Bonet. Madrid: Gredos, 1985.

Aristóteles. *Metafísica*. Edición trilingüe por Valentín García Yebra. Madrid: Gredos, 1987.

¹⁸ Siempre es de provecho consultar Héctor Padrón, "La urgencia de las humanidades". *Mikael*. 27: 3-34, 1981.

¹⁹ *Metafísica*, 993b 19-20.

²⁰ *Ética Nicomáquea*, 1179a 23-32.

Beauchamp, Tom y Childress, James. *Principles of Biomedical Ethics*. Oxford: OUP, 1979.

Ferry, Luc. *L'innovation destructrice*. Paris: Plon, 2014.

Foucault, Michel. *L'herméneutique du sujet: Cours au Collège de France*, 1981-1982, ed. Frederic Gros. Paris: Gallimard/Seuil, 2001.

Foucault, Michel. *Le Gouvernement de soi et des autres. Cours au Collège de France*, 1982-1983. Paris: Gallimard/Seuil, 2008.

Foucault, Michel. *Le courage de la vérité. Le gouvernement de soi et des autres II. Cours au Collège de France*. 1984. Paris: Gallimard/Seuil, 2009.

Jonas, Hans. *El principio de responsabilidad. Ensayo de una ética para la civilización tecnológica*. Barcelona: Herder, 1995.

Marcos, Alfredo. "Silencio tecnológico". *Scio. Revista de Filosofía* 15, 2018: 157-176.

Mársico, Claudia. "La noción de *parrhesía* en M. Foucault a la luz de los estudios sobre las filosofías del círculo socrático. *Aufklärung. Revista de Filosofía. Revista de Filosofía* 4, n° 3 (2017): 47-60.

Muñiz, Marta. La algor-ética (o los principios éticos de la inteligencia artificial), URL: <https://ethic.es/2022/11/la-algor-etica-o-los-principios-eticos-de-la-inteligencia-artificial/>

Padrón, Héctor. "La urgencia de las humanidades". *Mikael* 27: 3-34, 1981.

El autor es licenciado en Filosofía por la Facultad de Filosofía y Letras de la Universidad Nacional de Cuyo y Master of Arts y Dr. en Filosofía por la Université catholique de Louvain (Bélgica). Actualmente se desempeña como Profesor en la Universidad Gabriela Mistral, de Santiago (Chile). Ha publicado cinco libros y más de 70 artículos en revistas especializadas. Sus intereses son la filosofía medieval, la bioética y la filosofía política.