

Capítulo segundo

Dilemas éticos de la aplicación de la inteligencia artificial y la robótica en las Fuerzas Armadas

Jens Wegener

Este trabajo refleja exclusivamente los resultados de la investigación y la opinión del autor, pero no necesariamente los del Gobierno Federal, el Ministerio de Defensa o el Ejército del Aire alemán.

Resumen

Sobre la base de la literatura científica de idioma inglés y alemán de 2009 a 2019, el presente trabajo de investigación examina los dilemas éticos que plantea el uso de la robótica y la inteligencia artificial en las Fuerzas Armadas. El utilitarismo se establece como un paradigma ético que puede dar respuestas adecuadas a la mayoría de los dilemas planteados.

Palabras clave

Inteligencia artificial, robótica, Fuerzas Armadas, ética, utilitarismo.

Ethical dilemmas of the use of artificial intelligence and robotics in the Armed Forces

Abstract

This paper examines the ethical dilemmas of the use of robotics and artificial intelligence in the Armed Forces based on English and German language contributions to the scientific debate written between 2009 and 2019. Utilitarianism is presented as an ethical paradigm that can provide adequate answers to most of the dilemmas raised.

Keywords

Artificial intelligence, robotics, Armed Forces, ethics, utilitarianism

Introducción

Hoy en día, distintos Estados líderes en el ámbito militar, especialmente Estados Unidos, han identificado la automatización como la clave para las futuras capacidades¹. A partir de 2003 se observó un aumento exponencial de los sistemas no tripulados utilizados por las Fuerzas Armadas de dicho país². Desde 2009, la U.S. Air Force ha entrenado más pilotos de vehículos aéreos no tripulados (*unmanned aerial vehicles* o UAV, también llamados «drones») que pilotos de aviones de caza y de bombarderos combinados³. Este desarrollo se formalizó en la estrategia norteamericana de *Third Offset* de 2014. Previamente hubo otros dos Offset: la primera, sobre el uso de armas nucleares; y la segunda, sobre el uso de ordenadores, redes y armas de precisión para compensar la inferioridad cuantitativa convencional, en ambos casos contra la Unión Soviética en la Guerra Fría. En la *Third Offset* se considera que la inteligencia artificial (IA) y la autonomía se utilizarán para mejorar la efectividad de las Fuerzas Armadas⁴.

Recientemente, la Agencia Europea de Defensa también identificó la IA y los sistemas autónomos como futuras tecnologías militares⁵. Sin embargo, otros Estados occidentales, además de Rusia, China e Israel, llevan tiempo siguiendo el ejemplo de EE. UU. y están utilizando los vehículos aéreos y distintos tipos de vehículos terrestres no tripulados (*unmanned ground vehicles*, o UGV). El uso de estos dispositivos está destinado principalmen-

¹ HAGSTRÖM, Martin, «Military applications of machine learning and autonomous systems», en: BOULANIN, Vincent (ed.), *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk. Volume I. Euro-Atlantic Perspectives*, Stockholm International Peace Research Institute, Solna, 2019, p. 35.

² SINGER, Peter W., «War of the Machines: What Is The Real Story of Robotic Weaponry», en: SCHMIDT-RADEFELDT, Roman y MEISSLER, Christine (eds.), *Automatisierung und Digitalisierung des Krieges. Drohnenkrieg und Cyberwar als Herausforderungen für Ethik, Völkerrecht und Sicherheitspolitik*, Baden-Baden: Nomos Verlagsgesellschaft 2012, pp. 24-25.

³ WEISS, Lora G., «Autonome Roboter im Nebel des Krieges», en: MARSISKE, Hans-Arthur (ed.), *Kriegsmaschinen. Roboter im Militäreinsatz*, Hannover: Heise 2012, p. 164.

⁴ LEWIS, Larry, «Insights for the Third Offset: Addressing Challenges of Autonomy and Artificial Intelligence in Military Operations», 2017, pp. 2-6. https://www.cna.org/cna_files/pdf/DRM-2017-U-016281-Final.pdf [Fecha de la consulta: 20.11.2019].

⁵ PRIŠMANTAITĖ, Kristina, «Battlefield Digitalisation through Research and Innovation in the European Union», en: ROMANOV, Uģis y ANDŽĀNS, Māris (eds.), *Digital Infantry Battlefield Solution. Research and Innovation. Part III*, Tallinn: Milrem Robotics 2019, p. 20.

te a la desactivación de artefactos explosivos y para la defensa contra dispositivos explosivos improvisados, pero también para la defensa contra agentes de guerra nuclear, biológica y química; para el reconocimiento, con fines logísticos, de control, y para la seguridad⁶. Los vehículos navales de superficie y subacuáticos, por otra parte, son la variante menos utilizada de vehículos no tripulados hasta la fecha y, en la actualidad, se utilizan especialmente para la desactivación de minas⁷.

Aparte de algunas pocas publicaciones anteriores, *Wired for War* de Peter W. Singer (2009) puso el tema en el foco del público interesado. Desde entonces, la rápida difusión y el uso de sistemas robóticos en el contexto militar han sido objeto de críticas generalizadas⁸. Esto ya se aplica a los UAV tras su polémico uso por parte de Estados Unidos para combatir individuos en Pakistán, Yemen y Somalia⁹, pero el discurso ético y jurídico

⁶ Se puede encontrar una visión general de los sistemas actuales y en desarrollo, por ejemplo, en SINGER, PETER W., «[Wired for War]. The Robotics Revolution and Conflict in the Twenty-first Century», New York: The Penguin Press 2009, pp. 110-114; KRISHNAN, Armin, «Killer Robots. Legality and Ethicality of Autonomous Weapons», Farnham/Burlington: Ashgate 2009, p. 28; ROMANOV, Uģis y ANDŽĀNS, Māris, «Digital Infantry Battlefield Solution. Concept of Operations. Part II», Tallinn: Milrem Robotics 2017, pp. 45-111; BOISBOISSEL, Gérard de, «France», en: ROMANOV, Uģis y ANDŽĀNS, Māris (eds.), Digital Infantry Battlefield Solution. Research and Innovation. Part III, Tallinn: Milrem Robotics 2019, pp. 71-81, 88; QASRAWI, Yazan, BOUKHTOUTA, Abdeslem y GIZEWSKI, Peter, «Canada», en: ROMANOV, Uģis y ANDŽĀNS, Māris (eds.), Digital Infantry Battlefield Solution. Research and Innovation. Part III, Tallinn: Milrem Robotics 2019, p. 53; ROGERS, James y CLARK, Robert, «The United Kingdom», en: ROMANOV, Uģis y ANDŽĀNS, Māris (eds.), Digital Infantry Battlefield Solution. Research and Innovation. Part III, Tallinn: Milrem Robotics 2019, pp. 103-104; SMOLAREK, Mirosław, «Germany», en: ROMANOV, Uģis y ANDŽĀNS, Māris (eds.), Digital Infantry Battlefield Solution. Research and Innovation. Part III, Tallinn: Milrem Robotics 2019.

⁷ PETERMANN, Thomas, «Unbemannte Systeme als Herausforderung für Sicherheits- und Rüstungskontrollpolitik – Ergebnisse eines Projekts des Büros für Technikfolgen-Abschätzung beim Deutschen Bundestag», en: SCHMIDT-RADEFELDT, Roman y MEISSLER, Christine (eds.), Automatisierung und Digitalisierung des Krieges. Drohnenkrieg und Cyberwar als Herausforderungen für Ethik, Völkerrecht und Sicherheitspolitik, Baden-Baden: Nomos Verlagsgesellschaft 2012, p. 74.

⁸ Un órgano de críticos destacado es la «Campaña para detener a los robots asesinos», apoyada por Human Rights Watch, Amnistía Internacional y el Comité Internacional para el Control de Armas Robóticas. Ver SCHÖRNIG, Niklas y WEIDLICH, Christian, «Keine Macht den Drohnen! Warum Deutschland sich jetzt gegen autonom tötende Militärsysteme einsetzen muss», HSFK Standpunkte 8/2013, Frankfurt am Main: Hessische Stiftung Friedens- und Konfliktforschung 2013, p. 5.

⁹ Para una antología relevante, véase STRAWSER, Bradley J., «Opposing Perspectives on the Drone Debate», New York: Palgrave Macmillan 2014. Como ejemplo del debate parcialmente unilateral y crítico de Estados Unidos, Weber en su artículo sobre los horrores de la guerra de los UAV en Pakistán ni siquiera menciona el hecho de que los

internacional se centra en los sistemas de combate totalmente autónomos (*autonomous weapon systems*, AWS), sobre todo en *lethal autonomous weapon systems* (LAWS). A primera vista, este enfoque es sorprendente, ya que el uso de la robótica y la IA como sistemas de asistencia para humanos todavía es mucho más relevante militarmente¹⁰. Sin embargo, los LAWS ya se utilizan hoy en día en contextos simples y claramente estructurados como la supresión de las defensas aéreas enemigas, la defensa aérea y la defensa contra proyectiles (*counter rocket artillery mortar*, CRAM). En el futuro, el desarrollo de la IA permitirá su uso en un entorno complejo y dinámico¹¹.

Si bien la literatura anglosajona al respecto del objeto de estudio desde 2009 ha sido amplia (la mayoría de los argumentos empleados en el debate científico pertinente, y presentados a continuación, han sido desarrollados hasta 2013¹²), las implicaciones del uso de la robótica, la IA y los LAWS en el entorno militar, así como las cuestiones éticas pertinentes, apenas se han discutido en el ámbito alemán. Ni el Gobierno Federal ni el Parlamento han creado comisiones de expertos para debatir la utilización de los AWS en el Ejército¹³. La robótica militar no se menciona en el Libro Blanco de 2016; ni siquiera el «control humano» sobre los robots ha sido definido todavía en ningún documento estratégico alemán¹⁴. La Estrategia de IA del Gobierno solamente menciona que la IA es un tema importante para las Fuerzas Armadas, y que

talibanes sean reconocidos partidarios de terroristas que asesinan deliberadamente a civiles y aceptan víctimas civiles mezclándose con la población. Véase WEBER, Jutta, «Vorratsbomben am Himmel. Über digitalen Terror, unsichtbare Opfer und die Rhetorik der Präzision», en: MARSISKE, Hans-Arthur (ed.), *Kriegsmaschinen. Roboter im Militäreinsatz*, Hannover: Heise 2012, pp. 13-52.

¹⁰ DAHLMANN, Anja y DICKOW, Marcel, «Preventive Regulation of Autonomous Weapon Systems. Need for Action by Germany at Various Levels», *Stiftung Wissenschaft und Politik*, Berlin, 2019, p. 10.

¹¹ DAHLMANN y DICKOW, *op. cit.*, p. 5; SAUER, Frank, «Mensch, Maschine, Krieg», *Zur Sache Bw*, 1/2019a, p. 26.

¹² Para resúmenes cortos de numerosos problemas y soluciones propuestas que se discuten a continuación, véanse BARRETT, Edward, «Executive Summary and Command Brief», *Journal of Military Ethics*, 9(4), 2010; LEWIS, *op. cit.* (2017), pp. 33-34.

¹³ El 11 de mayo de 2020, en un panel en el Ministerio de Defensa, únicamente se discutió sobre la cuestión del empleo de drones armadas. Véase WIEGOLD, Thomas, «Dokumentation: Drohnendebatte im BMVg», *Augen geradeaus!*, 11 de mayo de 2020, <https://augengeradeaus.net/2020/05/dokumentation-drohnendebatte-im-bmvg/> [Fecha de la consulta: 12.05.2020].

¹⁴ DAHLMANN y DICKOW, *op. cit.*, pp. 6-7, 20-21.

se valorarán conscientemente las ventajas y desventajas de su empleo¹⁵.

El documento oficial alemán más detallado sobre la aplicación de la IA en las Fuerzas Armadas, un concepto de la Oficina de Desarrollo del Ejército de Tierra alemán¹⁶, sí indica que hay problemas relacionados con el derecho internacional, pero no los especifica, y no sugiere ningún problema ético. Al mismo tiempo, en el ámbito civil alemán se puede constatar un predominio de voces críticas hacia la robótica y la IA, quizás parcialmente porque no existe un foro científico, como por ejemplo en EE. UU., en el que los militares activos se expresen sobre las ventajas de los UAV y porque, en general, en Alemania es raro un intercambio científico entre civiles y militares¹⁷. Además, y no solamente en Alemania, el debate ético sobre los robots se ve obstaculizado por el hecho de que la mayoría de los expertos en robótica tengan una formación técnica pero no una formación en humanidades o, incluso más allá, holística. A menudo no entran en consideraciones morales por miedo a ser avergonzados en un campo científicamente desconocido¹⁸. Por el contrario, se puede suponer que el conocimiento de las posibilidades técnicas y tácticas de la robótica y la IA por parte de los defensores de las consideraciones éticas con frecuencia están también limitados.

Por esos factores, este trabajo analítico-sintético, basado en la literatura científica de lengua anglosajona y alemana de 2009 hasta 2019, va a ser de utilidad sobre todo para militares que carezcan de una visión holística de dichas tecnologías, para estudiosos de la ética que desconozcan sus ventajas y desventajas

¹⁵ Bundesregierung, «Strategie Künstliche Intelligenz der Bundesregierung», 2018, p. 32, https://www.ki-strategie-deutschland.de/home.html?file=files/downloads/Nationale_KI-Strategie.pdf [Fecha de la consulta: 20.11.2019].

¹⁶ Amt für Heeresentwicklung, «Künstliche Intelligenz in den Streitkräften. Ein Positionspapier des Amtes für Heeresentwicklung», 2019, <https://www.bundeswehr.de/resource/blob/156024/d6ac452e72f77f3cc071184ae34dbf0e/download-positionspapier-deutsche-version-data.pdf> [Fecha de la consulta: 08.03.2020].

¹⁷ KIESEWETTER, Roderich, «Kampfdrohnen für die Bundeswehr? Ja – zum Schutz der Soldaten!», *Sicherheit & Frieden*, 31(3), 2013, p. 177. El problema potencial, identificado por SINGER, Peter W., «The Ethics of Killer Applications: Why Is It So Hard To Talk About Morality When It Comes to New Military Technology?», *Journal of Military Ethics*, 9(4), 2010b, pp. 304-306, de que las entidades que defienden los sistemas autónomos están mejor equipadas financieramente y, por lo tanto, son más influyentes que sus oponentes, lo que podría conducir a un sesgo en el debate público, no pudo determinarse durante la investigación para el presente trabajo, y en absoluto para Alemania.

¹⁸ SINGER, *op. cit.* (2009), pp. 175, 422.

militares, y para lectores hispanohablantes sin conocimientos suficientes de la lengua alemana que deseen un tratamiento del tema que considere una parte significativa de las pocas obras científicas relevantes en alemán.

Para alcanzar sus objetivos, el texto inicialmente aclara y define unos conceptos clave de la discusión científica pertinente. Luego, se presentan las ventajas y desventajas militares esenciales de la IA y la robótica. Después, se examinan los posibles problemas de estas tecnologías en términos éticos. Se llegará a la conclusión de que el paradigma utilitarista¹⁹ ofrece soluciones satisfactorias para la mayoría de los dilemas identificados. Por razones de espacio, la importante polémica sobre los asesinatos selectivos ya mencionados con drones estadounidenses no se tratará por separado; más bien, los argumentos éticos relevantes se extraerán de la discusión y se generalizarán. Por la misma razón, al contrario de la versión de esta obra entregada como Trabajo Fin de Curso y Trabajo Fin de Máster en el XXI Curso de Estado Mayor de las Fuerzas Armadas, no se analizará por separado la amplia discusión sobre los aspectos legales y de seguridad asociados con dichas tecnologías. De acuerdo con esta reducción de contenido, algunas conclusiones han sido suprimidas.

Por la repetición ya indicada de los argumentos principales con relativamente pocas contribuciones originales desde 2013, y para conseguir resultados más equilibrados, habría sido interesante considerar la literatura científica de países no occidentales, especialmente de países asiáticos y de países musulmanes, que han sido las «víctimas» más importantes del uso de armas robóticas hasta la fecha. Los conocimientos idiomáticos del autor y el sitio disponible no lo han permitido; por lo tanto, una evaluación exhaustiva de los argumentos pertinentes queda como un objetivo propuesto para futuras investigaciones.

Delimitación conceptual

El término *robot* tiene su origen en la obra *Rossumovi Univerzální Roboti* (1921) de Karel Čapek, en la que describe a servidores mecánicos que se rebelan contra sus dueños humanos.

¹⁹ El trabajo asumirá, y por lo tanto no proporcionará, conocimientos básicos de los diferentes enfoques éticos. Para amplia información sobre el utilitarismo, véase STAFFORINI, Pablo, «Utilitarian Philosophers», 2020, <https://www.utilitarian.net/> [Fecha de la consulta: 05.04.2020].

La palabra *robot* significa «trabajo» o «trabajo forzado»²⁰. Hoy en día, como diferencia de otros tipos de máquinas, un robot se define como un ente con tres componentes decisivos: sensores que monitorean el medio ambiente y detectan cambios en él, un procesador o inteligencia artificial que decide cómo se responde a las condiciones ambientales, y efectores que vuelven a actuar sobre el medio ambiente basándose en esta información²¹.

El término *inteligencia artificial* fue acuñado a mediados de los años cincuenta por John McCarthy, quien lo entendió como la «ciencia y la ingeniería de la fabricación de máquinas inteligentes»²². Marvin Minsky expresó una opinión similar en 1968: «la Inteligencia Artificial es la ciencia de hacer que las máquinas hagan cosas que requerirían inteligencia si las hicieran los hombres»²³. Sin embargo, la siguiente definición se adecúa más al objeto de estudio del presente trabajo: «la IA es un término comodín que se refiere a un amplio conjunto de técnicas computacionales que permiten a los ordenadores y a los robots resolver problemas complejos, aparentemente abstractos, que antes solo habían cedido a la cognición humana»²⁴.

En la IA, a menudo se hace una distinción entre IA «débil», es decir, la IA que emula ciertas habilidades humanas como la observación, el lenguaje y el aprendizaje; y la IA «fuerte», o también llamada «inteligencia artificial general», una mente humana o similar a la humana creada artificialmente. Mientras que la IA «débil», desde los años sesenta, ha sido superior a la mente humana para escenarios estrictamente definidos y en áreas específicas, la IA «fuerte» seguirá siendo ciencia ficción al menos a corto plazo²⁵.

En 2009, los investigadores más optimistas, incluyendo la Agencia de Proyectos de Investigación Avanzada de Defensa de Esta-

²⁰ SCHÖRNIG, Niklas, «Die Automatisierung des Krieges: Eine kritische Bestandsaufnahme», en: SCHMIDT-RADEFELDT, Roman y MEISSLER, Christine (eds.), *Automatisierung und Digitalisierung des Krieges. Drohnenkrieg und Cyberwar als Herausforderungen für Ethik, Völkerrecht und Sicherheitspolitik*, Baden-Baden: Nomos Verlagsgesellschaft 2012, p. 33, n. 2; SINGER, *op. cit.* (2012), p. 24.

²¹ SINGER, *op. cit.* (2009), 67.

²² BOULANIN, Vincent, «Artificial Intelligence: A primer», en: BOULANIN, Vincent (ed.), *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk. Volume I. Euro-Atlantic Perspectives*, Solna: Stockholm International Peace Research Institute 2019a, p. 13.

²³ Citado en KRISHNAN, *op. cit.* (2009), p. 47.

²⁴ BOULANIN, *op. cit.* (2019a), p. 14.

²⁵ BOULANIN, *op. cit.* (2019a), pp. 13-14; DAHLMANN y DICKOW, *op. cit.*, p. 9.

dos Unidos (DARPA), asumieron que la verdadera autonomía con pensamiento humano no existiría experimentalmente antes de 2030²⁶. En este contexto, es esencial distinguir entre los sistemas de IA que funcionan *top-down*, es decir, debido a una anidación de reglas extremadamente compleja y a una derivación lógica y causal, y los ahora más prometedores sistemas *bottom-up*, que tienen una estructura básica relativamente sencilla, pero que mediante la evaluación masiva de datos y el reconocimiento de patrones construyen un conjunto de reglas inductivamente (*machine learning* o *deep learning*). Dentro de ambos paradigmas existen diferentes enfoques funcionales. Es importante para el presente trabajo que las acciones (y los fallos) de los sistemas *bottom-up* sean intrínsecamente incomprensibles y, por lo tanto, impredecibles (problema de la *black box*), lo que podría tener consecuencias catastróficas para los sistemas de armas²⁷.

El momento en el que el hombre desarrolle una IA que sea consciente de sí misma, y que también pueda desarrollarse a sí misma, fue llamado *singularidad* por Vernor Vinge en 1993. Análogamente a una singularidad cuántica, es decir, un agujero negro, las leyes y suposiciones anteriores sobre las capacidades, los límites y la velocidad de desarrollo de la inteligencia (artificial) perderían su validez en este punto: hemos establecido estas leyes sobre la base de nuestra propia inteligencia relativamente estática, pero esta inteligencia humana y su limitación biológico-genética serían entonces eliminadas de la ecuación del progreso. No está claro si se logrará o no la singularidad, pero la historia de la tecnología de la información nos ha sorprendido repetidamente con su evolución exponencial. La singularidad de Vinge durante la vida de la mayoría de los lectores de este trabajo es por lo tanto concebible²⁸.

Todavía no se dispone de una definición y una escala de autonomía que sean uniformes a nivel internacional. Sin embargo, se encuentran frecuentemente las referencias al OODA Loop (*observe, orient, decide, act*), desarrollado por John Boyd, coronel de la U.S. Air Force: si el hombre está *in the loop*, debe activar las funciones esenciales del bucle; si está *on the loop*, monitorea la ejecución independiente por un ordenador, y puede intervenir si

²⁶ KRISHNAN, *op. cit.* (2009), 44, 48.

²⁷ KRISHNAN, *op. cit.* (2009), pp. 48-53, 58; BOULANIN, *op. cit.* (2019a), pp. 14-21.

²⁸ SINGER, *op. cit.* (2009), pp. 94-108; KRISHNAN, *op. cit.* (2009), p. 48; SCHÖRNIG, *op. cit.* (2012), pp. 39-40. Para ejemplos actuales de los logros de la inteligencia artificial, véase LEWIS, *op. cit.* (2017), pp. 1-2.

es necesario; si, por otro lado, está *out of the loop*, solo puede influir en las acciones del ordenador después de la ejecución, o no puede influir en absoluto²⁹.

Una definición vinculante en virtud del derecho internacional de los AWS, de los LAWS o incluso de «autonomía», aún no existe³⁰. Son interesantes las definiciones adoptadas oficialmente por Estados Unidos y el Comité Internacional de la Cruz Roja (CICR), según las cuales los sistemas se consideran totalmente autónomos si son capaces, con la ayuda de sensores y software, de pasar por todo el ciclo de *targeting* (*find, fix, track, target, engage, assess, F2T2EA*) de forma independiente después de haber sido activados³¹. Cabe señalar que una prohibición de los LAWS según esta definición sin excepciones significaría la abolición de los sistemas CRAM ya mencionados —y en gran medida incontrovertidos—³². La Oficina de Desarrollo del Ejército de Tierra alemán ofrece la primera definición alemana de los LAWS: «Un [LAWS] es un sistema de armas que está diseñado principalmente para ejercer violencia mortal únicamente contra individuos y que, sin ninguna interferencia y control humano, percibe su entorno y condición interna, hace una evaluación de la situación, decide, actúa, evalúa y aprende de ella»³³. Esta definición no incluye, como explica la Oficina, sistemas cuyas armas no son intencionadamente letales, y armas que son diseñadas principal-

²⁹ SCHÖRNIG, Niklas, «Unmanned Systems: The Robotic Revolution as a Challenge for Arms Control», en: REUTER, Christian (ed.), *Information Technology for Peace and Security. IT Applications and Infrastructures in Conflicts, Crises, War, and Peace*, Wiesbaden: Springer Vieweg 2019, pp. 239, 245. Para una visión general desarrollada por la Administración Nacional de Aeronáutica y del Espacio (NASA) de EE. UU. de ocho grados diferentes de autonomía, relacionados con las subáreas del OODA Loop, véase PROUD, Ryan W., HART, Jeremy J. y MROZINSKI, Richard B., «Methods for Determining the Level of Autonomy to Design into a Human Spaceflight Vehicle: A Function Specific Approach», 2003, p. 4, <https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20100017272.pdf> [Fecha de la consulta: 08.03.2020]. Para escalas alternativas, véanse FINN, Anthony y SCHEDING, Steve, «Developments and Challenges of Autonomous Unmanned Vehicles. A Compendium», Berlin/Heidelberg: Springer 2010, pp. 42-43; Heinrich-Böll-Stiftung, «Autonomy in Weapon Systems. The Military Application of Artificial Intelligence as a Litmus Test for Germany's New Foreign and Security Policy», Berlin: Heinrich-Böll-Stiftung 2018, p. 42.

³⁰ SAUER, *op. cit.* (2019a), p. 25.

³¹ SAUER, *op. cit.* (2019a), p. 25; SAUER, Frank, «Military applications of artificial intelligence: Nuclear risk redux», en: BOULANIN, Vincent (ed.), *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk. Volume I. Euro-Atlantic Perspectives*, Solna: Stockholm International Peace Research Institute 2019b, pp. 86-87.

³² DAHLMANN y DICKOW, *op. cit.*, p. 17.

³³ Amt für Heeresentwicklung, *op. cit.*, p. 9. Traducción del autor.

mente para la lucha contra aviones, tanques y buques, aunque estén tripulados.

Si bien existen varios enfoques específicos, en este trabajo el paradigma utilitarista se entiende como el paradigma ético que aprueba o desaprueba toda acción, según su tendencia de aumentar o disminuir el interés, la ventaja o la felicidad de la parte (un individuo, grupo o una sociedad) que se está considerando³⁴.

Ventajas y desventajas militares de la inteligencia artificial y la robótica

Ventajas

Los sistemas robóticos, especialmente los UGV, pueden descargar a los soldados de los efectos físicos negativos de la guerra. Por ejemplo, ya se están utilizando dispositivos de transporte semiautónomos sobre ruedas o piernas, que tienen un alto beneficio militar debido a la gran vulnerabilidad de los convoyes. También se están desarrollando híbridos de vehículos y aeronaves para el rescate, la atención y la evacuación de los heridos. Estas tareas en el ámbito de la logística y la reposición seguirán siendo probablemente el mayor campo de aplicación de la robótica hasta el año 2040³⁵.

Con respecto a los vehículos aéreos no tripulados armados, a veces se argumenta que su permanencia sobre las fuerzas hostiles, la indefensión de estas fuerzas frente a ellos y la necesidad de arriesgar la propia vida para, en el mejor de los casos, destruir una máquina, desgastan psicológicamente al enemigo. Aparte del hecho de que este efecto solo pueda ocurrir, si es que ocurre, en conflictos asimétricos en los que el enemigo no tiene posibilidad de combatir directamente a los drones, se puede asumir que el enemigo se adaptará rápidamente y se acostumbrará a él³⁶. Un argumento más valioso es que la vigilancia permanente por parte de los UAV puede ayudar a reunir pruebas contra los insurgentes que luego pueden ser condenados en un juicio regular³⁷.

³⁴ BENTHAM, Jeremy, «An Introduction to the Principles of Morals and Legislation», BURNS, J. y HART, H. L. A. (eds.), 4.a ed., Oxford: Oxford University Press 2005, pp. 11-12.

³⁵ SINGER, *op. cit.* (2009), pp. 112-113; BOISBOISSEL, *op. cit.*, p. 79; QASRAWI, BOUKHTOUTA y GIZEWSKI, *op. cit.*, pp. 48, 51-56.

³⁶ SINGER, *op. cit.* (2009), pp. 297-301, 306.

³⁷ SINGER, *op. cit.* (2009), pp. 222-223.

Los sistemas autónomos ofrecen otras ventajas que no ofrecen los sistemas robóticos de control remoto³⁸. En primer lugar, el futuro campo de batalla y su *hyperwar*³⁹ probablemente se volverá tan rápido y confuso que el tiempo de reacción de un humano ya no será suficiente para ganar la batalla⁴⁰. Por otro lado, la adquisición de blancos y la respuesta al ataque por parte de robots autónomos pueden tener lugar tan rápidamente que cualquier enemigo que dispare al robot es inmediatamente combatido y eliminado; una característica que traería enormes ventajas tácticas, especialmente en los campos de batalla urbanos del futuro⁴¹. Al mismo tiempo, los AWS pueden actuar precisamente a distancia y, si se desea, preservar los órganos vitales. Un producto correspondiente llamado Snibot ya está en desarrollo⁴². Además, en un futuro próximo los sistemas de control remoto, cuyo número sigue creciendo, ya no serán controlables debido a las limitaciones de ancho de banda⁴³. Siguiendo con esta lógica, los UAV de combate modernos deberán funcionar de forma autónoma, al menos como opción, ya que es de esperar que la navegación GPS, la comunicación por satélite y otros enlaces de comunicación se interrumpan en un conflicto de alta intensidad⁴⁴. Los aviones hipersónicos, en cualquier caso, serán demasiado rápidos para ser controlados por los humanos⁴⁵.

Los soldados que estén físicamente presentes en combate y necesiten usar toda su atención para controlar un robot perderán la conciencia de la situación, estarán particularmente en riesgo ellos mismos y enlazarán fuerzas necesarias para su protec-

³⁸ En vista de estas ventajas, parece (dependiendo de la perspectiva) políticamente útil o hipócrita insistir en el concepto de «man in the loop» para reducir la resistencia al desarrollo presumiblemente inevitable de los LAWS. Véase SPARROW, Rob, «The Ethical Challenges of Military Robots», en: DABRINGER, Gerhard (ed.), *Ethical and Legal Aspects of Unmanned Systems. Interviews*, Wien: BMVLS 2010, pp. 97-98.

³⁹ Amt für Heeresentwicklung, *op. cit.*, pp. 5-6.

⁴⁰ SINGER, *op. cit.* (2009), p. 128.

⁴¹ GRAHAM, Stephen, «Sehende Kampfzonen. Urbane Kriegsführung und US-Militärtechnologie», en: MARSISKE, Hans-Arthur (ed.), *Kriegsmaschinen. Roboter im Militäreinsatz*, Hannover: Heise 2012, pp. 72-73.

⁴² BOISBOISSEL, *op. cit.*, pp. 77-78.

⁴³ SINGER, *op. cit.* (2009), pp. 202-203; KRISHNAN, *op. cit.* (2009), p. 38.

⁴⁴ SINGER, *op. cit.* (2009), p. 199; KRISHNAN, *op. cit.* (2009), pp. 38-39; SCHÖRNIG, *op. cit.* (2012), p. 39; BRONK, Justin, «The impact of unmanned combat aerial vehicles on strategic stability», en: BOULANIN, Vincent (ed.), *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk. Volume I. Euro-Atlantic Perspectives*, Solna: Stockholm International Peace Research Institute 2019, p. 102.

⁴⁵ KRISHNAN, *op. cit.* (2009), p. 62.

ción. Los operadores de múltiples robots alcanzan rápidamente sus límites cognitivos y pueden ser saturados por el enemigo a través de ataques masivos y la necesidad de autorizar todos y cada uno de los usos de armas. Aunque se podría utilizar personal menos cualificado para confirmar las soluciones de ataque ya optimizadas, ese personal podría verse desbordado por decisiones inesperadas de considerable importancia. Además, si cada vez fuera necesaria una orden de fuego explícita de un ser humano, el enemigo solo tendría que interrumpir el enlace de datos para inutilizar los robots. Por otro lado, el manejo de robots por humanos no supone prácticamente ningún ahorro de personal⁴⁶.

En general, los AWS pueden realizar ciertas tareas mucho más rápido que los humanos o los robots controlados por humanos, lo cual es particularmente atractivo para tareas de tiempo crítico como la defensa aérea, el combate aéreo o la ciberdefensa. Los AWS no requieren un enlace de comunicación permanente con el mando superior o su controlador, y de esa manera, al contrario de sistemas de control remoto, permiten reducir el número de operadores y analistas humanos. Son especialmente adecuados para tareas monótonas, penosas y peligrosas, ya que no presentan limitaciones de rendimiento humano como el agotamiento, el aburrimiento, el hambre o el miedo. Pueden utilizarse en enjambres, lo que sería imposible con sistemas controlados individualmente. Pueden penetrar en zonas inaccesibles o peligrosas para los soldados y los robots controlados por el hombre, como las zonas protegidas por sistemas anti-acceso/denegación de áreas (A2/AD), los océanos, los polos y el espacio⁴⁷. Por razones fisiológicas y de coste, el espacio como campo de batalla probablemente solo será accesible para los robots; la tecnología robótica desarrollada con este fin también podría utilizarse para el asentamiento civil del espacio y de las profundidades marinas⁴⁸.

La investigación en el campo de la IA y la robótica asegura el desarrollo tecnológico militar de los Estados más pequeños. Esto es, les permite poseer un ejército eficiente y promueve su desarrollo económico⁴⁹. Las nuevas tecnologías en estos ámbitos pueden reforzar la industria de defensa de la UE y, por tanto,

⁴⁶ SINGER, *op. cit.* (2009), pp. 75, 126-127; KRISHNAN, *op. cit.* (2009), pp. 36-37.

⁴⁷ BOULANIN, *op. cit.* (2019a), pp. 23-24.

⁴⁸ SINGER, *op. cit.* (2009), pp. 120-121; KRISHNAN, *op. cit.* (2009), pp. 22-23, 70.

⁴⁹ KRISHNAN, *op. cit.* (2009), p. 37; DILĀNS, Andis, «Foreword», en: ROMANOV, Uģis y ANDŽĀNS, Māris (eds.), *Digital Infantry Battlefield Solution. Research and Innovation. Part III*, Tallinn: Milrem Robotics 2019, pp. 10-11.

su autonomía estratégica⁵⁰. Una prohibición del desarrollo de los LAWS, por otra parte, podría obstaculizar la investigación civil, por ejemplo, de los automóviles autodirigidos; a este respecto, una restricción legal al *uso* de los LAWS podría ser más ventajosa que una prohibición completa del *desarrollo*, incluso si los LAWS pudieran ser utilizados ilegalmente en casos puntuales⁵¹. Como tecnología de *dual-use*, la proliferación de la tecnología robótica difícilmente puede evitarse de todos modos⁵².

No cabe duda de que pedir a los soldados humanos que arriesguen su vida por su presencia física en el campo de batalla, teniendo en cuenta la letalidad de las guerras futuras, la lucha contra enemigos que no respetan el derecho internacional y los actuales objetivos bélicos no ligados directamente a la supervivencia del Estado, podría ser una petición excesiva⁵³. Además de proteger la vida humana misma, la sustitución de los seres humanos por robots ayudaría a reducir fenómenos como el trastorno de estrés postraumático, el suicidio y otros acompañamientos psicopatológicos en soldados que tienen que matar o experimentar la muerte de compañeros⁵⁴.

Según Strawser⁵⁵, el uso de los vehículos aéreos no tripulados no solo está moralmente permitido, sino también debe ser obligatorio, debido al menor riesgo para la tripulación, siempre que la guerra en la que se utilizan sea justa y que su uso no dé lugar a una significativa reducción de capacidades militares (incluida la distinción de objetivos exigida por el Derecho Internacional Humanitario, o DIH) o a otras consecuencias negativas, como la incitación de la población civil⁵⁶. En este contexto, cabe señalar que la omisión del uso de UAV (o incluso de los LAWS) podría violar el párrafo 1 del artículo 31 de la Ley de militares alemanes

⁵⁰ PRIŠMANTAITĖ, *op. cit.*, p. 20.

⁵¹ ALTMANN, Jürgen y SAUER, Frank, «Autonomous Weapon Systems and Strategic Stability», *Survival*, 59(5), 2017, pp. 132-133; DAHLMANN y DICKOW, *op. cit.*, p. 16.

⁵² SCHÖRNIG, *op. cit.* (2012), p. 59.

⁵³ En el contexto del empleo de robots para reducir la letalidad del campo de batalla para seres humanos, parece irónico que KRISHNAN, *op. cit.* (2009), p. 123, identifique la contaminación por nanopartículas (es decir, nanobots; véase abajo) como un aspecto de dicha letalidad.

⁵⁴ KRISHNAN, *op. cit.* (2009), pp. 121-124; ARKIN, Ronald C., «The Case for Ethical Autonomy in Unmanned Systems», *Journal of Military Ethics*, 9(4), 2010a, pp. 336-337.

⁵⁵ STRAWSER, Bradley J., «Moral Predators: The Duty to Employ Uninhabited Aerial Vehicles», *Journal of Military Ethics*, 9(4), 2010, pp. 343-349, 351, 355.

⁵⁶ Sin embargo, dado que la obligación de emplear drones puede ser anulado por una «strong enough countervailing normative reason», que STRAWSER, *op. cit.* (2010), p. 350, no define, su argumento parece trivial.

(Soldatengesetz): «Por la relación de servicio y de lealtad entre militares y la Federación, la Federación velará por el bienestar de los soldados de compromiso permanente y de compromiso temporal, así como de sus familias, incluido el tiempo posterior a la terminación del servicio»⁵⁷. Además, todas las fuerzas armadas del mundo tienen el deber de proteger a su población respectiva y de hacer valer sus intereses contra las amenazas externas de la manera más eficaz y barata posible.

En este sentido, la renuncia a los sistemas robóticos o autónomos podría significar no solo un despilfarro de recursos humanos y financieros —y, por lo tanto, un daño indirecto a la propia población—, sino en casos extremos también permitiría la agresión o incluso el genocidio contra ella⁵⁸. Esta problemática se agravará cuando el enemigo despliegue los LAWS altamente desarrollados. En este momento, sería necesario desplegar igualmente los LAWS para evitar la derrota⁵⁹.

El ahorro económico a través de sistemas no tripulados, que a su vez puede ser utilizado para mejorar la calidad de vida de los ciudadanos, se debe, por un lado, a que pueden prescindir de componentes que sirven para proteger o apoyar una tripulación humana⁶⁰. Además, los robots no necesitan un salario ni una pensión⁶¹. En sociedades cada vez más envejecidas y menos saludables, son menos las personas que pueden asumir el papel clásico de los soldados. Los sistemas de control remoto o autónomos son particularmente útiles en este entorno y están en consonancia con la experiencia lúdica de la juventud de hoy⁶². Por otro lado, la formación constante a la que se somete el soldado se eliminaría con la utilización de los robots que aprenden inmediatamente a través de su programación, y cuyos conocimientos se conservan permanentemente⁶³. Por último, no debe olvidarse que la reducción de las fuerzas armadas mediante la introduc-

⁵⁷ Traducción del autor.

⁵⁸ KRISHNAN, *op. cit.* (2009), pp. 119-120.

⁵⁹ KRISHNAN, *op. cit.* (2009), p. 70; DAHLMANN y DICKOW, *op. cit.*, p. 11.

⁶⁰ SCHÖRNIG, *op. cit.* (2012), pp. 44-45.

⁶¹ KRISHNAN, *op. cit.* (2009), pp. 2, 35-36. En 2020, está previsto que 34,04% (= 45.000 millones de euros) del presupuesto de defensa alemán se gastan por el personal. Véase Bundesministerium der Finanzen, «Bundeshaushalt. Einzelplan 14», 2020, <https://www.bundeshaushalt.de/#/2020/soll/ausgaben/einzelplan/14.html> [Fecha de la consulta: 03.04.2020].

⁶² SINGER, *op. cit.* (2009), pp. 368-370.

⁶³ SINGER, *op. cit.* (2009), p. 64.

ción de sistemas autónomos contribuye a la protección del medio ambiente y al ahorro de combustibles fósiles⁶⁴.

Como alternativa al uso de robots, el uso de implantes u otras mejoras artificiales para los soldados (*human enhancement*) y, por lo tanto, la creación de los llamados «cíborgs» (organismos cibernéticos) emerge como una opción para aumentar la supervivencia de los militares y reducir la presencia humana en el campo de batalla. Ya hoy en día, las prótesis robóticas, algunas de las cuales están conectadas al sistema nervioso de los soldados, permiten reintegrarlos después de una lesión grave⁶⁵. Además, a partir de 2010, el ejército de Estados Unidos empezó a apoyar proyectos de investigación sobre la nutrición humana a base de hierba y otras plantas indigeribles, suplementos dietéticos, la mejora de la capacidad de aprendizaje, la resistencia al estrés, los exoesqueletos, la representación visual de amenazas percibidas inconscientemente, la mejora del sentido del olfato para reconocer agentes de guerra química y la comunicación del pensamiento⁶⁶.

Los enjambres de microrobots, los llamados «nanobots», que pueden manipular objetos a nivel atómico, pueden considerarse como la forma definitiva de desarrollo de sistemas autónomos. Estos podrían penetrar en el cuerpo humano y, dependiendo de su programación, curar o causar heridas⁶⁷. Los nanobots también podrían utilizarse en los campos de la criptografía, la tecnología de ocultación, la reparación de sistemas mecánicos, la tecnología de sensores y la detección nuclear, biológica y química, en ordenadores cuánticos o con fines de *human enhancement*. Los nanobots altamente desarrollados serían prácticamente indestructibles, ya que la destrucción de algunos especímenes proporcionaría materia prima con la que los demás nanobots suplirían sus bajas. Las nanoguerras podrían librarse de forma muy rápida y eficaz, ya que una nanofábrica simplemente tendría que trasladarse a territorio enemigo, donde crearía su propio ejército, posiblemente de robots autónomos. En general, la guerra podría miniaturizarse de esta manera, ya que solamente los sistemas más pequeños tendrían la posibilidad de pasar desapercibidos por los nanobots y no ser destruidos⁶⁸.

⁶⁴ KRISHNAN, *op. cit.* (2009), p. 121.

⁶⁵ SINGER, *op. cit.* (2009), pp. 374-377.

⁶⁶ LIN, Patrick, «Ethical Blowback from Emerging Technologies», *Journal of Military Ethics*, 9(4), 2010, pp. 317-318.

⁶⁷ SINGER, *op. cit.* (2012), p. 26.

⁶⁸ KRISHNAN, *op. cit.* (2009), pp. 83-86.

Aunque estos nanobots siguen pareciendo hechos del futuro, ya existen hoy en día consideraciones concretas de inundar los espacios urbanos con enjambres de microrobots que proporcionarían datos de objetivos a los sistemas de mayor alcance, como los drones, con los que, opcionalmente de forma autónoma, el espacio urbano puede ser despejado de enemigos prácticamente sin demora. Como alternativa, estos robots de reconocimiento podrían escanear las corrientes cerebrales (actitudes) de la población y realizar una selección de objetivos sobre esta base⁶⁹. No será posible controlar a estos pequeños robots individualmente: por un lado, debido a su gran número; por otro lado, porque su control directo causaría náuseas en la mayoría de los operadores, ya que sus movimientos repentinos son incompatibles con la percepción humana⁷⁰.

A medio plazo, es más probable que la inteligencia artificial se utilice en áreas tácticas que en áreas operativas y estratégicas⁷¹. Sin embargo, los sistemas para el nivel operacional (generación y juego de líneas de acción) ya empezaron a desarrollarse hacia 2008⁷². Sin apoyo de la IA, los Estados Mayores del futuro no serán capaces de superar la cantidad enorme de información disponible⁷³. En el contexto de la guerra de la información y la cibernética, en el que los seres humanos son incapaces de descubrir la identidad de un atacante, una IA podría decidir en última instancia quién es el enemigo y entrar incluso en el nivel estratégico⁷⁴.

Como aplicación estratégica adicional, los sistemas autónomos podrían permitir a los Estados democráticos poderosos (además del uso de tropas o mercenarios indígenas) librar guerras con fines de democratización, protección de los derechos humanos y prevención de una catástrofe humanitaria o del genocidio⁷⁵.

⁶⁹ SINGER, *op. cit.* (2009), pp. 289-291; KRISHNAN, *op. cit.* (2009), p. 84; GRAHAM, *op. cit.*, pp. 70-71. Incluso la Oficina de Desarrollo del Ejército de Tierra alemán (Amt für Heeresentwicklung, *op. cit.*, pp. 4-5) describe un escenario en cual el ejército emplea enjambres de drones pequeños para el reconocimiento de un terreno urbano y la lucha contra vehículos y drones enemigos.

⁷⁰ SINGER, *op. cit.* (2009), pp. 118-119.

⁷¹ DILĀNS, *op. cit.*, p. 10.

⁷² KRISHNAN, *op. cit.* (2009), pp. 54-55.

⁷³ Amt für Heeresentwicklung, *op. cit.*, p. 7.

⁷⁴ KRISHNAN, *op. cit.* (2009), p. 35.

⁷⁵ SCHÖRNIG, *op. cit.* (2012), pp. 46-48, 54-55; SCHÖRNIG, Niklas, «Unmanned Warfare. Towards a Neo-Interventionist Era?», en: KÜMMEL, G. y GIEGERICH, S. (eds.), *The Armed Forces: Towards a Post-Interventionist Era?*, Wiesbaden: Springer 2013, p. 228.

Operacionalmente, podría haber un ejército robótico de mantenimiento de la paz, estabilizando los lugares «olvidados» del mundo que se encuentran fuera del área de interés directo de los Estados occidentales. En vista del poder destructivo potencial de las armas autónomas o robóticas descritas anteriormente, estas podrían tener al mismo tiempo un efecto disuasorio tal que impedirían completamente la guerra al menos entre Estados dotados de ellas, y, por lo tanto, los robots serían armas preferibles desde el punto de vista moral. De hecho, el uso de robots en ambos lados podría hacer posible la guerra sin pérdidas humanas. Es decir, si el comportamiento de los robots puede ser simulado de forma fiable, los Estados podrían prescindir por completo de la confrontación física y decidir la guerra en una simulación. Una guerra que se libra a pesar de una pérdida simulada a lo mejor terminaría muy rápidamente y sería muy costosa y, por lo tanto, no racional. Alternativamente, las guerras reales se decidirían por la guerra cibernética, en concreto la repentina y completa toma de control de la infraestructura de mando del enemigo por medio de la ciberpiratería y, posiblemente, sin el uso de la fuerza física⁷⁶.

Desventajas

La idea de un ejército de intervención robótica que acaba de ser presentada puede ser contrarrestada por el hecho de que los robots y los drones no estén en condiciones de ganarse los *hearts and minds* de la población debido a la falta de humanidad, lo que podría ser perjudicial para las operaciones⁷⁷. Por el contrario, los robots que patrullan las calles pueden atemorizar o incluso molestar a la población y dar la impresión de que a Occidente no le vale la pena enviar verdaderos soldados⁷⁸.

Incluso dentro de las propias Fuerzas Armadas, la coexistencia de seres humanos y sus compañeros robóticos (altamente desarrollados) podría causar problemas. Los primeros podrían resultar perjudicados por el mal funcionamiento de sus supuestos amigos digitales⁷⁹. Precisamente por la dificultad, por definición,

⁷⁶ KRISHNAN, *op. cit.* (2009), pp. 118, 125-126.

⁷⁷ SCHÖRNIG, *op. cit.* (2012), p. 55.

⁷⁸ SINGER, *op. cit.* (2009), pp. 323-324; KRISHNAN, *op. cit.* (2009), pp. 126-127; SPARROW, *op. cit.* (2010), p. 93.

⁷⁹ KRISHNAN, *op. cit.* (2009), p. 127. Sin embargo, este temor contrasta con una notable estadística de la guerra de Irak de 2003, en la que el 17% de todas las bajas

de verificar *definitivamente* la funcionalidad de las máquinas *autónomas* antes de su uso, sería difícil para el soldado medio desarrollar la confianza en esos nuevos tipos de combatientes de cuyo comportamiento, a diferencia del de su colega humano, no puede sacar ninguna conclusión a través de la empatía, las preferencias compartidas y los valores. Este problema aumentará a medida que los robots se vuelvan más adaptables, más potentes y, por lo tanto, más impredecibles. En casos extremos, el soldado podría negarse a servir con los AWS o a utilizarlos, lo que reduciría la efectividad de la fuerza militar⁸⁰. Por otro lado, con miras a sus subordinados robóticos, los líderes militares podrían acostumbrarse al hecho de que sus subordinados siempre funcionen perfectamente y nunca cuestionen sus órdenes⁸¹.

Además, el uso de armas de control remoto o autónomas podría acarrear una disminución de los valores militares como el sacrificio, la resistencia física y mental, la valentía y la capacidad de mando⁸². Existe la posibilidad de dividir a los militares en soldados «clásicos» y «nuevos», para quienes el combate físico ya no forma parte del servicio, con consecuencias poco claras para el ethos común y la identidad militar⁸³.

Asimismo, la introducción de los cibernéticos podría formar una división de «castas» entre militares mejorados y sus compañeros o civiles no mejorados. En este caso, los veteranos mejorados que se incorporan al mercado laboral civil después de despedirse de las Fuerzas Armadas, dependiendo de la profesión, superarían con creces a sus competidores. De esta manera, y al igual que el dopaje deportivo, la creación de cibernéticos debilitaría la importancia de los logros y las cualificaciones personales⁸⁴. La prohibición del despido de los soldados mejorados parece a primera vista una

militares de Estados Unidos se debieron al fuego amigo, aunque se sabe que no se utilizaron robots «asesinos» durante el conflicto. Véase LEWIS, Larry, «Redefining Human Control. Lessons from the Battlefield for Autonomous Weapons», 2018, pp. 6-8, https://www.cna.org/CNA_files/PDF/DOP-2018-U-017258-Final.pdf [Fecha de la consulta: 20.11.2019].

⁸⁰ ROFF, Heather M. y DANKS, David, «'Trust but Verify': The Difficulty of Trusting Autonomous Weapons Systems», *Journal of Military Ethics*, 17(1), 2018, pp. 8-12.

⁸¹ KRISHNAN, *op. cit.* (2009), p. 137.

⁸² SINGER, *op. cit.* (2009), pp. 331-333; KRISHNAN, *op. cit.* (2009), pp. 134-135, 136-137.

⁸³ SINGER, *op. cit.* (2009), p. 370.

⁸⁴ Por otra parte, las ventajas en el mercado laboral civil podrían utilizarse explícitamente como argumento de reclutamiento y no necesariamente tendrían que ser percibidas como injustas por los civiles en vista de la especial calidad del servicio en las fuerzas armadas.

solución parcial, pero supondría una restricción considerable de la libertad personal de los afectados y los exponería a posibles abusos por parte de la organización militar. Equipar solo a los soldados de carrera o incluso a las fuerzas especiales con mejoras podría mitigar el problema, pero no eliminarlo. Alternativamente, los implantes podrían ser retirados tras la terminación del contrato (un problema en sí mismo), pero la restricción a intervenciones fundamentalmente reversibles presumiblemente impediría la colocación de los implantes más potentes. Las consecuencias psicológicas, fisiológicas y posiblemente patológicas de retirar o apagarlas (y ya de su instalación) serían en cualquier caso inicialmente poco claras. El tratamiento histórico de los afectados por las armas experimentales (por ejemplo, los ensayos nucleares, el agente naranja o el síndrome de la Guerra del Golfo) no augura nada bueno para los primeros cíborgs. Lo mismo ocurre con el peligro de que los implantes sean hackeados por el oponente. Por otra parte, es concebible que los cíborgs puedan hacer un mal uso de sus mejoras con fines delictivos, especialmente después de su despido. Por último, el aumento de la resistencia de los cíborgs a los daños físicos y mentales podría tener consecuencias adversas para el tratamiento de prisioneros de guerra normales⁸⁵.

Es probable que los sistemas con mayor autonomía se vuelvan más complejos y, por lo tanto, no solo más caros, sino también más susceptibles a las deficiencias internas⁸⁶. Además, con una complejidad creciente, es decir, con un desempeño ético y legal cada vez más adecuado, serían más sensibles a los ataques cibernéticos contra sus sensores, sistemas de control y centros de mando superiores⁸⁷. También habría un riesgo de pérdida de control debido a la interferencia electromagnética⁸⁸. Así, una

⁸⁵ SINGER, *op. cit.* (2009), pp. 377-381; KRISHNAN, *op. cit.* (2009), p. 75; LIN, *op. cit.*, pp. 322-323, 326; SHARKEY, Noel, «Moral and Legal Aspects of Military Robots», en: DABRINGER, Gerhard (ed.), *Ethical and Legal Aspects of Unmanned Systems. Interviews*, Wien: BMVLS 2010a, p. 51.

⁸⁶ PETERMANN, *op. cit.*, p. 78.

⁸⁷ KRISHNAN, *op. cit.* (2009), pp. 38-39; KLINCEWICZ, Michał, «Autonomous Weapons Systems, the Frame Problem and Computer Security», *Journal of Military Ethics*, 14(2), 2015, pp. 168-170; BOULANIN, *op. cit.* (2019a), p. 24; BOULANIN, Vincent, «Introduction», en: BOULANIN, Vincent (ed.), *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk. Volume I. Euro-Atlantic Perspectives*, Solna: Stockholm International Peace Research Institute 2019b, p. 3; DAHLMANN y DICKOW, *op. cit.*, p. 12.

⁸⁸ SINGER, *op. cit.* (2009), pp. 197-199.

fuerza armada basada enteramente en sistemas automatizados podría perder toda su capacidad de defensa de un solo golpe por ataques cibernéticos o pulsos electromagnéticos (EMP)⁸⁹.

Cuanto más autónomo, es decir, cuanto menor sea el intercambio de datos con sus propietarios, más difícil será para el enemigo poner un sistema autónomo bajo su control⁹⁰. Para evitar un ciberataque con seguridad, cualquier conexión entre el hombre y la máquina tendría que ser desconectada. El beneficio militar de tales máquinas en un campo de batalla complejo, sin embargo, sería probablemente bajo, ya que serían incapaces de reaccionar a cambios operacionales y estratégicos o a acciones imprevistas del enemigo en el alto nivel táctico⁹¹. Si, por otro lado, las inteligencias artificiales avanzadas desarrollan una comprensión estratégica u operativa, esta podría desviarse de la de los humanos⁹². Después de todo, históricamente se ha desarrollado un antídoto para cada tecnología de guerra. Los robots podrían llegar a ser demasiado predecibles rápidamente, por lo que parece posible un retorno a la guerra humana después de un período de guerras de robots⁹³.

Consideraciones éticas

Riesgo y equidad

En el discurso científico, principalmente en el contexto del UAV, algunos autores consideran que el riesgo de que a un soldado se lo mate en la guerra es un requisito básico para su derecho a matar a otros⁹⁴. Así, en el pasado, la conciencia de un riesgo mutuo había sido la fuente del ideal militar de coraje y caballeridad del que las fuerzas armadas de hoy se distanciarían utilizando la robótica⁹⁵. De tal manera, la guerra se aleja de la clásica

⁸⁹ SINGER, *op. cit.* (2009), pp. 199-201; BOISBOISSEL, *op. cit.*, pp. 87-88.

⁹⁰ KRISHNAN, *op. cit.* (2009), p. 39.

⁹¹ KLINCEWICZ, *op. cit.*, p. 173.

⁹² ROFF, Heather M., «The Strategic Robot Problem: Lethal Autonomous Weapons in War», *Journal of Military Ethics*, 13(3), 2014.

⁹³ CHAPA, Joe O. y BLAIR, Dave J., «The Just Warrior Ethos: A Response to Colonel Riza», *Journal of Military Ethics*, 15(3), 2016, pp. 173-174.

⁹⁴ SCHÖRNIG, *op. cit.* (2012), p. 56.

⁹⁵ SINGER, *op. cit.* (2009), pp. 432-433.

situación de duelo y recuerda más al «control de plagas»⁹⁶. Matar con robots incluso a combatientes enemigos completamente indefensos no parecería una batalla, sino una masacre ilegal⁹⁷.

De hecho, el uso de drones significa una transferencia de riesgo de militares amigos o del operador de drones al enemigo⁹⁸. La mayoría de las personas también pueden tener una reticencia intuitiva a matar a personas indefensas por control remoto. Es cuestionable, sin embargo, si esta incomodidad debe ser la base de una evaluación moral, o si dicha evaluación debe ser el resultado de criterios racionales. También hay que tener en cuenta que la lucha contra los combatientes indefensos (por ejemplo, dormidos) no está ni explícita ni implícitamente prohibida por el DIH, siempre que no se hayan rendido o estén heridos. La desagradable imagen del control de plagas también pierde su impacto bajo el supuesto de que la «plaga» ha atacado y matado previamente por su propia voluntad⁹⁹.

Además, parece dudoso que la legitimidad de un combatiente pueda establecerse sobre la base del riesgo personal que soporta, ya que siempre ha habido soldados cuya muerte y cuyas heridas eran más o menos probables, riesgo que se basa no solo en su especialización, sino también en las circunstancias del caso individual y en la naturaleza de la guerra. Además, ¿qué criterio debería determinar qué umbral de riesgo personal es suficiente para tener la autorización de librar una guerra? También cabe preguntarse si es el riesgo personal, o más bien la contribución al cumplimiento de la misión o a la protección de las propias tropas, el criterio de un combatiente legítimo. Este último argumento es a menudo más importante entre los pilotos de UAV que entre las tripulaciones de aeronaves convencionales. Por otro lado, ¿no es la idea de la perfección personal y la virtud a través de o durante la guerra una idea anticuada que va en contra del verdadero propósito de la guerra, a saber, la protección del interés nacional¹⁰⁰?

⁹⁶ MÜNKLER, Herfried, «Die neuen Kriege», Rowohlt, Reinbek, 2002, citado en SCHULZKE, Marcus, «Rethinking Military Virtue Ethics in an Age of Unmanned Weapons», *Journal of Military Ethics*, 15(3), 2016, p. 195.

⁹⁷ KRISHNAN, Armin, «Ethical and Legal Challenges», en: DABRINGER, Gerhard (ed.), *Ethical and Legal Aspects of Unmanned Systems*. Interviews, Wien: BMVLS 2010, p. 56.

⁹⁸ CHAPA y BLAIR, *op. cit.*, p. 182.

⁹⁹ STRAWSER, *op. cit.* (2010), pp. 357-358. Esto no significa que la mayoría de los soldados no tengan una aversión intuitiva a matar a un oponente dormido. Véase MORKEVICIUS, Valerie, «Tin Men: Ethics, Cybernetics and the Importance of Soul», *Journal of Military Ethics*, 13(1), 2014, pp. 5-7.

¹⁰⁰ CHAPA y BLAIR, *op. cit.*, pp. 172-173, 177-182.

¿No podrían ser los drones armas apropiadas para el «postheroísmo» de las sociedades occidentales tras dos guerras mundiales¹⁰¹? Lo que en el pasado se consideraba heroico —por ejemplo, el avance cerrado en una falange o la posición estática en una fila para disparar mosquetes— hoy solo puede considerarse una aberración tras la invención de la ametralladora¹⁰². En cualquier caso, el punto de la falta de riesgo y de «injusticia» ha sido superado desde hace mucho tiempo con el uso de misiles de crucero y cazas con armas de precisión, y no solo por los UAV o las armas autónomas en general. Por analogía con la lucha contra la delincuencia, nadie exigiría que la policía renunciara deliberadamente a los chalecos antibalas para permitir a los delincuentes una lucha justa. La exigencia de tal lucha podría calificarse como algo «akin to an archaic demand of military commanders in eighteenth century warfare to line up their troops across from one another for a “dignified battle”¹⁰³».

Incluso si el riesgo personal se acepta en principio como legitimación de la matanza en la guerra, hay que tener en cuenta que los UAV de la mayoría de los Estados, al contrario de EE. UU., no están controlados desde sus territorios respectivos. Incluso si este fuera el caso, los pilotos de UAV estarían expuestos a peligros físicos similares a los de otros soldados en una guerra simétrica, ya que sus centros de mando podrían ser atacados con armas de largo alcance. Por otro lado, en un contexto asimétrico, los pilotos de UAV pueden estar físicamente más seguros que los civiles en su propio país, ya que las bases de UAV están particularmente bien protegidas contra ataques terroristas¹⁰⁴.

No obstante, hay que considerar que incluso los operadores de UAV estacionados a una distancia continental del lugar de su empleo arriesgan contraer enfermedades psicológicas. Existen informes, principalmente no oficiales, de trastorno de estrés postraumático, automedicación, agotamiento y suicidio entre

¹⁰¹ MÜNKLER, Herfried, «Neue Kampfsysteme und die Ethik des Krieges», en: Heinrich-Böll-Stiftung (ed.), *High-Tech-Kriege. Frieden und Sicherheit in den Zeiten von Drohnen, Kampfrobotern und digitaler Kriegsführung*, Berlin: Heinrich-Böll-Stiftung 2013, pp. 13-14.

¹⁰² KIRKPATRICK, Jesse, «Drones and the Martial Virtue Courage», *Journal of Military Ethics*, 14(3-4), 2015, p. 213.

¹⁰³ STRAWSER, *op. cit.* (2010), pp. 355-357, cita en p. 357.

¹⁰⁴ SPARROW, Rob, «Martial and Moral Courage in Teleoperated Warfare: A Commentary on Kirkpatrick», *Journal of Military Ethics*, 14(3-4), 2015, p. 221.

las tripulaciones de UAV¹⁰⁵. Los estudios oficiales, que sí se han publicado, muestran que la incidencia de enfermedades mentales entre los operadores de UAV es al menos tan alta como entre las tripulaciones de aeronaves tripuladas, pero menor que entre otros soldados desplegados en tierra en la zona de operaciones¹⁰⁶.

Estos fenómenos pueden explicarse, por ejemplo, debido a la muerte accidental de civiles, por la observación de la baja de militares del propio bando o de acontecimientos, por ejemplo, ejecuciones o decapitaciones, que contradicen convicciones morales profundamente arraigadas. Ese problema se agrava si no hay posibilidad de intervención o si el soldado no ha cumplido una función protectora autoimpuesta, quizás debido a una falta de atención momentánea. También es concebible que el daño psicológico pueda ser causado por el registro constante, a veces la observación en vivo, de las propias acciones por parte de las autoridades superiores. Además, desempeñan un papel las propias acciones de asesinato, que son difíciles de procesar si no existe un riesgo físico directo de la propia persona, incluso en el caso de blancos legítimos. También existe el riesgo de perder la carrera y la seguridad financiera si se niega a cumplir las órdenes pertinentes¹⁰⁷.

Otra explicación para el número sorprendentemente alto de enfermedades psicológicas sin amenaza física es la falta de descansos mentales como el *beer call*¹⁰⁸ por parte de los pilotos de combate; así, las propias experiencias ni se comparten con los compañeros ni con la familia, esto último por razones de clasificación. Al igual que los retornados de misiones más largas en el extranjero, los pilotos de UAV pueden tener problemas para diferenciar el mundo civil del militar, lo que puede conducir, por ejemplo, a una alienación de los problemas familiares «no significativos» que pueden ser ignorados durante una misión regular en el extranje-

¹⁰⁵ CHAPA, Joe O., «Remotely Piloted Aircraft, Risk, and Killing as Sacrifice: The Cost of Remote Warfare», *Journal of Military Ethics*, 16(3-4), 2017, p. 262; BENTLEY, Michelle, «Fetishised data: counterterrorism, drone warfare and pilot testimony», *Critical Studies on Terrorism*, 11(1), 2018, p. 96; LEE, Peter, «The Distance Paradox: Reaper, the Human Dimension of Remote Warfare, and Future Challenges for the RAF», *Air Power Review*, 21(3), 2018, pp. 116-117.

¹⁰⁶ KIRKPATRICK, *op. cit.*, p. 209; CHAPA, *op. cit.*, pp. 262, 264.

¹⁰⁷ SINGER, *op. cit.* (2009), p. 346; KRISHNAN, *op. cit.* (2010), pp. 55-56; SINGER, *op. cit.* (2012), p. 29; KIRKPATRICK, *op. cit.*, pp. 208-213; CHAPA, *op. cit.*, pp. 262, 264; BENTLEY, *op. cit.*, pp. 96-97; LEE, *op. cit.*, pp. 113-114, 117-119, 123-125.

¹⁰⁸ La reunión tradicional en los escuadrones de caza por la noche, después de volver de las misiones.

ro o que tienen que ser resueltos por la pareja. Al mismo tiempo, la diferencia horaria entre el país de despliegue y el país de estacionamiento a veces impide una vida familiar normal a pesar de la presencia física en el propio país¹⁰⁹. Finalmente, la opinión pública, o incluso la opinión de conocidos y familiares, de que una tripulación de UAV es una máquina asesina sin corazón, o la crítica, al menos subliminal, de los compañeros de no ser un soldado «real» o de ganar dinero con los juegos de ordenador, puede llevar a cargas y riesgos psicológicos adicionales¹¹⁰.

Reducción de la contención de la violencia

La inhibición natural humana de la matanza y su capacidad de empatía también para el enemigo son dos de los factores más importantes que limitan la crueldad en la guerra¹¹¹. Las estadísticas muestran que, durante la Segunda Guerra Mundial y la Guerra de Corea, los soldados estadounidenses no mataron a sus oponentes en numerosas ocasiones, aunque tuvieron la oportunidad de hacerlo¹¹². Al mismo tiempo, la evidencia empírica muestra que la empatía y la compasión disminuyen al aumentar la distancia local y mental del enemigo, esta última, por ejemplo, a través de una designación deshumanizadora del enemigo¹¹³. En la opinión de algunos científicos, la asimetría de riesgo mencionada en el epígrafe anterior elimina el doble papel de perpetrador y víctima en el campo de batalla y reduce aún más la inhibición de la violencia¹¹⁴. Esto mismo se aplica a la alienación de las propias acciones debido a un arma que no está físicamente presente o

¹⁰⁹ SINGER, *op. cit.* (2009), pp. 346-347; CANNING, John y DABRINGER, Gerhard, «Ethical Challenges of Unmanned Systems», en: DABRINGER, Gerhard (ed.), *Ethical and Legal Aspects of Unmanned Systems. Interviews*, Wien: BMVLS 2010, p. 14; SHARKEY, Noel, «Saying “No!” to Lethal Autonomous Targeting», *Journal of Military Ethics*, 9(4), 2010b, p. 372; SINGER, Peter W., «The Future of War», en: DABRINGER, Gerhard (ed.), *Ethical and Legal Aspects of Unmanned Systems. Interviews*, BMVLS, Wien, 2010a, p. 81; BENTLEY, *op. cit.*, p. 95.

¹¹⁰ BENTLEY, *op. cit.*, pp. 95, 97.

¹¹¹ KRISHNAN, *op. cit.* (2009), p. 130; MORKEVICIUS, *op. cit.*, pp. 14-15.

¹¹² ARKIN, *op. cit.* (2010a), p. 337.

¹¹³ SINGER, *op. cit.* (2009), pp. 395-396; SCHMIDT-RADEFELDT, Roman y MEISSLER, Christine, «Einführung», en: SCHMIDT-RADEFELDT, Roman y MEISSLER, Christine (eds.), *Automatisierung und Digitalisierung des Krieges. Drohnenkrieg und Cyberwar als Herausforderungen für Ethik, Völkerrecht und Sicherheitspolitik*, Baden-Baden: Nomos Verlagsgesellschaft 2012, p. 14.

¹¹⁴ SCHMIDT-RADEFELDT y MEISSLER, *op. cit.*, p. 11.

incluso autocontrolada: no mata el «yo», sino la máquina¹¹⁵. Matar por control remoto hace que el acto de matar y la guerra se conviertan en *business-as-usual*, y formen parte de la vida normal¹¹⁶.

La visión del adversario a través de las pantallas incluso podría llevar a que la realidad aparezca como ficción y a que se tenga la impresión de que nadie es realmente asesinado, por lo que no surge ningún arrepentimiento¹¹⁷. En este contexto, a veces se critica que la interfaz de usuario de los drones se asemeja a una consola de videojuegos¹¹⁸. De hecho, la familiaridad de muchos jóvenes con los juegos de ordenador los convierte en los operadores de UAV ideales a este respecto¹¹⁹. Sin embargo, surge la pregunta de por qué los militares no deberían aprovechar esta familiaridad y gastar dinero en el desarrollo de nuevas interfaces que las empresas de videojuegos ya han desarrollado de forma elaborada¹²⁰. A este respecto, no se ha encontrado literatura relevante sobre cómo deberían ser estas interfaces para ser más «morales».

Esto no quiere decir que, debido a la diferencia externa relativamente pequeña entre las consolas de videojuegos y, por lo tanto, la afición de no pocos jóvenes operadores de UAV, no haya peligro de perder la conciencia de que las acciones militares tienen consecuencias reales¹²¹. Según Sullins, este efecto secundario podría, sin embargo, contrarrestarse permitiendo que solo los oficiales, al menos el personal con formación ética, sirvan como operadores de UAV¹²². En cualquier caso, la tesis de una inhibición particularmente baja de la violencia por parte de los operadores de UAV es cuestionable. Hoy en día, ni los ballesteros de la Edad Media ni los operadores de misiles de crucero, ni siquiera de los intercontinentales, son considerados particularmente despiadados, aunque no podían o pueden ver al enemigo, en contraste con un operador de drones¹²³. Un

¹¹⁵ KRISHNAN, *op. cit.* (2009), p. 128.

¹¹⁶ SULLINS, John P., «Aspects of Telerobotic Systems», en: DABRINGER, Gerhard (ed.), *Ethical and Legal Aspects of Unmanned Systems. Interviews*, Wien: BMVLS 2010, pp. 165-166.

¹¹⁷ ARKIN, *op. cit.* (2010a), p. 336; SCHMIDT-RADEFELDT y MEISSLER, *op. cit.*, p. 14.

¹¹⁸ WEBER, *op. cit.*, p. 43.

¹¹⁹ SINGER, *op. cit.* (2009), pp. 360-367.

¹²⁰ SINGER, *op. cit.* (2009), p. 68.

¹²¹ SINGER, *op. cit.* (2009), p. 367.

¹²² SULLINS, *op. cit.*, p. 163.

¹²³ SPARROW, *op. cit.* (2010), pp. 90-91; STRAWSER, *op. cit.* (2010), p. 365, n. 32.

ejemplo citado con frecuencia en este contexto es la prohibición del papa Inocencio II del uso de ballestas contra los cristianos en 1139, porque, como dice Lin, esta arma, desde el punto de vista del papa, «made warfare into a depersonalized and dishonorable “point-and-click” affair»¹²⁴.

Además, la llamada «mentalidad del joystick» aún no ha sido probada. Más bien, los pilotos de drones parecen tener una proximidad mental especial a sus objetivos debido a la observación intensiva de los mismos, que podría ser corresponsable de los casos de trastornos psicopatológicos¹²⁵. Mientras que los artilleros o los pilotos de caza solo pueden percibir a sus enemigos (si es que los perciben) durante un corto periodo de tiempo, la permanencia de los UAV por encima del objetivo y sus cámaras de alta resolución permiten observar las secuelas del ataque. Antes, algunos operadores de UAV ven la vida diaria y familiar del objetivo. Por lo tanto, la deshumanización de la víctima y el distanciamiento psicológico asociado se dificultan¹²⁶. La distancia física entre la víctima y el soldado puede haber llegado a su máximo teórico mediante el uso de drones, pero la distancia emocional ya está disminuyendo y se reducirá aún más en el futuro debido a la mejora de la tecnología de los sensores¹²⁷.

Erosión de la dignidad humana

En contraste con la discusión anglosajona, la idea de que el uso de los LAWS violaría al menos potencialmente la dignidad humana, juega un papel prominente en el panorama de la investigación alemana. En este sentido, se argumenta que los robots no tienen una concepción de la mortalidad y la moralidad, por lo tanto, no entienden lo que significa matar a un ser humano, y ni siquiera son conscientes de que están llevando a cabo un acto de matar. Sin esta capacidad de reflexión y esta conciencia, los robots no estarían cualificados para acabar con una vida. Al contrario, esta decisión es automatizada y «procesada», tratando

¹²⁴ LIN, *op. cit.*, pp. 320-321, cita en p. 321. Véase también, por ejemplo, KRISHNAN, *op. cit.* (2009), p. 95.

¹²⁵ DAHLMANN y DICKOW, *op. cit.*, p. 14 n. 17.

¹²⁶ Sobre esto, véanse KRISHNAN, *op. cit.* (2009), p. 129; KIRKPATRICK, *op. cit.*, pp. 211-212; CHAPA, *op. cit.*, p. 263; BENTLEY, *op. cit.*, pp. 94-95, 102; LEE, *op. cit.*, pp. 115-116.

¹²⁷ CHAPA, *op. cit.*, pp. 265-267. Como ejemplo de este problema por los operadores ingleses del UAV MQ-9 Reaper, véase en detalle LEE, *op. cit.*

a las víctimas como si se tratara de meros puntos de datos, es decir, de meros objetos o medios para alcanzar un fin. Asesinar a una persona de esta manera ignora o niega su característica de sujeto moral. Además, la víctima no tiene la posibilidad de apelar a la humanidad de su oponente o de pedir misericordia y, por lo tanto, no tiene ni siquiera una pequeña oportunidad de influir en su propio destino. Estos problemas son independientes del desarrollo tecnológico de los LAWS¹²⁸. Un argumento similar puede hacerse para los UAV, ya que ayudan a matar sin necesidad de conocer el entorno, idioma y la cultura de la víctima¹²⁹.

Para la sociedad alemana, la muerte con una máquina despierta un recuerdo incómodo del Holocausto y del asesinato estatal e industrial en masa que lo acompañó. De hecho, el párrafo 1 del artículo 1 de la Constitución alemana, principalmente por la experiencia del genocidio durante la Segunda Guerra Mundial, postula la inviolabilidad de la dignidad humana, estando esto en línea con el concepto de dignidad de Kant¹³⁰. Para Sauer, uno de los críticos alemanes más activos de los LAWS, la intervención en la dignidad humana es un mal en sí mismo¹³¹. Como consecuencia, Sauer exige que Alemania y sus aliados de la UE y de la OTAN respeten y promuevan la dignidad humana a nivel internacional¹³².

Un artículo de la Fundación Heinrich Böll alemana afirma: «Las valoraciones legales y políticas de los AWS pueden diferir, pero la Constitución alemana proporciona el recordatorio ético *axiomático* de que la dignidad de todos los seres humanos, incluidos aquellos contra los que se dirige legítimamente la violencia militar, debe mantenerse intacta. La subcontratación de la selección y el compromiso de blancos a algoritmos en máquinas militares está fuera de discusión *para una sociedad que acepte este imperativo*»¹³³. Más adelante, se rechaza explícitamente contraponer

¹²⁸ KRISHNAN, *op. cit.* (2009), p. 133; SCHÖRNIG, *op. cit.* (2012), pp. 34, 55-56; JOHNSON, Aaron M. y AXINN, Sidney, «The Morality of Autonomous Robots», *Journal of Military Ethics*, 12(2), 2013, p. 134; ROFF, *op. cit.*, p. 214; Heinrich-Böll-Stiftung, *op. cit.*, pp. 32-33; DAHLMANN y DICKOW, *op. cit.*, p. 18.

¹²⁹ WEBER, *op. cit.*, p. 43.

¹³⁰ KANT, Immanuel, «Die Metaphysik der Sitten», en: Königlich Preußische Akademie der Wissenschaften (ed.), *Kant's gesammelte Schriften*. Band IV. Erste Abtheilung: Werke. Vierter Band, Berlin: Georg Reimer 1903, pp. 427-428.

¹³¹ SAUER, Frank, «Stopping "Killer Robots": Why Now Is the Time to Ban Autonomous Weapon Systems», *Arms Control Association*, 2016, p. 3, <https://www.armscontrol.org/printpdf/7713> [Fecha de la consulta: 20.11.2019].

¹³² SAUER, *op. cit.* (2019a), p. 29.

¹³³ Heinrich-Böll-Stiftung, *op. cit.*, p. 17. Traducción e itálica del autor.

las violaciones de la dignidad humana y las posibles consecuencias positivas de su violación¹³⁴. En este sentido, la discusión se esconde detrás de un consenso básico kantiano que evita la ponderación laboriosa de las ventajas y desventajas de la robótica. No obstante, el párrafo 1 del artículo 1 de la Constitución alemana, en función de la correspondiente interpretación del Tribunal Constitucional alemán, apoya como un argumento válido y posiblemente definitivo la no utilización de los LAWS por parte de Alemania.

Por otro lado, si se asume como correcto que el concepto de la dignidad humana es un concepto casi universalmente aceptado, un rechazo obligatorio de los LAWS por casi todos los países podría sostenerse en la cláusula Martens de la Orden de La Haya sobre la guerra terrestre de 1899/1907. Esta cláusula estipula que, en ausencia de normas específicas, la humanidad y la conciencia pública (en otras palabras, las convicciones morales de la gente corriente) deben determinar cómo tratar a los combatientes y a los civiles en la guerra¹³⁵. En este contexto, cabe señalar que, según un estudio de la Campaña para detener a los robots asesinos y Human Rights Watch, en 2017 y 2018 alrededor del 60% de los encuestados de 23 y 26 Estados, respectivamente, repudiaron el uso de armas totalmente autónomas¹³⁶.

Sin embargo, la aplicación del artículo 1 de la Constitución en caso de guerra es cuestionable a la luz de las normas pertinentes de los Convenios de Ginebra, que han sido ratificados también por Alemania. Estas normas, que pueden considerarse reflejo de las convicciones éticas de la comunidad internacional, permiten claramente sopesar la vida humana frente a una ventaja militar¹³⁷. Al mismo tiempo, la referencia a la Constitución alemana

¹³⁴ Heinrich-Böll-Stiftung, *op. cit.*, p. 33.

¹³⁵ Heinrich-Böll-Stiftung, *op. cit.*, pp. 33-34.

¹³⁶ SAUER, *op. cit.* (2019a), p. 28. Como señala KRISHNAN, *op. cit.* (2009), p. 114, una estrategia militar podría ser, por lo tanto, introducir los LAWS solo después de que la opinión pública se haya acostumbrado suficientemente a los robots en otras funciones bélicas menos controvertidas.

¹³⁷ El artículo 51, apartado 5, letra *b*; el artículo 57, apartado 2, letra *a*, inciso iii; y el artículo 57, apartado 2, letra *b* del Protocolo I adicional de los Convenios de Ginebra (PA I) exigen que se abstenga de realizar o que se interrumpan (*solamente*, comentario del autor) los ataques que previsiblemente causen pérdidas o daños a la población civil, «que serían excesivos en relación con la ventaja militar concreta y directa prevista». En cuanto a la discusión alemana del problema, hay que destacar dos puntos. En primer lugar, las versiones españolas, francesas e inglesas de esos artículos, que son auténticas en virtud del artículo 102 del PA I en contraste con la versión

probablemente no impedirá que otros Estados, como EE. UU., cuya Constitución no hace referencia directa a la dignidad humana, utilicen los LAWS.

En el contexto internacional, la referencia no especificada al riesgo de «abandonar los valores más básicos de la civilización y los principios humanitarios»¹³⁸, posiblemente bajo el supuesto de una comprensión moral deontológica compartida entre autor y lector, parece insuficiente. Sería más útil y necesaria una derivación lógica o empírica, al menos no axiomática, de por qué el hombre tiene dignidad, por qué es inviolable e incalculable, o por qué los postulados axiomáticos son preferibles a los postulados en sentido contrario.

Aparte de las dificultades paradigmáticas de justificación que se han planteado, cabe preguntarse hasta qué punto es prácticamente relevante para una persona asesinada que el responsable

alemana, expresan una mayor permisividad. Esta última habla de ataques que son «desproporcionados con respecto a la ventaja militar concreta e inmediata esperada» (Traducción del autor. Véase Bundesanzeiger Verlag, «Bundesgesetzblatt, Jahrgang 1990, Teil II, S. 1550-1649. Gesetz zu den Zusatzprotokollen I und II zu den Genfer Rotkreuz-Abkommen von 1949», 1990, 1589, http://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl290s1550.pdf [Fecha de la consulta: 20.11.2019]). Ambas formulaciones se encuentran en el discurso alemán. Para «desproporcionados» = «nicht im Verhältnis», véase, por ejemplo, WEBER, *op. cit.*, p. 34; similar ALTMANN, Jürgen, «Der Kriegsmaschine Grenzen setzen. Rüstungsbegrenzung für bewaffnete unbemannte Fahrzeuge», en: MARSISKE, Hans-Arthur (ed.), *Kriegsmaschinen. Roboter im Militäreinsatz*, Hannover: Heise 2012, p. 224; ARENDT, Rieke, «Der Einsatz autonomer Waffensysteme im Lichte des Verhältnismäßigkeits- und des Unterscheidungsgrundsatzes», en: FRAU, Robert (ed.), *Drohnen und das Recht. Völker- und verfassungsrechtliche Fragen automatisierter und autonomer Kriegführung*, Tübingen: Mohr Siebeck 2014, p. 30. Para la variante más común internacionalmente, véanse PLATEK, Olivia, «Autonome Kriegführung und legitime militärische Ziele», en: FRAU, Robert (ed.), *Drohnen und das Recht. Völker- und verfassungsrechtliche Fragen automatisierter und autonomer Kriegführung*, Tübingen: Mohr Siebeck 2014, p. 47; STROH, Philipp, «Das Menschenrecht auf Leben im ,entmenschlichten bewaffneten Konflikt», en: FRAU, Robert (ed.), *Drohnen und das Recht. Völker- und verfassungsrechtliche Fragen automatisierter und autonomer Kriegführung*, Tübingen: Mohr Siebeck 2014, p. 157. En segundo lugar, en Alemania, a falta de un fallo del Tribunal Supremo, no está claro si el asesinato de civiles por las Fuerzas Armadas en virtud del artículo 51, apartado 5, letra b del PA I es en absoluto compatible con la Constitución, independientemente de que se lleve a cabo con los UAV, sistemas autónomos o de otra manera. Véase BRUNNER, Manuel, «Grundrechtsfragen beim Einsatz von Kampfdrohnen durch die Bundeswehr im Ausland», en: FRAU, Robert (ed.), *Drohnen und das Recht. Völker- und verfassungsrechtliche Fragen automatisierter und autonomer Kriegführung*, Tübingen: Mohr Siebeck 2014, pp. 163-176.

¹³⁸ SAUER, *op. cit.* (2019a), p. 28.

sea un ser humano o una máquina¹³⁹. También parece absurdo que un hipotético combatiente enemigo, cuya finalidad —en el peor de los casos— consiste en acabar con la Constitución alemana, se «queja» de la violación de su dignidad humana protegida por esta misma Constitución al ser eliminado por un LAWS alemán. Además, se podría argumentar que en el caso de los LAWS con IA *top-down* el responsable de la muerte realmente fue un ser humano, o sea, el programador. Cuando la víctima decide no rendirse a un robot programado con las normas del DIH, incluso es realmente la víctima quien determina su propia muerte. El problema es más relevante con una IA *bottom-up* cuyos criterios para dar muerte al enemigo no sean completamente predecibles. Sin embargo, también podría alegarse en este caso que la decisión humana de desarrollar un robot de este tipo o de utilizarlo en un determinado contexto es la que, en última instancia, causa la muerte del objetivo¹⁴⁰. Como se explica en el siguiente epígrafe, el conjunto de problemas descrito podría desaparecer eventualmente por la creación de una IA «fuerte», consciente de sí misma. No parece, por tanto, muy adecuado que el Gobierno y las Fuerzas Armadas alemanas hayan decidido desarrollar exclusivamente la IA «débil»¹⁴¹.

Posibilidad e imposibilidad de la ética robótica

«It is not my belief that an autonomous unmanned system will be able to be perfectly ethical in the battlefield, but I am convinced that they can perform more ethically than human soldiers are capable of performing»¹⁴².

Los problemas enumerados anteriormente podrían ser resueltos, al menos parcialmente, por la programación de un concepto de moralidad en los LAWS. El representante más destacado de este enfoque es Ronald C. Arkin, quien quiere equipar a los robots no solo con subrutinas éticas y del DIH, sino también con remordimientos artificiales y sentimientos de culpa¹⁴³. El punto de partida de Arkin fue una investigación de las Fuerzas Armadas estadou-

¹³⁹ SAUER, *op. cit.* (2019a), p. 28. Sauer, al contrario del autor del presente trabajo, sí ve una diferencia significativa.

¹⁴⁰ ROFF, *op. cit.*, p. 214. Roff minimiza la validez de este último argumento.

¹⁴¹ Amt für Heeresentwicklung, *op. cit.*, pp. 6, 28.

¹⁴² ARKIN, *op. cit.* (2010a), p. 334.

¹⁴³ KRISHNAN, *op. cit.* (2009), pp. 108-109; ARKIN, *op. cit.* (2010a); ARENDT, *op. cit.*, p. 33.

nidenses en 2006, de la que se desprende una tendencia relativamente alta de los soldados a aprobar o al menos no denunciar el trato de los no combatientes contrario al DIH. En el caso de la utilización de los guerreros robóticos desprovistos de las emociones humanas, tales como el deseo de matar sin control o de poder, la venganza, el genocidio u otras similares, se evitaría la comisión de acciones contrarias al DIH, que sí cometen los seres humanos. Además, los daños accidentales a civiles inocentes también podrían llegar a ser mucho menos probables, ya que los robots pueden utilizar sensores que perciban la realidad con mayor precisión y objetividad que los humanos. Además, las vidas de civiles podrían ser programadas como más valiosas que la existencia de un LAWS, explotando así el hecho de que los robots, a diferencia de los humanos, pueden permitirse disparar después¹⁴⁴.

Aunque se puede asumir que los robots que disponen de las rutinas mencionadas son mejores que los robots que no las tienen¹⁴⁵, el enfoque de Arkin ha sido objeto de críticas extensivas. Krishnan considera que, si nadie exige la invención de minas «éticas», por qué debe exigírselo para los LAWS, ya que no parece que haya muchas diferencias entre morir por una mina o por un LAWS¹⁴⁶. En esta línea argumental, cabe destacar que a pesar de que las minas producen muchos más muertos que los LAWS, aquellas son percibidas por la gente corriente como mucho menos temibles que los «robots asesinos». Es posible que el miedo primario del hombre a los depredadores animales juegue un papel en esta falacia¹⁴⁷. Por tanto, la necesidad de la contención de los LAWS debería derivarse principalmente de su mayor peligro en comparación con las minas, al menos a largo plazo. Así, la referencia antes mencionada a las consecuencias indiferentes pierde su vigor. En este contexto, cabe destacar que los temores y peligros que emanan de la IA y la robótica —entendido «peligro» como la probabilidad de que en el futuro se produzca un sufrimiento colectivo o una reducción de la felicidad colectiva, y «temor» como el miedo de ese sufrimiento— son determinantes morales significativos en el paradigma utilitarista¹⁴⁸, pero no necesariamente en otros paradigmas éticos.

¹⁴⁴ SINGER, *op. cit.* (2009), pp. 394-395, 398; KRISHNAN, *op. cit.* (2009), p. 126; ARKIN, *op. cit.* (2010a), pp. 333-336, 338.

¹⁴⁵ KRISHNAN, *op. cit.* (2009), pp. 108-109.

¹⁴⁶ KRISHNAN, *op. cit.* (2009), pp. 110, 133-134.

¹⁴⁷ KRISHNAN, *op. cit.* (2009), p. 33.

¹⁴⁸ BENTHAM, *op. cit.*, pp. 143-144. STRAWSER, Bradley J., «Arguing in Good Faith about Drones», en: STRAWSER, Bradley J. (ed.), *Opposing Perspectives on the Drone*

No obstante, tampoco parece que eliminar las emociones humanas negativas, como el odio y la venganza, sea exclusivamente beneficioso, pues también se eliminan los sentimientos y virtudes positivas como la compasión, la empatía, la misericordia y la caballerosidad¹⁴⁹. Adicionalmente, los autores consideran que, en el mejor de los casos, los robots serían capaces de cumplir con las reglas morales, pero no de ejercer la supererogación, la excelencia moral o las hazañas heroicas¹⁵⁰. Sin embargo, estas objeciones pueden ser contrarrestadas por el hecho de que Arkin también señalara la necesidad de programar emociones positivas (si esto es técnicamente posible, pues aún queda por ver). La referencia a la imposibilidad de la supererogación se puede enfrentar con el argumento ya desarrollado de que la guerra no debe ser un foro para exhibir virtudes personales, y no debe ser más peligrosa de lo necesario, solamente para que esa exhibición sea factible¹⁵¹.

Un argumento más importante es la falta de claridad en cuanto a cómo se pueden probar las rutinas éticas programadas, en particular qué método se puede utilizar para determinar si un robot actúa realmente de forma más ética que los humanos, como Arkin espera¹⁵². También se plantea la cuestión de qué principios éticos deberían programarse, especialmente porque difieren de una nación y de una cultura a otra¹⁵³.

Por tanto, si se pretende unificar las normas éticas en el marco del derecho internacional, la cuestión sería qué órgano podría decidir al respecto en nombre de todos los Estados sin quedar paralizado políticamente. Después de todo, los países que imponen bajas exigencias morales sobre sus propios actos de guerra tendrían una ventaja militar insuperable en ciertos casos, como

Debate, New York: Palgrave Macmillan 2014, pp. 180-181, también considera el resultado de la discusión sobre los UAV como una evaluación empírica de sus beneficios y peligros potenciales. Para breves consideraciones sobre el uso ventajoso del «utilitarismo de las normas» según R. M. Hare para la interpretación de los términos «justo», «excesivo» y «necesario» en el contexto del DIH, véase JOHANSSON, Linda, «Ethical Aspects of Military Maritime and Aerial Autonomous Systems», *Journal of Military Ethics*, 17(2-3), 2018, pp. 149-151.

¹⁴⁹ KRISHNAN, *op. cit.* (2009), pp. 132-133; SCHÖRNIG, *op. cit.* (2012), p. 34; STROH, *op. cit.*, p. 160; Heinrich-Böll-Stiftung, *op. cit.*, p. 35.

¹⁵⁰ TONKENS, Ryan, «The Case against Robotic Warfare: A Response to Arkin», *Journal of Military Ethics*, 11(2), 2012, p. 151.

¹⁵¹ SCHULZKE, *op. cit.*, p. 196.

¹⁵² KRISHNAN, *op. cit.* (2009), pp. 110-111.

¹⁵³ CANNING y DABRINGER, *op. cit.*, p. 17; FINN y SCHEDING, *op. cit.*, pp. 199-200.

cuando dos robots técnicamente perfectos compiten entre sí. Los oponentes que ya son técnicamente inferiores no tendrían ningún incentivo para fortalecer aún más su inferioridad a través de la definición contractual de rutinas éticas complejas, o incluso su activación en combate¹⁵⁴.

Siguiendo con la argumentación, si existiera una claridad total sobre la programación ética y de derecho internacional de los robots, el oponente podría eludirla o explotarla deliberadamente. Esto se refiere, por ejemplo, al uso de niños soldados o de escudos humanos, frente a los cuales un robot programado *top-down* posiblemente detendría su ataque y podría ser combatido fácilmente. Por otra parte, un IA *bottom-up* podría reaccionar a esos abusos con el cambio de sus rutinas éticas: en futuras ocasiones similares también dispararía a niños o no combatientes¹⁵⁵. Es decir, si el enemigo obliga a los niños a inutilizar las cámaras de los robots de combate con pintura en aerosol, los robots adaptativos podrían reaccionar matando a los niños que lleven botellas de aerosol por defecto¹⁵⁶.

Por otro lado, desde el punto de vista de los países con los LAWS técnicamente muy desarrollados, los robots con restricciones éticas podrían no ser lo suficientemente fuertes para utilizarse en batallas simétricas de alta intensidad. Por tanto, ¿pueden los militares desear que un sistema aborte su misión o consulte a sus superiores humanos porque tiene la más mínima duda sobre la rectitud ética o de derecho internacional de sus acciones, o que rechaza una orden porque evalúa la situación de forma diferente a su operador? En este caso, el sistema sería de poca utilidad, tal vez de utilidad inferior a un soldado humano. Si, por otro lado, sus rutinas se aplican con demasiada laxitud, el sistema sería poco ético o incluso ilegal¹⁵⁷.

Si los humanos pueden anular la decisión moral del robot autónomo, entonces permanecerían los problemas fundamentales, es decir, la influencia de las emociones en decisiones morales y la falibilidad humana. Los robots podrían incluso ser utilizados específicamente

¹⁵⁴ BURKHARD, Hans-Dieter, «Lasst die Maschinen machen. Wie intelligent ist die Künstliche Intelligenz?», en: MARSISKE, Hans-Arthur (ed.), *Kriegsmaschinen. Roboter im Militäreinsatz*, Hannover: Heise 2012, p. 161; SCHÖRNIG, *op. cit.* (2013), pp. 232-233.

¹⁵⁵ KRISHNAN, *op. cit.* (2009), p. 110.

¹⁵⁶ SINGER, *op. cit.* (2009), p. 201. Por supuesto, el uso de tales tácticas por parte del enemigo sería particularmente censurable y legitimaría moralmente aún más la lucha contra él.

¹⁵⁷ KRISHNAN, *op. cit.* (2009), pp. 58-59, 109-110; BURKHARD, *op. cit.*, pp. 158-159.

para acciones inmorales¹⁵⁸. Los crímenes de guerra intencionados, ordenados por el mando superior, serían aún más fáciles de llevar a cabo, ya que los posibles «remordimientos» robóticos y las tendencias a negarse a obedecer órdenes, si es que existen, se pueden simplemente anular¹⁵⁹. Si los LAWS son programados *bottom-up* para aprender la acción ética de sus compañeros humanos, no se puede garantizar que no imiten sus acciones inmorales¹⁶⁰.

Otro argumento en contra de la propuesta de Arkin es que la promesa de una programación ética podría reducir la oposición política a los LAWS y conducir a su introducción sin haber adoptado realmente rutinas funcionales¹⁶¹. También se señala que los robots pequeños (especialmente los nanobots) no pueden tener una rutina moral eficaz porque les falta la capacidad de llevar los sensores necesarios. Finalmente, se advierte que, para tomar decisiones morales, no solo son relevantes las condiciones ambientales sensoriales, sino también la conciencia del contexto de la situación respectiva, de la que carecen los robots¹⁶².

Cabe destacar que Arkin es consciente de las dificultades técnicas de su proyecto (aplicación del derecho internacional en los algoritmos, problema de discriminación entre objetivos legales e ilegales, etc.)¹⁶³. Por eso, en la actualidad, Arkin limita el uso de los LAWS a la guerra abiertamente declarada y a las situaciones en las que no haya un número significativo de civiles¹⁶⁴.

Por el momento, podría ser más aconsejable concebir sistemas de asistencia ética para apoyar la toma de decisiones humanas, porque «humans are already very sophisticated ethical information processing systems»¹⁶⁵. En lugar de intentar que las armas autónomas se comporten éticamente, se podrían inten-

¹⁵⁸ TONKENS, *op. cit.*, pp. 157-160.

¹⁵⁹ SINGER, *op. cit.* (2009), p. 396.

¹⁶⁰ ROFF y DANKS, *op. cit.*, pp. 9-10.

¹⁶¹ KRISHNAN, *op. cit.* (2009), p. 109; SHARKEY, *op. cit.* (2010a), p. 49; ALTMANN, Jürgen, «Bewaffnete unbemannte Fahrzeuge – Beschränkungen dringend nötig», en: Heinrich-Böll-Stiftung (ed.), High-Tech-Kriege. Frieden und Sicherheit in den Zeiten von Drohnen, Kampfrobotern und digitaler Kriegsführung, Berlin: Heinrich-Böll-Stiftung 2013, p. 58.

¹⁶² KRISHNAN, *op. cit.* (2009), pp. 109-110.

¹⁶³ ARKIN, *op. cit.* (2010a), p. 339.

¹⁶⁴ ARKIN, Ronald C., «Governing Lethal Behaviour», en: DABRINGER, Gerhard (ed.), Ethical and Legal Aspects of Unmanned Systems. Interviews, Wien: BMVLS 2010b, p. 155.

¹⁶⁵ ASARO, Peter M., «Military Robots and Just War Theory», en: DABRINGER, Gerhard (ed.), Ethical and Legal Aspects of Unmanned Systems. Interviews, Wien: BMVLS 2010, pp. 108-110, cita en p. 108.

sificar los esfuerzos para fortalecer la moralidad de los soldados humanos¹⁶⁶. Entre las medidas posibles figuran la reducción de la duración de misiones en el extranjero, la mejora de la disponibilidad del tratamiento psicológico, un examen psicológico más exhaustivo de los soldados, una formación ética y del DIH más amplia, reglas de enfrentamiento más claras, sanciones más severas para las desviaciones de comportamiento, y una mejor supervisión. Puesto que lo normal es que los soldados no violen de hecho el DIH, la introducción de robots de combate autónomos, incluido su posible uso para violar (muy eficazmente) el DIH, podría generar en este momento más problemas morales de los que podría resolver¹⁶⁷. Además, es probable que existan áreas de la sociedad en las que los recursos para la creación de un robot de combate autónomo y moral puedan invertirse de forma más rentable¹⁶⁸.

Por otra parte, se argumenta que la inmoralidad de las acciones no debe basarse en sus resultados, sino en la cuestión de si una determinada acción moralmente incorrecta en sí misma se lleva a cabo de manera intencionada o al menos negligente a pesar de que estén al alcance las informaciones necesarias para tomar una decisión moralmente correcta. En este caso, la cantidad de información sensorial disponible, que presumiblemente es superior por los robots, en sí misma no es relevante¹⁶⁹. Además, se afirma que si bien los robots parecen *actuar* moralmente cuando siguen su programación, independientemente de su calidad no pueden *ser* morales, ya que no tienen una capacidad de toma de decisiones basada en valores, ni una capacidad autónoma de reflexión, ni conciencia, ni intuición moral que les permita poner en duda las reglas morales existentes¹⁷⁰. Del mismo modo, un robot no puede

¹⁶⁶ SHARKEY, *op. cit.* (2010a), p. 49; TONKENS, *op. cit.*, p. 151.

¹⁶⁷ TONKENS, *op. cit.*, pp. 160-162.

¹⁶⁸ TONKENS, *op. cit.*, pp. 166-167 n. 23.

¹⁶⁹ TONKENS, *op. cit.*, pp. 155-156.

¹⁷⁰ SINGER, *op. cit.* (2009), p. 425; KRISHNAN, *op. cit.* (2009), p. 132; CAPURRO, Rafael, «Wer ist der Mensch? Überlegungen zu einer vergleichenden Theorie der Agenten», en: MARSISKE, Hans-Arthur (ed.), *Kriegsmaschinen. Roboter im Militäreinsatz*, Hannover: Heise 2012, pp. 232-233; CAPURRO, Rafael y MARSISKE, Hans-Arthur, «Der Moment des Triumphs. E-Mail-Dialog über ein Bild», en: MARSISKE, Hans-Arthur (ed.), *op. cit.*, pp. 24-26; SCHMIDT-RADEFELDT y MEISSLER, *op. cit.*, p. 16; JOHNSON y AXINN, *op. cit.*, p. 135; PLATEK, *op. cit.*, p. 42. Una referencia de Johnson y Axinn a Kant muestra claramente que los autores están cerca del paradigma deontológico e intuicionista.

realmente *sentir* emociones, sino solo *imitarlas* para el observador externo¹⁷¹. Asimismo, se duda si la acción ética humana, como la acción robótica, puede reducirse en principio a relaciones causa-efecto¹⁷². Sparrow escribe: «It simply isn't possible to "algorithmatise" [ethics] —or at least no philosopher in human history has been able to come up with a formula that will determine what is ethical»¹⁷³.

Sin embargo, podemos considerar que esto sí es posible, ya que el utilitarista Jeremy Bentham, al menos en su propia opinión, ya encontró una fórmula ética alrededor de 1780¹⁷⁴. El estado actual de la investigación muestra que las personas basan sus decisiones morales principalmente en las emociones, las cuales son guiadas hacia ciertos caminos por normas religiosas, culturales y tradicionales. El trabajo sobre una moral utilitarista «artificial», que nos obliga a pensar con más autocrítica y racionalmente sobre nuestra moral existente, podría mejorar nuestras decisiones¹⁷⁵.

Esto es tan cierto como que la moral se desarrolló en un momento en que el hombre se enfrentaba a cuestiones éticas completamente diferentes (como la distribución justa del botín de caza en un clan)¹⁷⁶. Una moral «robótica», es decir, programada y objetiva, podría convertirse en una metamoralidad contra la cual se evalúa la moralidad humana convencional¹⁷⁷.

Al contrario, hay enfoques que rechazan expresamente la regulación excesiva o exclusiva de la moralidad y se refieren a la intuición y a la emoción humana como mejores criterios morales¹⁷⁸. Relacionado está el argumento de que los humanos podrían al menos recurrir a su intuición en ausencia de información clara, lo

¹⁷¹ VERUGGIO, Gianmarco y OPERTO, Fiorella, «Ethical and societal guidelines for Robotics», en: DABRINGER, Gerhard (ed.), *Ethical and Legal Aspects of Unmanned Systems. Interviews*, Wien: BMVLS 2010, p. 135.

¹⁷² DAHLMANN y DICKOW, *op. cit.*, p. 12.

¹⁷³ SPARROW, *op. cit.* (2010), p. 99.

¹⁷⁴ BENTHAM, *op. cit.*, pp. 38-40.

¹⁷⁵ SULLINS, *op. cit.*, p. 160.

¹⁷⁶ Sobre la crítica de la intuición como criterio moral decisivo en vista de su supuesta formación en tiempos humanos primitivos, que no puede ofrecer ninguna orientación para los problemas morales actuales, véanse BINMORE, Ken, «Egalitarianism versus Utilitarianism», *Utilitas*, 10(3), 1998, p. 353, y FREIMAN, Christopher, «Why Poverty Matters Most: Towards a Humanitarian Theory of Social Justice», *Utilitas*, 24(1), 2012, pp. 37-40.

¹⁷⁷ BEKEY, George, «Robots and Ethics», en: DABRINGER, Gerhard (ed.), *Ethical and Legal Aspects of Unmanned Systems. Interviews*, Wien: BMVLS 2010, p. 37.

¹⁷⁸ MORKEVICIUS, *op. cit.*, p. 16.

que los robots no podrían¹⁷⁹. Además, se podría argumentar que los juicios de valor utilitaristas, como el sopesar del valor de vidas humanas inocentes contra ventajas tácticas, también se basan al final solo en una agregación de los juicios de valor de un conjunto de personas que no han sido verificadas en cuanto a rectitud o calidad ética sobre la base de una norma externa¹⁸⁰.

En el presente trabajo no se puede dar una solución definitiva a este problema. Cabe preguntar, sin embargo, qué estándar externo debería representar el correctivo que se busca, si no es en última instancia el beneficio de un grupo determinado, la finalidad del utilitarismo clásico. Por otro lado, no está claro cómo afectaría negativamente a esta población el hecho de que los robots no *sean* morales, si al mismo tiempo *actúan* moralmente.

Un tratamiento especial merece la idea de que la moralidad está ligada a la existencia del libre albedrío. La acción robótica, que está determinada por algoritmos y de la cual el robot no es responsable, no tendría por definición ninguna cualidad moral¹⁸¹. El problema con este enfoque, sin embargo, es que los algoritmos de decisión, como el cerebro humano, se basan en fenómenos físicos describibles. En el fondo, se puede asumir que dicho cerebro es exactamente tan determinista como la operación de una computadora. Aunque la mayoría de las personas tienen la impresión subjetiva de poder determinar libremente sus acciones, no cabe duda de que su materialidad «configura sus creencias, necesidades y deseos (potenciales)»¹⁸², aunque todavía no hemos sido capaces de explicar plenamente el cómo. Además, si fuera posible reproducir artificialmente un cerebro humano a nivel submolecular, el ser artificial así creado probablemente también tendría la conciencia de sí mismo, el libre albedrío, su instinto de autopreservación, su inhibición de matar a sus semejantes, su empatía y su sentido de la moralidad. No es fácil comprender por qué un ser tal no debería ser capaz de actuar éticamente solo por su artificialidad. Incluso si fuera correcto que una copia hecha artificialmente de un ser humano no es un ser humano, sino otra especie¹⁸³, no estaría claro qué efectos prácticos tendría esto sobre su estatus ético. ¿Tratamos a un bebé probeta de forma diferente a un bebé concebido «normalmente»? ¿No sería

¹⁷⁹ SCHÖRNIG, *op. cit.* (2013), pp. 231-232.

¹⁸⁰ ASARO, *op. cit.*, pp. 116-117.

¹⁸¹ VERUGGIO y OPERTO, *op. cit.*, pp. 134, 146.

¹⁸² CAPURRO, *op. cit.*, p. 232. Traducción del autor.

¹⁸³ VERUGGIO y OPERTO, *op. cit.*, p. 134.

necesario tratar moralmente a una especie alienígena inteligente según nuestros estándares?

Además, parece cuestionable hasta qué punto los soldados adoc-trinados en un objetivo de guerra o en una política son más autónomos que los robots programados, y si es absolutamente necesario un nivel humano de autonomía, intencionalidad y responsabilidad para dar a una entidad la cualidad de actor moral¹⁸⁴. Más bien, esa condición podría entenderse como una monopolización inadmisibles del concepto de «moral». Además, a través de la digitalización y de la conexión humana en red con sus semejantes, el hombre se acerca a la esencia de un robot, que no puede permanecer sin consecuencias para su moral¹⁸⁵.

Hoy en día, en las culturas asiáticas conformadas por el sintoísmo o el budismo, como la japonesa o la coreana, los no humanos, y por lo tanto los robots, se perciben de forma más positiva y con menos desconfianza que en Occidente, entre otras cosas porque allí a los animales, las plantas o los objetos fabricados por los seres humanos se les atribuye tradicionalmente un espíritu o un alma; en este último caso, porque el espíritu del creador se transfiere de cierta manera a su producto. Por lo tanto, no es de extrañar que en Corea del Sur apenas se produzca un debate público sobre la moralidad de las armas autónomas, y que los robots, a diferencia de las culturas occidentales, tiendan a ser retratados positivamente o como héroes en la literatura¹⁸⁶. Es cierto que, además de este aspecto cultural, en los países mencionados existe una presión especial para utilizar la robótica, debido a las tasas de natalidad extremadamente bajas¹⁸⁷. A largo plazo, es probable que la globalización conduzca a una convergencia de valores, a menos que, por razones de seguridad, Occidente se vea obligado a actuar contra los robots por parte de Estados con menos reservas filosóficas.

También hay que tener en cuenta que, en la medida en que los robots se acercan a los seres humanos externamente y en términos de comportamiento, los hombres pueden, por razones emocionales, ver la necesidad de desarrollar un código ético hacia los robots o incluso concederles derechos, como ya ocurre con los animales en la actualidad. En primer lugar, el maltrato de los

¹⁸⁴ SULLINS, *op. cit.*, pp. 159-160.

¹⁸⁵ El problema se insinúa en CAPURRO, *op. cit.*, pp. 233-236.

¹⁸⁶ SINGER, *op. cit.* (2009), pp. 167-168; ARKIN, *op. cit.* (2010b), p. 151.

¹⁸⁷ SINGER, *op. cit.* (2009), pp. 242-243.

robots o su exterminio no provocado podría considerarse inmoral, ya que estos comportamientos dan evidencia de la personalidad de la persona o sobre cómo trataría a las demás. Más tarde, a los robots se les podría incluso conceder el derecho a la autodefensa contra los humanos y podría considerarse inmoral sacrificarlos por propósitos humanos¹⁸⁸. En vista de que evitar los sacrificios humanos es un argumento esencial para el uso de los AWS, no parece aconsejable diseñar a estos seres demasiado similares a los humanos; en cualquier caso, su mayor valor militar a través de una mayor inteligencia y autonomía tendría que ser compensado con su valor moral, que igualmente se incrementará¹⁸⁹.

Al final hay que señalar que los argumentos presentados contra la ética robótica pueden ser válidos en el estado actual de la técnica, pero deben ser reevaluados continuamente a medida que nos acercamos a la singularidad. No está claro si ya es necesario encontrar una respuesta moral a las implicaciones de una emergencia técnicamente distante de cualidades humanas como la conciencia de sí mismo, el libre albedrío y la dignidad, o si otros problemas éticos deberían tener prioridad por el momento¹⁹⁰. Esto mismo se aplica a la cuestión de si ya hoy necesitamos encontrar respuestas morales a los «oponentes que operan en un amplio frente con fuerzas totalmente autónomas», o si estas siguen siendo «pura especulación» y no deberían servir de punto de partida para nuestra acción política actual¹⁹¹. El rápido desarrollo de la inteligencia artificial parece prejuzgar una respuesta a esta pregunta.

Consideraciones específicas de interfaces neuronales

Una variante de la tecnología cibernética mencionada antes son las interfaces neuronales con las que un ser humano puede gestionar (varios) robots mediante el control mental, lo que combinaría las ventajas del *man-in-the-loop* y las de los robots. Éticamente problemática es la cuestión de cómo la personalidad y la identidad humana serían afectadas por tales interfaces. Tampoco está totalmente claro hoy en día hasta qué punto una persona

¹⁸⁸ SINGER, *op. cit.* (2009), pp. 403-407; KRISHNAN, *op. cit.* (2009), pp. 140-141.

¹⁸⁹ SINGER, *op. cit.* (2009), p. 418; PLÖGER, Paul G. y MARSISKE, Hans-Arthur, «Mythen werden Realität. Paul G. Plöger über Roboter im Kino und in der Wirklichkeit», en: MARSISKE, Hans-Arthur, *op. cit.*, p. 117.

¹⁹⁰ VERUGGIO y OPERTO, *op. cit.*, p. 139.

¹⁹¹ SAUER, *op. cit.* (2019a), pp. 28-29. Traducción del autor.

puede ser considerada legal o moralmente responsable de sus pensamientos o de sus acciones¹⁹². Finalmente, también existe el peligro de abuso, ya que las interfaces neuronales podrían ser invertidas y utilizadas para controlar a los humanos por otros humanos o máquinas, para el control totalitario del pensamiento o para el retardo mental. Una última forma de cibernéticos serían los cerebros humanos, posiblemente sin conciencia, que se crían específicamente para su uso en robots de combate con el fin de explotar sus ventajas sobre la inteligencia artificial, por ejemplo, en el reconocimiento de patrones. Los experimentos para integrar neuronas animales en los circuitos de los robots ya están en marcha¹⁹³. El estado moral de estos cerebros es hoy por hoy completamente incierto.

Conclusiones

A lo largo del texto se han desarrollado varios dilemas éticos del uso de la IA y la robótica y una cantidad de posibles problemas psicológicos, sobre todo en el contexto del uso de los UAV. Como se ha demostrado con varios ejemplos, el utilitarismo, con sus criterios clave de aumentar la felicidad y evitar el sufrimiento y el temor humano, así como el peligro de que estos ocurran, permite dar respuesta a casi todos los dilemas éticos expuestos.

Por ejemplo, el utilitarismo permite intrínsecamente sopesar las diversas ventajas y desventajas militares de la IA y la robótica. Dentro de estas ventajas destacan la mejor protección de los militares contra los negativos efectos de la guerra, la protección de la propia población, los ahorros financieros, la posible estabilización de regiones del mundo que de otro modo no serían interesantes para Occidente y una nueva cualidad de disuasión, que hace menos probable la guerra entre las principales potencias militares. Como desventajas, cabe mencionar la falta de idoneidad de los robots para interactuar con la población civil, la posible desconfianza de los soldados humanos en sus compañeros robóticos, el peligro de la pérdida de los valores militares y de la cohesión de las Fuerzas Armadas por el uso de los robots, y de sus vulnerabilidades tácticas.

¹⁹² SINGER, *op. cit.* (2009), pp. 378-379; KRISHNAN, *op. cit.* (2009), pp. 111-112; SHARKEY, *op. cit.* (2010a), p. 51.

¹⁹³ SINGER, *op. cit.* (2009), p. 172.

En vista de los numerosos riesgos psicológicos y, en menor medida, físicos a los que están expuestas las tripulaciones de los UAV, no puede aceptarse incondicionalmente la idea de que el empleo de los UAV no tenga riesgos, y que viole el ideal de la caballerosidad. El efecto «videojuego» que se puede observar en algunas tripulaciones no puede ser estadísticamente generalizado. Por el contrario, la proximidad mental del operador del UAV con la víctima, generada por el reconocimiento permanente de su objetivo, conduce a menudo a una identificación con este último y, por lo tanto, a barreras psicológicas para el uso de la violencia. Más bien, la robótica, la inteligencia artificial y su desarrollo ulterior ofrecen una oportunidad para reducir significativamente el sufrimiento de la población civil en la guerra, siendo un criterio básico utilitarista.

La posible violación de la dignidad por un «tratamiento» automático del ser humano mediante los sistemas autónomos representa la objeción ética más importante contra el uso de tales sistemas por parte de Alemania, ya que su Constitución somete esta dignidad a una protección especial. Sin embargo, se ha demostrado que existe cierta incompatibilidad entre la Constitución y la ponderación de las vidas de no combatientes frente a las ventajas militares, que está permitida por los Convenios de Ginebra. En cualquier caso, la referencia a la Constitución pierde peso en relación con las Fuerzas Armadas no alemanas y en culturas filosóficamente diferentes. Cuando no se puede asumir un sesgo kantiano compartido, una argumentación utilitarista parece más valiosa.

El intento de Ronald Arkin de equipar a los robots con rutinas éticas y de derecho internacional es de particular importancia en el debate científico. Por su carácter matemático, el paradigma utilitarista se ofrece como una posible solución internacionalmente aceptable de dotar a los robots de dichas rutinas. Este paradigma parece aún más aplicable si se asume que los seres humanos, como seres sujetos a las leyes de la naturaleza, estén al final tan determinados como los robots.

Una respuesta a varias objeciones contra LAWS a largo plazo, compatible con el hecho de que la moral utilitarista no sea necesariamente «humana», sería el desarrollo de una IA «fuerte», dotando a los robots con la conciencia similar a los animales y con un aspecto físico humano. Por otro lado, una moral estrictamente racional y robótica (como se puede considerar la ética

utilitarista) incluso podría representar una mejora de la moral humana, que se caracteriza por sus influencias irracionales.

Los cibernéticos, cuya propagación es concebible como una vía intermedia entre los sistemas autónomos y los teledirigidos, plantean sus propios problemas morales con limitaciones potenciales a la libertad física y mental de los militares mejorados, que el utilitarismo también podría solventar.

A la vista de estas consideraciones, resultan las siguientes conclusiones:

1. Deberían seguir desarrollándose los sistemas de armas por control remoto, cuya utilización permite al menos un veto humano efectivo e informado sobre el uso de la fuerza.
2. Es necesario regular la tecnología cibernética con leyes nacionales precisas. Desde la perspectiva del utilitarismo, básicamente no hay nada que decir contra el uso de la tecnología cibernética en este caso.
3. Se debe intensificar o comenzar a trabajar sobre las normas éticas acordadas internacionalmente para los LAWS. La cuestión de si la calidad moral de la decisión de un robot debe estar sujeta a un estándar más alto que el de un ser humano merece una consideración especial. También hay que tener en cuenta que un sistema normativo que no hace justicia a la realidad no se cumplirá a largo plazo.
4. Si no se llega a un acuerdo sobre fundamentos éticos internacionales comunes, parece aconsejable considerar que, después de todo, cualquier moral solo sobrevivirá si alguien vive y la transmite. Si no somos capaces al menos de defendernos contra los LAWS, el hecho de que incluso pensemos en la ética de los sistemas autónomos se vuelve potencialmente obsoleto, ya que nuestra voluntad de llevar a cabo este pensamiento podría ser destruido con nosotros.