

Deus ex machina

Cuando el uso de la inteligencia artificial y la robótica puede cambiar el mundo

CARLOS PÉREZ AMADOR
Teniente coronel
del Ejército del Aire

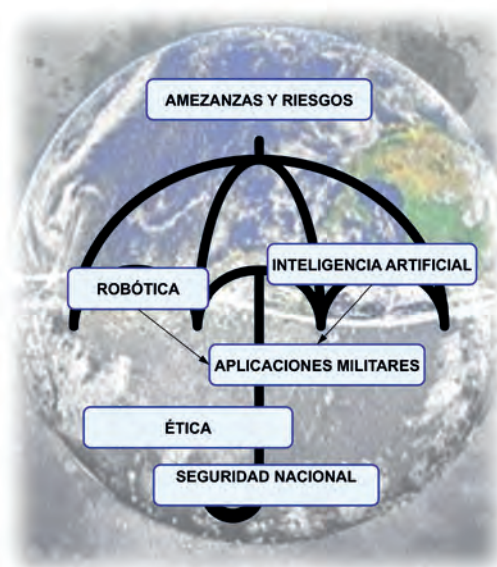
«El hombre necesita la máquina y la organización, pero tiene que dominarlas y humanizarlas en vez de resignarse a ser mecanizado y deshumanizado por ellas. El verdadero peligro para el hombre no está en los riesgos que corre la seguridad material, sino en el oscurecimiento del hombre mismo en su propio mundo humano»

RABINDRANATH TAGORE

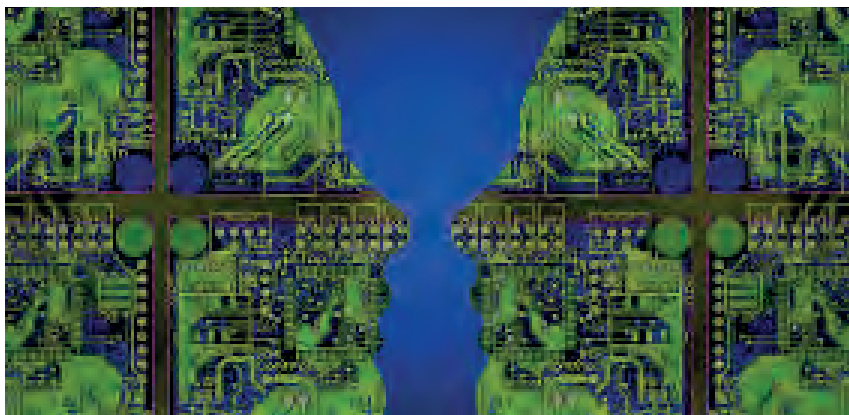
Cualquier tecnología disruptiva es susceptible de ser aplicada en el ámbito de la defensa, aunque su utilización puede desencadenar complejos dilemas éticos y problemas legales. Este fue en su día el paradigmático caso del armamento nuclear o el de los francotiradores, de la misma forma en que hoy en día ocurre con los sistemas de armas basados en la inteligencia artificial¹ (IA) y la robótica. Sucede en aplicaciones más conocidas y controvertidas como las armas letales autónomas (LAWS, por sus siglas en inglés), pero también en otros usos potencialmente perniciosos de la inteligencia artificial. La aplicación de estas tecnologías con fines militares es una realidad en el presente.

Su uso, lejos de reducirse, seguirá aumentando conforme futuros adelantos permitan una mayor autonomía, economía de medios, ventaja sobre el oponente o salvaguarda de vidas y bienes propios.

La lógica del conocimiento natural y el sentido común, así como los valores adquiridos a lo largo de la carrera miliar y las experiencias vitales nos sirven para realizar una primera aproximación a estos conflictos éticos y morales. Sin embargo, resulta pertinente tratar de abordar la problemática desde el enfoque diverso y bastante desconocido que nos puede proporcionar la filosofía: los sistemas filosóficos estimulan el pensamiento crítico que pueden orientar la respuesta a los desafíos éticos planteados a la comunidad internacional y a una nación como España, en la búsqueda de un entorno pacífico, seguro y justo. Este artículo trata de aportar alguna pista segura en esta encrucijada².



Mapa conceptual orientativo. (Imagen: elaboración del autor)



¿El hombre vs la máquina?

DILEMAS ÉTICOS DE LAS NUEVAS AMENAZAS Y RIESGOS

Como ya intuyera el hindú Rabin-dranath Tagore, Nobel de Literatura en 1913, uno de los principales peligros del desarrollo de la IA y la robótica es la propia deshumanización del hombre al dejar su destino en las manos de las máquinas.

Estos desarrollos tecnológicos, sin pretender caer en un alarmismo infundado, puede suponer un cambio de paradigma en las relaciones humanas, un giro de guión o un *deus ex machina* que puede producir resultados inesperados.

Aquí se origina uno de los principales peligros, del que numerosos autores alertan: la llamada superinteligencia artificial o singularidad tecnológica, que podría definirse de forma intuitiva como la capacidad de una máquina para realizar tareas propias de la inteligencia humana de forma más eficaz que una persona.

Así, el conocido astrofísico Stephen Hawking (1942-2018) afirmó en una entrevista en la BBC (2014): «El día que la inteligencia artificial se desarrolle por completo podría significar el fin de la raza humana. Funcionará por sí sola y se rediseñará cada vez más rápido. Los seres humanos, limitados por la lenta evolución biológica, no podrán competir con ella y serán superados».

Este fenómeno se podría denominar efecto caja de Pandora de la IA,

puesto que según Valls Estefanell (2018: 8) es probable que no tengamos una segunda oportunidad con esta tecnología y, consecuentemente, es imprescindible que sea desarrollada de tal forma que comparta nuestros objetivos.

Por otra parte, cuando a la inteligencia artificial se le añade la capacidad de realizar acciones letales, surgen nuevos dilemas éticos relacionados con su uso: el nivel de control humano sobre las máquinas, la ponderación entre sus ventajas frente a sus desventajas, la ausencia de honor al morir a manos de un LAWS, la dificultad de atribuir responsabilidad legal en el caso de que estas armas produzcan daños o su capacidad para cumplir con el DIH (Derecho Internacional Humanitario) en conflicto.

Los LAWS, denominados por algunos como robots asesinos, han despertado una fuerte polémica en el seno de la Organización de las Naciones Unidas (ONU), en alguna institución como Human Right Watch o, en el caso de la sociedad civil, en forma de la popular plataforma Stop Killer Robots, que aboga por su prohibición. Sin embargo, puede ser mucho más eficaz el empleo de la IA en el ámbito cognitivo, tal y como sostiene el coronel Gómez de Ágreda: «Los LAWS no son más peligrosos que otros sistemas de AI [...] el verdadero problema está en la capacidad para influir en percepciones y emociones, mucho más que en una forma más eficiente de matar. Hay que irse a lo cualitativo de la intrusión en el terreno de lo cognitivo más que a lo cuantitativo de cuántas personas puedes matar o a qué distancia».

LA FILOSOFÍA A NUESTRO SERVICIO

Gracias al conocimiento natural y las conductas morales que determinan nuestro comportamiento, podemos realizar una primera aproximación que permita determinar la bondad o maldad de la realidad planteada. Sin embargo, un estudio riguroso de estos dilemas implica la necesidad de subirse a hombros de gigantes para vislumbrar los conceptos e ideas desde la lente de algunos sistemas filosóficos.



Activistas de la plataforma Stop Killer Robots en la puerta de Brandenburgo (Berlín)

Se considera que no es posible dar respuesta a todos los dilemas éticos que se plantean mediante la aplicación de un único sistema o corriente filosófica puesto que puede dar lugar a la simplificación de una realidad compleja despreciando los matices necesarios que la perfeccionan.

Uno de los enfoques más naturales y cercanos al sentido común es el que proporciona el utilitarismo. Este planteamiento permite estudiar desde una perspectiva global, despreciando aquellos detalles de escasa relevancia práctica frente a la visión del beneficio sobre el grupo (Wegeener, 2019: 57).

Así, una crítica que se le puede hacer a la valoración utilitarista es que analiza las ventajas que produciría el uso de estas tecnologías para un contendiente, despreciando los efectos negativos que pueda tener en el enemigo o, incluso, en los daños colaterales. Sería algo equivalente a la máxima de que el fin justifica los medios, a la manera en la que parece aconsejar Maquiavelo a su Príncipe, sin tener en cuenta que estos medios pudieran carecer de ética.

En cambio, se proponen dos sistemas filosóficos –que se aplicarán a distintos ejemplos para su mejor comprensión– para dar respuesta a los dilemas planteados:

- La primera corriente filosófica moderna que resulta interesante aplicar es la del imperativo categórico de Kant. Immanuel Kant (1785), en su obra *Fundamentación de la metafísica de las costumbres*, sostiene que la prueba a la que hay que someter una acción o comportamiento para verificar si es o no ética, es decir buena o mala, es plantearse que ocurriría si dicha acción o comportamiento fuera generalizada³. En este caso, habrá que estudiar el supuesto de que estas tecnologías se aplicasen de forma universal⁴. Así, por ejemplo, un mundo en el que todos dispusieran de la capacidad de utilizar la inteligencia artificial en el ámbito cognitivo para manipular voluntades o de LAWS, sería sumamente hostil y quizás, ni siquiera habitable para el ser humano.

Además, para Kant: «El imperativo categórico sería el que representase una acción por sí misma, sin referencia a ningún otro fin, como objetivamente necesaria.» (Kant, 1785: 29). Es decir, al contrario del utilitarismo, el fin no puede de ninguna manera justificar el empleo de estas tecnologías.

- Por otra parte, las corrientes filosóficas subjetivistas del siglo XX, neopositivismo lógico, el existencialismo o el posmodernismo (Espinoza, 2016) justifican que los debates

éticos y morales carecen de sentido puesto que no se pueden validar científicamente. En contraposición a estas, surge la escuela de Fráncfort, de la que Jürgen Habermas es defensor⁵. Habermas ofrece una interesante perspectiva para analizar los dilemas éticos en estudio: la ética discursiva.

Para Habermas, «El lenguaje distingue y conecta la dimensión de la verdad y la de la justicia. [...], si participamos seriamente, nos comportamos argumentando y dejando argumentar como si todos, incluidos los hombres de cualquier tiempo, tuvieran acceso al diálogo en el que nos encontramos. Y cuando la discusión desemboca en un consenso de los participantes, el enunciado en el que todos coinciden es verdadero [...]». (Gabás, 2000: 220)

A pesar de lo utópico de este razonamiento, podría servir para realizar una aproximación a la solución de los dilemas analizando las declaraciones de los estados participantes en las convenciones sobre ciertas armas convencionales (CCW)⁶, auspiciadas por la ONU. Desde el punto de vista de la ética discursiva de Habermas, se puede convenir que los resultados recogidos en la tabla que se ofrece a continuación se acercan a las condiciones ideales necesarias para alcanzar una respuesta verdadera o práctica, en el consenso que pudiera lograrse en el seno de dichos CCW. Así, un gran número de personas, representados por los ponentes que participan en estas convenciones, de manera autónoma para expresar libremente aquellos argumentos que representan sus intereses y simétrica en cuanto al valor de sus argumentaciones, podrían alcanzar la respuesta ideal –a través de esta suma sinérgica– a los dilemas que plantea el uso de la inteligencia artificial y la robótica con fines militares.

Analizando las opiniones de los Estados participantes, una amplia mayoría estima que existen serias



Jürgen Habermas. (Imagen: Gorka Lejarcegi)

TABLA RESUMEN POSTURAS ESTADOS EN LOS ÚLTIMOS CCW*

PROHIBICIÓN	SÍ	45
	NO	24
	N/C	28
ÉTICA	SÍ	83
	NO	4
	N/C	10
LIMITACIÓN	HOL	76
	NO	5
	N/C	16
DESARROLLO	SÍ	6
	NO	8
	N/C	83

* (Conference on Certain Conventional Weapons).
Elaboración propia dentro del trabajo de fin del XXII
CEMFAS: IA y robótica en el Entorno Operativo 2035.

implicaciones éticas y se debería retener un control significativo sobre las acciones que ejecutan dichos sistemas de armas basados en la IA y la robótica. Esto es, que un ser humano se mantenga supervisando la cadena de decisiones que llevan a la máquina a realizar una acción potencialmente letal (Human On the Loop, HOL),

El mayor conflicto que aparece desde ambas perspectivas es el relacionado con los LAWS, puesto que su utilización universal crearía un mundo hostil y peligroso para los humanos (imperativo categórico), por otra parte, aplicando el diálogo internacional en el seno de la ONU (CCW), asimilable a la ética discursiva, queda expuesto que la mayor parte de los estados participantes los consideran un problema ético, y muchos abogan por su prohibición o al menos su regulación.

Además, no es la primera vez en la historia en que surge un problema ético similar: como se citó anteriormente, a principios del siglo XX el uso de francotiradores en combate también fue controvertido. Aunque, como sabemos su uso terminó aceptándose, inicialmente se generó un vacío legal que pudo afrontarse según la referencia definida por la Cláusula Martens: «Hasta que un Código más completo de las Leyes de guerra se emita, las Altas Partes Contratantes juzgan oportuno declarar que, en los casos no incluidos

en las disposiciones reglamentarias adoptadas por ellas, las poblaciones y los beligerantes quedan bajo la protección y el imperio de los principios del derecho internacional, tal como resultan de los usos establecidos entre naciones civilizadas, de las leyes de la humanidad y las exigencias de la conciencia pública». Convención de la Haya II, 29 de julio del año 1899

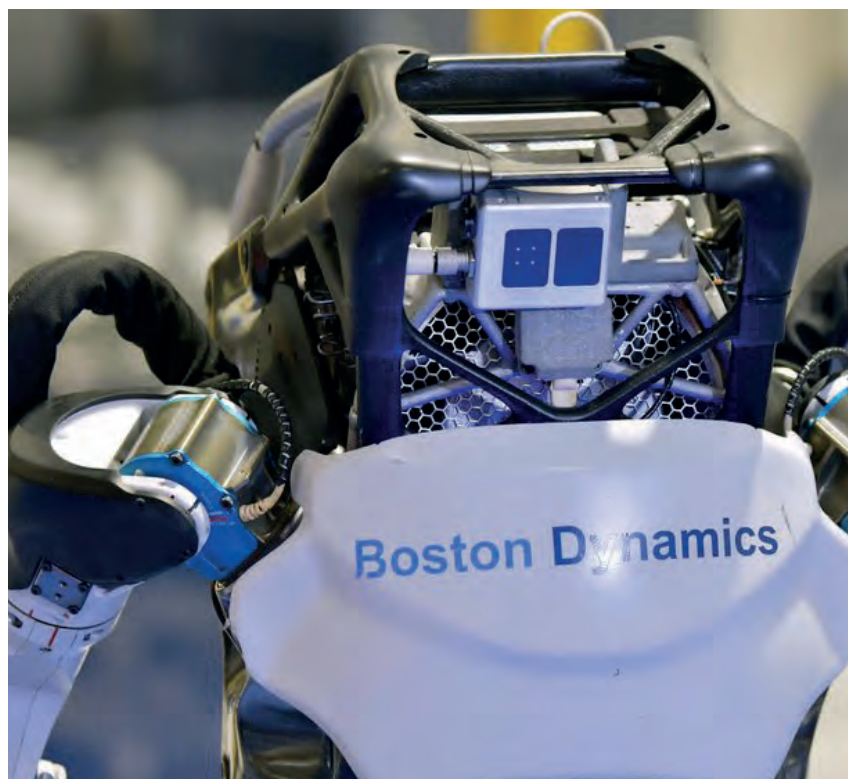
Así, puede considerarse que los usos y costumbres de la guerra han terminado aceptando estas armas y tácticas como moralmente aceptables.

En el caso que nos ocupa, uno de los principales problemas a los que se enfrenta la regulación internacional es la dificultad de alcanzar un consenso que le dé sentido a cualquier decisión y norma, y que facilite el compromiso de las partes. Como decíamos, no es la primera vez que esto ocurre, pero en esta ocasión se agrava por dos motivos fundamentales: por una parte, la rápida

evolución de la tecnología y las capacidades derivadas; y, por otra, debido a la dificultad de acordar las más básicas definiciones conceptuales relativas a la IA y la robótica.

Tomando como referencia las opiniones de los Estados y desde el enfoque que proporciona el razonamiento en la ética discursiva de Habermas, parece sensato pensar que la comunidad internacional debería regular o incluso prohibir el uso y desarrollo de los LAWS, y abrir el debate del uso de la IA en el ámbito cognitivo.

Sea como fuere, parece indudable que el devenir de los tiempos obliga a cualquier ejército moderno a prepararse para las amenazas presentes y futuras, por lo que se hace indispensable continuar avanzando en ambiciosos proyectos como el BAC-SI del Ejército del Aire (Base Aérea Conectada Sostenible e Inteligente), si se desea mantener la posición geopolítica actual.



Robot humanoide. (Imagen: Boston Dynamics)



Vehículos terrestres no tripulados cuadrúpedos.
(Imagen: Ghost Robotics)

La inteligencia artificial y la robótica han venido para quedarse entre nosotros, y su imbricación dentro de la humanidad y por ende de las fuerzas armadas va a ser cada vez mayor. Será necesario por lo tanto adaptar su uso de manera racional, para evitar que las máquinas oscurezcan el alma de las personas y deshumanicen sus relaciones sociales; incluso en los conflictos bélicos. ■

PALABRAS CLAVE

Ética, derecho humanitario, inteligencia artificial, robótica, defensa.

BIBLIOGRAFÍA

- BBC (2014) Stephen Hawking: *La inteligencia artificial augura el fin de la raza humana*. -BBC News Mundo. Disponible en: https://www.bbc.com/mundo/ultimas_noticias/2014/12/141202_ultnot_hawking_inteligencia_artificial_riesgo_humanidad_egn [Consulta: 23 de marzo de 2021].
- Espinosa, J. (2016) *Didáctica de la Filosofía*. Disponible en: <https://rafaelrobles.com/filo/index.php?title=Habermas> [Consulta: 18 de febrero de 2021].
- Gabás, R. (2000) *Escuela de Frankfurt*. Disponible en: [http://e-spacio.uned.es/fez/eserv/bibliuned:Endoxa-1999C97C7A3E-](http://e-spacio.uned.es/fez/eserv/bibliuned:Endoxa-1999C97C7A3E-24A6-309C-84D2-CD52F39079E4/escuela_frankfurt.pdf)

24A6-309C-84D2-CD52F39079E4/escuela_frankfurt.pdf [Consulta: 18 de febrero de 2021].

- Kant, I. (1785) *Fundamentación de la Metafísica de las Costumbres*. Disponible en: http://pmrb.net/books/kantfund/fund_metaf_costumbres_vD.pdf.

- ONU (n.d.). Disponible en: <https://www.un.org/disarmament/the-convention-on-certain-conventional-weapons/> [Consulta: 18 de febrero de 2021].

- Rouhiainen, L. (2018) *Inteligencia artificial*, p. 352. Disponible en: https://planetadelibrosar0.cdnstatics.com/libros_contenido_extra/40/39307_Inteligencia_artificial.pdf.

- *The Convention on Certain Conventional Weapons* (sin fecha). Disponible en: <https://www.un.org/disarmament/the-convention-on-certain-conventional-weapons/> [Consulta: 18 de febrero de 2021].

- Valls Estefanell, M. (2018) *La inteligencia artificial y su encaje en las Estrategias de Seguridad Nacional*. Instituto Español de Estudios Estratégicos. Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=6959963&info=resumen&idioma=SPA> [Consulta: 30 de julio de 2020].

- Wegener, J. (2019) *Dilemas de la aplicación de la inteligencia artificial y la robótica en las fuerzas armadas*. Trabajo Final del Curso de Estado Mayor de las FAS, CESEDEN, Madrid.

NOTAS

¹Según Rouhiainen (2018: 17): «Es la habilidad de los ordenadores para hacer actividades que normalmente requieren

inteligencia humana. Pero, para brindar una definición más detallada, podríamos decir que la IA es la capacidad de las máquinas para usar algoritmos, aprender de los datos y utilizar lo aprendido en la toma de decisiones tal y como lo haría un ser humano».

²Algunas de las ideas que se expondrán serán para algunos lugares comunes o temas recurrentes dentro de la literatura y los estudios relacionados con el uso de la IA y la robótica. No obstante, es necesario sacarlos a colación, aunque sea de forma superficial, para tratar de alcanzar ciertas conclusiones válidas dentro de la incertidumbre de los dilemas éticos que se suscitan.

³Kant entiende que la ética se puede entender con el análisis individual.

⁴«El imperativo categórico es, pues, único, y es como sigue: obra solo según una máxima tal que puedas querer al mismo tiempo que se torne ley universal.» (Kant, 1785: 35)

⁵Nacido en Düsseldorf (Alemania) en 1929, recibió el premio Príncipe de Asturias en 2003.

⁶«La Convención sobre Prohibiciones o Restricciones del Empleo de Ciertas Armas Convencionales que Puedan Considerarse Excesivamente Nocivas o de Efectos Indiscriminados, enmendada el 21 de diciembre de 2001, generalmente conocida como Convención sobre Ciertas Armas Convencionales o CAC, también se conoce como la Convención sobre Armas Inhumanas». (ONU, n.d.) (traducción del autor).