

# Dar un paso atrás: por qué la reglamentación multilateral de la autonomía de los sistemas de armas es difícil y, a la vez, imprescindible y realizable

**Frank Sauer**

Frank Sauer es investigador principal de la Universidad Bundeswehr de Múnich. Integra el Panel Internacional para la regulación de los sistemas de armas autónomos y el Comité Internacional para el control de las armas robóticas, que copatrocinó la campaña internacional Stop Killer Robots. Correo electrónico: [frank.sauer@unibw.de](mailto:frank.sauer@unibw.de).

## Resumen

*En este artículo, se explica por qué la regulación de los sistemas de armas autónomos, que implica la codificación de una obligación jurídicamente vinculante de conservar un control humano significativo en el uso de la fuerza, es una tarea complicada en el marco la Convención de las Naciones Unidas sobre ciertas armas convencionales. Se trata de una tarea complicada porque requiere un nuevo lenguaje diplomático y porque el valor militar de la autonomía de las armas es difícil de prever en esta época sombría en lo que atañe al control de armas. En este artículo, se sostiene que, a pesar de todo, la regulación es imprescindible, porque los riesgos estratégicos y éticos superan las ventajas militares de una autonomía irrestricta de las armas. Con este fin, se ofrecen ideas sobre cómo se puede acelerar la implementación de la regulación.*

**Palabras clave:** inteligencia artificial, sistemas de armas autónomos letales, época sombría en materia de control de armas, regulación, Convención sobre ciertas armas convencionales, estabilidad estratégica, dignidad humana.

\*\*\*

## Introducción

La Convención de las Naciones Unidas (ONU) sobre ciertas armas convencionales (CCAC) es el epicentro del debate mundial sobre la autonomía de los sistemas de armas. El propósito de la CCAC es “prohibir o restringir el uso de tipos específicos de armas que se considera que causan sufrimientos innecesarios o injustificados a los combatientes o que afectan indiscriminadamente a los civiles”<sup>1</sup>. En el lenguaje de la CCAC, se hace referencia al tema de la autonomía de las armas como “tecnologías emergentes en el ámbito de los sistemas de armas autónomos letales” (SAAL). En noviembre de 2019, los Estados Partes en la CCAC decidieron continuar, una vez más, sus deliberaciones sobre los SAAL. Sin embargo, por primera vez, las conversaciones, que anteriormente habían tenido lugar entre 2014 y 2016 en el contexto de reuniones informales y, a partir de 2017, en el marco de un órgano subsidiario de expertos denominado Grupo de Expertos Gubernamentales (GEG), tuvieron la finalidad de obtener un resultado específico. Durante diez días en 2020 y durante una cantidad aún no definida de días en 2021 (durante la próxima Conferencia de Examen de la CCAC), el GEG ha tenido y tendrá la tarea de aclarar y desarrollar “aspectos del marco normativo y operacional” sobre los SAAL<sup>2</sup>. Asimismo, en el anexo III de su informe de 2019, los Estados Partes aprobaron once principios rectores para tener en cuenta en el futuro<sup>3</sup>. Después de que se pospusiera y luego se realizara en modalidad híbrida la reunión de cinco días de 2020 debido a la pandemia de COVID-19, la segunda reunión debió suspenderse, y aún no se ha definido cuándo se reiniciarán las conversaciones y qué formato adoptarán.

Mientras que algunos Estados –en particular, Rusia– no han mostrado interés en desarrollar nuevos instrumentos jurídicos internacionales en el marco de la CCAC, aduciendo que “las preocupaciones que generan los SAAL pueden resolverse mediante la implementación rigurosa de las normas jurídicas internacionales en vigor”<sup>4</sup>, otros Estados –Alemania, por ejemplo– afirman que ya

1 Oficina de Asuntos de Desarme de las Naciones Unidas en Ginebra, “The Convention on Certain Conventional Weapons”, disponible en <https://www.un.org/disarmament/es/> (todas las referencias de internet fueron consultadas en diciembre de 2020).

2 ONU, *Meeting of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects: Revised Draft Final Report*, Doc. onu CCW/MSP/2019/CRP.2/Rev.1, Ginebra, 15 de noviembre de 2019 (Informe final de la reunión de la CCAC), p. 5, disponible en <https://tinyurl.com/y3gjy7mk>.

3 *Ibid.*, p. 10.

4 Federación de Rusia, *Potential Opportunities and Limitations of Military Uses of Lethal Autonomous Weapons Systems: Working Paper Submitted by the Russian Federation*, doc. ONU CCW/GGE.1/2019/WP.1, 15 de marzo de 2019, p. 5, disponible en <https://tinyurl.com/yx9op3n4>.

se ha conseguido un “hito importante” con el informe de 2019 antes mencionado, y hasta definen a los once principios rectores aprobados como una “regulación políticamente vinculante”<sup>5</sup>.

Entretanto, en el marco de la campaña internacional Stop Killer Robots (Campaña Killer Robots, CKR) se critica a la diplomacia de la CCAC por “moverse a paso de tortuga”, por tener pocas ambiciones y lograr resultados magros a pesar de la amplia oposición pública a los SAAL y al hecho de que alrededor de treinta países (veintiséis de los cuales son Estados Partes de la CCAC) reclaman la negociación inmediata de un nuevo instrumento jurídicamente vinculante en lugar de que continúen las conversaciones sobre marcos y principios, que la CKR tiende a considerar vagos y redundantes, respectivamente<sup>6</sup>.

Para empezar, es necesario observar que el término SAAL es, en sí, problemático. Después de todo ni la “letalidad” ni la “autonomía” son factores decisivos en el debate. La aplicación militar de la fuerza no letal también plantea problemas (un ejemplo es la prohibición de las armas láser cegadoras) y el término “autonomía”, desde una perspectiva filosófica, asigna, inadecuadamente, un carácter antropomórfico a máquinas con escasa agencia e incapaces de razonar, reflexionar o de asumir responsabilidades. No obstante, el término SAAL ya es de uso generalizado, por lo que también lo emplearé en este artículo. Asimismo, usaré el término “regulación” –en lugar de, por ejemplo, “prohibición”–, porque lo que normalmente se entiende que codificará un posible nuevo derecho internacional vinculante sobre esta cuestión no es una prohibición de una categoría de armas, sino una obligación positiva de conservar un control humano significativo en el uso de la fuerza militar. Y si bien podría argumentarse que garantizar un control humano significativo y prohibir las armas autónomas (también llamadas “robots asesinos”) son dos caras de la misma moneda, esas dos caras representan, sin embargo, distintas formas de abordar la cuestión, como intentaré explicar más adelante. Por último, usaré el término “difusión” en lugar de “proliferación” de la tecnología, porque este último sugiere una distribución desde uno o unos pocos puntos de partida (como es el caso de la proliferación nuclear), mientras que el primero indica una distribución multidireccional a partir de numerosas fuentes, una representación más apropiada para este caso que se caracteriza por la gran disponibilidad (a veces, incluso comercial) de software y hardware.

A continuación, explicaré primero por qué es tan difícil para todos los que participan en este debate tener una comprensión conceptual del tema y, para los Estados Partes en la CCAC, consensuar una regulación multilateral eficaz de los SAAL. Pienso que encontrar el lenguaje y el marco jurídico adecuados para

5 Oficina de Asuntos Exteriores de Alemania, “Foreign Minister Maas on agreement of guiding principles relating to the use of fully autonomous weapons systems”, nota de prensa, 15 de noviembre de 2019, disponible en [www.auswaertiges-amt.de/en/newsroom/news/maas-autonomous-weapons-systems/2277194](http://www.auswaertiges-amt.de/en/newsroom/news/maas-autonomous-weapons-systems/2277194).

6 CKR, “Las campanas de alarma suenan en robots asesinos”, 15 de noviembre de 2019, disponible en <https://www.stopkillerrobots.org/es/news/alarmbells/>; Richard Moyes, “Critical commentary on the ‘Guiding Principles’”, Article 36, noviembre de 2019, disponible en [www.article36.org/wp-content/uploads/2019/11/Commentary-on-the-guiding-principles.pdf](http://www.article36.org/wp-content/uploads/2019/11/Commentary-on-the-guiding-principles.pdf).

conservar un control humano significativo, a la luz del enorme valor militar asignado a la autonomía irrestricta de las armas, es lo que hace que sea sumamente difícil regular los SAAL. Luego, analizaré las consecuencias de la inacción a la vez que sostengo que es imprescindible conservar el control del uso de la fuerza dado que los riesgos estratégicos y éticos superan las posibles ventajas militares. Por último, presento algunas propuestas sobre cómo se podría promover la regulación en la práctica, a pesar del tremendo desafío de reunir la suficiente voluntad política entre los Estados Partes en la CCAC. Por último, ofrezco una breve conclusión.

## **Por qué es complicado regular la autonomía de las armas: dificultades conceptuales y política del poder**

Desde el Secretario General de la ONU, António Guterres, hasta miembros destacados de las comunidades de la inteligencia artificial y la tecnología<sup>7</sup>, amén de la mayoría de los Estados Partes en la CCAC, casi todos están de acuerdo en que los SAAL generan diversos interrogantes y dilemas jurídicos, estratégicos y éticos<sup>8</sup>. Aun así, entre los Estados Partes en la CCAC, todavía queda mucho por recorrer antes de que se logre el consenso acerca de nuevas normas internacionales jurídicamente vinculantes. Regular la autonomía de las armas en este foro multilateral no es un asunto sencillo. Tal como sostengo en este apartado, hay dos razones que explican esa dificultad. La primera es que la autonomía de las armas es una cuestión bastante escurridiza y difícil de conceptualizar. La segunda es que tiene un valor militar percibido especialmente alto, y el panorama geopolítico actual no es propicio al avance en el control de las armas nuevas.

Todo debate en torno a las dificultades conceptuales asociadas a la autonomía de las armas debe comenzar señalando un malentendido habitual: la falta de avances en la CCAC no puede atribuirse a que los Estados Partes aún no hayan llegado a un acuerdo sobre la definición de los SAAL<sup>9</sup>. Por el contrario, tiene más que ver con el hecho de que el intento por encontrar una definición de los SAAL estuvo mal concebido desde el principio, algo que merece un análisis más pormenorizado.

7 Future of Life Institute (FLI), “Autonomous weapons: An open letter from AI and robotics researchers”, 28 de julio de 2015, disponible en <https://futureoflife.org/open-letter-autonomous-weapons/>; FLI, “An open letter to the United Nations Convention on Certain Conventional Weapons”, 21 de agosto de 2017, disponible en <https://futureoflife.org/autonomous-weapons-open-letter-2017/>.

8 Mary Wareham, “As killer robots loom, demands grow to keep humans in control of use of force”, Human Rights Watch, 2020, disponible en <https://www.hrw.org/world-report/2020/country-chapters/global-0>.

9 La necesidad de lograr una definición común de los SAAL es una noción compartida entre los Estados Partes en la CCAC, y algunos aún la consideran un requisito para que las conversaciones lleguen a buen puerto. Como ejemplo de esta línea de pensamiento, v. el resumen del debate del GEG reunido en 2019 preparado por su presidente: “Algunas delegaciones eligieron abordar la cuestión de las definiciones, con distintas opiniones sobre la necesidad de contar con definiciones –prácticas o de otro tipo– para seguir avanzando en la labor del Grupo”. ONU, *Report of the 2019 Session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems: Chair’s Summary*, doc. ONU CCW/GGE.1/2019/3/Add.1, 8 de noviembre de 2019, p. 3, disponible en <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G19/320/72/PDF/G1932072.pdf?OpenElement>.

Los primeros dos o tres años del proceso de la CCAC sobre los SAAL estuvieron teñidos de confusión y polémicas en lo que respecta a la definición. Se requería un esfuerzo considerable para delinear el debate sobre los SAAL a partir de las disputas en torno a los vehículos aéreos no tripulados (drones), así como para evitar asignar un carácter antropomórfico a los SAAL al considerarlas un reemplazo de los soldados humanos individuales<sup>10</sup>. Todas las partes interesadas buscaban –y unas cuantas lamentaban su falta– una “posible definición de SAAL”, a veces deliberadamente a fin de justificar la lentitud en el terreno político. Los fundamentos eran que el control de las armas siempre requiere una categorización precisa del objeto en cuestión, por ejemplo, una mina terrestre, antes de que pueda emprenderse cualquier acción regulatoria.

No obstante, en el caso de los SAAL, la antigua costumbre de definir y después regular una categoría discreta de armamento militar no es aplicable<sup>11</sup>. Después de todo, cabe pensar que casi todos los sistemas de armas actuales y futuros pueden poseer funciones autónomas, y nadie podrá acertar el grado de dependencia humana de un sistema dado solo con observar el exterior. En el pasado, para el control bilateral de las armas nucleares entre Estados Unidos y la Unión Soviética, posteriormente Rusia, se implementaban controles de armas cuantitativos mediante la definición exacta y compartida de las normas de recuento y el empleo de dichas normas en regímenes de verificación<sup>12</sup>. Del mismo modo, en el ámbito del control multilateral de las armas convencionales, el Tratado sobre fuerzas armadas convencionales en Europa, hoy suspendido, dependía en gran medida de la definición y el recuento del armamento militar<sup>13</sup>. Sin embargo, las dificultades que plantean los SAAL no se resuelven intentando definir una categoría de sistema de armas –“SAAL”, separada de la categoría de “sistemas que no son SAAL” por una serie de criterios específicos– y luego contando y limitando su cantidad. De hecho, en la arquitectura de sistema de sistemas militar moderna, “algunos componentes de los SAA [sistemas de armas autónomos] son intangibles y pueden estar distribuidos geográficamente, [por lo tanto] no queda claro [...] dónde y cuándo empiezan y dónde y cuándo terminan los SAA”<sup>14</sup>. Así pues, en términos generales, la dificultad radica en elaborar una nueva norma para adaptar la relación entre humanos y máquinas a la guerra del siglo XXI. Es necesario adoptar un enfoque cualitativo en lugar de cuantitativo, lo que, a su vez, requiere un

10 Léonard Van Rompaey, “Shifting from autonomous weapons to military networks”, *Journal of International Humanitarian Legal Studies*, vol. 10, n.º 1, 2019, pp. 112-119, disponible en [https://brill.com/view/journals/ihtls/10/1/article-p111\\_111.xml](https://brill.com/view/journals/ihtls/10/1/article-p111_111.xml).

11 Elvira Rosert y Frank Sauer, “How (not) to stop the killer robots: A comparative analysis of humanitarian disarmament campaign strategies”, *Contemporary Security Policy*, 30 de mayo de 2020, disponible en <https://tinyurl.com/y23o8lo6>.

12 Jozef Goldblat, *Arms Control: The New Guide to Negotiations and Agreements*, SAGE Publications, Londres, 2002, cap. 5.

13 Tratado sobre fuerzas armadas convencionales en Europa, 19 de noviembre de 1990, disponible en <https://www.osce.org/files/f/documents/7/3/14092.pdf>.

14 Maya Brehm, *Defending the Boundary: Constraints and Requirements on the Use of Autonomous Weapon Systems Under International Humanitarian and Human Rights Law*, Geneva Academy Briefing n.º 9, mayo de 2017, pp. 15-16.

nuevo lenguaje diplomático que englobe los desarrollos tecnológicos subyacentes, lenguaje que aún no dominan ni los Estados Partes ni la sociedad civil.

Afortunadamente, el proceso de conceptualización de esta cuestión y su traducción a un lenguaje diplomático ya se ha iniciado, y se han logrado algunos avances. Después de casi seis años, la codificación de una obligación positiva de control humano sobre los sistemas de armas tiene un lugar cada vez más consolidado en el centro del debate. Esta noción general, que ha ido cobrando prominencia tras el llamamiento a un “control humano significativo”, propuesta realizada originariamente por la ONG Article 36<sup>15</sup>, está encontrando eco en la sociedad civil, así como en un número cada vez mayor de Estados Partes en la CCAC. En consecuencia, hoy en día, la conceptualización encuentra gran aceptación tanto en los textos académicos como en los debates diplomáticos, entre otras razones, porque Estados Unidos y el Comité Internacional de la Cruz Roja (CICR) la han adoptado. No se trata de una definición categórica de los SAAL (en oposición a los sistemas que no son SAAL) a través de un conjunto de criterios, sino de un conocimiento funcional del fenómeno<sup>16</sup>.

Desde una perspectiva funcionalista, la cuestión de los SAAL se comprende mejor como un tema de autonomía *en* un sistema de armas, es decir, la autonomía de una máquina en lugar de la de un humano que realiza determinada función (o funciones) durante el funcionamiento del sistema<sup>17</sup>. Toda operación militar que culmina en un ataque a un objetivo puede sistematizarse en una cantidad de pasos discretos que conforman una cadena de exterminio o un ciclo de selección y ataque de objetivos<sup>18</sup>. Entre esos pasos, se encuentran la búsqueda, la localización, el seguimiento, la designación, el ataque (y la evaluación posterior). Muchos sistemas de armas son capaces de realizar algunas de las funciones del ciclo de selección y ataque de objetivos sin la participación ni la supervisión humanas, por ejemplo, un dron que se traslada desde un punto de la trayectoria al siguiente guiándose por la navegación satelital, ejecuta la parte de “búsqueda” de la función sin necesidad de que se lo controle de forma remota. En cambio, un arma autónoma completa todo el ciclo de selección y ataque de objetivos –incluidas las etapas finales de designación y ataque– sin intervención humana. En el debate sobre los SAAL, la atención está puesta principalmente en las últimas dos funciones (a las que el CICR

15 Richard Moyes, “Key elements of meaningful human control”, Article 36, abril de 2016, disponible en [www.article36.org/wp-content/uploads/2016/04/MHC-2016-FINAL.pdf](http://www.article36.org/wp-content/uploads/2016/04/MHC-2016-FINAL.pdf). Article 36 es miembro de la CKR.

16 CICR, *Autonomous Weapon Systems: Implications of Increasing Autonomy in the Critical Functions of Weapons*, Ginebra, 2016; Departamento de Defensa de EE.UU. (DoD), directiva 3000.09, “Autonomy in weapon systems”, 2012 (modificada en 2017); Paul Scharre, *Army of None: Autonomous Weapons and the Future of War*, W. W. Norton, Nueva York, 2018.

17 Vincent Boulanin y Maaïke Verbruggen, *Mapping the Development of Autonomy in Weapon Systems*, Instituto Internacional de Estudios para la Paz de Estocolmo (SIPRI), Estocolmo, 2017, disponible en [www.sipri.org/sites/default/files/2017-11/siprireport\\_mapping\\_the\\_development\\_of\\_autonomy\\_in\\_weapon\\_systems\\_1117\\_0.pdf](http://www.sipri.org/sites/default/files/2017-11/siprireport_mapping_the_development_of_autonomy_in_weapon_systems_1117_0.pdf).

18 International Panel on the Regulation of Autonomous Weapons (iPRAW), *Focus on Human Control*, informe iPRAW n.º 5, agosto de 2019, disponible en <https://www.ipraw.org/publications/human-control/>.

denomina funciones “críticas”<sup>19</sup>), porque la mayoría de los efectos de la autonomía de las armas en cuestión provienen de abandonar el control humano y dejar la decisión del uso de la fuerza en manos de una máquina<sup>20</sup>.

Una peculiaridad del enfoque funcional es que nos recuerda que las armas con autonomía en sus funciones críticas ya existen. Transforma la cuestión en un tema del presente y no en una preocupación sobre la tecnología de armas del futuro. Dicho esto, hasta ahora la autonomía de las armas, incluso la de las funciones críticas de selección y ataque de objetivos, solo existe en aplicaciones militares limitadas. La munición israelí diseñada para merodear Harpy probablemente sea el mejor ejemplo de un sistema de armas existente que –si bien solo se emplea en el ataque a firmas de radar– selecciona y ataca objetivos sin supervisión ni control humanos<sup>21</sup>. Por lo tanto, Harpy puede considerarse un sistema de armas autónomo, pues completa el ciclo de selección y ataque de objetivos sin intervención humana. Otros ejemplos son los sistemas de defensa terminal que pueden disparar sin intervención humana, como el Phalanx o el Patriot.

En especial, en las primeras fases del proceso de la CCAC, los sistemas como el Phalanx hicieron que algunos Estados Partes trataran de clasificar como automáticos los sistemas de defensa terminal a fin de evitar que se los incluyera en el debate sobre la autonomía. En esta línea de razonamiento, los sistemas automáticos son estacionarios y están diseñados simplemente para repetir unas pocas acciones programadas previamente en caso de que les lleguen municiones y funcionan dentro de unos parámetros y marcos temporales estrictamente definidos en entornos estructurados y controlados. Los sistemas autónomos, en cambio, están diseñados para tener más tiempo y espacio operativo de maniobra<sup>22</sup>. Lamentablemente, desde el punto de vista de la ingeniería, no hay una distinción clara y consensuada entre automatismo y autonomía; de hecho, los términos “automático” y “autónomo” pueden usarse, y con frecuencia se usan, indistintamente para describir un proceso en el que una función es realizada por una máquina en lugar de un humano<sup>23</sup>.

Una concepción funcional torna superfluo cualquier intento de distinción entre automático y autónomo. Esto supone una ventaja en términos de claridad y sencillez conceptual. Asimismo, lo que dio inicio al debate sobre los SAAL fueron

19 CICR, nota 16 *supra*, p. 7.

20 Para las implicaciones de la autonomía en etapas anteriores del ciclo de selección y ataque de objetivos, que no se tratarán en este artículo, v. Arthur H. Michel, “The killer algorithms nobody’s talking about”, *Foreign Policy*, 20 de enero de 2020, disponible en <https://foreignpolicy.com/2020/01/20/ai-autonomous-weapons-artificial-intelligence-killer-algorithms-nobodys-talking-about/>.

21 Israel Aerospace Industries, “HARPY: Autonomous weapon for all weather”, disponible en [www.iaa.co.il/p/harpy](http://www.iaa.co.il/p/harpy). Una munición diseñada para merodear es un sistema de armas que “merodea” en una zona durante un período de tiempo prolongado esperando que aparezcan objetivos.

22 Frank Sauer, “Stopping ‘killer robots’: Why now is the time to ban autonomous weapons systems”, *Arms Control Today*, vol. 46, n.º 8, 2016, pp. 8-9.

23 Por dar un solo ejemplo, el estándar J3016 de niveles de automatización de vehículos autónomos elaborado por la Sociedad de Ingenieros de la Automoción (SAE) “establece seis niveles de conducción autónoma” y el nivel 5 corresponde a un “vehículo completamente autónomo”. SAE, “SAE standards news: J3016 automated-driving graphic update”, 7 de enero de 2019, disponible en [www.sae.org/news/2019/01/sae-updates-j3016-automated-driving-graphic](http://www.sae.org/news/2019/01/sae-updates-j3016-automated-driving-graphic).

las preocupaciones respecto de la selección autónoma de objetivos humanos, no de misiles, proyectiles de mortero u otras municiones<sup>24</sup>. En consecuencia, el punto decisivo de la cuestión, también respecto de la posible regulación, no es si un sistema ha de caracterizarse como automático o como autónomo, sino a qué objetivos se dirige. Retomaré esta línea de pensamiento en los apartados sobre por qué es imprescindible una regulación desde el punto de vista ético y cómo puede materializarse.

Otra ventaja de la perspectiva funcionalista es que nos permite, en gran medida, evitar tomar una posición respecto de la sofisticación o las características exactas de la tecnología de la que depende el sistema. Los sistemas de defensa terminales, para seguir con el mismo ejemplo, se han empleado durante décadas. Así, no necesariamente se requiere que técnicas como el aprendizaje automático (o lo que esté de moda en el amplio campo de la IA) otorguen autonomía (o automatismo, para el caso) a las funciones críticas de selección y ataque de objetivos. Con todo, la IA es un facilitador nuevo y poderoso. Así pues, la autonomía de las armas no es algo nuevo, pero las últimas innovaciones en materia de IA, como la visión artificial, permiten que los actores utilicen la autonomía de las armas a una escala mucho mayor. En efecto, no hace mucho que la selección de objetivos autónoma ha comenzado a abandonar sus antiguas aplicaciones militares acotadas y se está adoptando de manera generalizada.

La selección autónoma de objetivos que se usa solo para destruir municiones que se aproximan puede alimentar la preocupación general respecto del ritmo cada vez más veloz del combate<sup>25</sup>, pero no es de interés humanitario y, además, ayuda a proteger la vida de los soldados, un dato que habría que tomar en consideración en cualquier regulación posible de los SAAL. Por el contrario, la autonomía utilizada sin restricciones en todos los tipos de sistemas de armas, en diversos contextos operacionales y no solo contra municiones que se aproximan, sino *contra todos los objetivos, sin distinción*, incluidos los humanos, genera más riesgos que beneficios, como explicaré detalladamente más adelante.

En resumen, la razón principal por la cual la regulación de la autonomía de las armas no es una tarea sencilla deriva del hecho de que los Estados Partes en la CCAC tienen que afrontar no ya la búsqueda de una definición común de los SAAL, sino en la determinación, de manera conjunta, de cómo pueden elaborarse los procesos futuros de selección y ataque de objetivos para conservar el control

24 Esta idea fue aportada por una persona anónima que ha revisado este texto.

25 La noción general de la dinámica de acción-reacción que crea la autonomía creciente fue descrita por primera vez por Jürgen Altmann: “Debido a la rapidez de la acción y la reacción, los sistemas de armas autónomos generarían una fuerte presión para el ataque rápido si los dos oponentes los tienen”. Jürgen Altmann, “Military uses of nanotechnology: Perspectives and concerns”, *Security Dialogue*, vol. 35, n.º 1, 2004, p. 63.



humano en el uso de la fuerza militar<sup>26</sup>. Dicho de otro modo, el problema no radica en describir una categoría de armas específica, sino en regular en general cuándo una máquina debe tomar una determinada decisión o ejecutar una determinada función y cuándo debe hacerlo un humano, en particular, en las dos últimas etapas del ciclo de selección y ataque de objetivos.

Esa tarea se complica aún más por el hecho de que, según el contexto operacional y la naturaleza del objetivo, la manera en la que se implemente el control humano puede variar. El sistema de mando y control de combate marítimo de una fragata, por ejemplo, si está diseñado para atacar solo misiles antibuque que se aproximan y funciona en modo autónomo solo por períodos breves mientras la nave se encuentra en el entorno deshabitado del mar, puede considerarse que está bajo control humano “en diseño y uso”<sup>27</sup> aunque las funciones críticas de selección y ataque de objetivos se realicen de manera autónoma. Por el contrario, un arma basada en IA diseñada para acelerar la selección de objetivos en un tanque de guerra en un entorno urbano requeriría que cada uno de los proyectiles fuera disparado por un humano con el suficiente conocimiento de la situación para tomar decisiones informadas, de modo que se considere que el arma permanece bajo el control humano en un sentido real.

En resumen, no existe un único estándar de control humano significativo<sup>28</sup>, porque el control por diseño requiere un estándar mínimo de interacción entre el hombre y la máquina, mientras que en el uso el control se implementa según el caso<sup>29</sup>. Y si bien la defensa contra las municiones que se aproximan continúa siendo una aplicación válida de la autonomía en las funciones críticas de un arma, el debate acerca de los SAAL indicaría que todos los demás objetivos podrían requerir más participación y control humano. Esto convierte la cuestión de los SAAL en un tema más abstracto, complejo y problemático desde el punto de vista intelectual y diplomático que, por ejemplo, la conceptualización de las minas terrestres antipersonal.

La segunda razón por la cual es difícil regular la autonomía de los sistemas de armas es la gran importancia militar que se le asigna. Esto se inscribe dentro de los intereses de los cinco miembros permanentes del Consejo de Seguridad de la ONU, pero también de otros países con tecnología militar avanzada, como, por dar solo dos ejemplos, Israel y Australia. Desde luego, la dificultad en sí no

26 Maya Brehm, “Targeting people”, Article 36, noviembre de 2019, disponible en [www.article36.org/wp-content/uploads/2019/11/targeting-people.pdf](http://www.article36.org/wp-content/uploads/2019/11/targeting-people.pdf); Richard Moyes, “Autonomy in weapons systems: Mapping a structure for regulation through specific policy questions”, Article 36, noviembre de 2019, disponible en [www.article36.org/wp-content/uploads/2019/11/regulation-structure.pdf](http://www.article36.org/wp-content/uploads/2019/11/regulation-structure.pdf); Richard Moyes, “Target profiles”, Article 36, agosto de 2019, disponible en <https://t.co/HZ1pvMnIks?amp=1>; iPRAW, nota 18 *supra*; Vincent Boulanin, Neil Davison, Netta Goussac y Moa Peldán Carlsson, *Limits on Autonomy in Weapon Systems: Identifying Practical Elements of Human Control*, SIPRI y CICR, junio de 2020, disponible en [www.sipri.org/sites/default/files/2020-06/2006\\_limits\\_of\\_autonomy.pdf](http://www.sipri.org/sites/default/files/2020-06/2006_limits_of_autonomy.pdf).

27 iPRAW, nota 18 *supra*, pp. 12-13.

28 Daniele Amoroso y Guglielmo Tamburrini, *What Makes Human Control over Weapon Systems “Meaningful”?*, Comité Internacional para el Control de las Armas Robóticas, agosto de 2019, disponible en [www.icrac.net/wp-content/uploads/2019/08/Amoroso-Tamburrini\\_Human-Control\\_ICRAC-WP4.pdf](http://www.icrac.net/wp-content/uploads/2019/08/Amoroso-Tamburrini_Human-Control_ICRAC-WP4.pdf).

29 Agradezco a la persona anónima que ha revisado el texto y aportado esta aclaración.

es nueva. Se observa en otros procesos regulatorios del pasado reciente, como los de las minas terrestres, las municiones en racimo y las armas láser cegadoras, habiéndose alcanzado la regulación de estas últimas en el marco de la CCAC<sup>30</sup>. Sin embargo, las armas láser cegadoras siempre representaron una capacidad acotada y exótica a la que los Estados podían renunciar sin una percepción asociada de un gran costo militar. Las minas terrestres y las municiones en racimo también tenían campos de uso específicos y, al menos en parte, eran sustituibles. Ese no es el caso de la autonomía de las armas, cuyo impacto genera la percepción de un cambio en las reglas del juego para las fuerzas armadas al menos en dos ámbitos de suma importancia.

En primer lugar, la autonomía de las armas promete un amplio abanico de ventajas operacionales y estratégicas al tornar obsoletos el control y la comunicación constantes. Según afirman quienes la proponen, las ventajas militares de esta innovación son múltiples. Permite un nuevo nivel de multiplicación de la fuerza (un único humano hace funcionar varios, decenas o cientos de sistemas a la vez), permite el uso de estrategias de enjambre (lo que abre nuevas posibilidades de abrumar al enemigo y evitar el contraataque)<sup>31</sup>, reduce los costos de personal e incrementa la indetectabilidad del sistema en el espectro electromagnético (lo que ofrece garantías contra la interrupción o el bloqueo de las comunicaciones). Y más importante aún, evita el retraso inevitable entre la orden remota de un operador humano y la respuesta del sistema. La mayor velocidad de reacción genera una ventaja táctica fundamental respecto de los sistemas del adversario, controlados de manera remota y, por lo tanto, de reacción más lenta. De hecho, la promesa de obtener una ventaja por completar el ciclo de selección y ataque de objetivos a la velocidad de las máquinas es, presumiblemente, la principal motivación de quienes proponen una mayor autonomía de las armas<sup>32</sup>. En segundo lugar, la autonomía de las armas promete ayudar a evitar algunas de las atrocidades que se cometen en la guerra y la vuelve más humana. Como las máquinas no tienen miedo, estrés ni fatiga y están desprovistas de emociones humanas negativas, nunca entran en pánico, ni reaccionan exageradamente ni buscan revancha, según se afirma. Como carecen de instinto de autoconservación, siempre pueden demorar la respuesta al fuego. Supuestamente, permiten no solo una mayor moderación, sino también –con el tiempo, cuando la tecnología lo permita– una mejor discriminación entre civiles y combatientes, con lo que se podría aplicar la fuerza militar de manera más acorde a las normas del derecho internacional humanitario (DIH). Esto se sumaría a un beneficio ético

30 E. Rosert y F. Sauer, nota 11 *supra*.

31 Paul Scharre, *Robotics on the Battlefield, Part II: The Coming Swarm*, Center for a New American Security (CNAS), octubre de 2014, disponible en <https://tinyurl.com/yy4gxs43>; Maaïke Verbruggen, *The Question of Swarms Control: Challenges to Ensuring Human Control over Military Swarms*, Consorcio de Grupos de Expertos independientes de la UE sobre la no proliferación y el desarme, Non-Proliferation and Disarmament Paper n.º 65, diciembre de 2019.

32 Michael C. Horowitz, “When speed kills: Lethal autonomous weapon systems, deterrence and stability”, *Journal of Strategic Studies*, vol. 42, n.º 6, 2019; Jürgen Altmann y Frank Sauer, “Autonomous weapon systems and strategic stability”, *Survival*, vol. 59, n.º 5, 2017.

más general, en un sentido utilitario<sup>33</sup>. En suma, el potencial transformador que se atribuye a la autonomía de las armas, y la cantidad y la calidad de las ventajas militares con las que se la asocia incrementan su importancia en comparación con categorías de armas específicas, como las minas terrestres o las municiones en racimo, que han sido objeto de desarme humanitario en el pasado reciente.

En vista de promesas tan tentadoras y del número creciente de iniciativas en curso en Estados Unidos y China (así como de Rusia, aunque en menor medida) para incorporar innovaciones civiles a fin de incrementar la autonomía de las armas<sup>34</sup>, actualmente no hay demasiado interés en esos Estados en renunciar a las ventajas militares asociadas a este paso, en su opinión, fundamental, para el progreso de la tecnología militar<sup>35</sup>. La firme posición de Estados Unidos, por ejemplo, consiste en continuar explorando el uso de la autonomía de las funciones críticas de los sistemas de armas conforme al DIH<sup>36</sup>. Algunas potencias menos influyentes no están a favor de la regulación ni siquiera en este aspecto. India, por ejemplo, ve la oportunidad de avanzar y cerrar la brecha con las grandes potencias militares del mundo en materia tecnológica<sup>37</sup>. De hecho, después de las campañas contra las minas terrestres y las municiones en racimo, y las iniciativas actuales de desarme en las áreas del comercio de armas (Tratado sobre el comercio de armas) y de las armas nucleares (Tratado sobre la prohibición de armas nucleares), algunos diplomáticos en Ginebra parecen estar sumamente molestos con la acometida de la CKR para que se apruebe un nuevo tratado sobre prohibición de armas.

Además, por lo general, la geopolítica no propicia los avances en materia de control de armas nuevas. El Tratado sobre la prohibición de armas nucleares, que entrará en vigor el 22 de enero de 2021, está considerado por algunos como la excepción a esta regla, pero sus consecuencias para la arquitectura multilateral

- 33 Ronald C. Arkin, "Ethical robots in warfare", *IEEE Technology and Society Magazine*, vol. 28, n.º 1, 2009; Ronald C. Arkin, "The case for ethical autonomy in unmanned systems", *Journal of Military Ethics*, vol. 9, n.º 4, 2010; Ronald C. Arkin, "Governing lethal behavior in robots", *IEEE Technology and Society Magazine*, vol. 30, n.º 4, 2011; Estados Unidos, *Implementing International Humanitarian Law in the Use of Autonomy in Weapon Systems: Working Paper Submitted by the United States of America*, doc. ONU CCW/GGE.1/2019/WP.5, 28 de marzo de 2019, disponible en <https://tinyurl.com/y4xe7tmc>.
- 34 Elsa B. Kania, "In military-civil fusion, China is learning lessons from the United States and starting to innovate", *The Strategy Bridge*, 27 de agosto de 2019, disponible en <https://thestrategybridge.org/the-bridge/2019/8/27/in-military-civil-fusion-china-is-learning-lessons-from-the-united-states-and-starting-to-innovate>; Elsa B. Kania, "AI weapons' in China's military innovation", Brookings Institution, abril de 2020, disponible en [www.brookings.edu/wp-content/uploads/2020/04/FP\\_20200427\\_ai\\_weapons\\_kania\\_v2.pdf](http://www.brookings.edu/wp-content/uploads/2020/04/FP_20200427_ai_weapons_kania_v2.pdf); Frank Sauer, "Military applications of artificial intelligence: Nuclear risk redux", en Vincent Boulanin (ed.), *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, SIPRI, Estocolmo, 2019.
- 35 Michael C. Horowitz, "Artificial intelligence, international competition, and the balance of power", *Texas National Security Review*, vol. 1, n.º 3, 2018, disponible en <https://tnsr.org/2018/05/artificial-intelligence-international-competition-and-the-balance-of-power/>; Zachary Davis, "Artificial intelligence on the battlefield: Implications for deterrence and surprise", *Prism*, vol. 8, n.º 2, 2019, pp. 117-121.
- 36 Estados Unidos, nota 33 *supra*. Quisiera agradecer a la persona anónima que ha revisado y subrayado esta cuestión.
- 37 Shashank R. Reddy, *India and the Challenge of Autonomous Weapons*, Carnegie Endowment for International Peace, junio de 2016, p. 12, disponible en [https://carnegieendowment.org/files/CEIP\\_CP275\\_Reddy\\_final.pdf](https://carnegieendowment.org/files/CEIP_CP275_Reddy_final.pdf).

del control de las armas nucleares aún no están claras. Y, al mismo tiempo, los acuerdos y los tratados multilaterales y bilaterales existentes se están erosionando e incluso algunos han dejado de existir, como el extinto Tratado de fuerzas nucleares de alcance intermedio, el titubeante Plan de acción integral conjunto con Irán, el Tratado de cielos abiertos, que lucha por su continuidad; y pronto, posiblemente, NewSTART, el único tratado bilateral de control de armas nucleares que sigue vigente entre Rusia y Estados Unidos. Concertar un nuevo instrumento internacional jurídicamente vinculante por fuera de la CCAC sería problemático en un entorno geopolítico normal, menos frío. La época actual, sombría para el control de armas en el mundo, hace que alcanzar ese acuerdo sea toda una proeza.

No obstante, es imprescindible regular la autonomía de las armas de manera de restringir la autonomía de las funciones críticas para que estas permanezcan bajo control humano. Después de todo, las consecuencias de la inacción serían problemáticas, porque los riesgos estratégicos y éticos de mediano y largo plazo de la autonomía irrestricta superan las ventajas militares de corto plazo mencionadas anteriormente. Explicaré esta cuestión en dos pasos más adelante, centrándome primero en una serie de implicaciones operacionales y estratégicas, y evaluando luego las implicaciones éticas de la autonomía de las armas en relación con la dignidad humana.

## Por qué es imprescindible regular la autonomía de las armas: implicaciones estratégicas

El posible impacto de la falta de regulación de la autonomía de las armas en las operaciones militares, así como en la paz y la estabilidad estratégica mundial en su conjunto, ha llamado la atención de los académicos durante bastante tiempo<sup>38</sup>. La bibliografía dedicada a este tema señala que las implicaciones de la inacción regulatoria y la consiguiente diseminación rápida de la tecnología que hace posible la autonomía incorporada a las armas abarcan desde las nuevas vulnerabilidades militares hasta el aumento del riesgo de inestabilidad y de las tensiones en los

38 V. J. Altmann, nota 25 *supra*; Armin Krishnan, *Killer Robots: Legality and Ethicality of Autonomous Weapons*, Ashgate, Farnham, 2009, cap. 6; Jean-Marc Rickli, *Some Considerations of the Impact of LAWS on International Security: Strategic Stability, Non-State Actors and Future Prospects*, presentación en la Reunión de expertos para debatir sobre los SAAL en el marco de la CCAC, Ginebra, 16 de abril de 2015, disponible en [https://docs-library.unoda.org/Convention\\_on\\_Certain\\_Conventional\\_Weapons\\_-\\_Informal\\_Meeting\\_of\\_Experts\\_\(2015\)/2015\\_LAWS\\_MX\\_Rickli\\_Corr.pdf](https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Informal_Meeting_of_Experts_(2015)/2015_LAWS_MX_Rickli_Corr.pdf); Paul Scharre, *Autonomous Weapons and Operational Risk*, CNAS Ethical Autonomy Project, Washington, febrero de 2016, disponible en [https://s3.amazonaws.com/files.cnas.org/documents/CNAS\\_Autonomous-weapons-operational-risk.pdf](https://s3.amazonaws.com/files.cnas.org/documents/CNAS_Autonomous-weapons-operational-risk.pdf); Wendell Wallach, "Toward a ban on lethal autonomous weapons: Surmounting the obstacles", *Communications of the ACM*, vol. 60, n.º 5, 2017, p. 31; Irving Lachow, "The upside and downside of swarming drones", *Bulletin of the Atomic Scientists*, vol. 73, n.º 2, 2017; J. Altmann y F. Sauer, nota 32 *supra*; Paul Scharre, "Autonomous weapons and stability", tesis doctoral, King's College London, marzo de 2020, disponible en [https://kclpure.kcl.ac.uk/portal/files/129451536/2020\\_Scharre\\_Paul\\_1575997\\_thesis.pdf](https://kclpure.kcl.ac.uk/portal/files/129451536/2020_Scharre_Paul_1575997_thesis.pdf).

niveles tanto operacional como estratégico<sup>39</sup>. En consecuencia, son especialmente las grandes potencias las que deberían darse cuenta de que detener esa cadena de efectos desestabilizadores no solo es su responsabilidad, sino que también favorece sus propios intereses<sup>40</sup>.

## Diseminación de la tecnología

Para tener una idea de la diseminación que puede esperarse de la tecnología en el campo de los SAAL, los drones pueden servir de ejemplo<sup>41</sup>. China, en particular, actúa como inversor y exportador en este sector<sup>42</sup>. Según datos recogidos por New America Foundation<sup>43</sup>, doce países han ejecutado ataques con drones y treinta y ocho poseen drones armados, al igual que varios actores no estatales, como Hamas, Hizbulá, los rebeldes hutíes y el denominado Estado Islámico.

La tecnología de drones se disemina relativamente rápido debido a su naturaleza de uso dual. La autonomía también es de uso dual. La autonomía de las armas –a condición de que la plataforma en cuestión contenga los sensores y los actuadores necesarios– se reduce principalmente a un software, que se puede transferir y reproducir a un costo desdénable y es vulnerable a los robos por medio de ataques cibernéticos<sup>44</sup>. En consecuencia, cabe esperar que la adopción de funciones autónomas habilitadas por software se disemine rápidamente en el ecosistema del armamento militar actual. Además, los principales agentes de innovación para la autonomía son las empresas de tecnología y las universidades, y no el sector de defensa, por lo que es dudoso que las fuerzas armadas de un país puedan ser las que “llevan la delantera”<sup>45</sup> como lo ve la defensa de Estados Unidos, por ejemplo. Después de todo, el gobierno estadounidense no es el único que incorpora tecnología civil con fines militares gracias a firmas como Google, Microsoft y Amazon. China, por ejemplo, hace lo propio con Tencent, Ali Baba

39 Para el apartado siguiente he consultado el texto de J. Altmann y F. Sauer, nota 32 *supra*; F. Sauer, nota 34 *supra*; Aaron Hansen y Frank Sauer, “Autonomie in Waffensystemen: Chancen und Risiken Für die US-Sicherheitspolitik”, *Zeitschrift für Außen- und Sicherheitspolitik*, vol. 12, n.º 2, 2019.

40 Para la argumentación general, v. Hedley Bull, *The Anarchical Society: A Study of Order in World Politics*, Macmillan, Londres, 1977. Para el caso de la IA, v. Elsa B. Kania y Andrew Imbrie, “Great powers must talk to each other about AI”, *Defense One*, 28 de enero de 2020, disponible en [www.defenseone.com/ideas/2020/01/great-powers-must-talk-each-other-about-ai/162686/?oref=d-river](http://www.defenseone.com/ideas/2020/01/great-powers-must-talk-each-other-about-ai/162686/?oref=d-river).

41 Frank Sauer y Niklas Schönig, “Killer drones: The silver bullet of democratic warfare?”, *Security Dialogue*, vol. 43, n.º 4, 2012; Matthew Fuhrmann y Michael C. Horowitz, “Droning on: Explaining the proliferation of unmanned aerial vehicles”, *International Organization*, vol. 71, n.º 2, 2017; Andrea Gilli y Mauro Gilli, “The diffusion of drone warfare? Industrial, organizational, and infrastructural constraints”, *Security Studies*, vol. 25, n.º 1, 2016.

42 Defense Science Board, *The Role of Autonomy in DoD Systems*, 2012, pp. 69-71.

43 New America, “World of drones”, disponible en [www.newamerica.org/in-depth/world-of-drones/](http://www.newamerica.org/in-depth/world-of-drones/).

44 Sydney J. Freedberg Jr., “Robot wars: Centaurs, Skynet, & swarms”, *Breaking Defense*, 31 de diciembre de 2015, disponible en <http://breakingdefense.com/2015/12/robot-wars-centaurs-skynet-swarms/>; Thomas G. Mahnken, *Technology and the American Way of War Since 1945*, Columbia University Press, Nueva York, 2008, p. 123.

45 Robert O. Work, “Robert Work talks NATO’s technological innovation and the DoD”, *CNAS Brussels Sprouts Podcast*, 11 de enero de 2018, disponible en <https://www.cnas.org/publications/podcast/robert-work-talks-natos-technological-innovation-and-the-dod>.

y Baidu<sup>46</sup>. Así pues, es muy improbable que exista un monopolio en este campo, como el que tuvo Estados Unidos en materia de tecnología de baja detectabilidad en el pasado.

## Nuevas vulnerabilidades operacionales

La otra cara del efecto multiplicador de la fuerza que las fuerzas militares buscan con la tecnología que se disemina con facilidad es la escalabilidad, que genera la posibilidad de que las partes más débiles cambien la dinámica de poder entre ellas y respecto de sus adversarios. La militarización de sencillos drones comerciales por parte del denominado Estado Islámico y el ataque a las instalaciones petroleras de Aramco en Arabia Saudí son anticipos no independientes de lo que puede suceder de aquí en más y demuestran que, a pesar de las capacidades avanzadas de defensa antiaérea, están surgiendo nuevas vulnerabilidades. Desde el punto de vista de las fuerzas de tierra estadounidenses, tener que hacer frente a amenazas serias que vienen del cielo después de décadas de dominio aéreo representa un cambio de paradigma<sup>47</sup>. En consecuencia, Estados Unidos ya está obligado a revisar sus capacidades de defensa antiaérea poniendo el acento en el desarrollo de láseres y microondas. Las soluciones convencionales, como los misiles Stinger, no solo son inadecuadas para defenderse de la multitud de drones pequeños, baratos y descartables que hoy en día hace posible la autonomía, sino que tampoco son convenientes desde el punto de vista económico. Todavía no está zanjada la cuestión de si los nuevos sistemas defensivos pueden remediar esa situación<sup>48</sup>. Basta con decir que la combinación de sistemas no tripulados económicos, la autonomía y la estrategia de enjambre crean nuevos riesgos en general para las tropas en el campo de batalla, para la infraestructura de mando y control, y para los jefes en las denominadas situaciones de acefalía<sup>49</sup>.

Tal como se ha afirmado anteriormente, la posible eliminación del vínculo por control remoto es un incentivo clave para contar con una mayor autonomía en los sistemas de armas, pero ceder el control a una máquina también abre nuevos vectores de ataque. El envío de datos de geolocalización falsos al sistema es un

46 Defense Science Board, *Summer Study on Autonomy*, 2016, p. 45; Elsa B. Kania, *Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power*, CNAS, Washington, noviembre de 2017, disponible en <https://s3.amazonaws.com/files.cnas.org/documents/Battlefield-Singularity-November-2017.pdf?mtime=20171129235805>; E.B. Kania, "In military-civil fusion", nota 34 *supra*.

47 Kelley Saylor, *A World of Proliferated Drones: A Technology Primer*, CNAS, Washington, 2015, p. 29.

48 Sebastien Roblin, "The U.S. Army needs more anti-aircraft weapons – and fast", *War is Boring*, 22 de enero de 2018, disponible en <https://warisboring.com/the-u-s-army-needs-more-anti-aircraft-weapons-and-fast/>.

49 David Barno y Nora Bensahel, "The drone beats of war: The U.S. vulnerability to targeted killings", *War on the Rocks*, 21 de enero de 2020, disponible en <https://warontherocks.com/2020/01/the-drone-beats-of-war-the-u-s-vulnerability-to-targeted-killings/>. Una situación de acefalía es aquella en la que un atacante procura destruir o desestabilizar el liderazgo del adversario y su estructura de mando y control para destruir o diezmar severamente su capacidad de represalia (nuclear).

ejemplo; en 2011, Irán consiguió, aparentemente, secuestrar de esta manera un dron estadounidense de navegación autónoma<sup>50</sup>.

Es más, los sistemas que dependen del aprendizaje automático que emplean redes neuronales profundas<sup>51</sup>, que en la actualidad representan lo más innovador en el ámbito de la visión computarizada, también son especialmente susceptibles a la manipulación. Una cinta reflectante sobre una señal de “pare”, por ejemplo, puede engañar al sistema de reconocimiento de imágenes de un vehículo autónomo. Esta susceptibilidad al error es un problema complejo, pero puede tener solución en una aplicación civil como los vehículos autónomos. Los datos para el entrenamiento de dichos vehículos son abundantes y se consiguen con facilidad, y los vehículos autónomos están diseñados para funcionar cooperativamente en un entorno estrictamente regulado. El campo de batalla es muy diferente: se caracteriza por la escasez de datos y por sus altos niveles de imprevisibilidad y vulnerabilidad<sup>52</sup>. Después de todo, una parte siempre tratará de engañar y manipular los sistemas de su adversario. Estudios de ejemplos de partes en conflicto<sup>53</sup> indican que la visión computarizada habilitará la interferencia de los sistemas de armas autónomos mediante la manipulación del entorno que perciben las máquinas<sup>54</sup> o incluso mediante el reentrenamiento en el caso de que el aprendizaje continúe durante su despliegue<sup>55</sup>. No sería difícil engañar al sistema de reconocimiento facial con fines de selección de objetivos, como demuestra el rápido desarrollo de medidas defensivas contra la vigilancia doméstica<sup>56</sup>.

A medida que crece complejidad del software para los sistemas de armas, aumenta asimismo la cantidad de fallas que contiene. Los errores de programación pueden tener efectos críticos, incluido el fuego amigo<sup>57</sup>. De acuerdo con la teoría de

50 Sydney J. Freedberg Jr., “Drones need secure datalinks to survive vs. Iran, China”, *Breaking Defense*, 10 de agosto de 2012, disponible en <https://breakingdefense.com/2012/08/drones-need-secure-datalinks-to-survive-vs-iran-china/>.

51 Para una revisión crítica, v. Gary Marcus, “Deep learning: A critical appraisal”, Universidad de Nueva York, 2 de enero de 2018, disponible en <https://arxiv.org/ftp/arxiv/papers/1801/1801.00631.pdf>.

52 Michał Klincewicz, “Autonomous weapons systems, the frame problem and computer security”, *Journal of Military Ethics*, vol. 14, n.º 2, 2015; Anh Nguyen, Jason Yosinski y Jeff Clune, “Deep neural networks are easily fooled: High confidence predictions for unrecognizable images”, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 427-436; Z. Davis, nota 35 *supra*, pp. 121-122.

53 Ivan Evtimov *et al.*, “Robust physical-world attacks on deep learning models”, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, disponible en <https://arxiv.org/pdf/1707.08945.pdf>.

54 V. el hoy famoso ejemplo de la tortuga confundida con un rifle, en Anish Athalye, Logan Engstrom, Andrew Ilyas y Kevin Kwok, “Synthesizing robust adversarial examples”, *Proceedings of the 35th International Conference on Machine Learning*, vol. 80, 2018, disponible en <https://arxiv.org/pdf/1707.07397.pdf>.

55 Defense Science Board, nota 46 *supra*, p. 28; Vadim Kozyulin, “International and regional threats posed by the LAWS: A Russian perspective”, PIR Center for Policy Studies, abril de 2016, disponible en <https://tinyurl.com/y4qslefc>; P. Scharre, *Autonomous Weapons and Operational Risk*, nota 38 *supra*, p. 14.

56 Melissa Hellmann, “Special sunglasses, license-plate dresses: How to be anonymous in the age surveillance”, *Seattle Times*, 12 de enero de 2020, disponible en <https://www.seattletimes.com/business/technology/special-sunglasses-license-plate-dresses-juggalo-face-paint-how-to-be-anonymous-in-the-age-of-surveillance/>.

57 P. Scharre, nota 38 *supra*, p. 21.

los accidentes normales<sup>58</sup>, los errores son, básicamente, inevitables. Suceden hasta en dominios con estándares de seguridad sumamente elevados, como las plantas nucleares o los viajes espaciales tripulados<sup>59</sup>. El sector informático puede reducir la cantidad de fallas a 0,1-0,5 errores cada mil líneas de código, lo que significa que los sistemas militares complejos con varios millones de líneas de código, como el software para el jet de combate F-35, contienen miles de errores de programación<sup>60</sup>. La inevitable realidad de tener que actualizar los sistemas con regularidad complica aún más la cuestión<sup>61</sup>; es una fuente potencial de nuevas fallas y nuevos errores que surgen de la interacción entre el software nuevo y el anterior. Los sistemas de aprendizaje automático generan dificultades específicas, porque se presentan como “cajas negras” que no pueden corregirse como el software convencional, lo que equivale a decir que los errores no pueden eliminarse selectivamente<sup>62</sup>.

Por último, la autonomía de las armas da origen a una nueva tendencia a los errores respecto de cualquier interacción restante con los operadores humanos. Aquí entran en juego los preconceptos respecto de la automatización, es decir, la confianza infundada y acrítica en el funcionamiento de un sistema<sup>63</sup>. Para decirlo de manera sencilla, los sistemas autónomos podrían funcionar incorrectamente durante largos períodos de tiempo sin que nadie lo notara<sup>64</sup>. Un humano que comete un error puede entender la situación y corregirlo, pero los SAAL sin supervisión no podrán entender ni reflexionar críticamente en tiempo real como lo haría un humano<sup>65</sup>. Así se generaría el riesgo de una escalada militar no intencional.

## Riesgo de escalada e inestabilidad de las situaciones de crisis

Los sistemas de armas que funcionan sin control humano generan no solo nuevas vulnerabilidades, sino también incertidumbre debido a las interacciones imprevisibles con el entorno, lo que a su vez entraña nuevos riesgos de escalada

58 Charles Perrow, *Normal Accidents: Living with High-Risk Technologies*, Basic Books, Nueva York, 1984 [trad. esp. *Accidentes normales: Convivir con las tecnologías de alto riesgo*, trad. J. M. Iranzo, Modus Laborandi, Madrid, 2009].

59 John Borrie, *Security, Unintentional Risk, and System Accidents*, Instituto de las Naciones Unidas de Investigación sobre el Desarme (UNIDIR), Ginebra, 15 de abril de 2016, disponible en [https://docs-library.unoda.org/Convention\\_on\\_Certain\\_Conventional\\_Weapons\\_-\\_Informal\\_Meeting\\_of\\_Experts\\_\(2016\)/2016\\_LAWS%2BMX%2BPresentations\\_SecurityIssues\\_John%2BBorrie.pdf](https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Informal_Meeting_of_Experts_(2016)/2016_LAWS%2BMX%2BPresentations_SecurityIssues_John%2BBorrie.pdf); P. Scharre, *Autonomous Weapons and Operational Risk*, nota 38 *supra*.

60 P. Scharre, *Autonomous Weapons and Operational Risk*, nota 38 *supra*, p. 13.

61 UNIDIR, *The Weaponization of Increasingly Autonomous Technologies in the Maritime Environment: Testing the Waters*, recursos UNIDIR n.º 4, Ginebra, 2015, p. 8.

62 G. Marcus, nota 51 *supra*, pp. 10-11.

63 Para el ejemplo del sistema de defensa antimisiles Patriot, v. John K. Hawley, *Patriot Wars: Automation and the Patriot Air and Missile Defense System*, CNAS, Washington, enero de 2017, disponible en [www.cnas.org/publications/reports/patriot-wars](http://www.cnas.org/publications/reports/patriot-wars).

64 P. Scharre, *Autonomous Weapons and Operational Risk*, nota 38 *supra*, p. 31; Noel Sharkey y Lucy Suchman, “Wishful mnemonics and autonomous killing machines”, *Proceedings of the AISB*, vol. 136, 2013, pp. 16-17.

65 Defense Science Board, nota 42 *supra*, p. 15.



no intencional e indeseada<sup>66</sup>. En ese sentido, es necesario tener en cuenta la interacción entre dos o más sistemas autónomos en particular. La negociación de alta frecuencia<sup>67</sup> ofrece una analogía útil, porque, con regularidad, tienen lugar procesos de interacción imprevistos e indeseados entre dos o más algoritmos de negociación de funcionamiento autónomo que, a veces, causan los fenómenos denominados *flash crash* y generan pérdidas financieras. La solución, hasta cierto punto, puede encontrarse en la regulación de los mercados financieros, pero sin una regulación internacional jurídicamente vinculante de la autonomía en el campo de batalla, las interacciones imprevisibles de los SAAL podrían terminar en un uso de la fuerza no intencional a la velocidad de las máquinas, e incluso en una guerra accidental antes de que los humanos puedan intervenir<sup>68</sup>. Este riesgo no pertenece a un futuro lejano. En el Dubai Airshow de 2019, el jefe de estado mayor de la Fuerza Aérea de EE. UU., general David Goldfein, presentó la simulación de un ataque a un buque enemigo por medio de una cadena de exterminio automatizada casi por completo. Primero, el buque fue seleccionado por medio de un satélite, luego se transmitieron datos del objetivo a los dispositivos de vigilancia aerotransportada y a los componentes del mando y control. Luego se asignó a un destructor de la Armada de EE. UU. la tarea de lanzar un misil, el único punto en el que el ciclo de selección y ataque de objetivos dependía de la decisión humana, mientras que el resto de la “cadena de exterminio [...] se realizó entre máquinas a la velocidad de la luz”<sup>69</sup>. Cualquier error de una máquina en un sistema de este tipo se propagaría rápidamente si no fuera corregido por un humano debido a la preferencia por la automatización. Es evidente que el error también se propagaría “a la velocidad de la luz” si la participación humana fuese nula. Un ejercicio de simulacro de guerra organizado recientemente por RAND Corporation pone de manifiesto los riesgos de la inestabilidad de las crisis y la escalada no intencional. En el ejercicio, se configuraron las simulaciones de las fuerzas “en modo ‘totalmente automático’ para comunicar las decisiones [...] en un caso, se causó una escalada involuntaria. Los sistemas configurados en modo autónomo reaccionaron usando la fuerza en una situación no prevista en la que los humanos no deseaban usar la fuerza”<sup>70</sup>.

66 André Haider y Maria Beatrice Catarrasi, *Future Unmanned System Technologies: Legal and Ethical Implications of Increasing Automation*, Joint Air Power Competence Centre, noviembre de 2016, p. 10, disponible en [www.japcc.org/wp-content/uploads/Future\\_Unmanned\\_System\\_Technologies\\_Web.pdf](http://www.japcc.org/wp-content/uploads/Future_Unmanned_System_Technologies_Web.pdf); CICR, *Views of the International Committee of the Red Cross (ICRC) on Autonomous Weapon System[s]*, Ginebra, 11 de abril de 2016, p. 3, disponible en [www.icrc.org/en/download/file/21606/ccw-autonomous-weapons-icrc-april-2016.pdf](http://www.icrc.org/en/download/file/21606/ccw-autonomous-weapons-icrc-april-2016.pdf).

67 Gary Shorter y Rena S. Miller, *High-Frequency Trading: Background, Concerns, and Regulatory Developments*, Congressional Research Service, 19 de junio de 2014, disponible en <https://fas.org/sgp/crs/misc/R43608.pdf>.

68 P. Scharre, *Autonomous Weapons and Operational Risk*, nota 38 *supra*, p. 53; J. Altmann y F. Sauer, nota 32 *supra*, pp. 128-132.

69 “Video: Here’s how the US Air Force is automating the future kill chain”, *Defense News*, 2019, disponible en [www.defensenews.com/video/2019/11/16/heres-how-the-us-air-force-is-automating-the-future-kill-chain-dubai-airshow-2019/](http://www.defensenews.com/video/2019/11/16/heres-how-the-us-air-force-is-automating-the-future-kill-chain-dubai-airshow-2019/).

70 Yuna H. Wong et al., *Deterrence in the Age of Thinking Machines*, RAND Corporation, 2020, p. xi, disponible en [www.rand.org/pubs/research\\_reports/RR2797.html](http://www.rand.org/pubs/research_reports/RR2797.html).

Los humanos son menos susceptibles que las máquinas a los errores masivos. Además, a pesar de ser más lentos y de que, a veces, cometen errores, gestionan las situaciones mejor que las máquinas. Tienen la capacidad de comprender situaciones inusuales y su contexto, así como de pensar en las decisiones, su génesis, sus implicaciones y la importancia de la responsabilidad que implican. En términos de gestión de crisis, todas estas características hacen que los humanos sean superiores a las máquinas, que, de momento, solo son capaces de reconocer patrones y de ejecutar acciones predefinidas, y que solo pueden alcanzar un desempeño sobrehumano en situaciones definidas con gran precisión para las que fueron específicamente entrenadas. Si se elimina el control humano, se pierde el papel distintivo de las personas como mecanismo de seguridad versátil.

El famoso caso del teniente coronel Stanislav Petrov es prueba de ello. El ejercicio de la OTAN Able Archer, realizado en 1983, fue malinterpretado por los soviéticos, que pensaron que se trataba de un ataque real con fuerzas nucleares tácticas. En ese momento, un satélite soviético de alerta temprana registró primero un lanzamiento de un misil intercontinental con capacidad nuclear por parte de EE. UU. y luego un par más. Petrov, el oficial que estaba a cargo del centro de alerta temprana en ese momento, determinó (correctamente) que debía tratarse de una falsa alarma y así lo comunicó a la cadena de mando, con lo que evitó una escalada potencialmente nuclear de una situación tensa. La decisión de Petrov no podría haberla tomado un sistema totalmente automatizado. Después de un tiempo, Petrov declaró que había tomado esa decisión siguiendo una intuición, preguntándose acerca de la naturaleza del supuesto ataque y recurriendo a su experiencia con el sistema de alerta temprana, al que no consideraba del todo confiable<sup>71</sup>. Si se eliminara el factor humano del destructor en la cadena de exterminio automatizada casi por completo presentada por el general Goldfein, con lo que se concretaría la ventaja fundamental de la autonomía de las armas en un combate que se desarrollaría a la velocidad de las máquinas, se perdería el “efecto Petrov”. Si bien, en esa situación convencional, eso no necesariamente implicaría el uso involuntario de armas nucleares, la estabilidad estratégica ya se está viendo afectada por la tendencia a aumentar la autonomía de los sistemas militares.

71 Bruce G. Blair, *The Logic of Accidental Nuclear War*, Brookings Institution, Washington, 1993, p. 181; Richard Rhodes, *Arsenals of Folly: The Making of the Nuclear Arms Race*, Simon & Schuster, Londres, 2008, pp. 165-166; David E. Hoffman, *The Dead Hand: The Untold Story of the Cold War Arms Race and Its Dangerous Legacy*, Doubleday, Nueva York, 2009, pp. 6-11, 94-95; Mark Gubrud, “Stopping killer robots”, *Bulletin of the Atomic Scientists*, vol. 70, n.º 1, 2014; Michael C. Horowitz, Paul Scharre y Alexander Velez-Green, *A Stable Nuclear Future? The Impact of Autonomous Systems and Artificial Intelligence*, documento de trabajo, diciembre de 2019, pp. 13-14, disponible en <https://arxiv.org/ftp/arxiv/papers/1912/1912.05291.pdf>; Paul Scharre, “A million mistakes a second”, *Foreign Policy*, 12 de septiembre de 2018, disponible en <https://foreignpolicy.com/2018/09/12/a-million-mistakes-a-second-future-of-war/>.

## Inestabilidad estratégica

Recientemente, se ha propuesto que la IA como asistente para la toma de decisiones por los humanos podría mejorar el funcionamiento de los sistemas de advertencia nuclear temprana y de mando y control, con lo que se reduciría el riesgo de falsas alarmas y de uso involuntario de armas nucleares<sup>72</sup>. No obstante, la presión para implementar la automatización total en el ámbito nuclear –es decir que los humanos cedan a las máquinas la capacidad de decidir sobre el uso de armas nucleares– es casi nula<sup>73</sup>. Pero, aunque aún no se ha dejado en manos de algoritmos la proverbial activación del botón, la urgencia por incrementar la autonomía de las aplicaciones militares y por automatizar los procesos militares aumenta el riesgo de inestabilidad nuclear<sup>74</sup>.

Para dar un ejemplo, las capacidades cada vez mayores de los sistemas de armas convencionales –incluida la autonomía de las armas– están empezando a afectar el nivel estratégico. Ese fenómeno se ha descrito como una “imbricación” cada vez mayor entre los ámbitos nuclear y convencional, lo que trae como consecuencia, por ejemplo, “amenazas no nucleares a las armas nucleares y sus sistemas asociados de mando, control, comunicaciones e información”<sup>75</sup>. Dicho de una manera sencilla, las capacidades convencionales avanzadas aumentan el riesgo para los componentes nucleares. La autonomía de los sistemas de armas convencionales es una de esas capacidades avanzadas, lo que contribuye al aumento de la imbricación y, a su vez, deteriora la estabilidad estratégica.

Un ejemplo concreto de esa dinámica es el despliegue de vehículos aéreos no tripulados indetectables y el uso de la estrategia de “enjambre”. Perdix es un programa de prueba de este comportamiento utilizado por la Fuerza Aérea de EE. UU. En el futuro, los enjambres de drones de este tipo podrían ser de ayuda en la búsqueda de lanzamisiles móviles dispersos. Otro ejemplo es el uso de sistemas marítimos autónomos para cazar submarinos de misiles balísticos con capacidad nuclear, conocidos por su sigla en inglés: SSBN. En el marco del programa de buque no tripulado de rastreo continuo de guerra antisubmarina, financiado por la DARPA, se desarrolló un trimarán autónomo denominado Sea Hunter, que actualmente está

72 M. Horowitz, P. Scharre y A. Velez-Green, nota 71 *supra*, p. 14; Philip Reiner y Alexa Wehsener, “The real value of artificial intelligence in nuclear command and control”, *War on the Rocks*, 4 de noviembre de 2019, disponible en <https://warontherocks.com/2019/11/the-real-value-of-artificial-intelligence-in-nuclear-command-and-control/>. Sobre las vulnerabilidades cibernéticas resultantes, v. James Johnson, “The AI-cyber nexus: Implications for military escalation, deterrence and strategic stability”, *Journal of Cyber Policy*, vol. 4, n.º 3, 2019.

73 Con excepción de Adam Lowther y Curtis McGiffin, “America needs a ‘dead hand’”, *War on the Rocks*, 16 de agosto de 2019, disponible en <https://warontherocks.com/2019/08/america-needs-a-dead-hand/>.

74 Edward Geist y Andrew J. Lohn, *How Might Artificial Intelligence Affect the Risk of Nuclear War?*, RAND Corporation, 2018, disponible en [www.rand.org/content/dam/rand/pubs/perspectives/PE200/PE296/RAND\\_PE296.pdf](http://www.rand.org/content/dam/rand/pubs/perspectives/PE200/PE296/RAND_PE296.pdf); Vincent Boulanin, Lora Saalman, Petr Topychkanov, Fei Su y Moa Peldán Carlsson, *Artificial Intelligence, Strategic Stability and Nuclear Risk*, SIPRI, Estocolmo, junio de 2020, disponible en [www.sipri.org/sites/default/files/2020-06/artificial\\_intelligence\\_strategic\\_stability\\_and\\_nuclear\\_risk.pdf](http://www.sipri.org/sites/default/files/2020-06/artificial_intelligence_strategic_stability_and_nuclear_risk.pdf).

75 James M. Acton (ed.), *Entanglement: Chinese and Russian Perspectives on Non-Nuclear Weapons and Nuclear Risks*, Carnegie Endowment for International Peace, 2017, p. 1, disponible en [http://carnegieendowment.org/files/Entanglement\\_interior\\_FNL.pdf](http://carnegieendowment.org/files/Entanglement_interior_FNL.pdf).

siendo probado por la Armada de EE. UU. Su capacidad para detectar y perseguir SSBN podría limitar las capacidades de un segundo ataque de otras potencias nucleares.

Esas capacidades aún están en sus albores, y ni Perdix, ni Sea Hunter, ni los sistemas que los sucedan desestabilizarán por sí solos el orden nuclear mundial. A la vez, la hipótesis de que sistemas como Sea Hunter harán que los mares sean “transparentes”<sup>76</sup>, cancelando prácticamente la utilidad de las armas nucleares lanzadas desde el mar como elementos fiables de segundo ataque, es objeto de intensos debates. No obstante, la mera percepción de las capacidades nucleares como susceptibles a nuevos riesgos del ámbito convencional seguramente sembrará la desconfianza entre adversarios que poseen armas nucleares. Más aún, sistemas como Sea Hunter demuestran que las tecnologías de armas autónomas están acelerando la ejecución completa del ciclo de selección y ataque de objetivos, lo que pone al adversario bajo una presión adicional y posiblemente produzca situaciones del tipo “úsalo o piérdelo” referidas a la ejecución de un segundo ataque nuclear.

El problema de la imbricación, al que contribuye la autonomía de las armas, se ve agravado por la voluntad política cada vez más pronunciada de usar medios nucleares como represalia a ataques de naturaleza no nuclear contra los sistemas de alerta temprana y control o contra las armas propiamente dichas. Una reseña de la postura nuclear del gobierno de Trump<sup>77</sup> señala que Estados Unidos puede, a partir de ahora, responder con medios nucleares a ataques estratégicos importantes de naturaleza no nuclear (apartándose del enfoque de disuasión como “único propósito” para las armas nucleares). Rusia ya sostuvo esta posición durante algún tiempo debido a la ventaja tecnológica de Estados Unidos en materia de armas convencionales. Esto no es un buen presagio para la estabilidad entre las dos principales potencias nucleares.

En conclusión, la autonomía de las armas no solo promete ventajas militares, sino que también genera nuevas vulnerabilidades y, lo que es más importante aún, contribuye a la acumulación general de inestabilidad y riesgos estratégicos. Incrementar la velocidad operacional más allá de la capacidad cognitiva humana deja fuera de juego a los humanos como valioso mecanismo de seguridad contra una escalada no deseada.

## **Por qué es imprescindible regular la autonomía de las armas: implicaciones éticas**

El debate sobre los SAAL en la CCAC se inclina hacia el DIH y sus implicaciones jurídicas, como se observa, por ejemplo, en los once principios rectores aprobados en el informe de la reunión de los Estados Partes en la CCAC de 2019<sup>78</sup>. En el preámbulo a la lista de principios, el informe estipulaba que “el

76 Sebastian Brixey-Williams, “Will the Atlantic become transparent?”, noviembre de 2016, disponible en <https://britishpugwash.org/wp-content/uploads/2016/11/Will-the-Atlantic-become-transparent-.pdf>.

77 DoD, *Nuclear Posture Review 2018*, 2018, p. 21, disponible en <https://tinyurl.com/yc7lu944>.

78 Informe final de la reunión de la CCAC, nota 2 *supra*, p. 10.

derecho internacional, en particular, la Carta de las Naciones Unidas y el derecho internacional humanitario, así como las perspectivas éticas pertinentes, debían guiar permanentemente la labor del Grupo”. No obstante, solo cinco de los once principios rectores son de naturaleza jurídica, y ninguno contiene referencias a las implicaciones éticas. El aspecto jurídico del debate es importante, sin duda, en especial porque permite examinar sistemáticamente la afirmación de que la autonomía de las armas hace que mejore el cumplimiento del DIH en la guerra. En las circunstancias actuales, la tecnología no es capaz de cumplir la promesa de un mayor respeto del DIH<sup>79</sup>, aunque con el tiempo podría llegar a ser capaz de hacerlo. Más allá de lo jurídico, desde el punto de vista ético, la cuestión de los SAAL tiene raíces más profundas que el mero cumplimiento del DIH, porque tiene que ver con normas fundamentales que trascienden las leyes de la guerra<sup>80</sup>. Las implicaciones éticas fueron consideradas más sistemáticamente en toda su dimensión al principio del debate sobre los SAAL en la ONU. En 2013, cuando surgió el tema en el Consejo de Derechos Humanos, el Relator Especial Christof Heyns<sup>81</sup> se opuso a los SAAL, afirmando que violaban la dignidad humana.

## Dignidad humana universal

La postura de que el uso de los SAAL vulneraría la dignidad humana fue adoptada por numerosos académicos en el campo de la filosofía moral y la tecnología<sup>82</sup>. La noción encontró eco en Human Rights Watch<sup>83</sup> y, después, en el CICR<sup>84</sup>. La oposición a la autonomía de las armas por motivos relativos a la dignidad

79 Frank Sauer, Daniele Amoroso, Noel Sharkey, Lucy Suchman y Guglielmo Tamburrini, *Autonomy in Weapon Systems: The Military Application of Artificial Intelligence as a Litmus Test for Germany's New Foreign and Security Policy*, Heinrich Böll Stiftung Publication Series on Democracy, vol. 49, 2018, pp. 23-32, disponible en [www.boell.de/sites/default/files/boell\\_autonomy-in-weapon-systems\\_v04\\_kommentierbar\\_1.pdf](http://www.boell.de/sites/default/files/boell_autonomy-in-weapon-systems_v04_kommentierbar_1.pdf).

80 El apartado siguiente se basa en Elvira Rosert y Frank Sauer, “Prohibiting autonomous weapons: Put human dignity first”, *Global Policy*, vol. 10, n.º 3, 2019.

81 Christof Heyns, *Informe del Relator Especial sobre las ejecuciones extrajudiciales, sumarias o arbitrarias*, doc. ONU A/HRC/23/47, 2013, p. 19, disponible en [https://www.ohchr.org/sites/default/files/Documents/HRBodies/HRCouncil/RegularSession/Session23/A-HRC-23-47\\_sp.pdf](https://www.ohchr.org/sites/default/files/Documents/HRBodies/HRCouncil/RegularSession/Session23/A-HRC-23-47_sp.pdf).

82 Peter Asaro, “On banning autonomous weapon systems: Human rights, automation, and the dehumanization of lethal decision-making”, *International Review of the Red Cross*, vol. 94, n.º 886, 2012; Robert Sparrow, “Robots and respect: Assessing the case against autonomous weapon systems”, *Ethics & International Affairs*, vol. 30, n.º 1, 2016.

83 Human Rights Watch, “Construyendo el caso: Los peligros de los robots asesinos y la necesidad de una prohibición preventiva”, 6 de marzo de 2020, disponible en <https://www.hrw.org/es/report/2020/03/06/construyendo-el-caso/los-peligros-de-los-robots-asesinos-y-la-necesidad-de-una>.

84 CICR, *Ethics and Autonomous Weapon Systems: An Ethical Basis for Human Control?*, Ginebra, 3 de abril de 2018, disponible en [www.icrc.org/en/download/file/69961/icrc\\_ethics\\_and\\_autonomous\\_weapon\\_systems\\_report\\_3\\_april\\_2018.pdf](http://www.icrc.org/en/download/file/69961/icrc_ethics_and_autonomous_weapon_systems_report_3_april_2018.pdf).

humana ha sido objeto de análisis<sup>85</sup> y la supuesta “incomodidad”<sup>86</sup> de este punto de vista normalmente se apoya en el argumento de que existen distintos significados del término “dignidad” y de que no hay una única definición consensuada.

Sin embargo, el hecho de que sea difícil de definir, pero relevante y hasta decisivamente importante, es una característica de muchos conceptos normativos, muchos de los cuales están codificados jurídicamente. Entre los ejemplos, se encuentran conceptos fundamentales del DIH tales como la condición civil, a la que se define solo por la negativa, o la proporcionalidad, que es imposible de cuantificar y solo puede evaluarse caso por caso<sup>87</sup>. La dignidad humana también figura en numerosos documentos jurídicos internacionales. La Declaración Universal de los Derechos Humanos hace mención de este concepto en su preámbulo, al igual que la Carta de la ONU. También se la invoca en ordenamientos jurídicos nacionales, así como en fallos judiciales. El ejemplo clave es el artículo 1(1) de la Ley Fundamental de Alemania, que determina que la dignidad humana es inviolable y prohíbe tratar a los seres humanos como objetos o como medios para conseguir un fin y al que se hace referencia en un fallo histórico que el Tribunal Constitucional de Alemania pronunció en 2006. Los jueces dejaron sin efecto una ley federal que habría permitido que la Fuerza Aérea Alemana derribara un avión secuestrado que los secuestradores quizá quisieran usar como arma para matar personas que se encontraban en tierra. El Tribunal declaró inconstitucional el uso de los pasajeros de la aeronave como meros instrumentos para intentar alcanzar otro objetivo, aunque fuera respetable<sup>88</sup>.

La implicación ética fundamental de la autonomía de las armas en las funciones críticas de un sistema de armas es, entonces, que permitir que un algoritmo tome decisiones sobre la muerte vulnera la dignidad humana, porque la víctima queda reducida a un objeto, un mero dato que se proporciona a una máquina asesina automática caracterizada por la indiferencia.

Merece la pena aclarar que la objeción es válida incluso si los civiles (u otros no combatientes) no resultan heridos. Después de todo, al acotar el foco exclusivamente al hecho de que los SAAL podrían no ser capaces de hacer distinciones adecuadas –o mejores– entre combatientes y civiles –uno de los pilares del argumento jurídico en contra de los SAAL analizado en la CCAC– se pierde de vista el hecho de que los combatientes también poseen dignidad humana. Dicho de otro modo, la autonomía de las armas genera preocupaciones más fundamentales que las que pone de relieve el aspecto jurídico del debate sobre los SAAL, porque “distinguir efectivamente a los combatientes de los no combatientes no es, en absoluto, el único problema”<sup>89</sup>.

85 Amanda Sharkey, “Autonomous weapons systems, killer robots and human dignity”, *Ethics and Information Technology*, vol. 21, n.º 2, 2019.

86 Deane-Peter Baker, “The awkwardness of the dignity objection to autonomous weapons”, *The Strategy Bridge*, 6 de diciembre de 2018, disponible en <https://thestrategybridge.org/the-bridge/2018/12/6/the-awkwardness-of-the-dignity-objection-to-autonomous-weapons>.

87 Quisiera agradecer a la persona que hizo la revisión por especificar esas propiedades de la condición civil y la proporcionalidad.

88 F. Sauer *et al.*, nota 79 *supra*, p. 33.

89 Heather M. Roff, “The strategic robot problem: Lethal autonomous weapons in war”, *Journal of Military Ethics*, vol. 13, n.º 3, 2014, p. 219.

Por regla general, el uso de SAAL contra humanos puede considerarse una violación inaceptable de la dignidad humana, porque delegar la decisión de matar en un algoritmo devalúa la vida humana<sup>90</sup>. Las excepciones a la regla solo serían concebibles si explícitamente no se hicieran sobre la base de la comparación entre vidas desnudas y luego la selección deliberada de una matanza determinada por un algoritmo. Un ejemplo de ese caso límite podría ser la confianza de un marino en la autonomía de un arma en una situación cercana al límite de la defensa propia desesperada. Si la fragata<sup>91</sup> antes mencionada estuviera bajo un ataque de saturación lanzado por misiles antibuque y, posiblemente, también por vehículos aéreos tripulados, entonces poner en peligro involuntariamente vidas humanas al dejar en manos de fuego defensivo autónomo la supervivencia de la nave y su tripulación podría considerarse aceptable *a posteriori*.

En términos generales, que una persona muera como consecuencia de una decisión algorítmica concierne a esa persona, porque una máquina que se cobra una vida no tiene conciencia de lo que significan sus acciones: “En ausencia de una decisión de usar la violencia que tenga una intención y un sentido, las muertes resultantes son arbitrarias y no tienen sentido”<sup>92</sup>. Dicho de otro modo, lo menos que podemos hacer al matar a otro ser humano en una guerra es reconocer la muerte de un miembro de nuestra especie y cargar con el peso de esa decisión en nuestra conciencia. La falta de conciencia de las máquinas que matan personas basándose en un software priva a las personas del derecho a ser reconocidas como seres humanos al momento de su muerte. Esto también concierne a la sociedad en su conjunto. La guerra contemporánea, en especial en las democracias, disocia a las sociedades del combate en términos de costos políticos y financieros<sup>93</sup>. Una sociedad que se desentiende de los costos morales porque ya no se involucra en el acto de matar, porque ya no hay combatientes individuales cuyas psiquis carguen con la responsabilidad concomitante, cruza un límite moral. Corre el riesgo de perder contacto con valores humanitarios fundamentales, como el derecho a una vida digna y el respeto a los seres humanos<sup>94</sup>.

En suma, mientras que el veredicto jurídico acerca de si la autonomía de las armas mejora el cumplimiento del DIH aún no se ha alcanzado, y llevará un tiempo alcanzarlo, la objeción moral fundamental contra los SAAL continúa siendo válida sobre la base de límites deontológicos.

90 Christof Heyns, “Autonomous weapons in armed conflict and the right to a dignified life: An African perspective”, *South African Journal on Human Rights*, vol. 33, n.º 1, 2017, pp. 62-63.

91 V. el texto vinculado con la nota 27 *supra*.

92 Peter Asaro, “*Jus nascendi*, robotic weapons and the Martens Clause”, en Ryan Calo, A. Michael Froomkin e Ian Kerr (eds.), *Robot Law*, Edward Elgar, Cheltenham, 2016, p. 385.

93 F. Sauer y N. Schörnig, nota 41 *supra*; Sarah E. Kreps, “Just put it on our tab: War financing and the decline of democracy”, *War on the Rocks*, 28 de mayo de 2018, disponible en <https://warontherocks.com/2018/05/just-put-it-on-our-tab-21st-century-war-financing-and-the-decline-of-democracy/>.

94 Denise Garcia, “Killer robots: Toward the loss of humanity”, *Ethics and International Affairs*, abril de 2015, disponible en [www.ethicsandinternationalaffairs.org/2015/killer-robots-toward-the-loss-of-humanity/](http://www.ethicsandinternationalaffairs.org/2015/killer-robots-toward-the-loss-of-humanity/).

## Cómo se puede regular la autonomía de las armas: promover una norma para el control humano

Estados Unidos y China están prestando atención a los riesgos estratégicos que entraña la autonomía irrestricta de las armas. La directiva estadounidense sobre la autonomía de las armas<sup>95</sup>, a pesar de que persigue el objetivo de lograr lo imposible, es decir, usar la autonomía y descartar los riesgos asociados, puede interpretarse en ese sentido. China ha acuñado el término “singularidad del campo de batalla”, una situación aterradora en la que la guerra que se libra a la velocidad de las máquinas es demasiado rápida y la cognición humana no puede seguirle el ritmo<sup>96</sup>. No obstante, la rivalidad actual entre las grandes potencias –Estados Unidos, China y Rusia–, cada una de las cuales trata de superar a las demás en el campo de la IA militar, no propicia la regulación de la autonomía de las armas. Con Trump, Xi Jinping y Putin en el poder, no puede esperarse un avance en el futuro inmediato. Pero la voluntad política para regular la autonomía de las armas puede generarse también desde abajo.

### Voluntad política que avanza desde las bases

En las encuestas, se observa sistemáticamente que, en todas partes del mundo, la opinión pública se opone a los SAAL. La oposición mundial aumentó del 56 % en 2016 al 61 % en 2018, según una encuesta de la CKR<sup>97</sup>. Estos resultados coinciden con los de encuestas anteriores realizada por Open Roboethics Initiative<sup>98</sup> y por Heather Roff a través de IPSOS<sup>99</sup>. La oposición a este tipo de armas en Estados Unidos, China y Rusia alcanza el 52 %, el 59 % y el 60 %, respectivamente<sup>100</sup>. En Europa, las cifras van desde el 60 % en Finlandia hasta el 81 % en Irlanda<sup>101</sup>.

Los resultados de las encuestas también indican que la oposición de la opinión pública no se fundamenta principalmente en cuestiones jurídicas ni en las preocupaciones acerca del aumento de la tensión o de la inestabilidad que acompaña a las crisis, sino en la noción de que delegar las decisiones sobre la vida y la muerte en el campo de batalla implica cruzar los límites morales aceptados<sup>102</sup>. Así pues, si bien el debate filosófico sobre la ubicuidad cultural de la dignidad humana como concepto

95 DoD, nota 16 *supra*.

96 E. B. Kania, *Battlefield Singularity*, nota 46 *supra*.

97 CKR, “Global poll shows 61% oppose killer robots”, 22 de enero de 2019, disponible en [www.stopkillerrobots.org/2019/01/global-poll-61-oppose-killer-robots/](http://www.stopkillerrobots.org/2019/01/global-poll-61-oppose-killer-robots/).

98 Open Roboethics Institute, “The ethics and governance of lethal autonomous weapons systems: An international public opinion poll”, 9 de noviembre de 2015, disponible en [http://www.openroboethics.org/wp-content/uploads/2015/11/ORi\\_LAWS2015.pdf](http://www.openroboethics.org/wp-content/uploads/2015/11/ORi_LAWS2015.pdf).

99 Heather M. Roff, “What do people around the world think about killer robots?”, *Slate*, 8 de febrero de 2017, disponible en <https://slate.com/technology/2017/02/what-do-people-around-the-world-think-about-killer-robots.html>.

100 CKR, nota 97 *supra*.

101 CKR, “New European poll shows public favour banning killer robots”, 13 de noviembre de 2019, disponible en <https://www.stopkillerrobots.org/news/new-european-poll-shows-73-favour-banning-killer-robots/>.

102 CKR, nota 97 *supra*.



y su importancia para la cuestión de los SAAL desde el punto de vista utilitario en contraposición con la perspectiva deontológica es, sin duda, necesario e interesante, el problema, tal como se lo presenta en el apartado anterior, encuentra eco en la opinión pública. La noción de que hay algo que está esencialmente mal en que la vida de los seres humanos quede en manos de máquinas sin conciencia es apropiada para ejercer presión desde abajo sobre los gobiernos a fin de alimentar la voluntad política al respecto. Así lo reconocen hasta los escépticos respecto del argumento de la dignidad humana en general: “Las campañas podrían dar buenos resultados. Decir que algo va en contra de la dignidad humana provoca una fuerte respuesta visceral”<sup>103</sup>.

### Promover el desarrollo de normas en la CCAC

Los SAAL siguen captando la atención de los medios de comunicación en todo el mundo<sup>104</sup>. La creciente presión de la opinión pública y el mayor foco de atención actuarán como importantes incentivos para que los Estados Partes en la CCAC produzcan resultados concretos para la Conferencia de Examen de 2021. Los “aspectos del marco normativo y operacional” que deben continuar desarrollándose durante el año 2021 podrían empezar a concretarse en tres etapas.

En primer lugar, parece posible lograr un consenso sobre el lenguaje común que ha de adoptar la visión funcionalista de la autonomía de las armas, que en la actualidad goza de gran aceptación, así como un acuerdo común de que se necesita alguna forma de obligación positiva y de afirmación del principio del control humano de los sistemas de armas<sup>105</sup>. En el principio rector (b) de la CCAC, ya se advierte esta tendencia: “[e]l ser humano debe mantener la responsabilidad por las decisiones que se adopten sobre el uso de los sistemas de armas, ya que la obligación de rendir cuentas no puede transferirse a las máquinas”<sup>106</sup>. El Foro de apoyo al GEG sobre SAAL que tuvo lugar en abril de 2020 y adoptó la modalidad de una transmisión virtual gestionada por el Ministerio Federal de Asuntos Exteriores de Alemania con trescientos veinte participantes registrados en representación de sesenta y tres Estados Partes en la CCAC subrayó la importancia de seguir trabajando en la conceptualización del componente humano. Podría decirse que la posibilidad de controlar las armas ya es una protonorma<sup>107</sup>, y sería posible encontrar una terminología común –por ejemplo, “control humano significativo” u otra formulación–, a fin de estipular en un sentido general cuándo ha de ejecutarse una función en el ciclo de selección y ataque de objetivos un humano y cuándo ha de ejecutarla una máquina. Recientemente, el CICR y el Instituto Internacional

103 A. Sharkey, nota 85 *supra*, p. 9.

104 R. Charli Carpenter, “Lost” Causes: Agenda Vetting in Global Issue Networks and the Shaping of Human Security, Cornell University Press, Ithaca, Nueva York, 2014, pp. 88-121.

105 Stephen D. Goose y Mary Wareham, “The growing international movement against killer robots”, *Harvard International Review*, vol. 37, n.º 4, 2016, disponible en [www.jstor.org/stable/26445614](http://www.jstor.org/stable/26445614).

106 Informe final de la reunión de la CCAC, nota 2 *supra*, p. 10.

107 Sobre la noción de codificación del control humano como un principio del DIH en general, v. Elvira Rosert, *How to Regulate Autonomous Weapons*, PRIF Spotlight 6/2017, Instituto de Estudios para la Paz de Frankfurt, 2017, disponible en [www.hsfk.de/fileadmin/HSFK/hsfk\\_publicationen/Spotlight0617.pdf](http://www.hsfk.de/fileadmin/HSFK/hsfk_publicationen/Spotlight0617.pdf).

de Estudios para la Paz de Estocolmo (SIPRI) presentaron un marco conceptual que puede servir de base a la iniciativa de operacionalizar el control humano –es decir, aclarar “el quién, el qué, el cómo y el cuándo” del control de las armas y de la imposición de límites a su autonomía<sup>108</sup>.

En segundo lugar, como no existe un único estándar de control humano significativo, el intercambio de prácticas idóneas y, lo que es más importante aún, de estudios de caso de sistemas de armas y contextos operacionales específicos podría permitir a los Estados Partes en la CCAC alcanzar una comprensión conceptual profunda y compartida de las complejidades que implica la implementación del control humano en el diseño y el uso de los SAAL. El GEG está especialmente preparado para facilitar esa comprensión profunda a partir de análisis de distintas partes interesadas y diversas opiniones jurídicas, éticas y operacionales. Grupos de expertos más pequeños, como el International Panel on the Regulation of Autonomous Weapons (iPRAW) y la comisión sobre uso responsable de las tecnologías en el Sistema franco-alemán de combate aéreo del futuro, ya han comenzado a organizar su propia investigación en ese sentido.

En tercer lugar, podría elaborarse un plan de implementación diferenciada que concibiera el control humano como un ejercicio dependiente del contexto, es decir, supeditado al sistema de armas, el entorno de la misión, los “perfiles de objetivos”<sup>109</sup> y otros factores, como la duración de la misión<sup>110</sup>. En ese plan, se podrían prescribir normas mínimas para las capacidades de control en el diseño, por ejemplo, con respecto a la ergonomía de la interfaz humano-máquina, y determinar “niveles de control de supervisión humana”<sup>111</sup> en el uso, es decir, las tácticas, las técnicas y los procedimientos necesarios para mantener la responsabilidad y el control humanos intactos durante el funcionamiento del sistema.

Actualmente, aunque se completaran esos tres pasos, parece improbable que el proceso de la CCAC produzca más que “normas indicativas”, como una declaración política conjunta o un catálogo de prácticas idóneas. De hecho, cabe la posibilidad de que fracase todo el proceso de la CCAC en Ginebra. Pero incluso si la CCAC no llega a ser el ámbito de donde surja una norma jurídicamente vinculante para la autonomía de las armas, ha sido un centro de información y de formulación de normas durante los últimos seis años, y continuará siéndolo. Considerando, sobre todo, las consecuencias de la crisis provocada por la pandemia de COVID-19 en la agenda de reuniones en todo el mundo, todavía es muy pronto para decidir si otros foros pueden y deben tomar el relevo de la regulación donde la deje la CCAC en 2021 para continuar elaborando y codificando la norma del control humano como norma vinculante del derecho internacional y, de ser así, cuáles serían esos foros.

108 V. Boulanin *et al.*, nota 26 *supra*. V. también Ilse Verdiesen, Filippo Santoni de Sio y Virginia Dignum, “Accountability and control over autonomous weapon systems: A framework for comprehensive human oversight”, *Minds and Machines*, 2020, disponible en <https://link.springer.com/article/10.1007/s11023-020-09532-9>.

109 Moyes, “Target profiles”, nota 26 *supra*.

110 Para este enfoque general, así como para una enumeración de las variables que se han de tener en cuenta, v. V. Boulanin *et al.*, nota 26 *supra*, pp. 30-33.

111 F. Sauer *et al.*, nota 79 *supra*, pp. 42-45.

## Conclusión

La regulación multilateral de la autonomía de los sistemas de armas –es decir, la formulación de una obligación jurídicamente vinculante a fin de conservar un control humano significativo sobre el uso de la fuerza– es difícil de alcanzar, pero es imprescindible hacerlo. Los riesgos estratégicos graves, así como los riesgos éticos de mediano y largo plazo, por ejemplo, la escalada no intencional del conflicto a la velocidad de las máquinas y la violación de la dignidad humana, superan cualquier ventaja militar de corto plazo. En este análisis, se ha mostrado que es posible regular la autonomía de las armas en un proceso de tres pasos para facilitar la posibilidad de echarse atrás al llegar al punto crítico. El primer paso consiste en promover el consenso incipiente relativo a la noción de que contar con una obligación positiva para conservar el control humano sobre los sistemas de armas es una medida prudente que se necesita con urgencia; el segundo paso consiste en continuar elaborando la idea de que no existe un único estándar para el control humano significativo; y el tercer paso consiste en diseñar programas de control humano diferenciados y dependientes del contexto para los sistemas de armas. Dados el panorama geopolítico actual y la falta de voluntad política para participar en iniciativas de control de armas, avanzar en el proceso será como correr una maratón y no una carrera de velocidad. Después de todo, el valor militar percibido de la autonomía de las armas es excepcionalmente elevado, y el tema en sí mismo es esquivo, por lo que es necesario adoptar un enfoque cualitativo e innovador para el control de las armas.

No obstante, la historia deja ver con claridad que las grandes potencias no carecen de sensibilidad frente a la acumulación de riesgos colectivos; de no ser así, el control de las armas nucleares, químicas y biológicas nunca se habría concretado. Las tecnologías emergentes del siglo XXI ofrecen la oportunidad de demostrar que la humanidad ha aprendido las lecciones de la historia antes de que los riesgos se manifiesten en toda su dimensión. Los seres humanos hacen cosas terribles a sus semejantes en la guerra y no hay solución tecnológica para eso. Pero la comunidad internacional puede, al menos, establecer normas para frenar la escalada descontrolada y el cruce de líneas morales fundamentales. Si no lo logramos, no solo perderemos el espacio necesario para meditar y elaborar las respuestas<sup>112</sup>, un requisito fundamental de la gestión política de los conflictos, como muestra la crisis de los misiles de Cuba<sup>113</sup>, sino que también transformaremos la “indignidad extrema” de la guerra en una “muerte por algoritmo”<sup>114</sup>.

112 Frank Sauer, *Atomic Anxiety: Deterrence, Taboo and the Non-Use of U.S. Nuclear Weapons*, Palgrave Macmillan, Londres, 2015, pp. 91-92.

113 Z. Davis, nota 35 *supra*, p. 122.

114 Robert H. Latiff y Patrick J. McCloskey, “With drone warfare, America approaches the Robo-Rubicon”, *Wall Street Journal*, 14 de marzo de 2013, disponible en <https://tinyurl.com/y2t7odsh>.