



**Instituto Universitario
General Gutiérrez Mellado**

Investigación sobre la Paz, la Seguridad y la Defensa

MÁSTER UNIVERSITARIO EN PAZ, SEGURIDAD Y DEFENSA

Curso 2019/2020

**SISTEMAS DE ARMAS AUTÓNOMAS LETALES:
¿AUTÓNOMAS O AUTOMATIZADAS?**

**LETHAL AUTONOMOUS WEAPON SYSTEMS:
AUTONOMOUS OR AUTOMATED?**

Autor: Andreas Westhues

Fecha de finalización del trabajo: 31/08/2020

Tutor: Dr. D. Vicente Garrido Rebolledo

Esta página se ha dejado intencionadamente en blanco.

Andreas Westhues, graduado en Ciencia Política y de la Administración. Formador de alemán e inglés. Actualmente cursando el Máster Universitario Paz, Seguridad y Defensa en el Instituto Universitario Gutiérrez Mellado.

Andreas Westhues is graduated in Political Science and Administration. German and English language trainer. He is currently studying the MA Peace, Security and Defence at the Instituto Universitario Gutiérrez Mellado.

Resumen:

Los sistemas de armas de última generación están adquiriendo mayor autonomía gracias a la inteligencia artificial y es muy probable que dentro de los próximos años estas armas alcancen su plena autonomía en el campo de batalla. Esta evolución ha originado varios debates en la academia y la opinión pública. Un primer debate trata de diferenciar sistemas de armas automáticas de armas autónomas. Otro debate trata de explorar las implicaciones para la seguridad internacional, especialmente respecto a la carrera armamentística y la proliferación. Un tercer debate se articula sobre las implicaciones éticas y jurídicas de armas autónomas. Por último, varios actores de la sociedad internacional debaten sobre la necesidad de prohibir o de regular la plena autonomía en las armas. Se concluye este trabajo dando preferencia a la regulación de estas cuestiones en un nuevo marco regulatorio, en aras de mantener los beneficios de las nuevas tecnologías para el uso civil.

Abstract:

The latest generation weapons systems are gaining greater autonomy thanks to artificial intelligence and it is very likely that within the next few years these weapons will achieve their full autonomy on the battlefield. This evolution has originated several debates in academia and public opinion. A first debate tries to differentiate automatic weapons systems from autonomous weapons. Another debate intends to explore the implications for international security, especially regarding the arms race and proliferation. A third debate is articulated on the ethical and legal implications of autonomous weapons. Finally, several actors in international society debate the need to prohibit or regulate full autonomy in arms. This work is concluded giving preference to the regulation of these issues in a new regulatory framework, in order to maintain the benefits of new technologies for civil use.

Palabras Clave: sistemas de armas autónomas; inteligencia artificial; carrera armamentística; proliferación; marco regulatorio

Keywords: autonomous weapon systems; artificial intelligence; arms race; proliferation; regulatory framework

Esta página se ha dejado intencionadamente en blanco.

ABREVIACIONES USADAS EN ESTE TRABAJO

AIDP: Artificial Intelligence Development Plan

CAICT: China Academy of Information and Communications Technology

CCW: Convention on Certain Conventional Weapons

CEO: Chief Executive Officer

CIA: Central Intelligence Agency

C-RAM: Counter-rocket, artillery and mortar

DARPA: Defense Advanced Research Projects Agency

DoD: Department of Defense

DIH: Derecho Internacional Humanitario

EDIDP: European Defence Industrial Development Programme

EGE: European Group on Ethics

EI: Estado Islámico

GGE: Group of Governmental Experts

HRW: Human Rights Watch

IA: Inteligencia Artificial

IAI: Israel Aerospace Industries

INF: Intermediate Range Nuclear Forces

ISR: Intelligence, Surveillance and Reconnaissance

LIDAR: Light Detection and Ranging o Laser Imaging Detection and Ranging

MTCR: Missile Technology Control Regime

MUM-T: Manned-Unmanned Teaming

OODA: Observación, Orientación, Decisión, Actuación

POE: Procedimiento Operativo Estandarizado

PORM: Procedimiento de Obtención de Recursos Materiales

RMA: Revolution in Military Affairs

RPAS: Remotely Piloted Aircraft Systems

SAAL: Sistema de Armas Autónomo Letal

SIPRI: Stockholm International Peace Research Institute

START: Strategic Arms Reduction Talks

UAV: Unmanned Aerial Vehicle

VUCA: Volatility, uncertainty, complexity and ambiguity

Tablas e ilustraciones

Tabla 1: Los enfoques humanos y racionales en la inteligencia artificial	20
Tabla 2: Rasgos de autómatas y dispositivos dotados de inteligencia artificial	22
Tabla 3: Desglose de vehículos no tripulados en uso y en desarrollo	45
Ilustración 1: El bucle OODA.....	24
Ilustración 2: La semiautonomía.....	25
Ilustración 3: La autonomía supervisada	26
Ilustración 4: La autonomía completa	28
Ilustración 5: Lanzamiento y recogida de los Gremlins por transportadores C-130..	35
Ilustración 6: El modelo Blowfish A2.....	46
Ilustración 7: Imagen del UAV Altius-U.....	58
Ilustración 8: MQ-1 Predator, dotado con un AGM 114 Hellfire.....	65

Tabla de contenido

Introducción	9
1 Teoría, conceptos y metodología	10
2 Características, ventajas y riesgos de sistemas de armas autónomas	14
2.1 Ventajas de sistemas de armas con mayor autonomía de acción	15
2.2 Riesgos de los sistemas de armas autónomas	17
3 La relación entre el hombre y el arma en la acción.....	18
3.1 Delimitación de la inteligencia artificial de los autómatas.....	19
3.2 Los grados de autonomía y la implicación humana en ella.....	23
3.2.1 El modo semiautónomo (Human-in-the-loop weapons)	25
3.2.2 El modo de autonomía supervisada (Human-on-the-loop weapons)	26
3.2.3 El modo de autonomía completa (Human-out-of-the-loop weapons)...	28
3.3 Un área particular: la actuación en bandada	30
3.3.1 Arquitecturas de control y comando para las bandadas	33
3.4 Conclusión parcial.....	36
4 Implicaciones para la seguridad internacional.....	37
4.1 Una nueva carrera armamentística.....	37
4.2 La proliferación.....	39
4.3 Las posiciones oficiales de las grandes potencias.....	43
4.3.1 La posición de la República Popular China.....	44
4.3.2 La posición oficial de Estados Unidos	52
4.3.3 La posición oficial de la Federación Rusa.....	57
4.4 Conclusión parcial.....	60
5 Las implicaciones éticas y jurídicas de armas plenamente autónomas	61
5.1 Los principios de distinción, proporcionalidad y precaución	62

5.1.1	La cuestión de la responsabilidad legal	64
5.1.2	La falibilidad de los sistemas	66
5.2	Conclusión parcial.....	67
6	Instrumentos jurídicos del Derecho Internacional Humanitario actual.....	67
6.1	La cláusula de Martens	68
6.2	El examen jurídico de nuevas tecnologías.....	70
6.3	La Convención sobre ciertas armas convencionales	71
6.4	El Régimen de Control de Tecnología de Misiles	73
6.5	Conclusión parcial.....	74
	Conclusiones finales.....	76
	Referencias bibliográficas.....	82
	Anexo bibliográfico.....	84

INTRODUCCIÓN

Por la tarde del 2 de enero de 2020 dos misiles alcanzaron en los exteriores del aeropuerto de Bagdad a un vehículo en el que viajaban el comandante de una milicia iraquí Abu Mehdi al Muhandis, y Quasem Soleimani, líder de la Fuerza Quds, unidad élite de la Guardia Revolucionaria Islámica. Fuentes oficiales de Washington confirmaron horas más tarde la autoría del ataque (Guimón, 2020, 3 de enero).

Este golpe militar fue significativo no solamente porque se produjo en un momento de máxima tensión política entre Irán y Estados Unidos, sino también por su *modus operandi*. Primero, el ataque fue ejecutado con un dron *MQ-9 Reaper*, un vehículo aéreo pilotado a remoto desde bases aéreas que se encuentran normalmente lejos de la zona de operación. Segundo, la zona de la operación, el aeropuerto de Bagdad se encuentra en un país, con el que Estados Unidos oficialmente no está en guerra. Tercero, las víctimas del ataque fueron líderes militares a los que Estados Unidos consideran enemigos que, según las mismas fuentes, estaban preparando ataques contra el personal diplomático y militar estadounidense (ibid.). Se trataba, por tanto, de un ataque selectivo.

Este ataque es uno más de una larga serie que Estados Unidos está ejecutando desde la primera década del nuevo milenio en el exterior, principalmente en países de Oriente Medio y forma parte de la doctrina de guerra preventiva, cuya artífice es la anterior Secretaria de Estados de Estados Unidos, Condoleezza Rice. Dada la falta de transparencia de las autoridades estadounidenses sobre el número de las víctimas militares y civiles en estos ataques, solamente existen estimaciones, basadas principalmente en los relatos de la población civil. A pesar de esta carencia de datos, parece que la precisión de los ataques ha incrementado desde sus inicios en Pakistán.

Desde la publicación del informe del Christof Heyns, Relator Especial sobre ejecuciones extrajudiciales, sumarias o arbitrarias explica en su informe, del 9 de abril de 2013, se han multiplicado las voces que advierten de los peligros que podrían manifestarse en el futuro, cuando los drones adquieren tanta autonomía que podrían elegir los objetivos y ejecutar los ataques sin la intervención del hombre. El Secretario General del Consejo de Seguridad Naciones Unidas, advierte en su informe *Protección de civiles en conflictos armados* de la necesidad de que se regule sistemas de armas autónomas

letales (SAALs en adelante) en cumplimiento con el Derecho Internacional Humanitario.¹

En la actualidad, ya son nueve países que están desarrollando armas autónomas letales. Siete de ellos son: Estados Unidos, China, Rusia, Reino Unido, Francia, Israel y Corea del Sur (Report: Kill the idea of ‘killer robots’ before they kill us, Center for a New American Security, ápod Aznar, 2020, p. 26). A estos hay que añadir Irán, cuyos drones se han encontrado en los ataques a las instalaciones de Saudi Aramco, y Turquía que está usando sus drones en Libia. Este dato revela la importancia del área y no ha pasado de forma desapercibida en el mundo académico y la opinión pública.

En España, se utilizan distintos modelos de drones: el *Drone Searcher MK* y el *Black Hornet* del Ejército de Tierra, el *Scan Eagle* en la Armada, y desde 2019, también cuatro unidades del *MQ-9 Predator B* del Ejército de Aire (XL Semanal, 2018). Todos estos modelos se usan para operaciones de inteligencia, vigilancia y reconocimiento (en adelante ISR por sus siglas en inglés). La cuestión que ha surgido durante los últimos años es si se puede y se debe dotar a estos sistemas con mayor autonomía. El modelo *Predator* es capaz de llevar armas y se puede utilizar para ataques de precisión. Muchos expertos pronostican que, gracias a los avances en la inteligencia artificial, estos sistemas podrían operar con plena autonomía en un futuro no muy lejano. En este escenario surgen las siguientes preguntas de investigación: ¿Cómo deberían responder los decisores políticos a los retos que emanan de armas autónomas? ¿Cómo se define sistemas de armas autónomas? ¿Qué implicaciones se pueden esperar para la seguridad internacional? ¿Cuáles son los aspectos jurídicos y éticos para su desarrollo y su empleo? Y, por último, teniendo en cuenta que ya se han realizado primeros intentos de regular estos aspectos: ¿Qué probabilidad de éxito se puede esperar de la regulación?

1 TEORÍA, CONCEPTOS Y METODOLOGÍA

La teoría de los regímenes internacionales nos aportará las herramientas conceptuales para esta investigación sobre sistemas de armas autónomas y los marcos regulatorios. Esta teoría emergió en la estela de la publicación de *Power and Interdependence: World Politics in Transition*. Sus autores, Keohane y Nye, trataron de superar la restricción conceptual del enfoque estatocentrista del realismo clásico, agregando a su análisis las

¹ UN, Security Council, Report of the Secretary-General. Protection of civilians in armed conflict. S/2020/366, 06/05/2020.

relaciones transnacionales.

Krasner define los regímenes internacionales como principios, normas, reglas y procedimientos, alrededor de los cuales convergen las expectativas de los actores en un área determinada (García, 2009, 150). Los principios representan creencias o valores comunes, las normas pautas de comportamiento definidos, mientras las reglas son las prescripciones particulares en un grado inferior a las normas. Los procedimientos hacen referencia a los procesos de decisión colectiva (ibid.). Volker Rittberger añadió más tarde los criterios de efectividad y de duración. La efectividad de un régimen se muestra en la ausencia de otras opciones de comportamiento, resultado de las sanciones, cuando no se cumpla lo estipulado. El criterio de duración indica que un régimen internacional puede instituir un ejemplo de uso consuetudinario, cuando sobrepasa las limitaciones de un acuerdo de mera conveniencia mutua (ibid.). Hasenclever, Mayer y Rittberger señalan el Tratado de No Proliferación Nuclear y la Convención Europea de Derechos Humanos como dos ejemplos destacados de regímenes internacionales (Hasenclever, Mayer, Rittberger, 99, 500).

Estos autores presentaron a finales de los años noventa propuestas para una síntesis que consisten básicamente en completar las restricciones conceptuales del enfoque neoliberal y del enfoque realista. Lo hicieron integrándolos con un tercer componente, al que llaman cognitivismo débil. Este último enfoque incluye el valor de las ideas y la capacidad de aprendizaje y trata de completar así el predominio del pensamiento racionalista en los otros dos enfoques (Hasenclever, Mayer, Rittberger, 1999, 508-510). Esta síntesis nos ofrece un marco suficientemente amplio para analizar un futuro marco regulatorio para los sistemas de armas autónomas, porque nos permite incluir las ideas en el análisis.

Basándonos en lo expuesto anteriormente, se utilizará el siguiente diseño de investigación. El marco teórico se fundamentará en la teoría de los regímenes internacionales. La perspectiva realista enfatiza los Estados como principales actores, mientras el institucionalismo recalca los actores transnacionalistas de la sociedad civil. En nuestro análisis, los actores son, por tanto: los Estados individuales, organizaciones internacionales como la Organización de Naciones Unidas, organizaciones no gubernamentales como Human Rights Watch (HRW en adelante) y Cruz Roja, la Campaña para la Prohibición de Robots Asesinos, pero también grupos y personas

individuales procedentes de la academia.

El estado de la cuestión se articula en tres debates. Un primer debate gira alrededor de la denominación adecuada de las armas autónomas: ¿Se trata de sistemas automatizados o armas autónomas? ¿Dónde está la línea delimitadora entre la automatización y la autonomía? Parece que la inteligencia artificial está dando un cambio significativo al desarrollo de armas, y varios autores registran aquí una nueva era. Jordán y Baqués señalan que ellos no entienden «una revolución como un cambio rápido si no, más bien, como un cambio profundo. Su principal atributo no es, por lo tanto, la velocidad, sino su capacidad para cambiar las cosas» (Jordán y Baqués, 2014).

Un segundo debate se preocupa de las implicaciones para la seguridad internacional. Nueve países están trabajando en el desarrollo de armas autónomas, y una treintena de países dispone ya de armas autónomas, en su mayoría drones. El primer dato apunta hacia una nueva carrera armamentística, mientras el segundo dato muestra la proliferación de estas tecnologías. Esta tendencia podría acarrear cambios en la polaridad del sistema internacional. Todo depende de los avances de cada país en esta área, pero también del acceso de otros países o actores no estatales a tecnologías, cuya proliferación es más fácil que componentes de armas nucleares, por ejemplo.

El tercer debate se desarrolla en el ámbito jurídico y ético. ¿Cómo afectarían armas autónomas en los ámbitos del Derecho Internacional Humanitario y el Derecho de Conflictos Armados? ¿Quién es el responsable en el uso de armas autónomas letales y de las bajas por un ataque? ¿Qué ocurre cuando un arma actúa de forma tan autónoma que para un comandante ya no es posible frenar el ataque?

Estas cuestiones se contextualizan en un entorno VUCA (Acrónimo de *Volatility, Uncertainty, Complexity y Ambiguity*) donde han desaparecido los referentes de la Guerra Fría y su aparente estabilidad geopolítica. El nuevo entorno internacional se caracteriza por un creciente número de Estados fallidos y de guerras asimétricas.

En el contexto de los debates ha surgido la reivindicación de regular el desarrollo y el uso de estas armas. En la actualidad, el régimen que regula la exportación de tecnología utilizable para los drones es el Régimen de Control de Tecnología de Misiles (*Missile Technology Control Regime*, en adelante MTCR).

Una propuesta de origen más cercano concierne al Convenio de ciertas Armas Convencionales, en vigor desde el 2 de diciembre de 1983, que consiste en cinco

protocolos adicionales a los Convenios de Ginebra. Su objetivo es la proscripción de diferentes clases de armas que pueden considerarse excesivamente nocivas o que puedan producir efectos indiscriminados (CICR, 2002). Este convenio, una iniciativa basada en la resolución 32/152 de la Asamblea General de las Naciones Unidas, de 19 de diciembre de 1977 es el régimen internacional que ha asumido el debate institucional sobre las armas autónomas, para crear un nuevo marco regulatorio ex profeso.

El objeto de esta investigación son los sistemas de armas autónomas. El Departamento de Defensa de Estados Unidos (Department of Defense, en adelante, DoD) define en su directiva 3000.09 un sistema de armas autónomas como un sistema de armas que, una vez que se haya activado, puede elegir un objetivo y actuar sobre él sin la consiguiente intervención de un operador humano, si bien el hombre siempre puede anular la operación (DoD, 2012, 13). Esta definición será la base conceptual para los sistemas de armas en este trabajo. A pesar de su amplia difusión entre organizaciones no gubernamentales y otros actores del sector civil, no se usará el concepto de robots asesinos (*killer robots* en inglés), porque la complejidad y diversidad de los sistemas desaconsejan el uso de un concepto tan restrictivo y a la vez valorativo.

Con el motivo de evitar repeticiones conceptuales durante la lectura, se usarán en este trabajo de forma equivalente el concepto anterior junto con las denominaciones armas autónomas, armas autónomas letales y la abreviatura SAALs (Sistemas de armas autónomas letales). Ejemplos de esta clase de armas podrían ser el sistema de defensa de buques *AEGIS*, pero también vehículos aéreos no tripulados como los que se usan en el Ejército español.

Trabajaremos con una hipótesis principal y dos derivadas.

Hipótesis principal: ¿Hasta qué punto es la inteligencia artificial determinante en la definición y el desarrollo de sistemas de armas autónomas?

Primera hipótesis derivada: Si la inteligencia artificial es determinante, ¿hasta qué punto obstaculiza su definición llegar a una regulación efectiva?

Segunda hipótesis derivada: La regulación efectiva de sistemas de armas autónomas requiere nuevos marcos regulatorios.

El objetivo de la investigación es comprobar las hipótesis y discutir las implicaciones de los sistemas de armas autónomas en las áreas de seguridad internacional, así como sus

aspectos éticos y jurídicos.

La relevancia de esta investigación está en explorar qué variables influyen en la delimitación de un marco regulatorio para sistemas de armas autónomas.

El método que se usa en este trabajo es analítico-discursivo. Por último, el carácter transversal de la inteligencia artificial impone delimitar el objeto de esta investigación. Aunque se comentarán para señalar el potencial del uso de la IA, no se analizarán estas áreas de su aplicación: el uso de la inteligencia artificial en armas nucleares, biológicas y químicas. El dominio cibernético es un campo donde la inteligencia artificial tiene muchas aplicaciones como la realización de ataques cibernéticos y la divulgación masiva de noticias falsas (fake news), pero no se estudiará en este trabajo. No cabe la menor duda que la inteligencia artificial es un componente con un potencial peligroso en estas tecnologías, pero su prolijidad excede las limitaciones de este trabajo. Asimismo, este trabajo analizará tampoco el escenario de una superinteligencia superior a la inteligencia humana y de sus implicaciones éticas, cuando esta supuesta superinteligencia tomara el control sobre todo desarrollo posterior, tal como planteó Bostrom (Bostrom, 2003).

La pandemia provocada por el virus Sars-CoV-2 podría provocar cambios significantes en el escenario geopolítico mundial y la gestión de las crisis internacionales. Aún es pronto para pronosticar la forma y el alcance de estas transformaciones, por lo que en este trabajo no se tendrán en cuenta la pandemia y posibles cambios geopolíticos.

2 CARACTERÍSTICAS, VENTAJAS Y RIESGOS DE SISTEMAS DE ARMAS AUTÓNOMAS

Antes de entrar en el análisis de los SAALs, veremos brevemente las características, las ventajas y los riesgos de armas autónomas.

Actualmente, existe cierta ambigüedad conceptual sobre este tipo de armas. El Grupo de Expertos Gubernamentales sobre las Tecnologías Emergentes en el Ámbito de los Sistemas de armas autónomas Letales (Group of Governmental Experts, en adelante GGE) diferencia entre los siguientes sistemas de armas autónomas que ya existen en la actualidad:

- sistemas de armas de defensa aérea con modos o funciones autónomos,²

² Algunos autores definen la propulsión propia de los vehículos uno de los rasgos destacado, por lo que los sistemas de defensa aérea y naval quedarían excluidos de los SAALs. Como se verá en capítulo 3, la clasificación de diferentes modos de autonomía aconseja seguir la clasificación del GGE.

- misiles con modos o funciones autónomos,
- sistemas de armas de protección activa con modos o funciones autónomos,
- armas merodeadoras con modos o funciones autónomos,
- minas navales o terrestres con modos o funciones autónomos, y, por último,
- armas centinela con modos o funciones autónomos (Convention on Certain Conventional Weapons [CCW]/GGE.1/2019/3).

Como se aprecia en esta clasificación, existen ya muchos sistemas de armas en los dominios tierra, mar y aire con diferentes grados de autonomía. Entre ellos, los vehículos aéreos no tripulados (UAV en adelante) ocupan una posición destacada, porque han sido empleado ya ampliamente en Pakistán, Irak y Yemen, entre otras regiones.

2.1 Ventajas de sistemas de armas con mayor autonomía de acción

Un aspecto técnico es la latencia de la señal entre el puesto de control y el vehículo, resultado de la larga distancia que la señal tiene que recorrer a través de un satélite y quizás puntos de relé en otros países. La latencia puede llegar a sumar varios segundos, por lo que el pilotaje de un RPAS en tiempo real no es posible. Este es el motivo por el cual estos vehículos disponen ya de suficiente autonomía para las maniobras de despegue y aterrizaje. El *Northrop Grumman X-47B* ya puede despegar y aterrizar en un portaviones. Otro dron, el *Taranis*, desarrollado por BAE Systems, es capaz de regresar de forma autónoma a la base aérea, sin la intervención de un operador de vuelo (Scharre, 2018, 109).

La comunicación puede interrumpirse por condiciones geográficas, meteorológicas o por interferencias del enemigo. En estos casos, mayor autonomía significa menor dependencia de la comunicación. Los vehículos no tripulados pueden continuar con su misión.

Christof Heyns, explica en su informe las «enormes ventajas» que ofrecen los sistemas de armas autónomas (Heyns, 2013, 11).

La distancia entre operadores de RPAS y el teatro de operaciones reduce los riesgos de bajas en las propias filas. La muerte de un militar supone hoy en día no solamente la vida de una persona, sino también la pérdida de una especialista en su área, altamente entrenado y, por ende, no tan fácilmente reemplazable como soldados en las grandes

contendidas de los siglos pasados. Además, desde un punto de vista mediático, la pérdida de un militar reduce el respaldo ciudadano a las operaciones militares. González señala que en la Revolución de Asuntos Militares (RMA en sus siglas en inglés) no solamente se manifiesta en nuevas formas de combatir, sino en un cambio de las preferencias en las sociedades occidentales, donde uno de estos cambios es el rechazo social al dolor y a la destrucción. González señala que «en Occidente hoy cada uno de los hijos encarna una porción mucho mayor de capital emocional que antes (González, 2014, 48).»

Sistemas no tripulados pueden ampliar el campo de batalla, entrando con mayor facilidad detrás de las líneas enemigas, sin arriesgar la vida de los propios soldados. Estos sistemas se pueden mantener durante mucho tiempo en el aire, sin que factores físicos como las fuerzas gravitacionales limiten la acción, caso que ocurre en aeronaves pilotadas (Heyns, 2013, p.11). La distancia espacial entre el área de operación y el puesto de pilotaje reduce también los costes personales y materiales, limitándolos a equipos de mantenimiento en las bases aéreas desde donde se despliegan los drones (Heyns, 2013, 11).

En cuanto al tipo de misiones, se está proyectando una gran variedad de posibilidades. Dada su gran capacidad de vigilancia en misiones ISR, sistemas autónomos aéreos ofrecen la opción de ser desplegados para fines humanitarios, por ejemplo, la localización de desplazados internos lo que facilitaría su recuperación por equipos especializados.

Drones como el *Black Hornet* facilitan vigilar una zona de lucha en terreno urbano puede aportar informaciones valiosas a los propios soldados. El dron y los militares desplegados pueden formar una red que, junto con la geolocalización, facilita dispersar la niebla de la guerra y proteger a los soldados.

Otra ventaja, referenciada a menudo, es la racionalidad de sistemas artificiales. La ausencia de emociones, tales como venganza, odio o miedo, ayudaría a limitar los daños colaterales entre civiles, y aportaría una ventaja táctica en la batalla.

Sistemas de defensa como *Phalanx* y *Mantis* actúan a una velocidad y precisión muy alta, sin que el hombre tenga que intervenir. Ofrecen así una ventaja indisputable cuando operan en modo de autonomía supervisada.

Se espera también una reducción de costes relacionada con el desarrollo de sistemas de armas autónomas por el menor despliegue de efectivos humanos. Drones como el *Loyal Wingman* serán considerablemente más baratos que cazabombarderos comparables

de la quinta generación, a pesar de su tecnología avanzada (Defense News, 2020, 4 de mayo).

Por último, los defensores de armas autónomas esperan que la capacidad de autoaprendizaje (*machine learning* o *deep learning* en inglés) les ayude con mayor seguridad a operar en entornos poco estructurados o caóticos.

2.2 Riesgos de los sistemas de armas autónomas

Todos los sistemas manifiestan cierta probabilidad de fallar, y si no existe ningún control - sea otro sistema o el hombre – las consecuencias pueden ser dramáticas. Uno de los ejemplos más notorios es el del Teniente Coronel Stanislav Petrov quien evitó una guerra nuclear el 26 de septiembre de 1983, cuando confió más en su intuición en lugar del sistema alerta temprana del que era responsable. El nuevo satélite *Oko* había interpretado los reflejos de la luz solar en las nubes como el lanzamiento de misiles intercontinentales. Petrov decidió contrastar la información del sistemas con la vigilancia aérea y sus superiores en lugar de iniciar el contraataque que exigía el protocolo de acción (Scharre, 2018, 2).

¿Qué hubiera ocurrido si un sistema automatizado hubiera actuado sin la intervención humana? No es probable que en un escenario de estas envergaduras el hombre esté fuera del bucle, pero ¿quién puede garantizar que esto no cambie cuando un conflicto se escale? ¿Quién puede garantizar que un país no despliegue estas armas en el modo de autonomía supervisada para responder con más rapidez a un ataque masivo de enemigo? Críticos de las armas autónomas alegan también una posible mayor predisposición para la entrada en conflicto, debido al menor riesgo para las personas.

Gutiérrez y Cervell advierten que el incremento de armas autónomas podría ir acompañado por un mayor desinterés de la opinión pública, lo que a los gobiernos daría más libertad de actuar al margen de la legalidad (Gutiérrez, Cervell, 2013, 31).

La razón que alegan muchos críticos es la ingente cantidad de ejecuciones extrajudiciales que han estado ejecutando las fuerzas de seguridad estadounidenses durante los últimos años en Oriente Medio. El informe de Heyns, antes citado, estaba motivado precisamente por este tipo de ataques selectivos. Jordán y Baqués señalan que Estados Unidos llevó a cabo en diez años más de 300 ataques aéreos contra combatientes

talibanes y cuadros de Al Qaeda en zonas fronterizas de Pakistán (Jordán, Baqués, 2014)³. Pero estos ataques no fueron operaciones militares stricto sensu. Al contrario, se trataba de una operación encubierta de la Agencia Central de Inteligencia (en adelante CIA por sus siglas en inglés). Según un recuento de la *New America Foundation*, entre 2004 y marzo de 2013, los ataques con drones acabaron con la vida de entre 1535 y 2660 militantes de talibanes (New America Foundation, ápod Montoya, 2014). La misma fuente señala que solamente el dos por ciento de las víctimas fueron cuadros de Al Qaeda o de otros grupos terroristas, mientras más de 80 por ciento eran militares rasos o colaboradores que apoyaban la yihad de forma no combativa (ibid.).

3 LA RELACIÓN ENTRE EL HOMBRE Y EL ARMA EN LA ACCIÓN

Una de las cuestiones esenciales es el debate sobre la automatización y la autonomía en los sistemas de armas. Las armas no son un invento de las últimas décadas. Losada señala que las invenciones para la guerra han sido uno de los motores de la historia (Losada, 2014). Prácticamente, desde el principio de la humanidad, la especie humana se encuentra en lucha unos contra otros por cuestiones territoriales. El gran cambio se produjo con la emergencia de los primeros imperios que se formaron a base de guerras de expansión, en las cuales determinadas invenciones cruciales aportaron una ventaja estratégica. Por ejemplo, los inventos de ingeniería para la marina de guerra proporcionaron una ventaja importante al imperio romano, frente a los pueblos conquistados. (Losada, 2014). Siglos más tarde, se produjo en Europa un cambio disruptivo con la Ilustración, que sirvió de trasfondo intelectual para el salto científico y tecnológico asociada a ella. La automatización industrial se reprodujo también en las armas, que se sofisticaron cada vez más.

Las minas terrestres son un ejemplo nefasto de armas automatizadas. Han matado a miles de personas, entre ellas a muchos niños. Una vez que se hayan escondido en el terreno, se activan a partir de un determinado peso. Un sistema similar existía en el lado oriental de la frontera entre las dos Alemanias. Se trataba de dispositivos automáticos de disparo que se activaron cuando una persona trataba de huir hacia el lado occidental de la frontera. Ambos sistemas, las minas terrestres y los dispositivos automáticos de disparo son sistemas que funcionan de forma automática, es decir, el soldado que los ha

³ Se menciona solamente el año de publicación cuando la versión electrónica no contiene paginación.

desplegado no tiene que activarlos. Lo hace la persona que va a ser la víctima de esta arma.

Las minas tienen en común un rasgo con las armas autónomas: son capaces de actuar sin la intervención directa de las personas. Pero no son capaces de buscar objetivos. Una excepción son minas de torpedo encapsuladas que lanzan un torpedo en lugar de explotar cuando hacen contacto con un buque. Este torpedo, una vez lanzado, busca el buque en cuestión y muestra por tanto cierta autonomía (Scharre, 2018, 51).

Viendo estos ejemplos, se puede sostener que una de las grandes diferencias entre el modo automatizado y la autonomía es la existencia de otro componente no tangible. Un sistema informático no solamente requiere un hardware⁴, sino también un conjunto de instrucciones, los algoritmos, o el software. Los algoritmos no son nada nuevo en el armamento. En realidad, muchos sistemas modernos no podrían funcionar sin un software consolidado, resultado de décadas de desarrollo. Muchos sistemas actuales actúan de forma automática. Sistemas de defensa antiaérea como el estadounidense *Phalanx*, el alemán *Mantis* y el sistema israelí *Iron Dome* son los ejemplos más conocidos de sistemas capaces de detectar un ataque y responder con un contraataque de forma automática. Aun así, algoritmos bien desarrollados que permiten a un sistema de dar una respuesta tan rápida que un humano no podría efectuar en el mismo espacio de tiempo, también son una variante de automatización, aunque ya con mayor independencia del hardware.

Aquí entre en juego la inteligencia artificial, cuyo desarrollo empezó simultáneamente con la computación moderna. Y como se verá más adelante, un rasgo predominante de la inteligencia artificial es su adaptabilidad a entornos variables. Otra diferencia entre sistemas de armas automáticas y sistemas basadas en inteligencia artificial es la capacidad de aprendizaje. Este es un paso significativo en las nuevas generaciones de armamento, aunque en la actualidad aún sea difícil prever los futuros desarrollos.

3.1 Delimitación de la inteligencia artificial de los autómatas

Russel y Norvig ofrecen en su obra *Artificial Intelligence – A modern approach* la base conceptual de la IA para este trabajo. Los autores señalan que este campo de investigación y desarrollo comenzó después de la Segunda Guerra Mundial, y el nombre

⁴ Tanto el término hardware como su complemento software son anglicismos, pero la Real Academia de la Lengua Española ha admitido ambos.

mismo fue acuñado en el año 1956 (Russel y Norvig, 2016, 1).

En su libro comparan ocho diferentes definiciones sobre la inteligencia artificial, agrupadas en cuatro diferentes campos. En la tabla 1 se pueden comparar cuatro de ellas, elegidas solamente con fines ilustrativos, una para cada enfoque. La columna izquierda muestra el enfoque humano, con los campos pensar y actuar. Los dos primeros enfoques tratan de emular al ser humano en su forma de pensar y la forma de actuar. La columna derecha muestra el enfoque racional, dividido también en los campos pensar y actuar. Así, la primera y segunda definición toman al ser humano como referencia y se ajustan al pensamiento y la actuación humanos.

Las otras dos definiciones se alejan del enfoque humano. Sus defensores parten del supuesto que la computación ofrece unas ventajas sobre el ser humano, debido a la rapidez del procesamiento y por la ausencia del componente emocional. Los autores subrayan la importancia de desarrollar justamente esta diferencia en sus trabajos sobre computación e inteligencia artificial (Russel y Norvig, 2016, 2).

Tabla 1: Los enfoques humanos y racionales en la inteligencia artificial

Enfoque humano	Enfoque racional
Pensar de forma humana	Pensar de forma racional
«[La automatización de] ... actividades que asociamos con el pensamiento humano: actividades tales como tomar decisiones, solucionar problemas, aprender ... (Bellman, 1978).» ⁵	«El estudio de la computación que haga posible de percibir, razonar y actuar (Winston, 1992).»
Actuar de forma humana	Actuar de forma racional
«El estudio de cómo conseguir que los ordenadores hagan cosas, en las que, por el momento, las personas son mejores (Rich y Knight, 1991).»	«Inteligencia computacional es el estudio del diseño de agentes inteligentes (Poole et al., 1998).»

Fuente: Tabla adaptada de Russel y Norvig, 2016, 2.

Bellman publicó su definición en 1978, una década antes de aparecer los primeros ordenadores para uso doméstico. Su referencia son las actividades intelectuales del humano, en particular las de toma de decisiones, solución de problemas y la capacidad de aprendizaje. El rasgo distintivo de esta definición es la automatización de estos procesos.

Rich y Knight parten de una posición modesta, cuando robots trataban emular de

⁵ Traducciones propias de las definiciones originales.

forma tosca los movimientos de los humanos. De ahí que el objetivo de esta definición es simplemente desarrollar aquellas aptitudes, en las que el humano es mejor.

Las otras definiciones se fundamentan en el enfoque racional, donde la computación tiene las dos grandes ventajas mencionadas, rapidez en los procesos y ausencia del componente emocional. Winston incluye en su definición las capacidades de percibir, razonar y de actuar. La definición de Poole es la más reciente y amplia en esta categorización, definiendo la inteligencia artificial como el estudio del diseño de agentes inteligentes. (Russel y Norvig, 2016, 2).

Todas las definiciones aportan algún aspecto importante al estudio de la inteligencia artificial. Estados Unidos está usando el nuevo concepto algoritmo de guerra, con la intención de obtener más opciones y decisiones algorítmicas para asegurarse la superioridad tecnológica (Fojón, 2018). El proyecto *Algorithmic Warfare Cross-Functional Team* es uno de los ejemplos de este nuevo concepto. En este proyecto, creado en abril de 2017 y más conocido por el nombre *Maven*, participó Google hasta que la dirección de la compañía decidió no renovar el contrato de cooperación después de marzo de 2019. *Maven* consistía en transferir los conocimientos que Google había adquirido en el etiquetaje para Google Mapas a la evaluación automatizada de imágenes obtenidos por drones de reconocimiento (Conger, 2018, 1 de junio). Este procedimiento usa la capacidad de autoaprendizaje, en inglés conocida como *machine learning* (Fojón, 2018)⁶. El autoaprendizaje es un enfoque de la inteligencia artificial que emula el funcionamiento del cerebro animal creando una red, compuesta por diferentes capas neuronales. Estas redes se alimentan continuamente con nuevos datos, que se procesan creando nuevas conexiones entre sus capas, de forma similar a las neuronas en el cerebro animal. Los desarrolladores diseñan la red y los algoritmos que la capacitan para el procesamiento autónomo de los datos (Scharre, 2018, 86-87). Los procesos concretos son difícilmente reproducibles, incluso para los propios desarrolladores.

Aunque el objetivo de *Maven* es el tratamiento automatizado de imágenes obtenidas por los servicios de inteligencia, el proyecto se caracteriza por tres rasgos intrínsecos de la IA: su adaptabilidad a la entrada de material topográfico cada vez nuevo, y su capacidad de autoaprendizaje que, por su parte, requiere de la capacidad de adaptarse. El tercer rasgo

⁶ Otro concepto usado ampliamente es el de *deep learning*. Ambos conceptos se usan para la descripción de redes neuronales (*deep neural networks*).

es la autonomía, que consiste en la capacidad de realizar una serie de tareas que suelen ser predefinidas. Pero su realización depende de los otros dos rasgos: la adaptabilidad y el autoaprendizaje.

Una mina terrestre, sea antipersonal o anticarro, tiene una adaptabilidad predefinida y por tanto más limitada, que consiste normalmente en un peso predeterminado que activa su explosión. Una mina no tiene ninguna capacidad de aprendizaje porque se trata de un dispositivo cuya actuación se basa en un funcionamiento mecánico.

Tabla 2: Rasgos de autómatas y dispositivos dotados de inteligencia artificial

Dispositivos automáticos	Dispositivos dotados de Inteligencia Artificial
<ul style="list-style-type: none"> • adaptabilidad predefinida y limitada 	<ul style="list-style-type: none"> • adaptabilidad variable
<ul style="list-style-type: none"> • sin capacidad de aprender 	<ul style="list-style-type: none"> • capacidad de autoaprendizaje
<ul style="list-style-type: none"> • autonomía predefinida y limitada 	<ul style="list-style-type: none"> • autonomía ampliable

Fuente: Elaboración propia

Google abandonó el proyecto tras las protestas de sus empleados. El proyecto *Maven* es un ejemplo de lo difícil que es delimitar con nitidez el amplia área de inteligencia artificial para usos civiles de su empleo como arma. ¿Cuándo empieza una herramienta a ser un arma? Los desarrolladores de Google lo tenían claro: 4.000 empleados exigieron en una carta abierta al CEO Sundar Pichai el abandono del proyecto.⁷ Su principal preocupación consistía en que su producto podría facilitar la selección de blancos en el campo de batalla. ¿Qué ocurriría si un dron de reconocimiento armado tuviera la capacidad de decidir de forma autónoma a lanzar un ataque, basado sobre las imágenes que el mismo dron haya obtenido durante un vuelo de reconocimiento? En la actualidad, estas decisiones toman operadores militares desde bases aéreas, pero nadie puede garantizar que los drones no dispongan de la capacidad decisoria en el futuro. Entonces, ¿qué parte en la cadena sería el agente? ¿El algoritmo que analiza las imágenes o el dron dotado con inteligencia artificial capaz de decidir sobre un ataque? Ambas partes forman parte de un conjunto, y desde este punto de vista, la preocupación de los desarrolladores de Google es comprensible.

Para nuestra definición de sistemas de armas autónomos letales conviene seguir a Russel y Norvig y su enfoque de inteligencia artificial como agente racional, que es un

⁷ Letter to Google C.E.O. La carta se encuentra en el anexo bibliográfico.

agente capaz de operar autónomamente, percibir y procesar datos de su entorno, mantenerse operativo durante un espacio de tiempo prolongado, adaptarse a cambios y perseguir objetivos. Un agente racional actúa de tal manera que alcance el mejor resultado, o, en situaciones de incertidumbre, el mejor resultado esperado (Russel y Norvig, 2016, 4).

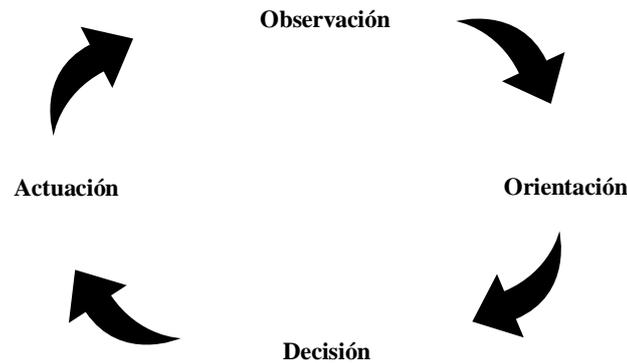
Teniendo en cuenta las consideraciones anteriores, definimos el objeto de esta investigación como sigue: Un sistema de arma autónomo letal es un agente racional, dotado con inteligencia artificial que le proporciona la capacidad de adaptarse a contextos cambiantes, de aprender durante el ciclo de su existencia y de realizar tareas con diferentes grados de autonomía.

3.2 Los grados de autonomía y la implicación humana en ella

Se define ciclo de actuación como el proceso de un sistema de defensa que observa un espacio definido, detecta un objeto y lo identifica como amigable u hostil. Este proceso continúa con la toma de decisión – neutralizar el objetivo o no – y la acción de neutralizar misma. Los sistemas de defensa mencionados anteriormente operan de esta manera. El proceso de *targeting* consiste en la detección y fijación del objetivo, después del que el sistema actúa contra la amenaza.

Esta operación se puede representar gráficamente con modelo que fue desarrollado por el coronel John Boyd, un aviador estadounidense. Su modelo nace del combate aéreo (*dogfight* en inglés), un entorno altamente inestable, donde el piloto debe analizar continuamente su propia posición respecto al otro y tomar decisiones en un espacio de tiempo muy reducido. El bucle OODA (*OODA-Loop* en inglés) describe este ciclo, basado en los cuatro pasos observación, orientación, decisión y acción. Aunque este ciclo fue desarrollado en el dominio aire, se puede trasladar a sistemas de defensa de otros dominios. En situaciones complejas como un combate aéreo o la defensa contra un ataque de saturación a los propios sistemas de defensa, los diferentes pasos se suceden a un ritmo tan elevado, que el humano no pueda responder de forma óptima.

Ilustración 1: El bucle OODA



Fuente: Elaboración propia basada en Scharre (2018)

La observación del entorno y la posición propia como la de otros combatientes forman parte de este ciclo en el que se perciben continuamente nuevos datos. La decisión de atacar o no hacerlo corresponde al humano, por ejemplo, al piloto en combates aéreos o al comandante de un sistema de defensa antiaéreo. Sistemas modernos realizan la fijación de un objetivo de forma automática, de manera que el comandante se puede concentrar más en los pasos decisión y actuación. Sin embargo, se pueden producir situaciones muy complejas, en las que el humano quizás no pueda actuar con la rapidez necesaria, especialmente cuando se produce un ataque de saturación de defensa. Por esto, los sistemas mencionados cuentan ya con diferentes modos de automatización. Conviene analizar la posición del militar dentro del ciclo de actuación de armas automáticas.

Scharre diferencia el rol del humano dentro del bucle OODA con relación a la función que va a desempeñar. El humano puede actuar: a) como operador esencial de un sistema de armas, b) como un seguro (*fail-safe* en el texto original), o c) como un agente moral. Cuando el humano actúa como un operador esencial del sistema de armas, los pasos de decisión y actuación se encuentran en las manos del humano y el sistema no puede actuar sin él. En el segundo caso, el arma puede actuar de forma autónoma, pero el humano supervisa el bucle para poder intervenir, cuando las circunstancias lo requieran y el ataque ya no es la opción adecuada. Cuando el humano actúa como agente moral, es él que toma decisiones basadas en valores sobre lo apropiado que puede ser el uso de la fuerza o no (Scharre, 2018, 322).

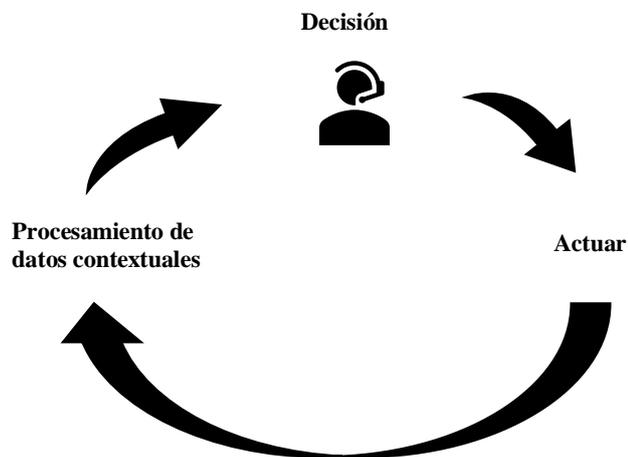
En base a la posición del humano en el proceso de fijación y seguimiento del objetivo se pueden crear tres grupos. El primer grupo abarca los sistemas en los que el humano es una parte esencial para su funcionamiento, mientras en el segundo grupo la función del

humano se limita a las funciones supervisión e intervención. El tercer grupo comprende sistemas de armas, donde el humano ya no forma parte del sistema, y describe, por tanto, las verdaderas armas autónomas, el objeto de este estudio.

3.2.1 *El modo semiautónomo (Human-in-the-loop weapons)*

En el primer modelo (ilustración 2) el hombre forma parte integral del bucle. Cuando los sensores de un sistema de armas detectan un objeto en su espacio, trata de identificarlo y seguir su trayectoria. Hemos juntado del bucle OODA los pasos observación y orientación en un único paso llamado procesamiento de datos contextuales que comprende tanto la posición del propio sistema de vigilancia como la posición del objeto.

Ilustración 2: La semiautonomía



Fuente: Elaboración propia basada en Scharre (2018)

En este gráfico se observa que el hombre tiene una función clave porque es él que toma la decisión de actuar. El sistema realiza la tarea de observación y orientación de forma autónoma, pero espera la orden del humano para actuar (Scharre, 2018, 29).

Un ejemplo de este tipo de sistema es el sistema de defensa *C-RAM* (*Counter-rocket, artillery and mortar*). La parte de detección, persecución y fijación del objetivo está automatizada, pero es el comandante que da la orden de actuar. *C-RAM* cuenta con una red de radares interconectados que identifican automáticamente objetivos dentro de su espacio. El sistema se despliega frecuentemente para la protección de bases aéreas donde circulan aeronaves amigo. Para evitar fratricidios, el sistema crea de forma automática una zona de no atacar alrededor de los propios aviones una vez que está correctamente identificados. El resultado es un sistema altamente automatizado, pero el operador del C-

RAM aún toma la última decisión de verificar el objetivo y autorizar el ataque defensivo (Scharre, 2018, 323-324).

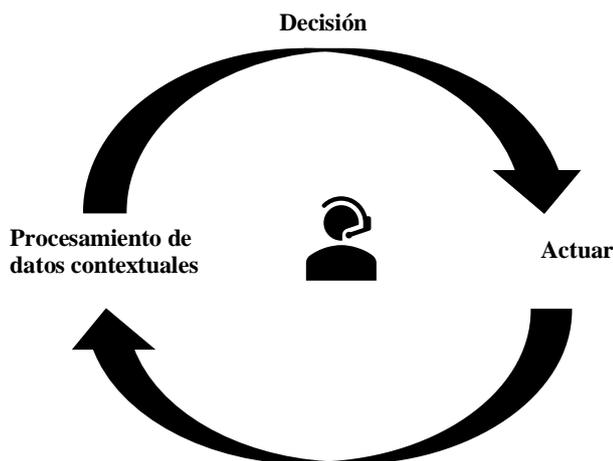
Otro ejemplo es el sistema antiaéreo *MANTIS*, desarrollado por la empresa alemana *Rheinmetall Air Defence*. El sistema detecta y combate con rapidez drones, misiles o granadas. Según un vídeo del ejército alemán, el comandante de la posición tiene un máximo de 20 segundos para dar la orden de disparar. Aquí, el humano también se encuentra dentro del bucle, es decir, él toma la última decisión (Bundeswehr, 2012, 27 de septiembre).

No obstante, en escenarios muy disputados en los que, por ejemplo, se produce un ataque de saturación de defensa, puede ser que el operador humano sea precisamente el factor que ralentice la respuesta, cuando tiene que evaluar cada objeto sospechoso. En este caso, la rapidez es el factor que impone otro modo, en el que se evita este tipo de inconveniente (Scharre, 2018, 325).

3.2.2 *El modo de autonomía supervisada (Human-on-the-loop weapons)*

Este segundo grupo de armas comprende sistemas que son capaces de actuar de forma autónoma, desde la detección e identificación de un objetivo como hostil, hasta el ataque para neutralizarlo. El comandante ya no toma la decisión de atacar.

Ilustración 3: La autonomía supervisada



Fuente: Elaboración propia basada en Scharre (2018)

En la representación gráfica de este modelo, el operador del sistema está situado en el centro del círculo. El sistema de defensa actúa de forma autónoma: observa el entorno,

detecta e identifica objetivos dentro de su espacio vigilado, y los neutraliza con un contraataque, siguiendo las reglas de enfrentamiento que se han establecido previamente. El operador supervisa todo este proceso e interviene solamente, cuando lo considere necesario.

Scharre señala que más de treinta países emplean ya armas de defensa autónomas en situaciones en las que la velocidad de actuación es demasiado alta para el humano. Estos sistemas se usan para defender buques y bases militares contra ataques de misiles y granadas, pero siempre en modo supervisado (Scharre, 2018, 4-5). Ejemplos son defensas navales como el sistema de combate *AEGIS* y el sistema de defensa *Phalanx Close-In Weapon System* (CIWS), ambos americanos.

AEGIS (*AEGIS Combat System, ACS*) es un sistema altamente complejo. En la actualidad, está instalado en 84 buques estadounidenses (US Navy, 2019, 10 de enero). Estos sistemas pueden disponer de diferentes modos de operación, lo que permite un modo supervisado u otro modo semiautónomo. En todo caso, el hombre establece qué tipo de amenazas deberían ser atacadas, y cuáles deberían ser ignoradas, así como diferentes reglas para amenazas en función de su dirección, ángulo y velocidad (Scharre, 2018, 46).

Phalanx es un sistema de armamento de proximidad, desarrollado por la empresa estadounidense *General Dynamics*. Se emplea en buques de guerra como sistema de defensa contra misiles antibuque. El sistema es capaz de detectar, evaluar, seguir y combatir automáticamente objetos rápidos y pequeños como aeronaves tripulados y no tripulados. En todo momento, se trata de un proceso autónomo al que el operador supervisa y en el que interviene, cuando haga falta. El sistema ofrece funciones de evaluación para realizar ataques letales (*kill assessment*) (US Navy, 2019, 15 de enero).

El sistema israelí *Trophy* y el sistema ruso *Arena* son otros ejemplos de sistema autónomos que actúan de forma totalmente autónoma, aunque siempre con supervisión humana (Scharre, 2018, 45-46).

Trophy fue desarrollado por las empresas *Rafael Advanced Defense Systems* y el grupo *ELTA* de *Israel Aerospace Industries* (IAI). Se trata de un sistema de protección activa (*Active Protection System*) que se monta en carros de combate, tanto para su uso en campo abierto como en territorio urbano. El sistema cuenta con sensores (antenas de radar) que detectan misiles y granadas propulsadas (Defensa.com, 2019, 22 de octubre).

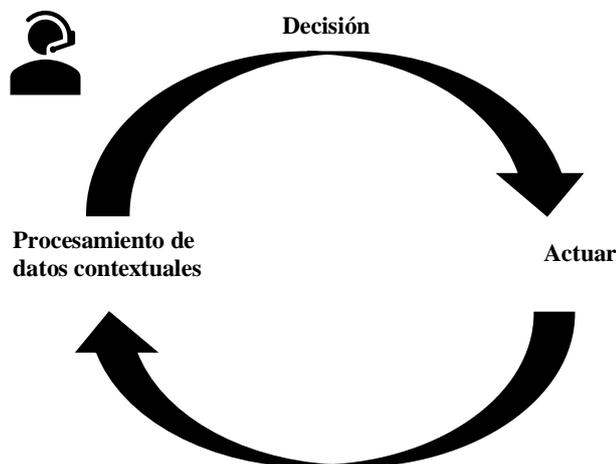
Arena es otro sistema de protección activa, pero de origen ruso, con el mismo objetivo: la protección contra misiles antiblindaje.

En resumen, la autonomía supervisada ofrece algunas ventajas respecto al modo semiautónomo. El sistema puede reaccionar más rápido a ataques porque la persona pasa a otro plano sin obstaculizar el proceso. Esto es especialmente importante para escenarios en el futuro en los que para un ataque se realice en bandada, usando la actuación en bandada (*swarm-technology*). La ventaja de rapidez compensa en este escenario la pérdida de control humano sobre la actuación del sistema, pero, por otra parte, se plantea aquí otra cuestión: ¿Quién es responsable cuando el sistema se equivoca y comete un fratricidio? Scharre se muestra optimista en esta cuestión cuando alude que el hecho que se ha usado el sistema *AEGIS* durante décadas hace pensar que este riesgo es manejable (Scharre, 2018, 325).

3.2.3 *El modo de autonomía completa (Human-out-of-the-loop weapons)*

Los sistemas de armas completamente autónomos funcionan sin el hombre. No hay comunicación entre el arma y el humano, lo que podría ser interesante para operaciones donde el sistema de arma queda incomunicado.

Ilustración 4: La autonomía completa



Fuente: Elaboración propia basada en Scharre (2018)

Israel Aerospace Industries presentó en la feria aérea de París en 2007 el dron *Harpy*, descrito habitualmente como munición merodeadora (*loitering munition*), cuyo objetivo principal es la supresión de defensas aéreas (*SEAD = Suppression of Enemy Air Defenses*). *Harpy* está dotado con una ojiva explosiva de 32 kilogramos y con un sensor

que durante su vuelo escucha permanentemente la señal de radar enemiga. Cuando encuentra una estación de radar, compara la señal percibida con una biblioteca de emisiones hostiles incorporada y en caso de coincidencia, procede al ataque de la estación. Se lanza en vuelo vertical sobre el objetivo y hace explotar la carga explosiva a algunos metros encima de la estación, para causar de esta manera el mayor radio destructivo posible (Aznar, 2020, 23-24). Este sistema ha sido vendido a varios países y China ha producido su propia versión mediante la ingeniería inversa (Scharre, 2018, 5).

En 2019, la misma empresa presentó un sucesor, bajo el nombre *Mini-Harpy*, basado en el mismo concepto de munición merodeadora. En un vídeo promocional, se lanzan dos drones desde una plataforma móvil, un pequeño camión en este caso. Ambos drones rastrean su zona de operación en búsqueda de objetivos. El *Mini-Harpy* está dotado con un sensor de radiación que le permite detectar estaciones de radar. El primer objetivo seleccionado en el vídeo promocional es una estación de radar rusa, usada en los sistemas de defensa aérea S-300 y S-400. El *Mini Harpy* se lanza de forma vertical encima de la estación y hace explotar su ojiva explosiva a pocos metros encima del objetivo. El segundo *Mini-Harpy* ataca una batería antiaérea montada encima de un vehículo todoterreno. En el vídeo, un operador da la orden de atacar al vehículo, pero el sistema impide la explosión cuando detecta a varias personas supuestamente civiles en los alrededores del objetivo. Sin embargo, el dron es capaz de volver a atacar, para lo cual elige un edificio donde se esconden dos personas armadas con un lanzacohetes portátil. En este caso, el ataque se realiza sin la intervención del operador (Israel Aerospace Industries, 2019, 17 de febrero).

Este vídeo es interesante por dos aspectos. Primero, la elección de los objetivos ya no se limita a estaciones de radar. Su capacidad de detectar a personas permite al sistema de atacar de forma directa a combatientes, lo que supone una ventaja para el combate en zonas urbanas. El segundo aspecto es la implicación del humano durante la operación. En el ataque contra la estación de radar es el operador que da la orden de atacar pulsando un botón en su tableta. El segundo dron recibe también la orden del operador, no obstante, aborta el ataque cuando detecta a personas supuestamente no armadas. Tras un reposicionamiento en el aire y la detección de nuevos objetivos – el edificio donde se esconden dos combatientes – el dron realiza su ataque de forma autónoma. Esto significa que en el ataque contra la estación de radar se trata de un ataque semiautónomo, mientras la actuación en zona urbana (abortar el ataque contra el vehículo y atacar el edificio) se

realiza con completa autonomía. Sus sensores le permiten evaluar los posibles daños tras un golpe (*Battle Damage Assessment*).

Otro ejemplo de un sistema que ya ha alcanzado plena autonomía es el robot centinela *Samsung SGR-AI*, desarrollado conjuntamente por la compañía Samsung y la Universidad de Corea. Corea del Sur emplea este robot en la zona desmilitarizada para la vigilancia. El sistema está dotado de cámaras termográficas y es capaz de detectar e identificar a soldados de Corea de Norte y disparar sin intervención humana. Sin embargo, en la actualidad, el robot avisa a un operador humano que decide cómo se va a actuar (NBC News, 2014, 15 de mayo).

3.3 Un área particular: la actuación en bandada

Un área especial, donde se están llevando a cabo muchos proyectos de investigación, es la tecnología en bandada (*swarm technology* en inglés). Este tipo de tecnología, que aún se encuentra en una fase inicial, consiste en la actuación conjunta y sincronizada de una multitud de vehículos no tripulados. Estos vehículos suelen ser en su mayoría vehículos aéreos – drones de tamaño reducido – pero también existen ya proyectos de investigación sobre vehículos acuáticos.⁸

Este modelo de actuación toma las bandadas como referencia, porque la forma de desplazamiento se parece mucho a la de insectos, aves o peces. Al tratarse de grandes números de vehículos, su gestión y control supone nuevos retos. Dado su incipiente estado de desarrollo, muchas de las funciones de actuación en bandada se debaten aún de forma hipotética en las publicaciones. A pesar de ello, se vislumbran algunas posibilidades para su despliegue, tales como la búsqueda de combatientes enemigos o de rehenes de las propias fuerzas, la exploración de edificios en combates urbanos, o la ejecución de ataques de saturación sobre posiciones antiaéreas enemigas. En el ámbito subacuático, su uso podría desarrollarse para la vigilancia de zonas costeras o la exploración de aguas en zonas enemigas. Uno de sus objetivos principales parece ser la actuación en contextos, donde la comunicación con la comandancia de la operación se ha interrumpido por la falta de infraestructuras o por interferencias hostiles, lo que requiere un grado de autonomía alto en la actuación de las bandadas. Esta es la razón por la que muchos actores

⁸ Publicaciones en español usan mayormente el término de enjambre, pero dada la existencia de drones subacuáticos se usa aquí bandada que comprende tanto grupos de aves como bancos de peces.

de la sociedad civil y actores internacionales han dirigido su atención hacia esta tecnología.

Un estudio realizado por el *United Nations Institute for Disarmament Research* (UNIDIR) ha investigado las implicaciones para el control humano sobre las bandadas artificiales. Sus autores, Ekelhof y Persi desarrollaron para ello un modelo que analiza la interacción humana con la máquina a través del comando, las diferentes de arquitecturas de control y una serie de formas de cooperación. El conjunto de estos factores – C3 (comando, control y comunicación) – sirve para aproximarse de forma más analítica a un área de investigación que, dado su estado de desarrollo incipiente y el potencial amplio de uso, aún no se ha explorado de forma satisfactoria.

El trabajo de Ekelhof y Persi parte de la concepción insatisfactoria del control humano significativo (*meaningful human control*), concepto en uso desde el primer encuentro en el marco de la Convención sobre Ciertas Armas Convencionales en 2014. A pesar de usarse ampliamente, los expertos no han podido ponerse de acuerdo sobre el significado exacto del concepto y las publicaciones se limitan normalmente a la acción directa ente el operador y el arma en el momento de ejecutar el ataque. Los autores sostienen que este enfoque es demasiado estrecho porque desatiende varios pasos anteriores a la ejecución del ataque (Ekelhof y Persi, 2020, 12-13). Basándose en los términos usados por la OTAN y Naciones Unidas, los autores describen los términos comando y control de la siguiente manera.

El comando hace referencia a la autoridad individual de las Fuerzas Armadas para la dirección, coordinación y control de las fuerzas militares. Los comandantes pueden asignar misiones o tareas a los comandantes subalternos o fuerzas bajo su control (Ekelhof y Persi, 2020, 13)⁹. El control se refiere a la autoridad de un comandante sobre una parte de las actividades de organizaciones subalternas, u otras organizaciones que normalmente no se encuentran bajo su mando, y que engloba la responsabilidad de implementar órdenes o directivas. El control puede referirse a los niveles operacional o táctico (Ekelhof y Persi, 2020, 13)¹⁰.

Este aspecto es sustancial, porque en la mayoría de las Fuerzas Armadas las variables

⁹ Traducción propia de la definición de Ekelhof y Persi.

¹⁰ Propio resumen de la definición que ofrecen los autores en su informe.

comando y control se desglosan en tres niveles: el estratégico, el operacional y el táctico. En el nivel estratégico, se decide la dirección global y la coordinación de las fuerzas. Además, las consideraciones en este nivel sirven para asesorar a decisores políticos nacionales e internacionales. En el nivel operacional, se trata de transformar los objetivos definidos en el nivel superior en objetivos operacionales para los teatros de operaciones. En un nivel más bajo, el táctico, se concreta el uso específico de las fuerzas militares en las operaciones (Ekelhof y Persi, 2020, 14).

Los autores recalcan esta división del comando y control en los tres niveles para mostrar la complejidad de los procesos decisorios y del control de las acciones en el ámbito militar. El control humano significativo no se limita a apretar un botón porque esta acción se da en un nivel táctico. En los niveles superiores se toma la decisión de desplegar el sistema de armas y se definen las reglas de enfrentamiento.

La comunicación es el tercer factor del concepto C3. La comunicación bidireccional incluye diferentes componentes, desde radios tradicionales hasta complejos sistemas a través de satélite (Ekelhof y Persi, 2020, 17). Curiosamente, los autores descartan para su trabajo este tercer componente y se limitan a analizar los factores comando y control para especificar apropiadamente el concepto del control humano significativo. Sin embargo, la comunicación entre los componentes de una bandada es esencial como se verá más adelante.

Un aspecto importante es el número elevado de los componentes de una bandada. Muchos escenarios están basados en bandadas numerosas, como los modelos en la naturaleza. Ekelhof y Persi definen los componentes de una bandada militar como pequeños robots, dotados con las habilidades de recibir informaciones del entorno, de decidir y de actuar (Ekelhof y Persi, 2020, 24). Estas fases son idénticas con las del bucle OODA que hemos visto anteriormente.

Una bandada, tanto en la naturaleza como en los proyectos militares se caracteriza por la habilidad de actuar como un ente individual, aunque complejo. Se diversifican diferentes funciones, tales como liderar, recabar informaciones sobre el entorno y transmitir órdenes. Pero lo más visible de una bandada es su movimiento en conjunto. Aún no está del todo claro si en la naturaleza existen líderes dentro de las formaciones, o si los animales individuales forman una especie de cerebro común, pero sí se puede postular que todos los componentes de una bandada se mueven siempre orientándose a

los acompañantes más cercanos.

Los proyectos de bandadas militares tratan de emular este comportamiento colectivo. Para ello, hay que desarrollar algoritmos que capacitan a los drones a observarse mutuamente y a comunicarse entre ellos. La comunicación puede desarrollarse usando la señal inalámbrica, Bluetooth, Wi-Fi, radio, LIDAR¹¹, infrarrojo o una mezcla de estas formas, pero todas estas formas de comunicación están expuestas al riesgo de cortarse o de ser interferida por el enemigo (Ekelhof y Persi, 2020, 27). Una vez que los componentes de una bandada tienen la capacidad de observarse mutuamente y sus chips contengan una base de librerías de órdenes incorporadas, las bandadas pueden alcanzar un grado de autonomía que les permita actuar como un conjunto. La mejor organización de la bandada sería la descentralizada, pero el control humano sería el menos eficaz, porque controlar a cada uno de los componentes de forma individual conllevaría un coste material y personal tan alto que la involucración humana ya no tendría sentido.

3.3.1 Arquitecturas de control y comando para las bandadas

La relación entre el hombre y las bandadas se materializa en diferentes arquitecturas de control, cuyas diferencias detallan Ekelhof y Persi en su informe. Conforme las estructuras se vayan descentralizando, la persona pasa a un segundo plano.

La primera de estas estructuras es la centralizada donde la orden humana va dirigida a un robot controlador que distribuye a continuación instrucciones a cada componente individual de la bandada. Ekelhof y Persi comparan esta arquitectura con la del control de tráfico aéreo que regula todo desplazamiento de los aviones desde la posición de aparcamiento del aeropuerto de origen hasta el estacionamiento en aeropuerto del destino. Esta centralización ofrece la ventaja que la transmisión de la orden humana hasta la realización de la tarea dentro de la bandada es directa e inmediata. En cambio, esta estructura es altamente vulnerable y cuando se interrumpe la conexión entre la persona y la bandada, la última se quede inoperable (Ekelhof y Persi, 2020, 35).

En la segunda estructura, se sustituye al robot controlador por un grupo de líderes de escuadrón (*squad leaders* en el texto original), que constituyen un nivel intermedio entre el operador humano y la bandada. Cuando la bandada es demasiado grande para la

¹¹ LIDAR (Light Detection and Ranging o Laser Imaging Detection and Ranging) es un sistema de medición de distancias basado en el reflejo de un haz de luz, similar al rayo de laser que usan los sistemas radar.

transmisión directa, la segunda arquitectura ofrece la ventaja de garantizar la comunicación entre el operador y los robots individuales. Este modelo ofrece mayor estabilidad organizacional para bandadas pequeñas, según informa un experto anónimo al que entrevistaron los autores (Ekelhof y Persi, 2020, 36).

La tercera estructura es un ente con relativa independencia del control humano. El operador da la orden a la bandada en conjunto, tras lo cual los robots individualmente deciden sobre cómo llevar a cabo esta orden. Es una estructura descentralizada, y por la ausencia de un controlador central, no existe un punto vulnerable para fallos. Este tipo de organización estaría pensado para entornos complejos que requieren un alto grado de adaptabilidad y cuando el ancho de banda para la comunicación entre los diferentes elementos sería muy limitado. Al ser una arquitectura descentralizada, su vulnerabilidad es menor que la del control centralizado o jerárquico (Ekelhof y Persi, 2020, 37).

En la cuarta arquitectura cada robot individual tiene una librería de comportamientos preprogramados que se activan cuando la bandada en su conjunto recibe la orden humana. Esta arquitectura con el nombre control conductual (*behavioural control* en inglés) es también una estructura descentralizada y ofrece, por tanto, sus ventajas características, especialmente cuando la comunicación es muy restringida o interrumpida. Las librerías de comportamientos pueden comprender diferentes acciones, tales como la búsqueda de objetos o personas y su persecución, y también la fijación como blanco para su neutralización. Los autores comparan el control conductual con los procedimientos operativo estandarizado (POE), que usa el personal militar en operaciones rutinarias para asegurar resultados eficientes y uniformes (Ekelhof y Persi, 2020, 38). Esta arquitectura operaría de forma completamente autónoma, y, por tanto, sin control humano.

El Instituto Tecnológico de Massachussets desarrolló el proyecto *Perdix*, que consiste en una bandada de pequeños drones de tamaño de seis pulgadas. Se lanzan desde el dispensador de bengalas de un caza y se dirigen posteriormente de forma autónoma hacia posiciones enemigas (Gómez, 2019, 132-133). En un experimento en *China Lake* en California se lanzaron un total de 103 de estos drones desde dos cazas F-18 que se movieron después como un enjambre hacia su destino. Según la fuente *AIIR Source Military*, el objetivo de estos drones es confundir las defensas enemigas y bloquear señales de radar (AIIR Source Military, 2017, 10 de enero). Ekelhof y Persi, en cambio, definen como principal objetivo de *Perdix* realizar operaciones inteligencia, vigilancia y

reconocimiento (Ekelhof y Persi, 2020, 50). El vídeo de *AIIR Source* muestra solamente el lanzamiento desde los cazas F-18 y el seguimiento de los drones en un área acotado, por lo que cabe suponer que el objetivo de esta prueba era simplemente el comportamiento en formación.

El programa *Gremlins* es uno de los proyectos que gestiona la agencia DARPA. Como en el caso de *Perdix*, estos UAV también pueden ser lanzados desde otros aviones. A diferencia de los *Perdix*, los *Gremlins* serían recuperables y se podrían preparar en un espacio de 24 horas para su nuevo despliegue. El lanzamiento y la recogida en aire se realizaría con un transportador aéreo como el C-130 (DARPA, s. f.).

Ilustración 5: Lanzamiento y recogida de los Gremlins por transportadores C-130



Fuente: DARPA

Ambos proyectos serían ejemplos de MUM-T (*Manned-Unmanned Teaming* = formación de equipos tripulados-no tripulados), que ofrecen la ventaja de proteger a los efectivos militares porque se acercan solamente hasta un determinado punto en la zona de operaciones, mientras los vehículos no tripulados se adentran en la zona.

La Unión Europea ha lanzado también un proyecto de desarrollo dentro del Programa Europeo de Desarrollo Industrial en materia de Defensa (EDIDP por sus siglas en inglés). El programa *Dronedge-E*, bajo el liderazgo francés y la participación de Eslovenia, Bulgaria y Alemania, trata de diseñar una plataforma de computación para el control

autónomo de drones en tiempo real que permita la generación de algoritmos automatizada con la ayuda de la inteligencia artificial. Según la descripción del proyecto, el objetivo es el desarrollo de una arquitectura de comando novedosa. Este proyecto, dotado con 1.949.439 de Euros, tiene una duración de 24 meses desde su iniciación (European Commission, 2020).

3.4 Conclusión parcial

Esta primera parte de la investigación ha intentado aclarar la diferencia entre sistemas de armas automatizadas y armas autónomas. Basado en la definición de Russel y Norvig que conciben la inteligencia artificial como un agente racional, se ha ampliado este concepto con los rasgos adaptabilidad, la capacidad de aprender y la autonomía en el desempeño de tareas. Un arma autónoma letal es un agente racional, dotado con inteligencia artificial que le proporciona la capacidad de adaptarse a contextos cambiantes, de aprender durante el ciclo de su existencia y de realizar tareas con autonomía.

El criterio de inteligencia artificial como agente racional ofrece también la opción de contrastar los diferentes grados de autonomía en el bucle OODA y la posición del humano dentro de este bucle, desde el modo semiautónomo, hasta la autonomía supervisada y la autonomía completa. La mayoría de los sistemas de armas funcionan en la actualidad de forma semiautónoma, donde el operador humano toma la decisión de atacar, y, de una forma más sofisticada, la autonomía supervisada, en la que el arma ya toma estas decisiones bajo la supervisión del hombre. A pesar de estas limitaciones, algunas armas como *Harpy*, *Mini-Harpy* y el *SGR-AI*, ya están diseñados para actuar con plena autonomía, si bien en la actualidad no se hace uso de este potencial.

Las bandadas de robots suponen un nuevo hito en el campo de armas autónomas. Su potencial de actuar con considerable autonomía los convierte en una herramienta con mucho potencial, incluso letal, cuando su función es la de atacar. Las bandadas ofrecen muchas ventajas, entre las que la posibilidad de actuar sin arriesgar la vida del propio soldado es indudablemente la mayor fuerza motriz para su desarrollo.

La diferenciación entre los diferentes niveles para el comando y el control muestra la complejidad de delimitar la autonomía en los sistemas de armas. Por una parte, esta complejidad supone un reto para la cuestión de si se regulan estas armas o si se prohíben, como reivindican algunos actores. Por otra parte, el potencial de autonomía ofrece unas

ventajas que favorecerían su empleo, razón por la que muchos países han impulsado su desarrollo. Ambos aspectos se analizarán en los siguientes capítulos.

4 IMPLICACIONES PARA LA SEGURIDAD INTERNACIONAL

El segundo debate tiene como objeto las implicaciones de los sistemas de armas autónomas para la seguridad internacional. Como se ha visto, dotar los sistemas de armas brinda una serie de ventajas que los hace muy atractivos para los Estados. Las implicaciones que se analizan en este capítulo se concentran en dos áreas: la carrera armamentística y la proliferación. En este capítulo se trabaja con la hipótesis que los mayores esfuerzos en el desarrollo de sistemas de armas autónomas acarrearán nuevos riesgos para la seguridad internacional. Estos riesgos se expresarían en una nueva carrera armamentística entre las grandes potencias y un mayor riesgo de proliferación de tecnología entre Estados rebeldes y grupos subestatales.

4.1 Una nueva carrera armamentística

Kasperson advirtió en 2016 en un artículo para el Foro Económico Mundial de una nueva carrera armamentística que tendría lugar en el campo de la inteligencia artificial, pero la autora aún consideraba la posibilidad que se pudiese contener esta carrera (World Economic Forum [WEF], 2016, 15 de junio). Su propuesta consiste en crear una nueva plataforma global cuyo objetivo sería monitorizar todos los desarrollos en el área de la inteligencia artificial y formular recomendaciones sobre las implicaciones de tecnologías emergentes para la seguridad internacional (ibid.). De momento, es la plataforma Inteligencia Estrategia y Robótica (*Strategic Intelligence and Robotics* en inglés) que ejerce esta función, bajo el patrocinio del Foro Económico Mundial y la gestión de la Universidad de Cambridge (WEF, s. f.). Sin embargo, parece que este proyecto no ha frenado la evolución en absoluto porque no solamente las grandes potencias, sino también otros países como Israel, Corea del Sur y Turquía han incrementado sus esfuerzos.

Cuatro años después de la publicación del artículo de Kasperson, esta carrera se está perfilando con más dinamismo que antes. El informe *Global Risk Report 2020*, del mismo organismo, advierte de nuevo de este peligro. El documento cita al presidente ruso Putin que pronosticó ya en 2017, que él que liderara en el área de la inteligencia artificial, dominaría el mundo. (WEF, 2020, 14). China ha elevado esta área como prioridad para la seguridad nacional, mientras Estados Unidos ha triplicado el presupuesto para el Centro Común de Inteligencia Artificial (*Joint Artificial Intelligence Center*) del Departamento

de Defensa, llegando a un importe de 268 millones de dólares estadounidenses. La Administración de Trump justificó esta decisión con el desarrollo imparable de las capacidades en IA en China y Rusia (WEF, 2020, 14).

Como en los tiempos de la Guerra Fría, los decisores políticos se encuentran en un nuevo escenario de incertidumbre sobre las intenciones de otros países. Este problema, conocido como el dilema de seguridad, enfrenta a los gobiernos con la necesidad de evaluar apropiadamente la evolución en el desarrollo y la adquisición del armamento de los países contrincantes, sin conocer exactamente las verdaderas intenciones de sus decisores. John Herz, experto en Relaciones y Derecho Internacionales, desarrolló este concepto en su libro *Political Realism and Political Idealism*, de 1951, para describir la situación de enfrentamiento entre dos bloques que estaban emergiendo en la incipiente Guerra Fría. El entorno internacional se caracteriza por un estado anárquico, donde los países están en un permanente estado de preocupación por su seguridad. La trampa de Tucídides es el ejemplo paradigmático de un conflicto entre una potencia establecida que observa con reojo el surgimiento de un nuevo competidor. Tucídides describió así el conflicto entre Esparta y su nuevo contrincante Atenas que acabó inevitablemente en una guerra entre los dos.

En la actualidad, los decisores políticos en Estados Unidos se enfrentan a un escenario similar, en el que Rusia y China tratarían de arrebatarse el dominio. La única manera de escapar del dilema de seguridad sería facilitar transparencia y controles mutuos como en el caso de los tratados bilaterales del repliegue nuclear que se acordaron entre Estados Unidos y la antigua Unión Soviética. En la actualidad, en la lucha por la hegemonía, se ha apuntado China como nuevo actor mundial, mientras Rusia se ha erguido de nuevo, buscando reanudar el papel que jugaba anteriormente la Unión Soviética. Esta competición por la hegemonía se despliega nuevamente en los dominios armas convencionales y nucleares, pero además en las nuevas tecnologías, donde la inteligencia artificial juega un papel destacado.

Klare señala que esta carrera armamentística ha ganado velocidad. Las grandes potencias Estados Unidos, Rusia y China han intensificado sus esfuerzos en el área de las nuevas tecnologías, como muestran los documentos de estrategia emitidas por estos países evidencian esta tendencia (Klare, 2019a). El desenlace, bien un sistema multipolar compuesto por las tres potencias geopolíticas mundiales de ahora, bien otro sistema

multipolar, en el que países más pequeños comparten el dominio con las grandes potencias, dependerá en parte de los progresos en el desarrollo de la inteligencia artificial en las armas autónomas dentro de los próximos años.

4.2 La proliferación

Gaub señala en el informe *Global Trends to 2030*, que la disciplina de previsiones nació del deseo de saber dónde, cuándo y cómo se desarrollarían guerras en el futuro, citando a Lawrence Freedman (Gaub, 2019, 25). A pesar del interés que suscita esta disciplina, la autora sostiene que las predicciones no han sido siempre muy certeras. Esto se debe sobre todo a nuestra tendencia de sobrevalorar el progreso tecnológico, en detrimento de la complejidad de otras claves para el desenlace de los conflictos. En cambio, subestimamos la letalidad de sistemas de armas poco sofisticados (Gaub, 2019, 25). Es cierto que los avances tecnológicos albergan el riesgo de servir de patrón para previsiones de conflictos, pero esto depende también de los costes de implantación que suponen estos avances. La adquisición de un avión de combate de la quinta generación dotado con los últimos avances tecnológicos conlleva un coste que muchos países no pueden asumir. Si, además, tenemos en cuenta que la gran mayoría de los conflictos en la actualidad, son conflictos intraestatales, disponer de armas de última generación no estará al alcance de muchos actores. Gaub espera que el número de guerras intraestatales permanecerá o se incrementará dentro de los próximos años. En la actualidad son unos 40 conflictos en curso, impulsados por factores como el cambio climático, desigualdad, represión estatal, entre otros (Gaub, 2019, 25).

Es cierto que un submarino o un vehículo aéreo no tripulado de la categoría del *Global Hawk* no estarán al alcance de la mayoría de pequeños grupos armados, pero el tráfico ilegal con armas pequeñas y ligeras muestra que la facilidad de transporte podría ser una ventaja comparativa frente a sistemas de armas pesadas y de grandes dimensiones. Y la tecnología de pequeños drones está accesible para cualquier persona, sea por la compra en un mercado de ocio, o por fabricación propia. En la red se pueden encontrar instrucciones y tutoriales para construir sus propios drones.¹² De esta manera, atentados como aquel que perpetraron rebeldes hutíes el 10 de enero de 2019 en un desfile militar en Yemen, son fácilmente reproducibles. La tecnología de impresión 3D permite producir

¹² El uso de dos búsquedas devuelve múltiples resultados en la plataforma Youtube. Los términos de búsqueda eran *diy (do it yourself) drone* e *improvised explosive device*.

pequeños drones a costes muy bajos, cuando se quiere disponer de un mayor número de ellos. Mediante un software instalado en sus chips y con un sistema de comunicación como *bluetooth*, estos drones podrían formar una bandada para guerra electrónica o incluso para lanzar ataques sincronizados.

Hoy en día, se pueden encontrar múltiples recursos en la red, y muchos de ellos de forma legal, bajo la licencia de código abierto. Scharre menciona como ejemplo la plataforma *TensorFlow* que ofrece librerías con algoritmos para el desarrollo de redes neuronales, así como tutoriales para iniciarse en la su programación. El material y las tutoriales proceden de la compañía Google que publicó en esta plataforma los resultados de sus propias investigaciones en redes neuronales en 2015, para el acceso legal a todos interesados (Scharre, 2018, 128-129). Uno de los usuarios más notorios es el Departamento de Defensa de Estados Unidos, que usó *TensorFlow* dentro del proyecto *Maven*, mencionado en el capítulo tres, y cuyo objetivo era asistir en el reconocimiento automatizado de posibles blancos en imágenes obtenidas por drones de reconocimiento (The Guardian, 2018, 7 de marzo).

Tanto estas tecnologías no tangibles como los componentes para la fabricación son fáciles de adquirir en el mercado, porque en muchos casos se trata de componentes que se usan en otros productos. Un teléfono inteligente de las últimas generaciones dispone de componentes que se pueden usar también para pequeños drones, fabricados en talleres mediante una forma fabricación aditiva moderna, la impresión 3D. Esta forma de fabricación se ha extendido durante esta década sobre todo para la elaboración de prototipos en la industria, pero también para la producción de piezas de recambio. Ofrece la posibilidad de elaborar piezas únicas a base de un plano, pero también facilita la imitación de casi cualquier pieza, cuando se usa un equipo de cámaras y un software para el tratamiento de imágenes para la medición de la pieza. Esta versatilidad le convierte en una tecnología de doble uso, que puede ser empleada tanto para usos civiles como usos militares. El control de estas tecnologías y productos se muestra como algo muy difícil porque su prohibición total impediría también su uso pacífico.

También el conocimiento teórico y su aplicación pueden compartirse de forma totalmente legal. El acceso libre a ello es un rasgo identificativo de sociedades abiertas y uno de los impulsores del progreso. Empresas y universidades forman parte de un tejido donde la información fluye de forma continua entres sus componentes. Están disponibles

también, aunque en menor grado, para otros actores de sector privado, tales como organizaciones y personas particulares. El transfer del conocimiento teórico y de sus aplicaciones puede producirse a través de manuales de instrucciones, publicaciones académicas, cursos, foros dedicados o el contacto personal, entre muchas otras. Bromley and Maletta (SIPRI, 2018) diferencian en un informe dedicado al transfer de tecnología y software entre el transfer de tecnologías tangibles y el de tecnologías no tangibles. Contrario a muchos productos de doble uso sujetos a controles de exportación de armas, la tecnología basada en algoritmos puede distribuirse de forma tangible o no tangible. La exportación de un ordenador o de un CD-ROM corresponde al transfer de bienes tangibles, mientras que la transferencia intangible de tecnología (*intangible transfer of technology* en inglés) podría producirse por el envío de datos técnicos por correo electrónico, por descargas y subidas a servidores y a través servicios de procesamiento de datos en la nube (Bromley y Maletta, 2018, 5).

El procesamiento de datos en la nube (*cloud computing* en inglés) es un negocio que ha sido posible gracias a enormes inversiones por proveedores grandes, principalmente Amazon Web Services, Google y Microsoft, pero existen empresas más pequeñas que ofrecen ahora el almacenamiento legal de datos de cualquier tipo, cuando se contrata o adquiere uno de sus servicios. Uno de los grandes problemas es el control de los datos que se almacenan en la nube, y con quién se compartan estos datos. Los operadores de telecomunicaciones se sustraen de la responsabilidad penal con las cláusulas en las condiciones de uso, según las cuales el usuario se compromete a no compartir contenidos ilegales, pero una vez que el usuario los haya aceptado con su confirmación y decida compartir instrucciones para construir drones y artefactos explosivos improvisados, nadie podrá controlar esta acción. La nube como depósito de tecnología es prácticamente inaccesible a controles de exportación en sociedades abiertas.

El carácter asimétrico de muchos conflictos actuales conlleva que muchos actores implicados son no estatales. El Estado Islámico, Boko Haram o las milicias agrupadas en torno al Ejército Nacional Libio (ENL en adelante) son unos pocos ejemplos de grupos no estatales que compensan su falta de armamento pesado con tácticas de guerrilla en las que los drones juegan un papel cada vez más importante. Según un informe de *Defense One*, el Estado Islámico ha utilizado drones de forma creciente. Los miembros del EI han usado diferentes tipos de artefactos de propia fabricación, incluyendo cuadricópteros para el ocio, adquiribles en cualquier negocio dedicado. Se han usado estos artefactos para la

vigilancia desde el aire como para el lanzamiento de granadas y morteros (Watson, 2017, 12 de enero).

El 14 de septiembre de 2019, las instalaciones de la empresa petrolera Saudi Aramco en Abqaiq sufrieron ataques con un total de 15 misiles y 10 drones. El grupo Huti reivindicó la autoría de los ataques. El año anterior, los hutíes afirmaron haber usado drones del modelo Quasef-1, de origen iraní, para atacar las instalaciones de la misma compañía en Jazan. (Middle East Eye, 2018, 11 de abril). Este dron es una variante de munición merodeadora de producción doméstica.

Según el corresponsal Osama Bin Javid, estos drones escapan de los sistemas de defensa aérea por su tamaño reducido y porque no producen suficiente calor para la detección con sensores de calor (Al Jazeera, 2019, 14 de septiembre).

Otros actores no estatales interesados en adquirir sistemas de armas autónomas podrían ser las compañías de seguridad, tales como el grupo ruso *Wagner*, o la empresa *Academi*, sucesora de la antigua *Blackwater*. Como en el caso de milicias, estos actores operan a menudo al margen de la legalidad internacional.

El analista Patrick Tucker pronosticó ya en 2014 que dentro de un espacio de 10 años todos los países dispondrían de las tecnologías de drones armados (Tucker, 2014, 6 de mayo). Zeng Yi, el alto directo de *Norinco*, afirmó que en 2025 el uso de armas autónomas letales sería la normalidad.¹³ Dos de las tres grandes potencias que se han analizado anteriormente, son desarrolladores de drones armados. A estos países se añadirían otros como Israel o Turquía que también han hecho grandes inversiones en el desarrollo de nuevas armas. El analista José Luis Calvo informa que Turquía envió a principios de 2020 las primeras tropas a Libia, donde apoya al único gobierno reconocido internacionalmente, el Gobierno de Acuerdo Nacional. La novedad es que Turquía está usando dos drones armados de propia fabricación: el *Bayraktar TB2* y el *Anka-S*. Las tropas turcas aprovechaban la ausencia del Ejército Nacional Libio en el espacio aéreo libio para la supresión de sistemas antiaéreos del ENL. Aunque Turquía perdió varios aparatos, la pérdida es más asumible que la de aviones tripulados (Calvo, 2020). Cabe señalar que Rusia, que apoya al ENL, ha respondido a este envite con el envío no oficial de varios cazabombarderos *MIG-29* y *SU-24*, lo que muestra una vez más que el

¹³ Véase el epígrafe dedicado a la posición oficial de la República Popular China.

desarrollo de drones de combate propios aún está en los inicios y que los sistemas de contramedidas, principalmente la defensa antiaérea *Pantsir*, tampoco son tan eficaces contra la supremacía aérea de los drones turcos (CNN, 2020, 27 de mayo).

Muchos de los países que invierten en el desarrollo de nuevas armas son a la vez exportadores. Los beneficios de la venta se invierten de nuevo en la continuación del desarrollo. Para algunos países exportadores, el sistema político y la existencia de reglas democráticas o su falta no interesan a la hora de vender. De esta manera, países cuyas importaciones han experimentado restricciones tienen acceso a sistemas de armas sin la necesidad de desarrollarlos ellos mismos. Fuhrmann y Horowitz señalan que más de dos docenas de países poseen ya drones armados (Fuhrmann y Horowitz, 2017, 397–398). Además, de los tres sistemas políticos que estos investigadores han relacionado con la posesión de drones armados, las autocracias han mostrado mayor predisposición para la adquisición de esta clase de plataformas, cuando no las desarrollan ellas mismas. La probabilidad pronosticada para las autocracias se sitúa en 25 por ciento, frente a 18 en el caso de democracias (ibid., 411). Este es un dato preocupante porque el control democrático es limitado cuando no inexistente en las autocracias.

La exportación no controlada de armas a Estados rebeldes puede provocar la escalada de un conflicto en curso. Tucker informa que varios fabricantes chinos exportan drones con capacidad letal, que operan con poca o ninguna supervisión humana. Tucker se refiere a afirmaciones que hizo el Secretario de Defensa Mark Esper al respecto. Según el Secretario de Defensa, los fabricantes chinos promocionan sus drones con la capacidad de actuar con plena autonomía, incluyendo la capacidad de realizar ataques dirigidos letales (Tucker, 2019, 5 de noviembre).

4.3 Las posiciones oficiales de las grandes potencias

En esta parte del trabajo se analizará la posición de la inteligencia artificial en las estrategias oficiales de tres grandes potencias: China, Estados Unidos y Rusia. Mearsheimer, quien desarrolló la teoría del realismo ofensivo, sostiene que las grandes potencias tratarán siempre de ampliar su dominio para evitar que otros competidores lo hagan. Mearsheimer se distancia en este punto de Waltz quien sostiene que los poderes aceptarían el statu quo en un determinado momento, si este estado garantizara la estabilidad en el sistema internacional. La teoría del realismo ofensivo, de la que el dilema de seguridad es uno de los componentes centrales, sirve de trasfondo analítico para

formular la siguiente pregunta: ¿Conducen los esfuerzos en la investigación de inteligencia artificial a una nueva carrera armamentística entre las grandes potencias? Se plantea además una segunda pregunta: ¿Conducen estos esfuerzos a un enfrentamiento ineludible entre China y Estados Unidos, si se confirma la hipótesis de la carrera armamentística en la inteligencia artificial?

Un estudio de *Absolute Market Insights* ha realizado un estudio sobre el volumen de mercado de robótica militar para 2027, una década más tarde para el año base del cálculo. El estudio partió de un mercado de 17,34 mil millones de dólares estadounidenses en 2018 y se espera alcanzar 53,93 mil millones de dólares en el año 2027, sobre una tasa de crecimiento anual compuesto de 13,5 por ciento. Este incremento tan espectacular se debe a la integración de la inteligencia artificial en la robótica militar (Cision, 2020, 6 de enero).

Este dato ilustra el potencial tecnológico y económico de la inteligencia artificial en los próximos años, y de lo difícil que podría ser para todos los actores de sustraerse de esta evolución.

4.3.1 La posición de la República Popular China

La mayoría de los expertos coincide en pronosticar mayor protagonismo chino en el área de las nuevas tecnologías. El ritmo al que China está alcanzando el nivel tecnológico en áreas como la electrónica de consumo y tecnologías de punta, es visible para observadores atentos.

McDonald y Howell presentan una imagen actualizada de la competición entre las grandes potencias en vehículos aéreos no tripulados y la inteligencia artificial. Para su estudio han clasificado los sistemas de armas en nueve diferentes categorías. La siguiente tabla refleja el estado actual de las plataformas desarrolladas y los proyectos que aún se encuentran en desarrollo. Se han tenido en cuenta solamente las participaciones de Estados Unidos y China a nivel global, para ilustrar la competición entre ellos. Ambos países están representados en todas las categorías, mientras Rusia, Israel, Turquía, Letonia, Corea del Sur, Francia e India participan, pero no en todas las categorías. Israel ostenta en la categoría munición merodeadora con el 40 por cien el liderazgo mundial, pero no se ha incluido en esta tabla por la misma razón.

Tabla 3: Desglose de vehículos no tripulados en uso y en desarrollo

Categoría	Desarrollado		En desarrollo	
	Participación global en %		Participación global en %	
	Estados Unidos	China	Estados Unidos	China
UAV de combate	37	31	Datos no disponibles	Datos no disponibles
UAV de alturas y rango altos	67	33	55	33
UAV de alturas medias y rango alto	15	26	12,5	12,5
UAV tácticos	33	8,3	50	Datos no disponibles
Sistemas aéreos tácticos no tripulados pequeños	34	10	Datos no disponibles	Datos no disponibles
Munición merodeadora	12,5	12,5	Datos no disponibles	Datos no disponibles
UAV de ala giratoria grandes	15	35	Datos no disponibles	Datos no disponibles
UAV de ala giratoria pequeños	51	Datos no disponibles	Datos no disponibles	Datos no disponibles
UAV de carga y transporte	51	12.5	70	20

Fuente: Elaboración propia basada en McDonald y Howell, 2019, 105-113.

Es muy llamativa la poca distancia en cuanto a representación global entre Estados Unidos y China, mostrando incluso una clara ventaja en las categorías UAV de alturas medias y rango alto y UAV de ala giratoria grandes. Esta imagen probablemente se ajusta más a la brecha tecnológica cada vez menor entre Estados Unidos y China. Gilli y Gilli analizaron en un estudio la brecha tecnológica entre ambos países que, según ellos no se recortaría simplemente aplicando ingeniería inversa. Para su estudio compararon el desarrollo de buques de guerra de Gran Britania y la Alemania imperial, y el estado de desarrollo de cazas de la quinta generación entre Estados Unidos y China. Tienen razón cuando concluyen que para Alemania era más fácil imitar los buques británicos y que, para China, es más difícil imitar los cazas de quinta generación estadounidenses (Gilli y Gilli, 2019, 187). China tiene dificultades de alcanzar el alto nivel tecnológico de Estados Unidos, pero en las categorías de vehículos no tripulados que requieren menos hardware sofisticada, pero más inteligencia artificial, la distancia entre ambas potencias se está acortando.

No sin razón, los autores McDonald y Howell titulan su estudio “Matándome

suavemente”¹⁴, para ilustrar el carácter lento pero imparable de este proceso. Sirva como ejemplo el dron de ala giratoria *Blowfish*, del fabricante chino Ziyan (Ziyan, s. f.). Sus tres modelos desarrollados *Blowfish A1*, *A2* y *A3* son pequeños helicópteros dotados con un cañón que, por su tamaño, se presta para operaciones en zonas urbanas. El modelo *A2* tiene la capacidad de transportar una carga explosiva de 12 kilogramos y de soltarla a una zona de impacto muy reducida desde una altura elevada. Según informaciones del fabricante, este modelo es capaz de ejecutar misiones de combate más complejas, que incluyen la detección de blancos y ataques de precisión.

La empresa ha vendido este modelo a Emiratos Árabes Unidos e inició en noviembre de 2019 negociaciones con Arabia Saudita y Pakistán (Allen, 2019, 7).

Ilustración 6: El modelo *Blowfish A2*



Fuente: Developing Pakistan (s. f.)

Kania ofrece una observación más detallada del papel que tiene la inteligencia artificial en las innovaciones militares chinas. En su último informe explica como las autoridades chinas conciben los sistemas de armas modernos. Kania interpreta como una buena señal que entre los científicos e investigadores militares chinos esté empezando un debate sobre temas relacionados con la seguridad, así que sobre consideraciones legales y éticas (Kania, 2020, 1). Una de las principales observaciones de este informe es que los teóricos chinos no se plantean la cuestión de autonomía en los sistemas de armas de la misma manera como lo hacen sus homólogos en occidente. Los comunicados oficiales

¹⁴ El título original es *Killing me softly - Competition in Artificial Intelligence and Unmanned Aerial Vehicles*.

chinos son ambiguos y rehúyen un posicionamiento claro sobre las armas autónomas. Por ejemplo, en octubre de 2018, la delegación china expresó en una discusión temática sobre el control de armas convencionales de la Asamblea General de Naciones Unidas, que «considera necesario llegar a un acuerdo legalmente vinculante sobre armas letales con plena autonomía para prevenir la aparición de máquinas automáticas asesinas (Kania, 2020, 5).»¹⁵ La delegación china repitió aquí su concepción de sistemas de armas letales con plena autonomía que había presentado un año antes en la reunión del Grupo de Expertos Gubernamentales en Sistemas de armas autónomas Letales.

El énfasis en la plena autonomía es lo que por lo visto justifica la necesidad de exigir un instrumento legalmente vinculante. Kania recalca que en la actualidad parece que los militares no quieren prescindir del control humano (Kania, 2020, 5). Como se ha visto en el capítulo 3 de este trabajo que, aunque los sistemas *SGR-A1* y *Harpy* son dos ejemplos que ofrecen esta opción, ambos operan siempre en modo semiautónomo, manteniendo el control humano durante todo el proceso.

¿Cómo es posible que los representantes chinos defienden por una parte un acuerdo legal sobre lo que ellos conciben como máquinas automatizadas asesinas, mientras por otra parte el país está trabajando en alcanzar a su gran competidor en el desarrollo tecnológico de armas? Kania ha observado que las publicaciones chinas tratan el tema de las nuevas tecnologías de forma diferente a las publicaciones occidentales. En las publicaciones chinas predomina el concepto de informatización militar.¹⁶ Interpretándolo así, se podría concluir que los militares chinos no tienen ninguna intención de ceder el control de un operador humano en el bucle OODA y que, por tanto, no tienen grandes inconvenientes de exigir un instrumento legal para la autonomía plena. Nadie podría criticarles por incumplir sus propias premisas cuando no tienen intención alguna de dotar sus sistemas de armas con plena autonomía.

La investigadora recuerda la falta de experiencia operacional del Ejército Popular de Liberación. Carece especialmente del conocimiento de primera mano de la “niebla de la guerra”, lo que podría conducir a errores y expectativas irreales sobre las perspectivas de

¹⁵ La traducción es propia.

¹⁶ Kania traduce el concepto original al inglés, llamándolo *military intelligentization*. Traducirlo como informatización militar parece lo más indicado cuando se entiende que este concepto involucra el uso de inteligencia artificial en los niveles estratégico, operacional y táctico.

la tecnología en el campo de batalla (Kania, 2020, p. 6). Estados Unidos tiene gracias a sus numerosas intervenciones militares una ventaja, sobre todo cuando se trata de subsanar errores fatales con daños colaterales, tales como los fratricidios que ocurrieron durante la operación Tormenta del Desierto con el sistema de defensa antiaérea *Patriot*. El posterior análisis de los fratricidios trajo importantes mejoras en los sistemas de identificación (Scharre, 2018, 164-171). Las Fuerzas Armadas chinas no cuentan con el campo de batalla como laboratorio.

Kania señala que los expertos chinos conocen las publicaciones occidentales sobre esta materia en detalle, algo que no ocurre al revés. Por eso, Kania recomienda atender más a las publicaciones chinas y traducirlas sin excepciones (Kania, 2020, 8).

El Plan de Desarrollo de la Inteligencia Artificial (*AIDP* en sus siglas en inglés) de la próxima generación da pistas provechosas sobre las ambiciones del gobierno chino en el campo de inteligencia artificial.¹⁷ El documento, publicado en julio de 2017 por el Consejo de Estado chino admite, que se tienen que acometer aún algunos esfuerzos para alcanzar el nivel de los países desarrollados. La brecha se da sobre todo en la teoría y algoritmos básicos, así como en partes del hardware, como chips de gama alta y materiales (ADIP, 2017, 4). Pero China lideraría ya el mercado global en la tecnología de reconocimiento de voz y el reconocimiento visual. Los objetivos para los próximos trece años, desde la publicación del plan, son ambiciosos.

Para el año 2020, el Consejo de Estado prevé que el nivel general de tecnología y aplicación de la IA habría alcanzado niveles avanzados globales, y que la industria de la inteligencia artificial se estaría convirtiendo en un impulsor de crecimiento económico importante, facilitando que China entrara en la categoría de las naciones más innovadoras que permita a la vez alcanzar el objetivo de una sociedad moderadamente próspera (AIDP, 2017, 5). En el contexto de este trabajo de investigación hay cuatro campos importantes en los que China quiere alcanzar progresos importantes: macrodatos, inteligencia de bandadas, inteligencia híbrida (integración humano-máquina) y sistemas de inteligencia autónoma. (AIDP, 2017, 6).

Para el año 2030, China quiere haber alcanzado el liderazgo global en el campo de la

¹⁷ Las informaciones son traducciones del documento *A Next Generation Artificial Intelligence Development Plan*, traducido este mismo del chino al inglés por un grupo de trabajo de New America. El documento original fue publicado en julio de 2017 por el Consejo de Estado.

inteligencia artificial, en concreto en las áreas de inteligencia inspirada en el cerebro humano, además en las áreas ya mencionadas antes. Pero no solamente esto: China prevé para este año haber creado regulaciones y leyes en IA más generales, así como un sistema de normas éticas y políticas (AIDP, 2017, 6-7).

Los objetivos son ambiciosos, pero los miembros del Consejo de Estado son también conscientes que su sistema político y la gestión centralizada les otorgan una ventaja comparativa respecto a sus competidores occidentales. El documento menciona repetidas veces que el camino hacia el liderazgo global en el área de la inteligencia artificial está vinculado al progreso de la sociedad china. Asimismo, se recalca varias veces la preparación de infraestructuras industriales (clústeres) y de la fuerza laboral para alcanzar sus objetivos. China dejará definitivamente de ser el taller para el mundo desarrollado y se convertirá en el líder global en inteligencia artificial y el hardware. El documento menciona también la aplicación de la IA en el campo militar, pero nunca se limita a ello. Todo el progreso en la IA tendrá lugar también en la sociedad civil.

La lectura que hace Allen de la estrategia china está más relacionada con la seguridad. Este investigador publicó en febrero de 2019 el informe *AI Weapons in China's Military Innovation*. En él confirma la impresión de la investigadora Kania de un debate inicial sobre los riesgos relacionados con el uso de armas con mayor autonomía. Allen acudió a cuatro conferencias en China donde se confirmó su impresión. Las conferencias eran puntos de encuentro para los sectores diplomáticos, militar y privados, todas ellas dedicadas a la inteligencia artificial. El resultado de su participación en las conferencias es el informe, en el que detalla en 16 puntos las posiciones académicas y la actuación oficial en el campo de la IA.

Allen cita a Fu Ying, vicepresidenta del Comité de Asuntos Exteriores del Congreso Popular Nacional que exigió en una charla que China debería mostrarse más cooperador para prevenir de la amenaza de la inteligencia artificial, un punto de vista que por lo visto comparten muchos otros diplomáticos y teóricos de laboratorios de ideas chinos (Allen, 2019, 5). Una de las mayores preocupaciones entre estos expertos es que el mayor uso de sistemas con IA podría facilitar interpretaciones erróneas y una escalada de conflicto no intencionada, debido a la falta de normas bien definidas (Allen, 2019, 5-6).

Otra institución, la Academia China de Tecnologías de Información y Comunicaciones (*China Academy of Information and Communications Technology, CAICT*) advirtió en

el *Libro Blanco sobre Seguridad en la Inteligencia Artificial* al gobierno chino de evitar una carrera armamentística en IA entre países (Allen, 2019, 6).

El hecho de que los representantes de algunas instituciones académicas se expresen de forma tan sincera dice probablemente menos sobre nuevos hábitos democráticos, sino que más bien ilustra un conflicto de intereses entre académicos y decisores. El objetivo del conseguir el liderazgo del en la inteligencia artificial y de conseguir de forma paralela la independencia tecnológica de los países competidores, es una de las prioridades del líder Xi Jinping.

Este dato casa mejor con otras voces que defienden una posición más asertiva, cuando no agresiva. Allen cita como ejemplo a Zeng Yi, alto directo de *Norinco*, la tercera empresa en equipamiento de defensa más grande en China. Zeng Yi pronostica que en el año 2025 las armas autónomas letales ya serían algo normal y que el mayor uso militar de IA sería inevitable (Allen, 2019, 6-7). Los datos de la tabla 3 en este capítulo confirman los esfuerzos del gobierno chino y de las compañías. La generación actual de los UAV chinos son dirigidos de forma remota, pero los oficiales chinos esperan mayor autonomía para el futuro (Allen, 2019, 7).

La asertividad china respecto a su posición de liderazgo dentro de los próximos años tiene su origen en parte en la convicción que el desarrollo del país se ajusta a la teoría del salto de etapas (*leapfrog development* en inglés)¹⁸. Esta teoría postula que un país que muestra retrasos de desarrollo en ciertas tecnologías puede saltarse un paso cuando apuesta por una tecnología incipiente y renuncia en su lugar a las fases anteriores del desarrollo. En el caso del desarrollo militar, esto supondría saltarse el paso de las armas convencionales. Los líderes chinos son conscientes del retraso en esta categoría de armas, pero en el documento AIDP identificaron la inteligencia artificial como una oportunidad histórica para saltarse los pasos de desarrollo y alcanzar el liderazgo en las tecnologías de seguridad nacional (AIDP, 2017, 4).

Un análisis de los objetivos que ha publicado del Consejo de Estado Chino en el documento AIDP, y de las innovaciones que están saliendo de los laboratorios chinos, induce casi de forma automática a la conclusión que China está buscando la hegemonía.

¹⁸ Como en muchas áreas de la economía, se usa el término original en inglés, por lo que por lo visto no existen traducciones al español.

Ejerce su influencia económica en varias regiones en América del Sur y en África. Además, su proyecto de la una Ruta de Seda actualizada (*One Belt One Road*) une Asia con los continentes africano y europeo mediante rutas marítimas y terrestres. Ofrece por una parte a países sin acceso al mar como Kazajstán la oportunidad de progreso económico - un salto de etapas en toda regla, pero, por otra parte, incrementa la influencia china en todas las regiones que se adhieren a la Nueva Ruta de Seda. El avance económico y la nueva asertividad china en la política internacional preocupan a muchos analistas y decisores, a pesar de que China no haya mostrado una política exterior agresiva antes del mandato de Xi Jinping.

La República Popular esconde su creciente potencial militarista, pero ejerce un agarre cada vez más agresivo en su interior. La organización no gubernamental *Freedomhouse* dio a China en su última evaluación para 2019 un resultado de 10 sobre una escala de 100. Describe el país como no libre y registra para los derechos políticos incluso una bajada de un punto respecto al año anterior. La actuación en las protestas civiles de Hong-Kong y la filtración de la noticia sobre centros de reeducación para las etnias uigur, kazajos y hui fueron las claves para este descenso (*Freedomhouse*, 2020). A esta imagen hay que añadir la postura china respecto a la República Taiwán que el gobierno chino considera una región disidente.

Entonces, ¿cómo se debe interpretar la política exterior del gobierno chino? La postura oficial trata de evitar cualquier matiz agresivo en sus publicaciones. Sí, se detecta una nueva asertividad. El Consejo de Estado chino publicó en septiembre de 2019 el documento *China y el mundo en la nueva era*. Como en la mayoría de los documentos con proyección internacional está disponible su traducción al inglés. Las autoridades chinas expresan en este documento su posición acerca de temas económicos y de desarrollo, pero también sobre la geopolítica. Dedicar un epígrafe entero a afirmar que China nunca buscará la hegemonía, a pesar de haber sufrido abusos de parte de los poderes occidentales desde la mitad del siglo XIX. A pesar de estas experiencias dolorosas, la «nación china no tiene el gen de invadir a otros y de dominar el mundo.»¹⁹ (*The State Council of the People's Republic of China*, 2019, 27). China se declara comprometido con el desarrollo pacífico y que nunca se desarrollará a costa de otros, pero tampoco renunciará a sus legítimos derechos e intereses (*ibid.*, 28).

¹⁹ La traducción del documento en inglés es propia.

El discurso de los documentos oficiales aquí citados evita cualquier matiz que puede inducir a ser interpretado como agresivo. Especialmente cuando el Consejo de Estado afirma que nunca buscará la hegemonía, ni que hará sufrir lo que este país había sufrido en el pasado. No obstante, la actuación china en un grupo de regiones cada vez mayor es un indicio de la estrategia que mejor conoce: la influencia económica. Su tono contenido contrasta con la actuación soberbia que muestran algunos representantes de países occidentales, especialmente de Estados Unidos. Por otra parte, las recientes participaciones en las maniobras conjuntas con Rusia e Irán tampoco se pueden interpretar como un signo de pacifismo.

Entonces, ¿cómo se debe interpretar la postura china en el mundo? Este es el meollo del dilema de seguridad. Nadie se atreve afirmar con certeza que China siga siendo un país que va a evitar enfrentamientos bélicos en el futuro. Si fuera así, ¿por qué trata de incrementar su potencial militar a toda costa?

4.3.2 *La posición oficial de Estados Unidos*

Silvela y León señalan que, desde el final de la Segunda Guerra Mundial, la actuación de Estados Unidos ha seguido tres diferentes estrategias. Con las llamadas estrategias de compensación (*offset strategy* en inglés), este país ha tratado de responder a los retos en la carrera armamentística. La primera estrategia de compensación emergió a decisión del presidente Eisenhower durante el inicio de la Guerra Fría. Estados Unidos debió desarrollar el poder nuclear para responder a la superioridad de las armas convencionales soviéticas. La segunda estrategia consistió en la Iniciativa de Defensa Estratégica (*Strategic Defense Initiative* en inglés), impulsada por la Administración Reagan, para compensar la ventaja numérica soviética en su armamento. Dicha estrategia apostó por intensificar el desarrollo tecnológico. La tercera estrategia de compensación nació en 2014. El secretario de defensa durante la Administración Obama, Chuck Hagel, propuso una nueva estrategia para reaccionar a la inminente paridad económica entre su país y China. La nueva estrategia debería intensificar la innovación en los campos de robotización, sistemas de armas autónomas, miniaturización, macrodatos, y la fabricación aditiva (Silvela y León, 2020, 318).

La propuesta no es un giro repentino, sino responde a una tendencia que se ha perfilado durante varios años. La Administración Obama había intensificado el uso de RPAS en Pakistán, y desde 2015 también en Yemen. El secretario de defensa sintetizó en su

propuesta solamente una tendencia, resultado de una creciente preocupación por la situación geopolítica en el mundo.

Dos años antes, el 21 de noviembre de 2012, el Departamento de Defensa había publicado la Directiva 3000.09, en la que se detallan los procesos de verificación y validación de las armas, las pruebas y la evaluación de los sistemas de armas, así como las responsabilidades que derivan de su empleo. Se trata, por tanto, de un documento completo por los aspectos tratados, y a la vez, transparente. Su finalidad es establecer las políticas del DoD y asignar responsabilidades para el desarrollo y el uso de funciones autónomas y semiautónomas en los sistemas de armas, entre los que se incluyen plataformas tripuladas y no tripuladas. Además, establece normas para minimizar la probabilidad y las consecuencias de fallos de sistemas de armas que operan con diferentes grados de autonomía, y que podrían llevar a combates no intencionados. Como se ha mencionado en el capítulo uno, la directiva concibe solamente los sistemas de armas semiautónomos y con autonomía supervisada, en la que incluye la categoría de munición guiada que es capaz de elegir e identificar objetivos de forma independiente (DoD, 2012, 1).

La directiva estipula que los sistemas de armas sean sometidos a un proceso riguroso de comprobación y evaluación para el hardware y software. Estos exámenes se realizarán durante el desarrollo de nuevos sistemas de armas y en las fases de pruebas operacionales, donde el análisis de comportamiento no anticipado en entornos operacionales complejos representa uno de los puntos más críticos. Asimismo, se comprobará la permanencia de las características de seguridad durante las pruebas operacionales, y se introducirán más cambios en los sistemas, si los resultados lo requieren (DoD, 2012, 6).

Para validar que no se hayan degenerado ciertas características críticas de seguridad, se ejecutarán pruebas de regresión automatizada, para identificar así posibles nuevos estados operacionales y cambios. Estos nuevos estados se someterán de nuevo a los procesos de pruebas y evaluación (DoD, 2012, 6).

La directiva establece que los sistemas de armas, los operadores humanos y los comandantes cumplan con el derecho de guerra, los tratados aplicables, las reglas de seguridad de las armas y las reglas de enfrentamiento aplicables (DoD, 2012, 11-12). El cumplimiento de los sistemas de armas con el *ius in bello* fue una de las primeras preocupaciones que surgió con el empleo de las nuevas tecnologías en los sistemas de

armas.

Un punto crucial tratado en la directiva es la pérdida de la comunicación entre el arma y el operador, con la consiguiente pérdida de control. El documento establece claramente que los sistemas de armas semiautónomos a bordo de plataformas no tripuladas deben ser diseñado de tal forma que no seleccionen y ataquen objetivos individuales o grupos específicos que no hayan sido autorizados anteriormente por un operador humano (DoD, 2012, 3).

Los sistemas de armas supervisados por un operador pueden ser usados solamente para la defensa local contra ataques de saturación o en situaciones donde el tiempo de reacción es crucial. En estos casos, las armas con autonomía supervisada no pueden elegir a personas como objetos (DoD, 2012, 3). Las intenciones del comandante y del operador son un concepto importante en esta directiva. Así, el documento estipula que el sistema de armas debe actuar dentro de un espacio de tiempo en consonancia con las intenciones del comandante y el operador, y en caso de no ser posible esto, prescindir del ataque o buscar la autorización humana antes de continuar (DoD, 2012, 7). Una vez más, esto puede ser un punto sustancial en entornos con una comunicación inestable o interferida por el enemigo.

En el documento no se concreta la plena autonomía para los sistemas de armas. Al contrario, tanto los sistemas autónomos como semiautónomos deben incorporar siempre «las necesarias capacidades que permiten al comandante y al operador ejercer niveles apropiados de juicio humano en el uso de la fuerza (DoD, 2012, 7).»

Como se puede comprobar, no se prevé que un sistema de armas actúe con plena autonomía. Cuando el tiempo lo permite, el hombre forma parte activa dentro del ciclo OODA en el modo semiautónomo, y cuando no sea posible, el hombre aún supervisa la actuación del arma. Además, la prohibición de elegir objetivos humanos individuales o en grupo refuta la imagen que evocan actores de la sociedad civil con términos como robots asesinos.

Lo más probable es que un militar no quiera perder el control en la batalla. Un sistema de armas fuera de control podría ser peligroso para el enemigo, pero también para las propias fuerzas. Por tanto, es preferible que se pueda detener el arma siempre que el comandante lo considere necesario.

La Administración de Trump continúa en la línea de reforzar la investigación y el

desarrollo en las nuevas tecnologías. La Estrategia de Seguridad Nacional 2017 menciona las áreas de ciencia de datos, encriptación, tecnologías autónomas, edición genética, entre otras. El documento hace hincapié en la autonomía, tanto en vehículos autónomos como en armas autónomas. Contrario a China, Estados Unidos no apuesta solamente por fomentar el potencial humano procedente del propio país, sino también desde otros lugares (White House, 2017, 30).

Un año más tarde, se publicó la Estrategia de Defensa Nacional. Este documento recalca la creciente complejidad del entorno global, así como la búsqueda china y rusa por la hegemonía regional en sus respectivos vecindarios. Por una parte, China trataría de dominar los países vecinos y la región del Indico-Pacífico, por otra parte, Rusia trabajaría en cambiar las estructuras de seguridad y economía en Europa oriental y Medio Oriente (DoD, 2018, 2).²⁰

Los cambios geopolíticos son acompañados por rápidos avances tecnológicos y el carácter cambiante de la guerra (DoD, 2018, 2). El departamento de defensa constata un impulso incesante para el desarrollo de nuevas tecnologías, que se expande hacia más actores, para quienes la barrera de entrada es más baja. La computación avanzada, el análisis de macrodatos, la inteligencia artificial, la autonomía, la robótica, la energía dirigida, la velocidad hipersónica y la biotecnología son las tecnologías con las que el departamento de defensa se considera capaz de luchar y ganar las guerras del futuro (DoD, 2018, 3).

Los sistemas autónomos avanzados representan un área, en el que el departamento invertirá ampliamente para la aplicación militar. Asimismo, se va a incluir la rápida aplicación de logros comerciales, para ganar ventajas competitivas en el área militar (DoD, 2018, 7).

En febrero de 2020, después de 15 meses de consultas con expertos de la industria, la academia, el gobierno y el público americano, el Departamento de Defensa adoptó cinco principios éticos para la inteligencia artificial (DoD, 2020). Estos principios forman parte de la Iniciativa Americana de Inteligencia Artificial que trata de promover el uso innovador de la IA, mientras protege las libertades civiles, la privacidad y los valores

²⁰ El análisis de la estrategia nacional de defensa se basa en el resumen ejecutivo, porque el documento entero no está disponible en la página del departamento de defensa.

americanos. Según este documento, la inteligencia artificial debe cumplir los siguientes principios.

1. Responsable: El personal del Departamento de Defensa será responsable del desarrollo, el empleo y el uso de las capacidades de la IA, manteniendo niveles apropiados de juicio y precaución.
2. Equitativo: El departamento procurará minimizar sesgos no intencionados en las capacidades de IA.
3. Trazable: El personal involucrado deberá disponer el entendimiento apropiado de la tecnología y de los métodos operacionales que se pueden aplicar en las capacidades de la IA. Estas incluyen metodologías transparentes y auditables, las fuentes de datos, el diseño de los procesos y de la documentación.
4. Fiable: Las capacidades de la IA tendrán usos bien definidos y explícitos, y la seguridad y eficiencia de estas capacidades. Estas serán objeto de pruebas de seguridad durante todo su ciclo de vida.
5. Gobernable: Con este principio, el departamento se reserva el derecho de desconectar o desactivar sistemas desplegados que muestran un comportamiento no intencionado (DoD, 2020, 24 de febrero).

Si bien estos principios no suponen ningún cambio de estrategia, es conveniente recalcar el control que quiere asegurarse el DoD en todos los pasos durante el desarrollo hasta el despliegue de las capacidades de inteligencia artificial en los sistemas. Especialmente el último principio de gobernabilidad es digno de subrayar, porque el DoD se reserva el derecho de desactivar sistemas cuya autonomía manifiesta comportamientos no intencionados. Interpretándolo de forma optimista, se puede esperar que el hombre se mantiene siempre en el ciclo, al menos en función supervisora. Es un aspecto, al que los documentos oficiales chinos no hacen ninguna referencia explícita. Se debe interpretar el nuevo documento como la actualización del contenido de la Directiva 3000.09, condensada en cinco principios éticos, que son conceptualmente manejables y a la vez dan suficiente espacio interpretativo por su contenido abstracto.

Ya se ha visto antes que Estados Unidos lleva años trabajando en ampliar su potencial de vehículos aéreos no tripulados, tanto para operaciones ISR como ejecutar ataques. Los fabricantes más grandes de aeronaves *Boing*, *Lockheed Martin*, *Northrop Grumman* y *General Atomics* cuentan con varias décadas de experiencia y conocimiento tecnológico,

lo que les ayuda a desarrollar plataformas nuevas como el *MQ-25 Stingray*, un dron de reabastecimiento con capacidad de operar desde un portaaviones.

4.3.3 La posición oficial de la Federación Rusa

La Federación Rusa destacó ya en la doctrina militar de 2010 que en las acciones militares aparecen cada vez más drones y embarcaciones autónomas, así como armas y equipamiento militar guiados robotizados (Russian Federation, 2010, 6.). Pero el documento no mencionaba explícitamente qué medidas iba a tomar el gobierno ruso para responder a esta tendencia.

En 2014 se actualizó la doctrina y hasta la fecha es el documento vigente, hasta su actualización, prevista para finales de 2020. Como en la doctrina de 2010, este documento no hace ninguna mención explícita a nuevos sistemas de armas autónomas. El documento señala en primer lugar a la Organización del Tratado del Atlántico Norte como una de las principales amenazas (Russian Federation, 2014, 1-2) para la seguridad rusa, que se articula sobre todo en el despliegue de tropas de la alianza en territorios en países limítrofes de la Federación Rusa.

Otra amenaza sería la creación y el despliegue de sistemas estratégicos de defensa antimisiles que socavarían la estabilidad global y el equilibrio de fuerzas en la esfera de misiles nucleares. En el mismo punto se mencionan las armas de precisión (ibid.).

Más adelante, el documento describe la guerra moderna que se caracteriza por el uso masivo de sistemas de armas y de tecnología militar de precisión y la guerra electrónica, entre otras. En este punto se menciona también el uso de vehículos aéreos y marítimos no tripulados, así como sistemas de armas y equipamiento militar con control robotizado (ibid., 4).

El documento dedica varios puntos (43-47) al equipamiento de las Fuerzas Armadas y otras tropas. Uno de los puntos (46/e) detalla la creación de nuevos tipos de armas de precisión y de los medios para su manejo. Otras áreas serían sistemas de defensa aérea y del espacio, comunicaciones, inteligencia y control, la guerra electrónica, además del desarrollo de robótica para vehículos aéreos no tripulados (ibid., 11).

Rusia constata también la necesidad de crear sistemas de gestión de información y de su integración con sistemas de control y sistemas de armas en los niveles estratégico, operacional y táctico (ibid.).

Estos proyectos requieren una planificación prolongada y una inversión que supera en muchos casos las capacidades de una sola empresa, por lo que Rusia apuesta por la mejora de su complejo militar-industrial. Este complejo se dedicará al desarrollo y la producción a largo plazo, para asegurar la independencia tecnológica de la Federación Rusa (ibid., 12).

Hasta 2019, Rusia no publicó ninguna estrategia dedicada a la inteligencia artificial. Es posible que el país haya estado trabajando ya desde hacía años al desarrollo de armas autónomas que usan inteligencia artificial, pero que haya querido ocultar estos esfuerzos. La política del Kremlin resulta a veces tan opaca como la del otro competidor en esta carrera armamentística, China. Pero lo más probable es que Rusia hasta la fecha no ha alcanzado el nivel de desarrollo de Estados Unidos ni de China.

El desarrollo de plataformas no tripuladas, dotadas con sistemas de inteligencia artificial muy sofisticados, requiere un ciclo de desarrollo bastante extenso. El Ministerio de Defensa ruso publicó el 20 de agosto de 2019 un vídeo de un vuelo de prueba del dron *Altius-U*, que es interpretado como la respuesta rusa al *RQ-4 Global Hawk*, usado para vuelos de reconocimiento. Dos semanas antes, se había lanzado por primera vez el dron *Hunter* (Business Insider, 2019, 20 de agosto).

Ilustración 7: Imagen del UAV Altius-U



Fuente: Business Insider (20/08/2019)

Según el Ministerio de Defensa, este dron podría mantener el vuelo no tripulado durante más de 24 horas y puede ser controlado por satélite lo que permite su radio de empleo en cualquier zona (TASS, 2019, 20 de agosto). El hecho que este vídeo se publicara en el mismo otoño que la Estrategia Nacional para el Desarrollo de Inteligencia

Artificial para el Periodo hasta 2030²¹ parece confirmar la tesis que Rusia se ve obligada a alcanzar el nivel de Estados Unidos y China para no perder la participación en la carrera armamentística global.

Este documento traza de forma detallada los pasos que va a acometer la Federación Rusia en esta área de desarrollo. No hace ninguna referencia expresa a fines militares, solamente se puede interpretar sus objetivos en cuanto a la parte intrínseca de las armas autónomas, que es la inteligencia artificial. El documento toma dos años como puntos de referencia para el desarrollo de la IA en Rusia: 2024 y 2030. Como el documento equivalente chino, el AIDP, Rusia busca su posición global entre los primeros para el año 2030, sin buscar el liderazgo, que es el objetivo crucial en la estrategia china. Llamam la atención una serie de aspectos.

Primero, Rusia quiere apoyar el desarrollo del software en librerías de código abierto (Center for Security and Emerging Technology [CSET], 2019, 5). En esta área, Rusia sigue la estrategia de China que basa todo el desarrollo de software en tecnologías basadas en código abierto para alcanzar la independencia tecnológica de las grandes compañías informáticas que, en su mayoría, son estadounidenses.

Segundo, Rusia promoverá las áreas clásicas de la inteligencia artificial, tales como macrodatos, etiquetaje automatizado, autoaprendizaje, entre otros. Pero el documento menciona también la toma de decisiones en sistemas biológicas, tales como los enjambres de abejas u hormigueros (CSET, 2019, 11). En la actualidad, se desconoce si Rusia está trabajando en estas tecnologías para el uso militar, al estilo del proyecto *Perdix* de Estados Unidos. Pero, es de esperar, que Rusia también fomentará esta tecnología en el futuro por el motivo de reducir los enormes costes que supone la continua actualización del armamento.

Tercero, el documento recalca la protección de los derechos y las libertades humanos garantizados por las leyes rusas e internacionales como uno de los principios básicos para el desarrollo y el uso de las tecnologías de inteligencia artificial (CSET, 2019, 7). Referirse a la ley internacional y los derechos y las libertades humanas es indudablemente

²¹ Traducción propia del título en inglés. El documento en inglés, por su parte, es una traducción realizada por el Center for Security and Emerging Technology (CSET en adelante), un laboratorio de ideas, situado en Georgetown. La traducción del CSET referencia el documento original en la cabecera del documento.

digno de elogio, pero basta contrastar estos principios con el informe de *Freedomhouse* que certifica que Rusia no es libre en absoluto. Alcanza un resultado de 20 sobre 100, que se desglosa entre un 5 sobre 40 en los derechos políticos y un 15 sobre 60 en las libertades civiles. Este país registra para el año 2019 el decimocuarto año consecutivo de descenso en el nivel de libertades (Freedomhouse, 2020).

Rusia tratará probablemente también exportar sus tecnologías, una vez que sus productos hayan alcanzado su madurez tecnológica. Como afirma Viktor Kladov, director de cooperación internacional y política regional de *Rostec* en una entrevista, Rusia exporta su tecnología en igualdad de condiciones, y sin imponer condiciones políticas (TASS, 2019, 27 de noviembre). Rusia es indudablemente un exportador interesante para países que no tienen acceso a tecnología americana o europea por restricciones en la venta de armas. La Corporación Estatal *Rostecnologii* fue creada a finales de 2007, con el fin de centralizar una parte de los activos industriales del país (Rostec, 2020). Otro ejemplo es la compañía *Sukhoi* que anuncia en un informe de 2019 su intención de entrar en el mercado de vehículos aéreos no tripulados. La empresa prevé su exportación a países situados en la región de Asia-Pacífico, Sudeste Asiático y África del Norte (Bulgarian Military, 2020, 16 de junio).

4.4 Conclusión parcial

Los riesgos relacionados con el empleo de sistemas de armas autónomas son varios. Independientemente de si trata de sistemas muy sofisticados, resultado de un ciclo de desarrollo extenso e intenso, o, por otro lado, de drones de fabricación casera con explosivos improvisados, los riesgos son difícilmente controlables.

Por una parte, el peligro de una nueva carrera armamentística no es algo que queda relegado al futuro. Al contrario, esta carrera ya está en pleno curso. Aunque el enfoque en este capítulo ha estado en las tres grandes potencias geopolíticas, otros actores como Israel y Turquía podrían cambiar la polaridad del sistema internacional, al menos cuando se analiza su potencial tecnológico. Parece que Rusia aún no ha alcanzado el nivel tecnológico de China y de Estados Unidos, pero no se debe desatender sus esfuerzos. Para los decisores políticos es difícil evaluar con certeza el potencial de sus competidores, así como sus verdaderas intenciones. Por tanto, nos encontramos con una nueva edición del clásico dilema de seguridad. De momento, todo apunta hacia una competición tecnológica entre China y Estados Unidos. Si ambas potencias no pacifican su enfrentamiento en

acuerdos bilaterales o multilaterales, la trampa de Tucídides parece inevitable y el desencuentro podría transformarse en un conflicto abierto, algo que Mearsheimer había pronosticado ya en 2001 (Mearsheimer, 2001).

Por otra parte, la proliferación descontrolada es uno de los mayores riesgos en la actualidad. La exportación de tecnología tangible como las plataformas de armas no tripuladas son un negocio lucrativo para todos los países que han alcanzado un nivel tecnológico alto. Con la venta de armas no solamente se amortiza la inversión en el desarrollo, sino que aportará grandes beneficios.

La transferencia de tecnología intangible es aún más difícil de controlar, debido a la facilidad de compartir software y conocimiento técnico a través de la red y canales personales. Nuevas formas de fabricación como la impresión 3D facilitan fabricar drones que, añadiendo la opción de adquirir aparatos más sofisticados también en mercados de ocio, facilitan a milicias armarse a un coste relativamente bajo.

5 LAS IMPLICACIONES ÉTICAS Y JURÍDICAS DE ARMAS PLENAMENTE AUTÓNOMAS

Tanto el desarrollo como el uso de sistemas de armas autónomas conllevan una serie de implicaciones éticas y jurídicas. Las primeras se refieren a la moralidad de dejar a un sistema la decisión de tomar la vida a una persona. Un militar puede abortar siempre un ataque en última instancia cuando considera que las circunstancias humanitarias lo exijan. Por ejemplo, si combatientes yihadistas buscaran refugio en un campo de desplazados internos, un militar podría decidir no arriesgar la vida de los civiles y abortar el ataque. Un robot no tendría esta consideración. No es un agente moral, porque desconoce las categorías éticas. Se les podría insertar ciertas reglas de conducta en el campo de batalla, pero a fin de cuentas sigue siendo una máquina sin vínculo emocional con las personas.

Las implicaciones jurídicas se refieren a la depuración de responsabilidades legales. Mientras Estados Unidos ha establecido una serie de principios éticos para el uso de la inteligencia artificial, no queda en absoluto claro que naciones como Rusia o China adopten un compromiso similar. En el caso de grupos terroristas, el uso de principios éticos es algo inexistente, teniendo en cuenta su forma de actuar.

Los principales marcos regulatorios bilaterales y multilaterales para la regulación de armas se crearon cuando las armas aún no disponían de la autonomía actual. Su objeto de regulación eran armas químicas (Convención sobre Armas Químicas, de 1993), armas biológicas (Convención sobre las Armas Biológicas de 1975), armas nucleares (Tratado

de No Proliferación Nuclear de 1968 y el Tratado sobre Fuerzas Nucleares de Rango Intermedio de 1987) y armas en el espacio (Tratado sobre el Espacio Ultraterrestre de 1967). De fechas más recientes son la Convención sobre la prohibición de minas antipersonales, de 1997 y el Tratado de comercio de armas de 2014. Ninguno de estos tratados regula sistemas de armas autónomas. Una excepción sería el Régimen de Control de Tecnología de Misiles, de 1987, que regula al menos la exportación de vehículos aéreos no tripulados. Lo analizaremos más adelante.

5.1 Los principios de distinción, proporcionalidad y precaución

El Diccionario del español jurídico define el derecho de los conflictos armados (*ius in bello*) como «rama del derecho internacional humanitario (DIH) que regula las hostilidades y establece límites a los medios y métodos de guerra.» (Diccionario del español jurídico, 2020). El Comité Internacional de la Cruz Roja reúne en sus publicaciones tanto el cuerpo teórico como la experiencia adquirida en incontables conflictos armados desde su fundación en 1863. Una de las publicaciones más recientes es el libro *Derecho Internacional Humanitario – Una introducción*. En él, su autor Nils Melzer explora las fuentes y los principios fundamentales del DIH.

La finalidad del DIH es la protección de las víctimas de conflictos armados, así como la regulación de las hostilidades, teniendo en cuenta un equilibrio entre la necesidad militar y las consideraciones humanitarias (Melzer, 2019, 16). Uno de los principios fundamentales es el de la no reciprocidad que obliga a las partes beligerantes cumplir con sus obligaciones humanitarias en todas las circunstancias e independientemente de la actuación del enemigo, y de la naturaleza u origen del conflicto (ibid.). Este principio excluye, por tanto, el maltrato de civiles o prisioneros de guerra, motivado por mera venganza. Uno de los principales dilemas del DIH es que, si bien la mayoría de los Estados lo aceptan formalmente con la firma de los tratados, la asimetría en la guerra y el protagonismo de grupos armados que actúan fuera de la legalidad, obstaculizan considerablemente su aplicación y la protección de las personas para las que fue concebido.

Las fuentes del derecho internacional humanitario son el derecho convencional y el derecho consuetudinario. El primero lo componen los cuatro Convenios de Ginebra de 1949, más los dos Protocolos adicionales de 1977. Otra fuente es la Convención de la Haya de 1954 para la protección de bienes culturales y sus dos protocolos. Forman parte

de este cuerpo también los tratados sobre armas, tales como la Convención de 1972 sobre armas bacteriológicas, la Convención de 1980 sobre ciertas armas convencionales y sus cinco protocolos, la Convención de 1993 sobre armas químicas, el Tratado de Ottawa de 1997 sobre las minas antipersonal y la Convención sobre municiones en racimo de 2008. El derecho convencional constituye un cuerpo legal bastante sólido y vinculante para las partes firmantes (ibid., 22).

La segunda parte de las fuentes – el derecho consuetudinario - se origina en las prácticas de los Estados y existe de forma independiente del derecho convencional. Pese a su principal desventaja – la ausencia de un acuerdo escrito, el derecho consuetudinario es ampliamente reconocido como fuente del DIH, porque su carácter interpretativo permite también la actualización continua de su aplicación en conflictos internos (ibid., 23).

En teoría, las partes beligerantes tienen que atenerse a una serie de principios que nacen de la finalidad del DIH, es decir, la protección de las personas que no participan en el combate, y una serie de restricciones que conciernen a los medios y los métodos de guerra. Respecto a los últimos, el derecho internacional humanitario reconoce una serie de principios que deben tener en cuenta las partes combatientes. Estos principios representan unos referentes jurídicos para el uso de armas autónomas.

El principio de la distinción obliga a los militares distinguir entre combatientes y todas las personas que no participan en los combates, hayan participado en los combates anteriormente o no. Esto significa que los combatientes lleven algún distintivo mínimo que les diferencie de combatientes, sea un uniforme o cualquier otra insignia. En una guerra convencional de los siglos pasados, esta distinción era más fácil que hoy día. En cambio, cuando una de las partes es un grupo armado, esta distinción es difícil porque en muchos casos sus miembros no suelen llevar informe. Al contrario, sus combatientes se reclutan muchas veces de la población civil o de ex miembros de fuerzas de seguridad. En enfrentamientos en áreas urbanas es más difícil diferenciarlos de civiles cuando se entremezclan con la población civil. Muchos expertos dudan de la capacidad de la inteligencia artificial de diferenciar en estos casos civiles y combatientes. Si bien los sistemas de armas modernos disponen de objetivos ópticos muy eficientes, la interpretación correcta de las imágenes aún es una tarea en la que la inteligencia artificial muestra fallos graves, a pesar de los últimos avances en las redes neuronales (Scharre,

2018, 180-184). Por tanto, no debería ser la base de decisión autónoma de un sistema de armas, cuando se trata de identificar a personas como combatientes o civiles.

El segundo principio atañe a la proporcionalidad y exige que los responsables militares eviten daños y sufrimientos innecesarios. Es probablemente el principio que más se quebranta en los conflictos internos, porque las acciones y reacciones se retroalimentan durante un conflicto en una espiral de venganza mutua. Klare se pregunta en este contexto, cómo se podría garantizar la proporcionalidad en un ataque autónomo a un edificio residencial, desde cuyo tejado un francotirador ataca a tropas amigas (Klare, 2019b). Responder a esta amenaza con un misil desde el *Predator* ya es cuestionable por la cercanía de civiles, pero ¿qué ocurriría cuando esta decisión no tomara un operador táctico sino un dron autónomo? ¿Cómo se puede garantizar en este caso la correcta distinción entre combatientes y civiles, por una parte, y la proporcionalidad en el ataque, por otra parte?

El tercer principio es el de precaución, según el cual se deben tomar todas medidas necesarias para proteger a personas y bienes civiles en ataques militares. Este principio incumbe sobre todo a la preparación de un ataque. Los responsables deben tomar todas las medidas necesarias para proteger a la población civil del contrincante. Esto podría realizarse mediante previo aviso por los medios o la evitación de ataques a objetivos donde se esperan aglomeraciones, como colegios, lugares para el culto religioso o mercadillos. Es una obligación que existe también para la parte que tiene que defenderse de un ataque, que en este caso debe tomar todas las medidas, tales como la evacuación previa. En la lucha urbana, este principio es también un reto para las partes beligerantes, por la mezcla de civiles y combatientes en zonas concurridas. Por esta misma razón, en el hipotético caso de que un sistema de armas actuara con plena autonomía, por ejemplo, un dron merodeador como el *Harpy*, no se podría garantizar ninguno de los tres principios aquí expuestos.

5.1.1 La cuestión de la responsabilidad legal

En una guerra interestatal, son los Estados implicados que responden jurídicamente por los actos cometidos por sus fuerzas armadas, al menos en la teoría. Existe una cadena de mando, y corresponde al Estado asumir la responsabilidad por los actos que se hayan cometido en el conflicto. Pero esta cuestión sería menos clara cuando se emplearan armas completamente autónomas. Por definición, el comandante ya no puede detener la

actuación del sistema de armas, de lo contrario, sería el modo de autonomía supervisada.

Aquí surge la cuestión de la responsabilidad legal, cuando el arma cause daños entre la población civil. ¿Podría ser este el caso, cuando un dron como el *Predator* actuara con plena autonomía y lanzara un misil *Hellfire* a un edificio residencial porque el sistema lo ha identificado como el escondite de una célula de combatientes yihadistas?

Ilustración 8: MQ-1 Predator, dotado con un AGM 114 Hellfire



Fuente: Wikipedia (2020)

Conforme la Directiva 3000.9, de 2012 y la posterior publicación de los principios éticos para la inteligencia artificial, de 2020, Estados Unidos se reserva siempre el derecho de desconectar o desactivar un sistema del que ha perdido el control.²² Resulta entonces poco probable, que este país permitiera la plena autonomía de sus sistemas de armas.

En Europa existe también una posición clara. El Grupo Europeo sobre Ética en la Ciencia (*European Group on Ethics*, EGE en adelante) subraya que el hombre debe mantener siempre un control significativo. Este grupo de asesores para cuestiones éticas ha creado una serie de principios para el desarrollo y empleo de inteligencia artificial y sistemas autónomos, tanto para usos civiles como para militares. A pesar de defender la misma posición sobre el control humano como Estados Unidos, los ocho principios difieren conceptualmente de los principios formulados por el Departamento de Defensa,

²² Véase el capítulo 4.

debido a su carácter más filosófico-abstracto. Conceptos como dignidad, justicia, equidad o solidaridad expresan los valores europeos y centran su alcance en el individuo que debe disfrutar libertad de movimiento, integridad física y moral (European Group on Ethics [EGE], 2018, 16-19). Respecto a los sistemas de armas autónomas, la declaración del grupo cita el concepto de control humano significativo, procedente de las deliberaciones que se han producido en el entorno del Grupo de Expertos Gubernamentales.

La mayoría de las sociedades occidentales comparten en su acervo cultural valores como la libertad, igualdad y justicia. Creen que solamente un sistema político que respete estos valores y que base sus prácticas políticas respetando el derecho, puede prosperar. Existe un debate sobre las implicaciones éticas y jurídicas sobre la autonomía en las armas. ¿Pero se puede esperar el mismo grado de reflexiones oficiales en otros países? Los documentos oficiales de China y Rusia no hacen ninguna referencia al respecto, y de otros países con un sistema político menos democrático se puede esperar que a sus gobernantes importa más el valor estratégico de las nuevas tecnologías que cuestiones jurídicas. Grupos terroristas, ya de entrada, no se plantean estas cuestiones.

5.1.2 La falibilidad de los sistemas

Otro peligro proviene de los sistemas mismos, en concreto de su falibilidad potencial. Gutiérrez y Cervell citan un incidente que ocurrió con el sistema de combate naval AEGIS, en uso desde 1983. Es un sistemas de defensa que opera con autonomía supervisada. Inicialmente, la confianza en su supuesta infalibilidad fue muy alta, pero el derribo del vuelo 655 de *Iran Air* el 3 de julio de 1988 causó la muerte a 290 civiles. El fallo era que el sistema confundió el vuelo 655 con un supuesto segundo caza F-14 iraní que en realidad acompañaba al avión de pasajeros. El F-14 evitó el impacto del misil con un viraje *en extremis*, pero el vuelo de *Iran Air* no pudo realizar esta maniobra. Aunque la tripulación del buque *USS Vincennes* tenía dudas sobre la identidad de la aeronave, se dio finalmente la orden de derribar el supuesto segundo F-14. El incidente mostró a las autoridades que la identificación de los objetivos aún no funcionaba con la seguridad necesaria. (Gutiérrez y Cervell, 2013, 30).

Otro ejemplo es el ya mencionado incidente con el sistema de satélites *Oko* que alertó erróneamente de un ataque de misiles intercontinentales. Como se señaló antes, fue la prudencia del Teniente Stanislav Petro quien desconfió del sistema y evitó así un

desenlace devastador entre las dos superpotencias.²³

Los dos incidentes conducen a dos conclusiones importantes: primero, un sistema no es infalible, a pesar de su estado de desarrollo. La identificación de un objeto no es siempre infalible, y cuando condiciones meteorológicas adversas dificultan la tarea, hace falta más prudencia aún. La segunda conclusión es la consecuencia de la primera: es necesario mantener el hombre dentro del ciclo OODA, al menos como supervisor. La interpretación de los datos que aportan los sensores de un sistema es un aspecto que aún no ha alcanzado el mismo estado de madurez que la identificación.

5.2 Conclusión parcial

El desarrollo y el uso de sistemas de armas autónomas plantean algunas cuestiones de carácter ético. Un sistema de armas, por muy sofisticado y autónomo que sea, no puede ser un agente moral porque no tiene ningún vínculo emocional con los hombres. Es por este motivo por el que la mayoría de los expertos rechazan la idea que un sistema deba tener la opción de decidir sobre vida o muerte.

Las implicaciones jurídicas se refieren al respeto de algunos principios claves en el Derecho Internacional Humanitario, en concreto los principios de distinción, proporcionalidad y precaución. Como señalaron Ekelhof y Persi, existen varios niveles donde se toman decisiones.²⁴ Especialmente en la planificación de un ataque existe la posibilidad de respetar estos principios.

Un sistema no funcionará nunca con plena seguridad. Esta es la principal razón por la que el hombre siempre debería mantener el control sobre el sistema, reivindicación que repiten numerosos críticos de los sistemas de armas autónomas. Los dos incidentes descritos en este capítulo muestran el peligro cuando un sistema malinterpreta una información e induce a activar un protocolo de contraataque.

6 INSTRUMENTOS JURÍDICOS DEL DERECHO INTERNACIONAL HUMANITARIO ACTUAL

La organización no gubernamental *Human Rights Watch* publicó en 2020 el informe *Cuidado con el vacío - La falta de responsabilidad con respecto a los robots asesinos*

²³ Véase el capítulo 2.2.

²⁴ Véase capítulo 3.3.

(Human Rights Watch [HRW], 2020).

El informe explora la cuestión de la responsabilidad para la actuación de sistemas de armas con plena autonomía, alegando que «las armas completamente autónomas no pueden sustituir a los humanos responsables como acusados en ningún procedimiento legal que busque lograr la disuasión y el castigo (HRW, 2020, 1).»

El sujeto responsable debe ser siempre la persona, porque un robot no tiene conciencia del bien y del mal, por lo que no se le puede considerar un agente moral. Sobre esta premisa, se reclama la rendición de cuentas en tres esferas: la responsabilidad personal, la responsabilidad criminal y la responsabilidad civil. La cuestión detrás es siempre esta: ¿a quiénes pueden pedir responsabilidades las víctimas que han sufrido un ataque por parte de un sistema de armas autónomo? ¿Hay que pedir responsabilidades al operador táctico o al comandante del sistema? ¿O hay que pedir responsabilidades al fabricante del arma o al programador quien ha diseñado los algoritmos? En el caso de que un sistema de armas actuara de forma autónoma, causando daños a personas y bienes civiles, ambas partes podrían alegar que esta acción no había sido intencionada. La posición del fabricante sería rechazar la responsabilidad sobre acciones militares, mientras la parte militar alegaría que ella no es responsable del funcionamiento de los algoritmos. Los militares se encargan de la planificación de la operación, y aunque respetaran el principio de precaución, una vez que el sistema actuara con plena autonomía, ni siquiera los mandos militares tendrían la posibilidad de detener la acción.

Para las víctimas sería muy costosa la demanda judicial, y probablemente resultaría insatisfactoria su reclamación de resarcir los daños personales y materiales que ha causado un sistema de armas autónomo (ibid., 1-6).

Por estas razones, HRW recomienda prohibir totalmente «el desarrollo, la producción y el uso de armas completamente autónomas a través de un instrumento internacional legalmente vinculante» (ibid., 6).

6.1 La cláusula de Martens

¿Qué base legal ofrece el Derecho Internacional para implantar la prohibición? Human Rights Watch propuso recurrir a la cláusula de Martens en su informe *Heed the call - A Moral and Legal Imperative to Ban Killer Robots* (HRW, 2018). Este documento es la aportación de la organización para la reunión del Grupo de Expertos Gubernamentales del Convenio sobre Ciertas Armas Convencionales. La cláusula data de

1899 y fue introducida a petición del delegado de Rusia en la Primera Conferencia de Paz de La Haya. Posteriormente fue adaptada en 1977 en el Protocolo adicional I (en adelante PA I) del Convenio de Ginebra. El artículo 1.2 reza:

En los casos no previstos en el presente Protocolo o en otros acuerdos internacionales, las personas civiles y los combatientes quedan bajo la protección y el imperio de los principios del derecho de gentes derivados de los usos establecidos, de los principios de humanidad y de los dictados de la conciencia pública.²⁵

El carácter subsidiario de esta fórmula daría seguridad a las personas incluso en situaciones para las cuales aún no existen acuerdos internacionales, y, por tanto, su mandato sería ineludible. Los autores del informe no hicieron ninguna referencia al Régimen de Control de Tecnologías de Misiles (*Missile Technology Control Regime*, MTCR en adelante) que regula la exportación de drones, probablemente porque su interés está embocado más en la prohibición preventiva de armas que podrían actuar algún día con plena autonomía.

El principio de humanidad es el referente ético para medir las acciones sujetas al Derecho Internacional Humanitario. Un sistema de arma autónomo puede demostrar su autonomía táctica en el campo de batalla, pero esta clase de autonomía no le convierte en un agente moral. Esta función queda relegada a las personas, porque ellas son los sujetos que por la naturaleza han sido dotadas de conciencia de sí mismo y de reconocer a sus congéneres. Los principios éticos del DIH son fruto de la humanidad de las personas, que en última instancia son capaces de abortar un ataque, de proteger a las personas civiles en la medida de lo posible, especialmente a aquellas que han sido heridas en la batalla. Este rasgo humanitario es algo que le caracteriza al hombre, a la par con su propensión para la violencia que le empuja a acometer atrocidades inimaginables. Un arma autónomo no es humanitario ni atroz, simplemente es un artefacto que actúa con plena libertad, pero sin estar atado por preceptos éticos.

La conciencia pública, la otra referencia de la cláusula de Martens, se basa en el supuesto que la mayoría de las personas comparten ciertos valores comunes por los que se rechazan ampliamente determinados actos de extrema crueldad, al ser considerados inhumanos.

²⁵ Protocolo adicional I a los Convenios de Ginebra de 1949 relativo a la protección de las víctimas de los conflictos armados internacionales, 1977.

6.2 El examen jurídico de nuevas tecnologías

El artículo 36 del mismo protocolo ofrece otro instrumento jurídico para nuestra cuestión. En concreto, obliga a los Estados a analizar y comprobar la compatibilidad de nuevas armas con los principios del DIH.

Cuando una Alta Parte contratante estudie, desarrolle, adquiera o adopte una nueva arma, o nuevos medios o métodos de guerra, tendrá la obligación de determinar si su empleo, en ciertas condiciones o en todas las circunstancias, estaría prohibido por el presente Protocolo o por cualquier otra norma de derecho internacional aplicable a esa Alta Parte contratante.²⁶

Los Estados signatarios de este acuerdo están obligados a someter al examen jurídico todas las armas nuevas para garantizar que no incumplan las prohibiciones del Protocolo o cualquier otra norma del derecho internacional, y que respeten los principios de distinción, de proporcionalidad y de precaución.

El artículo 1.2 del Protocolo I y el artículo 36 del mismo documento ofrecen, por tanto, las referencias éticas en el primer caso, y la obligación de un examen jurídico para nuevas armas, en el segundo caso. No sustituyen ningún tratado dedicado a sistemas de armas autónomas, pero aportan al menos una orientación temporal, hasta que la comunidad internacional regule el tema de forma clara.

Como se ha mencionado antes, Estados Unidos realiza un examen jurídico previo al desarrollo de nuevos sistemas de armas, a pesar de que no sea estado signatario del Primer Protocolo adicional del Convenio de Ginebra.

Por otra parte, España publicó su posición en mayo de 2019 en el documento *Papel de la Ciencia y la Tecnología en el Contexto de la Seguridad Internacional y Desarme*: los sistemas de armas deben permitir adjudicar la responsabilidad legal en todo momento a un operador humano, incluso en los sistemas de defensa aérea. Estos «no se operan en las Fuerzas Armadas en modo autónomo de forma que se puedan producir acciones letales no decididas por una persona.»²⁷

Para el examen jurídico, España dispone de la instrucción SEDEF 67/2011, de 15 de septiembre, por la que se regula en la etapa de previabilidad²⁸ operativa el Procedimiento

²⁶ Protocolo I adicional a los Convenios de Ginebra de 1949 relativo a la protección de las víctimas de los conflictos armados internacionales, 1977.

²⁷ Asamblea General de Naciones Unidas, Resolución 73/32, del 5 de diciembre de 2018, sobre el Papel de la Ciencia y la Tecnología en el Contexto de la Seguridad Internacional y Desarme. Posición de España.

²⁸ La Real Academia de lengua española no ha registrado este término, pero el documento referenciado lo usa. Para evitar confusiones, no se ha sustituido por otro en este trabajo.

de Obtención de Recursos Materiales (PORM). El documento recalca que «el control humano significativo sobre el uso de armas y sus efectos es esencial para asegurar que el uso de un arma sea éticamente justificable y legal». (ibid.)

6.3 La Convención sobre ciertas armas convencionales

Varios actores de los sectores civil y público han intensificado sus esfuerzos en concienciar la opinión pública sobre los riesgos relacionados con las armas autónomas. La Campaña de Prohibición de Robots Asesinos es una coalición internacional de 165 organizaciones no gubernamentales en 65 países que inició su trabajo en octubre de 2012. Esta alianza cuenta con el apoyo del Secretario General de Naciones Unidas, además de diferentes parlamentos y personajes de la vida pública (Stop Killer Robots, 2020).

Otro actor es el Grupo de Expertos Gubernamentales, creado en 2017 por las Altas Partes Contratantes de la CCW que decidieron asumir el debate sobre las implicaciones de los sistemas de armas autónomas para el Derecho Internacional Humanitario y el Derecho Internacional de Derechos Humanos. El GGE trata de para explorar posiciones y acordar sugerencias relacionadas con las armas autónomas (Meza, 2018, 6). En el grupo se encuentran representantes del sector militar y de la sociedad civil, de la industria y de la academia. Las cuestiones que se tratan en los diferentes grupos de trabajo son aspectos tecnológicos, efectos militares y cuestiones de índole jurídico-ético (ibid., 9).

Es interesante reseñar algunas posiciones de los países. Un grupo de trabajo, compuesto por Países Bajos y Suiza recuerda los mecanismos de revisión de armas al artículo 36 del Protocolo adicional I a los Convenio de Ginebra. El documento señala que lo establecido en este artículo no solamente abarca los métodos y medios de guerra existentes, sino además sistemas de armas futuros (ibid., 12-13). Estados Unidos, pese a no formar parte de este acuerdo, adopta la misma posición (ibid., 13-14).

A pesar de esta postura común, se manifiestan también mayores dificultades a la hora de acordar un código de conducta políticamente vinculante. Rusia se ha revelado como uno de los países que rechazan a priori comprometerse de forma vinculante y defiende que el cuerpo normativo vigente sería suficiente (ibid., 15).

En marzo y agosto de 2019, se celebraron en Ginebra dos sesiones del GGE donde se presentaron unos primeros resultados de los grupos de trabajo que consistieron en una serie de recomendaciones y la reiteración de mantener el control humano en todo momento.

La interacción entre el ser humano y la máquina, que puede adoptar diversas formas y puede tener lugar en distintas etapas del ciclo de vida de un arma, debe garantizar que el posible uso de los sistemas de armas basados en tecnologías emergentes en el ámbito de los sistemas de armas autónomas letales sea compatible con el derecho internacional aplicable, en particular al derecho internacional humanitario.²⁹

La cadena de mando debe garantizar que se respeten los principios de distinción, proporcionalidad y precaución, entre otros.³⁰ Asimismo, queda aún mucho trabajo aclaratorio sobre «el tipo y grado de interacción entre el hombre y la máquina, incluidos los elementos de control y de juicio, en las diferentes etapas de ciclo de vida de un arma».³¹

La vaguedad de los resultados fue decepcionante para muchos observadores y participantes que habían esperado conclusiones más concretas. Se constata entre algunos países la falta de voluntad política de asumir acuerdos vinculantes. En el caso de los miembros permanentes del Consejo de Seguridad de Naciones Unidas, el derecho del veto dificultaría adoptar una posición común si los intereses particulares se ven obstaculizados.

Por este motivo, algunos actores proponen impulsar un acuerdo paralelo entre países voluntarios, tal y como ocurrió en la Convención de Ottawa sobre la prohibición de minas antipersonales y la Convención de Oslo sobre municiones en racimo (Meza, 2018, 18).

Pero no todos los autores se posicionan claramente a favor de una prohibición total de armas que operan con plena autonomía. Anderson y Waxman defendieron ya en 2013 la regulación de las SAALs en lugar de prohibir su uso e incluso su desarrollo. Estos autores señalan que todo desarrollo se produce de forma progresiva, también en áreas tan recientes como la de las tecnologías intangibles. Esto facilita que las normas se pueden actualizar siempre con suficiente tiempo, y, además, adaptándose a los cambios concretos. Anderson y Waxman defienden en su análisis que la directiva 3000.09, que el Departamento de Defensa publicó en 2012, contiene ya las normas necesarias para las nuevas tecnologías, así como la práctica de tomar el artículo 36 del PA I como referencia para el examen jurídico.

Anderson y Waxman advierten que no deberían rechazarse desarrollos tecnológicos, máxime cuando traen ventajas en el campo de batalla, gracias a las mejoras en precisión

²⁹ Informe del período de sesiones de 2019 del Grupo de Expertos Gubernamentales sobre las Tecnologías Emergentes en el Ámbito de los Sistemas de armas autónomas Letales, del 25 de septiembre de 2019, CCW/GGE.1/2019/3.

³⁰ Ibid., 4.

³¹ Ibid., 6.

e identificación de objetivos, a las que se añade también la racionalidad de las decisiones que tomarían armas dotadas con inteligencia artificial. La ausencia de miedo, venganza u otras emociones ayudaría cumplir con el Derecho Internacional Humanitario. Asimismo, una mayor dosificación del fuego evitaría daños colaterales innecesarios y cumplir de esta manera con el principio de proporcionalidad (Anderson y Waxman, 2013, 15).

6.4 El Régimen de Control de Tecnología de Misiles

Los vehículos aéreos no tripulados son en la actualidad los sistemas de armas que mayor autonomía y madurez tecnológica han alcanzado. La posible vinculación tecnológica de los UAV con el etiquetaje automatizado en los mapas y los progresos en el autoaprendizaje convierten estos sistemas en los primeros candidatos para transformarse en armas autónomas. Existe un acuerdo internacional donde se regulan los UAV, pero la cuestión es, si este acuerdo sirve para tratar las cuestiones éticas y jurídicas que hemos estudiado antes, motivo por el que exploraremos brevemente el Régimen de Control de Tecnología de Misiles.

Este acuerdo, en vigor desde abril de 1987, regula la exportación de misiles balísticos y de otros sistemas de transporte no tripulados. Aunque el MTCR estaba pensado para sistemas de transporte para armas de destrucción masiva, se incluyen en la actualidad también los UAV por sus características de autonomía en el vuelo. Este acuerdo cuenta con 35 países miembros, contando el ingreso de India en 2016 (MTCR, 2017, i-iii). Entre los miembros se encuentran Turquía, la Federación Rusa y Estados Unidos, pero China e Israel, ambos países exportadores de UAV, no forman parte del acuerdo. Este dato revela la importancia de que todos los países exportadores participen en un régimen donde, según la definición de Krasner, convergen las expectativas de los países involucrados. Cuando países cruciales para el control sobre la exportación de tecnologías no participan en el régimen, su eficacia se ve afectada seriamente.

Los drones se encuentran en la categoría I del Anexo de Materiales y de Tecnología, cuya última versión data de 2017. Este catálogo incluye sistemas completos de misiles, cohetes y vehículos aéreos no tripulados. La exportación de sus componentes se restringe normalmente con la cláusula «fuerte presunción de denegación», al tratarse de sistemas completos y, salvo en casos excepcionales, prohibidos para la exportación según las directrices del MTCR (ibid.).

En el pasado, el gobierno de Estados Unidos había restringido la exportación de drones

basándose en la mencionada cláusula de denegación. Sin embargo, la administración Trump decidió en julio de 2020 relajar unilateralmente las restricciones sobre sistemas aéreos no tripulados que vuelan a velocidades menores de 800 kilómetros por hora. Con este paso mejorarían considerablemente las oportunidades de venta para *General Atomics* y *Northrop Grunman* porque sus productos estrella, el *MQ-9 Reaper* y el *RQ-4 Global Hawk*, no alcanzan esta velocidad. Estados Unidos podría exportar así dos de sus drones más desarrollados y probados en el campo de batalla (Defense News, 2020, 24 de julio).

Las directrices del MTCR estipulan que ningún país puede vetar las exportaciones de otro miembro, por lo que a los otros 34 países no les queda otro remedio que aceptar la decisión unilateral que ha tomado Estados Unidos. Queda demostrado así que el MTCR no es el foro adecuado para regular sistemas de armas autónomas, ni por su objetivo ni por su baja eficacia.

6.5 Conclusión parcial

El Derecho Internacional Humanitario exige a las partes combatientes que sus métodos y medios de guerra cumplan con los principios de distinción, de proporcionalidad y de precaución.³² Asimismo, existen dos mecanismos más que son aplicables para el desarrollo y el uso de cualquier arma: los artículos 1.2 y 36, ambas partes del Protocolo Adicional I al Convenio de Ginebra relativo a la protección de las víctimas de los conflictos internacionales, de 1977.

El primer mecanismo, conocido como cláusula de Martens, exige el respeto de los principios de humanidad y de la conciencia pública para todas las personas civiles y combatientes, cuando no exista ningún acuerdo específico. Se suele citar esta cláusula para exigir la protección de las personas mientras no se haya alcanzado otro acuerdo internacional. La Corte Internacional de Justicia expresó en 1996 en un procedimiento consultivo su opinión sobre la legalidad de la amenaza o el empleo de armas nucleares. En el examen de los “principios cardinales” del DIH, la Corte reconoce que la cláusula de Martens ha sido «un medio eficaz para reflejar la rápida evolución de la tecnología militar»³³. A pesar de este importante respaldo jurídico, la margen interpretativa de la cláusula deja solamente claro que hay que respetar los principios del DIH, pero, dado su

³² Véase el capítulo 5.1.

³³ NNUU, Asamblea General, Opinión consultiva de la Corte Internacional de Justicia sobre la legalidad de la amenaza o el empleo de armas nucleares, A/51/218, 19 de julio de 1996, párr. 78.

carácter genérico, no da ninguna recomendación de cómo hacerlo.

El segundo mecanismo, el artículo 36, exige un examen jurídico para los nuevos desarrollos de armas y métodos de guerra. Ambos artículos, ofrecerían ya una base mínima para la regulación de sistemas de armas autónomas. Fleischer señala en un artículo motivado por el debate de armar RPAS del ejército alemán que el uso y empleo de armas autónomas son menos una cuestión del Derecho Internacional Humanitario, sino más bien una cuestión ética y política (Fleischer, 2020). Defensores de una prohibición total de estas armas, incluido su desarrollo, fundamentan su reivindicación en el desarrollo avanzado que han alcanzado ya muchos sistemas de armas. No se percibe la voluntad política de parar este proceso entre los países que van a la cabeza del desarrollo.

Dadas las ventajas que ofrecen sistemas de armas con semiautónoma u autonomía supervisada, sería más recomendable su regulación en lugar de buscar la prohibición que a estas alturas ya se presenta como una opción inalcanzable. El problema es: ¿Cuál sería el foro para buscar un acuerdo? La inoperatividad del Consejo de Seguridad de Naciones Unidas, fruto del permanente choque entre los antiguos bloques de la Guerra Fría sugiere que se requiere otro foro. Las propuestas más pragmáticas señalan en este contexto dos ejemplos exitosos, donde países voluntarios alcanzaron acuerdos: el Tratado de Ottawa de 1997 sobre las minas antipersonales y la Convención sobre municiones en racimo de 2008. Este forma de proceder serviría al menos para juntar a todas las partes dispuestas a regular sistemas de armas autónomas y señalarían de forma incuestionable a aquellos países que no firmarían este tipo de tratados.

Por otra parte, el mecanismo de voto por consenso dificulta considerablemente que se llegue a un acuerdo vinculante. Y en el caso de que se llegue a un acuerdo algún día, surgen dudas sobre su aplicabilidad práctica. ¿Cómo se debería verificar y comprobar el cumplimiento de este acuerdo? El MTCR ofrece esta posibilidad al menos para sistemas completos y parciales. Aunque su objetivo no es la regulación de futuros sistemas de armas autónomas, la categoría I ya incluye vehículos aéreos no tripulados, así como la tecnología y el software necesario para el uso de los drones (MTCR, 2017, V-VI). Un sistema para incluir en la categoría I podrían ser futuros submarinos autónomos con misiles a bordo.

La regulación de armas autónomas es seguramente la solución preferible a la prohibición total, porque permite aprovechar las ventajas que ofrecen las nuevas

tecnologías. Regular el desarrollo y empleo de armas autónomas requerirá un trabajo preparativo mayor que la prohibición total, pero a la vez se crearía un marco regulador más detallado y provechoso.

CONCLUSIONES FINALES

La implantación de la inteligencia artificial en los sistemas de armas se produce de forma progresiva, pero sus implicaciones en la seguridad internacional y en los ámbitos ético y jurídico son, sin duda alguna, disruptivas.

La hipótesis principal de este trabajo era mostrar que la inteligencia artificial es el rasgo distintivo entre sistemas de armas automáticas y sistemas de armas autónomas.

Sistemas de armas modernos como vehículos no tripulados disponen de una serie de características que les diferencia claramente de armas automáticas. En concreto, la capacidad de orientarse en entornos no estructurados y la capacidad de aprender son dos rasgos destacados. El paso de la automatización a la autonomía no se produce de forma abrupta sino gradual: en el caso de RPAS empieza con capacidad de despegar y aterrizar de forma autónoma, llegando hasta la capacidad de volver a la base sin el control por el operador táctico.

Partiendo de la definición de Russel y Norvig, que tratan la inteligencia artificial como un agente racional, hemos elaborado una definición de un sistema de arma autónomo letal. Un sistema de arma autónomo letal es un agente racional, dotado con inteligencia artificial que le proporciona la capacidad de adaptarse a contextos cambiantes, de aprender durante el ciclo de su existencia y de realizar tareas con diferentes grados de autonomía. Su empleo conlleva una serie de ventajas por las que los decisores políticos impulsan su desarrollo, pero también algunas desventajas, relacionadas sobre todo con el vacío legal dentro del que operarían estos sistemas en modo de plena autonomía.

Se ha mostrado antes que la creciente autonomía de sistemas de armas modernos es acompañada por diferentes grados de implicación del hombre en el ciclo OODA, desde el semiautónomo, pasando por el modo de autonomía supervisada, hasta la plena autonomía, que en la actualidad en la mayoría de los sistemas no está disponible. El modo de autonomía supervisada es el que usan hoy en día los sistemas de defensa aéreos y navales, por sus ventajas en rapidez y precisión. Pero no hay que desatender las fases previas, especialmente la planificación en los niveles estratégico, operacional y táctico. Especialmente en la planificación y las consideraciones del despliegue de armas y las

normas de enfrentamiento se pueden tener en cuenta los principios rectores del Derecho Internacional Humano.

Como se ha mencionado antes, la posible introducción de bandadas en el campo de batalla representa nuevos retos para la implicación del hombre en el ciclo. Cuanto menor sea la participación humana en la organización de las bandadas, mayor será su eficacia. Considerando el ritmo al que se producen los avances en la inteligencia artificial, especialmente en el área de autoaprendizaje, el paso del modo de autonomía supervisada al modo de autonomía completa se va a producir pronto, porque es especialmente lo que se busca. Para lograr mayor eficiencia en la actuación, cualquier vehículo no tripulado, sea de tierra, mar o el aire, requiere protección contra las inferencias electrónicas del enemigo y, por tanto, mayor autonomía.

Algunos países, entre ellos Estados Unidos, están llevando a cabo ensayos con este tipo de operaciones que, debido a la necesidad de organizarse y operar de forma eficiente, requieren un grado de autonomía aún mayor que los sistemas de armas que están en uso ahora. Existen diferentes arquitecturas de control y comando, desde la centralizada donde el hombre ocupa una posición clave, hasta la estructura del control conductual donde la bandada actúa con plena autonomía.

Por lo expuesto anteriormente se puede confirmar la hipótesis principal. La inteligencia artificial es el rasgo distintivo entre armas automáticas y armas autónomas. Su implantación se produce de forma gradual y las armas han adquirido más autonomía durante los últimos años.

La primera hipótesis derivada preguntaba si la existencia de la inteligencia artificial podría obstaculizar una regulación efectiva. Hemos visto que en los grupos de trabajo del Grupo de Expertos Gubernamentales existen discrepancias sobre la autonomía, y, por tanto, sobre la definición de sistemas de armas autónomas. La inteligencia artificial funciona en estas concepciones como variable independiente, y el aumento de las capacidades, especialmente en el autoaprendizaje, revelan un escenario preocupante para el futuro. Mientras muchos países continúan con el desarrollo, parece que algunos países como China y Rusia aprovechan las discusiones conceptuales en el seno del GGE para rehuir verdaderos compromisos que limitarían sus propios desarrollos. Podemos confirmar, por tanto, la primera hipótesis derivada.

La segunda hipótesis derivada postula que los sistemas de armas autónomas requieren

nuevos marcos regulatorios. La investigación nos ha mostrado que muchos marcos regulatorios actuales son insuficientes porque su regulación concierne a clases de armas de destrucción masiva. Otros tratados como el Convenio de Ottawa sobre la prohibición de minas antipersonales y el Convenio de Oslo sobre municiones en racimo nos aportan formas de acuerdos voluntarios, pero su objeto de regulación no son las armas autónomas.

El Régimen de Control de Tecnologías de Misiles aporta al menos un mecanismo para controlar la proliferación de una clase de sistema de armas que funcionan de momento en modo semiautónomo, con la intervención del operador táctico: los drones. Esta clase de armas tiene por su desarrollo avanzado el potencial de alcanzar algún día la plena autonomía. Pero el MTCR es un acuerdo voluntario diseñado solamente para controlar la proliferación y no trata las cuestiones éticas ni las que conciernen al Derecho Internacional Humanitario. Además, su carácter voluntario impide un control realmente eficaz, cuando un país miembro relaja unilateralmente las restricciones que se habían acordado anteriormente. Este es, según la nueva interpretación de Estados Unidos, el caso de los drones que vuelan a velocidades menores de 800 kilómetros por hora.

El Derecho Internacional Humanitario clásico ofrece unos principios cardinales que prohíben el empleo de armas que actúan contra estos principios, especialmente el de la no discriminación entre combatientes y no combatientes. Este precepto es la base conceptual de la mayoría de los tratados internacionales dedicados al control o la prohibición de armas, pero en ningún caso ofrece la garantía ni los mecanismos para que las víctimas puedan reclamar resarcir los daños sufridos en un ataque indiscriminado.

La ya citada definición de los regímenes internacionales de Krasner nos aportó la herramienta analítica del MTCR. Este régimen nació en las postrimerías de la Guerra Fría. Junto con el Tratado sobre Fuerzas Nucleares de Rango Intermedio (*Intermediate Range Nuclear Forces Treaty*, INF en adelante) se crearon en un ambiente de incipiente confianza mutua. No cabe ninguna duda que los mecanismos de control reforzaron este nuevo espíritu colaborativo en el desarme y el control.

Empero, el entorno actual es diferente. Se caracteriza por un incremento de desconfianza, una política exterior cada vez más agresiva de las grandes potencias, pero también de países más pequeños, y una transformación geopolítica en muchas zonas. Teniendo en cuenta la situación actual y las limitaciones del MTCR podemos concluir entonces que los sistemas de armas autónomos sí requieren un nuevo marco regulatorio,

dedicado a este objeto, y, además, de forma urgente.

¿Qué perspectivas para un marco regulatorio se presentan en la actualidad? Analizando las circunstancias se perfilan tres escenarios como los más probables.

Escenario 1: Se inicia una nueva ronda de negociaciones entre las superpotencias actuales. El final del INF y el próximo fin del Tratado de Reducción de Armas Estratégicas (*Strategic Arms Reduction Talks III*, START III en adelante) en febrero de 2021 podrían iniciar una nueva ronda de negociaciones multilaterales entre los países con potencial nuclear. Es posible que el presidente Trump esté provocando el final de los tratados de control de armas bilaterales, para incluir la nueva potencia China en una futura ronda de negociaciones. En la actualidad, China no está atado por ninguna clase de restricciones contractuales y ha podido incrementar su potencial militar durante las últimas décadas. A pesar de cultivar un discurso que evita un tono agresivo, el potencial armamentístico de China ha quedado patente. El presidente estadounidense expresó su interés de incluir a China en una nueva ronda de negociaciones sobre el control de armas, una vez que venza START III (Holland, 2020, 28 de febrero). En el caso de que China finalmente accediera a participar en estas negociaciones, se podría abrir una ventana de oportunidad para negociar controles sobre la regulación de armas autónomas.

Escenario 2: Los actores internacionales que apoyan la Campaña para la Prohibición de Robots Asesinos consiguen que gracias a la presión de la opinión pública un grupo de países voluntarios firmen un tratado multilateral donde se comprometen a no desarrollar ni emplear sistemas de armas autónomas. En la actualidad, los miembros de la campaña se reclutan casi exclusivamente de organizaciones de la sociedad civil de un total de 30 países. Aunque la campaña cuenta también con el apoyo de cargos políticos del entorno internacional, parece que ningún país productor de armas modernas se va a adherir de forma oficial a la campaña. Un régimen internacional que encuentre suficientes países voluntarios como las Convenciones de Ottawa y Oslo, no parece una solución probable por el momento.

Escenario 3: Las negociaciones dentro del marco del GGE se mantienen sin avances visibles, debido a la falta de voluntad política de los países miembros. No se llega a ningún acuerdo sobre la regulación de sistemas de armas autónomas. Los países que desarrollan armas autónomas no quieren renunciar a su desarrollo por motivos de seguridad estratégica frente a competidores regionales e internacionales. Tampoco les

interesa detener unilateralmente el desarrollo y exponerse así a una asimetría tecnológica que difícilmente podrían recuperar más tarde. Países como Turquía o Irán que muestran poco respeto por las reglas internacionales podrían alcanzar una ventaja tecnológica si los países occidentales detuvieran su desarrollo. La opacidad de muchos regímenes autoritarios alimenta un dilema de seguridad ante el cual la renuncia unilateral de los países occidentales a armas autónomas sería una solución muy arriesgada.

A los países exportadores de armas autónomas no les interesa someterse a restricciones que les impida la venta de modernos sistemas de armas y perder una cuota de mercado importante. La venta de armas es para los países desarrolladores una fuente de ingresos importante para compensar las inversiones que han hecho instituciones e industria en esta área. Además, un país exportador cuenta con una ventaja respecto al comprador que consiste en la dependencia tecnológica debida al umbral de inversión inicial y la necesidad de futuras actualizaciones.

No hemos incluido la regulación a través del MTCR en estos escenarios porque este régimen considera solamente el control de las tecnologías. No obstante, es posible que, tras el anuncio de la administración Trump de relajar las restricciones a la venta de drones el debate adquiera un nuevo impulso. Pero esta posibilidad es bastante remota, incluso en el caso de que se produjera un cambio en la presidencia en noviembre de 2020. Los presidentes de Estados Unidos han mantenido en la mayoría de los casos la línea de sus predecesores en las políticas de seguridad y defensa. Por otro lado, la cuestión de plena autonomía no se plantea para los decisores de la Casa Blanca. Otros actores, como la alianza atlántica y la Unión Europea también han recalcado reiteradamente que el hombre debe estar siempre dentro del proceso. China presentó en dos ocasiones una postura similar, mientras Rusia no se ha expresado al respecto.

Según la teoría de los regímenes internacionales, los intereses de los actores convergen en una determinada área que, en este caso, sería mantener al operador humano siempre dentro del bucle OODA. El hombre debe mantener siempre el control sobre los sistemas de armas que emplea, tal y como subrayan los documentos oficiales de Estados Unidos, la Unión Europea, Alemania y también de la OTAN. Esto implica que se mantenga siempre el modo de autonomía supervisada, no solamente para los RPAS, vehículos no tripulados terrestres y navales, sino también para las bandadas donde el control humano, como hemos visto, entorpece la organización interna de sus

componentes.

Lo deseable sería que los países democráticos dieran un paso adelante para formalizar esta posición en un nuevo tratado internacional dedicado a sistemas de armas autónomas letales, que incluya una serie de mecanismos de control mutuo como los conocemos de otros tratados como el INF o las diferentes versiones de START.

Los avances en la inteligencia artificial y las tecnologías intangibles auguran una serie de profundos cambios en el futuro que se van a producir tanto en la vida civil como en el sector militar. Como esta evolución tan sólo acaba de comenzar, se concibe este trabajo como una base para futuras investigaciones que profundicen en los aspectos que hemos esbozado aquí.

REFERENCIAS BIBLIOGRÁFICAS

- ANDERSON K. Y WAXMAN M. (2013). *Law and Ethics for Autonomous Weapon Systems: Why a Ban Won't Work and How the Laws of War Can*. Stanford University and The Hoover Institution.
- AZNAR LAHOZ, J.L. (02/2020). *La robótica en la guerra del futuro*. Instituto Español de Estudios Estratégicos, Documento Marco.
- BOSTROM, N. (2003). Ethical Issues in Advanced Artificial Intelligence. En *Cognitive, Emotive and Ethical Aspects of Decision Making in Humans and in Artificial Intelligence*, Vol. 2, 2003.
- BROMLEY, M. Y MALETTA, G. (2018). *The Challenge of Software and Technology Transfers to Non-Proliferation Efforts. Implementing and complying with Export Controls*. STOCKHOLM INTERNATIONAL PEACE RESEARCH INSTITUTE, APRIL 2018.
- FUHRMANN, M. Y HOROWITZ, M.C (2017). *Droning On: Explaining the Proliferation of Unmanned Aerial Vehicles*. En *International Organization*, Vol. 71, n° 2, pp. 397 – 418.
- GARCÍA PICAZO, P. (2009). *Teoría breve de relaciones internacionales*. Madrid, Tecnos.
- GILLI, A. Y GILLI, M. (2019). Why China Has Not Caught Up Yet: Military-Technological Superiority and the Limits of Imitation, Reverse Engineering, and Cyber Espionage., *International Security*, vol. 43, n ° 3, 2019, pp. 141-189.
- GÓMEZ DE AGREDA, A. (4/2019). La inteligencia artificial en el campo de batalla. En: Instituto Español de Estudios Estratégicos, *Usos militares de la inteligencia artificial, la automatización y la robótica (IAA&R)*, Documento de Trabajo.
- GONZÁLEZ MARTÍN, G. (2014) El impacto de las nuevas tecnologías y las formas de hacer la guerra en el diseño de las Fuerzas Armadas del futuro. En: Centro Superior de Estudios de la Defensa Nacional y Escuela de Altos Estudios de Defensa, *El impacto de las nuevas tecnologías y las formas de hacer la guerra en el diseño de las Fuerzas Armadas*. Documentos de Seguridad y Defensa, n° 61, 2014.
- GUTIÉRREZ C. Y CERVELL, M.J. (2013). *Sistemas de armas autónomas, drones y derecho internacional*. En *Revista del Instituto Español de Estudios Estratégicos*, n°2, 2013.
- HASENCLEVER, A., MAYER, P. Y RITBERGER, V. (1999). Las teorías de los regímenes internacionales: situación actual y propuestas para una síntesis. En *Foro Internacional* Vol. 39, No. 4 (158) (Oct. - Dec., 1999), pp. 499-526.
- HEYNS, C. (2013) *Informe del Relator Especial sobre las ejecuciones extrajudiciales, sumarias o arbitrarias (A/HCR/23/47)*.
- JORDÁN, J. Y BAQUÉS, J. (2014). *Guerra de drones: política, tecnología y cambio social en los nuevos conflictos*. Madrid: Biblioteca Nueva.
- LOSADA, J. (2014). *De la honda a los drones: la guerra como motor de la historia*. Barcelona: Pasado & Presente.
- MCDONALD, N. Y HOWELL, G. (2019). Killing Me Softly Competition in Artificial Intelligence and Unmanned Aerial Vehicles. En *National Defense University Press, Prism*, Vol. 8, No.3, 2019.
- MEARSHEIMER, J (2014). *The Tragedy of Great Power Politics*. Norton, New York. Versión actualizada.
- MELZER, N. (2019). *Derecho Internacional Humanitario. Una introducción*. CICR.
- MEZA RIVAS, M. (2018). *Los sistemas de armas autónomas: crónica de un debate internacional y prospectivo dentro de Naciones Unidas*. IEEE, Documento de opinión, 41/2018.

- MISSILE TECHNOLOGY CONTROL REGIME (2017). *Annex handbook*.
- MONTOYA, R. (2014). *Drones. La muerte por control remoto*. Madrid, ediciones Akal.
- RUSSEL, S. Y NORVIG, P. (2016). *Artificial Intelligence. A Modern Approach*. Pearson, 3ª edición.
- SCHARRE, P. (2018). *Army of None: Autonomous Weapons and the Future of War*. Norton & Company, New York.
- SILVELA CRIADO, E. Y LEÓN, G. (2020). Estrategia y tecnología. En Instituto Español de Estudios Estratégicos, *Panorama Estratégico 2020*, pp. 277-339.

ANEXO BIBLIOGRÁFICO

- ACNUR (19/06/2020). *Datos básicos*. <https://www.acnur.org/es-es/datos-basicos.html>
- AIIR SOURCE MILITARY (10/01/2017). *Perdix Drone Swarm – Fighters Release Hive-mind-controlled Weapon UAVs in Air*.
<https://www.youtube.com/watch?v=ndFKUKHfuM0&t=27s>
- AL JAZEERA (14/09/2019). *Drones hit 2 Saudi Aramco oil facilities*.
https://www.youtube.com/watch?v=E6Do_zVL8_A
- ALLEN, G.C. (06/02/2019). *Understanding China's AI Strategy. Clues to Chinese Strategic Thinking on Artificial Intelligence and National Security*. Center for a New American Security.
https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy?utm_medium=email&utm_campaign=Understanding%20Chinas%20AI%20Strategy%20White%20Paper%20Rollout%20Media&utm_content=Understanding%20Chinas%20AI%20Strategy%20White%20Paper%20Rollout%20Media+CID_eecd0d32a84c74e153b8409074cd04e&utm_source=Campaign%20Monitor&utm_term=Read%20the%20full%20paper
- BULGARIANMILITARY.COM (16/06/2020). *Russia is planning mass export of Su-57 and to enter the market of combat drones*. <https://bulgarianmilitary.com/2020/06/16/russia-is-planning-mass-export-of-su-57-and-to-enter-the-market-of-combat-drones/#:~:text=usContact%20us-Russia%20is%20planning%20mass%20export%20of%20Su%2D57%20and%20to,the%20market%20of%20combat%20drones&text=As%20Interfax%20reported%2C%20the%20Sukhoi,flew%20up%20to%20the%20sky>
- BUNDESWEHR (27/09/2012). *Mantis schießt scharf – die neue Flugabwehr der Bundeswehr*.
<https://www.youtube.com/watch?v=6xxcW8H1mfo>
- BUSINESS INSIDER (20/08/2019). *Russia's new high-altitude drone just flew for the first time, and they want to arm it with one ton of bomb*. <https://www.businessinsider.com/russias-new-high-altitude-drone-flew-for-the-first-time-2019-8?IR=T>
- CALVO ALBERO, J.L. (2020). *Global Strategy Report 31/2020: El auge de los drones turcos en Siria y Libia*. <https://global-strategy.org/el-auge-de-los-drones-turcos-en-siria-y-libia/>
- CAMPAIGN TO STOP KILLER ROBOTS (2020). *A growing global coalition*.
<https://www.stopkillerrobots.org/about/>
- CENTER FOR NEW AMERICAN SECURITY (December 2019). *The American AI Century: A Blueprint for Action*. <https://www.cnas.org/publications/reports/the-american-ai-century-a-blueprint-for-action>
- CENTER FOR SECURITY AND EMERGING TECHNOLOGY (28/10/2019) *National Strategy for the Development of Artificial Intelligence Over the Period Extending up to the Year 2030*.
<https://cset.georgetown.edu/wp-content/uploads/Decree-of-the-President-of-the-Russian-Federation-on-the-Development-of-Artificial-Intelligence-in-the-Russian-Federation-.pdf>
- CISION (06/01/2020). *Global Military Robotics Market was estimated to be US\$ 17340.00 Mn in 2018 and is expected to reach US\$ 53934.60 Mn by 2027 growing at a CAGR of 13.5% over the Forecast Period Owing to the Integration of Artificial Intelligence with Military Robotics, says Absolute Markets Insights*. <https://www.prnewswire.com/news-releases/global-military-robotics-market-was-estimated-to-be-us-17340-00-mn-in-2018-and-is-expected-to-reach-us-53934-60-mn-by-2027-growing-at-a-cagr-of-13-5-over-the-forecast-period-owing-to-the-integration-of-artificial-intelligence-w-300981624.html>

- CONGER, K. (01/06/2018). *Google Plans Not to Renew Its Contract for Project Maven, a Controversial Pentagon Drone AI Imaging Program.* <https://gizmodo.com/google-plans-not-to-renew-its-contract-for-project-mave-1826488620>
- CNN (27/05/2020). *US military accused Russia of deploying fighter aircraft to Libya.* <https://edition.cnn.com/2020/05/26/politics/russia-fighter-aircraft-libya/index.html>
- DARPA (s. f.) *Program Information: Gremlins.* <https://www.darpa.mil/program/gremlins>
- DEFENSA.COM (22/10/2019). *El primer sistema de protección activa “Trophy” de producción entregado al US Army.* <https://www.defensa.com/otan-y-europa/primer-sistema-proteccion-activa-trophy-produccion-entregado-us>
- DEFENSE NEWS (04/05/2020). *Boeing rolls out Australia’s first ‘Loyal Wingman’ combat drone.* <https://www.defensenews.com/air/2020/05/04/boeing-rolls-out-australias-first-loyal-wingman-combat-drone/>
- DEFENSE NEWS (24/07/2020). *Trump admin officially makes it easier to export military drones.* <https://www.defensenews.com/industry/2020/07/24/us-state-department-officially-makes-it-easier-to-export-military-drones/>
- DEPARTMENT OF DEFENSE. *Directive 3000.09. Autonomy in Weapon Systems.* <https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/DoD/300009p.pdf>
- DEPARTMENT OF DEFENSE (2018). *Summary of the National Defense Strategy of the United States of America. Sharpening the American Military’s Competitive Edge.* <https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf>
- DEPARTMENT OF DEFENSE (24/02/2020) *DOD Adopts Ethical Principles for Artificial Intelligence.* <https://www.defense.gov/Newsroom/Releases/Release/Article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/>
- DEVELOPING PAKISTAN (s. f.). *Chinese company Ziyan is negotiating sell Blowfish A2 to the governments of Pakistan and Saudi Arabia* <https://developing.pk/2019/11/23/chinese-company-ziyan-is-negotiating-sell-blowfish-a2-to-the-governments-of-pakistan-and-saudi-arabia/>
- DICCIONARIO DEL ESPAÑOL JURÍDICO (2020). *Derecho de los conflictos armados.* <https://dpej.rae.es/lema/derecho-de-los-conflictos-armados>
- EUROPEAN COMMISSION (2020). *Dronedeg-E.* FS/20/1079. https://ec.europa.eu/commission/presscorner/detail/en/fs_20_1079
- EUROPEAN GROUP ON ETHICS IN SCIENCE AND TECHNOLOGY (2018). *Statement on Artificial Intelligence, Robotics and Autonomous Systems.* http://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf
- FLEISCHER, J. (06/05/2020). *Bundesministerium der Verteidigung: Aspekte zum Einsatz bewaffneter Drohnen.* <https://www.bmvg.de/de/aktuelles/aspekte-einsatz-bewaffnete-drohnen-politisch-rechtlich-ethisch-254946>
- FOJÓN, E. (09/03/2018). *La cuarta revolución industrial, el algoritmo de guerra y su posible aplicación a la Defensa española.* http://www.realinstitutoelcano.org/wps/portal/rielcano_es/contenido?WCM_GLOBAL_CONTEXT=/elcano/elcano_es/zonas_es/defensa+y+seguridad/ari35-2018-fojon-4-revolucion-industrial-algoritmo-guerra-posible-aplicacion-defensa-espana
- FREEDOMHOUSE (2020). *China.* <https://freedomhouse.org/country/china/freedom-world/2020>
- FREEDOMHOUSE (2020). *Russia.* <https://freedomhouse.org/country/russia/freedom-world/2020>
- ISRAEL AEROSPACE INDUSTRIES (17/02/2019). *Mini HARPY: Multi-Purpose Tactical Loitering Munition.*

https://www.youtube.com/watch?time_continue=141&v=cBmhJ-jrzrc&feature=emb_title

GAUB, F. (04/2019). *Global Trends to 2030. Challenges and choices for Europe*. EU Institute for Security Studies.

https://www.iss.europa.eu/sites/default/files/EUISSFiles/ESPAS_Report.pdf

GUIMÓN, P. (03/01/2020). EE UU mata al poderoso general iraní Soleimani en un ataque con drones en el aeropuerto de Bagdad, *El País*.

https://elpais.com/internacional/2020/01/03/actualidad/1578010671_559662.html

HUMAN RIGHTS WATCH (2018). *Heed the Call. A Moral and Legal Imperative to Ban Killer Robots*. <https://www.hrw.org/report/2018/08/21/heed-call/moral-and-legal-imperative-ban-killer-robots>

HUMAN RIGHTS WATCH (2020). *Cuidado con el vacío. La falta de responsabilidad con respecto a los robots asesinos*. <https://www.hrw.org/es/report/2015/04/09/cuidado-con-el-vacio/la-falta-de-responsabilidad-con-respecto-los-robots-asesinos>

HOLLAND, S. (28/02/2020). *Trump willing to meet leaders of Russia, China, Britain, France on arms control*. <https://www.reuters.com/article/us-usa-trump-russia-summit/trump-willing-to-meet-leaders-of-russia-china-britain-france-on-arms-control-idUSKCN20M3CJ>

KANIA, E.B. (2020). *AI Weapons in China's Military Innovation*. Brookings, April 2020. <https://www.brookings.edu/research/ai-weapons-in-chinas-military-innovation/>

KLARE, M. (Marzo 2019a). *AI Arms Race Gains Speed*. En *Arms Control Association, Arms Control Today*. <https://www.armscontrol.org/act/2019-03/news/ai-arms-race-gains-speed>

KLARE, M (Marzo 2019b). *Autonomous Weapons Systems and the Laws of War*. En *Arms Control Association Control Today*. <https://www.armscontrol.org/act/2019-03/features/autonomous-weapons-systems-laws-war>

LETTER to Google C.E.O. <https://static01.nyt.com/files/2018/technology/googleletter.pdf>

MIDDLE EAST EYE (11/04/2018). *Houthi claim drone attack on Saudi Aramco oil refinery*. <https://www.middleeasteye.net/news/houthi-claim-drone-attack-saudi-aramco-oil-refinery>

NATO REVIEW (28/07/2017). *Autonomous military drones: no longer science fiction*. <https://www.nato.int/docu/review/articles/2017/07/28/autonomous-military-drones-no-longer-science-fiction/index.html>

NBC NEWS (15/05/2014). *Future Tech? Autonomous Killer Robots Are Already Here*. <https://www.nbcnews.com/tech/security/future-tech-autonomous-killer-robots-are-already-here-n105656>

ROSTEC (2020). *About, history*. <https://rostec.ru/en/about/history/>

RUSSIAN FEDERATION (05/02/2010). *The Military Doctrine of the Russian Federation*. https://carnegieendowment.org/files/2010russia_military_doctrine.pdf

RUSSIAN FEDERATION (25/12/2014). *The Military Doctrine of the Russian Federation*. <https://rusemb.org.uk/press/2029>

TASS (20/08/2019). *Russia's latest reconnaissance drone makes debut flight*. <https://tass.com/defense/1074237>

TASS (27/11/2019). *Director de Cooperación Internacional, Rostec: Rusia está lista para compartir tecnología*. <https://tass.ru/interviews/7203847>

- THE GUARDIAN (07/03/2018). *Google's AI is being used by US military drone programme*. <https://www.theguardian.com/technology/2018/mar/07/google-ai-us-department-of-defense-military-drone-project-maven-tensorflow>
- TUCKER, P (06/05/2014). *Every Country Will Have Armed Drones Within 10 Years*, Defense One, <https://www.defenseone.com/technology/2014/05/every-country-will-have-armed-drones-within-ten-years/83878/>
- TUCKER, P. (05/11/2019). *SecDef: China Is Exporting Killer Robots to the Mideast*. Defense One, <https://www.defenseone.com/technology/2019/11/secdef-china-exporting-killer-robots-mideast/161100/>
- STOCKHOLM INTERNATIONAL PEACE RESEARCH INSTITUTE (Abril 2019). *Trends in World Military Expenditure*, 2018. SIPRI Factsheet. <https://www.sipri.org/publications/2019/sipri-fact-sheets/trends-world-military-expenditure-2018>
- TUCKER, P. (4/1/2019). *Pentagon Seeks a List of Ethical Principles for Using AI in War*. En Defense One. <https://www.defenseone.com/technology/2019/01/pentagon-seeks-list-ethical-principles-using-ai-war/153940/>
- THE STATE COUNCIL OF THE PEOPLE'S REPUBLIC OF CHINA (20/07/2017). *A Next Generation Artificial Intelligence Development Plan*. Traducido al inglés por un grupo de trabajo de New America. <https://www.newamerica.org/cybersecurity-initiative/digichina/blog/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/>
- THE STATE COUNCIL OF THE PEOPLE'S REPUBLIC OF CHINA (September 2019). *China and the World in the New Era*. http://english.www.gov.cn/archive/whitepaper/201909/27/content_WS5d8d80f9c6d0bcf8c4c142ef.html
- WHITE HOUSE (2017). *National Security Strategy of the United States of America*. <https://www.whitehouse.gov/wp-content/uploads/2017/12/NSS-Final-12-18-2017-0905.pdf>
- UNITED STATES NAVY FACT FILE (15/01/2019). *MK 15 – Phalanx Close-in-Weapon System (CIWS)*. <https://www.navy.mil/Resources/Fact-Files/Display-FactFiles/Article/2167831/mk-15-phalanx-close-in-weapon-system-ciws>
- UNITED STATES NAVY FACT FILE (10/01/2019). *Aegis Weapon System*. <https://www.navy.mil/Resources/Fact-Files/Display-FactFiles/Article/2166739/aegis-weapon-system>
- WATSON, B. (12/01/2017). *The Drones of ISIS*, Defense One, <https://www.defenseone.com/technology/2017/01/drones-isis/134542/>
- WIKIPEDIA (2020). *Archivo:MQ-1 Predator, armed with AGM-114 Hellfire missiles*. https://es.wikipedia.org/wiki/Archivo:MQ-1_Predator,_armed_with_AGM-114_Hellfire_missiles.jpg
- WORLD ECONOMIC FORUM (15/06/2016). *We're on the brink of an artificial intelligence arms race. But we can curb it*. <https://www.weforum.org/agenda/2016/06/should-we-embrace-the-rise-of-killer-robots/>
- WORLD ECONOMIC FORUM (2020). *The Global Risks Report 2020 - 15th Edition*. <https://www.weforum.org/reports/the-global-risks-report-2020>
- WORLD ECONOMIC FORUM (s. f.) *Artificial Intelligence and Robotics*. <https://intelligence.weforum.org/topics/a1Gb000000pTDREA2?tab=publications>

XL SEMANAL (14/08/2018) *Los 'Top gun' españoles de los drones.*
<https://www.xlsemanal.com/conocer/tecnologia/20180814/drones-militares-en-espana.html>

ZIYAN UAV (s. f.). <http://www.ziyanuav.com/index.html>