

LA RESPONSABILIDAD DEL ESTADO POR EL USO DE ARMAS AUTÓNOMAS LETALES

*Premio de Defensa 2022, en su modalidad
Premio José Fernando Querol y Lombardero*

Laura González Calvache

Resumen

Las armas autónomas letales son un tipo de armas que, empleando la inteligencia artificial, son capaces de seleccionar objetivos y decidir si disparan o no. Este artículo examina la responsabilidad del Estado por su uso, pues deben estar sujetas a control humano para ser conformes a derecho. Así, Estados y organizaciones proponen dos posibles estándares de control: el control humano significativo y los niveles apropiados de control humano. También se examina la conformidad de las armas con el derecho internacional humanitario a través de los principios de distinción y proporcionalidad, concretamente si estas armas serían capaces de distinguir entre objetos y si sus ataques serían proporcionales. Finalmente, se analizan y proponen distintas vías de ejercer la responsabilidad del Estado.

Palabras clave: Armas autónomas letales, Control humano significativo, Derecho internacional humanitario, Drones.

Abstract

Autonomous weapon systems are a class of weapons that, use artificial intelligence to select targets and decide whether or not to fire. This article

examines the State's responsibility for their use; the weapons must be subjected to human control to be according to law. States and organizations therefore propose two possible standards, the meaningful human control, and the appropriate levels of human control. Their compliance with international humanitarian law is examined through the principles of distinction and proportionality, more precisely whether these weapons are capable of distinguishing between objects and whether their attacks can be proportionate. Finally, different ways of holding States accountable are analysed and proposed.

Keywords: Autonomous weapon systems, Meaningful human control, International humanitarian law, Drones.

Sumario

1. Introducción. 2. Definiendo las armas autónomas letales. 2.1. Autonomía y armas autónomas letales. 2.2. Otros conceptos técnicos referidos a las armas autónomas letales. 2.3. Límites tecnológicos de las armas autónomas. 2.4. In the loop, on the loop, out of the loop. 3. Manteniendo el control. 3.1. Control humano significativo (MHC). 3.2. Niveles apropiados de juicio humano sobre el uso de la fuerza (AHJ). 4. Las armas autónomas letales y el derecho internacional humanitario. 4.1. Las armas autónomas y los principios de derecho internacional humanitario. 5. La responsabilidad del Estado por el uso de armas autónomas letales. 5.1. Responsabilidad internacional del estado por hechos ilícitos. 5.2. Responsabilidad del estado por infracción del *due diligence*. 5.3. Responsabilidad civil por los daños ocasionados (*strict liability*). 5.4. Desafíos procesales. 6. Conclusiones. 7. Bibliografía.

ÍNDICE DE ABREVIATURAS

AHJ	Appropriate Leves of Human Jugment
API	Protocolo I adicional a los Convenios de Ginebra de 1949 relativo a la protección de las víctimas de los conflictos armados internacionales
ARSIWA	Articles on Responsibility of States for Internationally Wrongful Acts
AWS	Autonomous Weapons Systems
CCW	Convention on Conventional Weapons
CDEM	Collateral Damage Estimate Methodology
CIJ	Corte Internacional de Justicia
CNAS	Center for a New American Security
DIH	Derecho Internacional Humanitario
EE. UU.	Estados Unidos
GGE	Group of Governmental Experts
HRW	Human Rights Watch
IA	Inteligencia Artificial
ICRAC	International Committee for Robot Arms Control
ICRC	International Committee of the Red Cross
ILC	International Law Commission
LAWS	Lethal Autonomous Weapons Systems
MHC	Meaningful Human Control
ONU	Organización de las Naciones Unidas
SCIJ	Sentencia de la Corte Internacional de Justicia
STPIY	Sentencia del Tribunal Penal Internacional para la antigua Yugoslavia

1. INTRODUCCIÓN

Las armas autónomas letales son armas que seleccionan por sí mismas objetivos a los que detectan como potenciales amenazas y deciden si disparar o no.

Estados y organizaciones internacionales han pedido su prohibición por los riesgos que genera. Sin embargo, teniendo en cuenta que más de diez países las están desarrollando actualmente —entre los que se encuentran los cinco miembros permanentes del Consejo de Seguridad de la ONU—, es posible aventurar que las armas autónomas van a revolucionar los conflictos armados del futuro, siendo una tecnología aún en desarrollo.

Uno de los problemas principales que generan las armas autónomas letales es la disolución de la responsabilidad. La selección de objetivos y la decisión de disparar son realizadas por la máquina, lo que impide asegurar a quién compete exigir responsabilidades: ¿al fabricante, al ingeniero que programó los algoritmos que hacen operar el sistema, al operador que debía controlar la máquina o al superior que autorizó su empleo? Parte de la doctrina, incluyendo a Human Rights Watch, Amnesty International o ICRC apuestan por la inclusión de un régimen de responsabilidad individual que aporte garantías a posibles víctimas por los accidentes que puedan provocar. Este trabajo se ocupa de la responsabilidad internacional del Estado sin prejuzgar los otros tipos de responsabilidad que puedan derivarse del uso de armas autónomas letales, como la responsabilidad penal. La responsabilidad del Estado es importante para su aplicación a las armas autónomas letales por varios motivos:

- Son los Estados los que están invirtiendo en el desarrollo de las armas autónomas letales.
- Los Estados son actores políticos principales en los conflictos armados, tanto nacionales como transnacionales.
- La responsabilidad del Estado podría desincentivar su uso, o provocar un uso más restrictivo.
- El Estado tiene fondos suficientes como para satisfacer las compensaciones que se deriven, a diferencia de los individuos.

La responsabilidad del Estado responde a una vertiente tradicional, por hechos ilícitos, y a una vertiente más nueva, basada en un régimen de responsabilidad como consecuencia de la omisión de un deber de cuidado. Numerosos autores están proponiendo además un nuevo régimen de responsabilidad civil por daños, que permitiría imputarle responsabilidad al Estado con mayor facilidad que en los dos casos anteriores.

La responsabilidad del Estado por hechos ilícitos se produce por una infracción de derecho internacional, que puede ser una acción o una omisión, y esta conducta debe serle atribuible al Estado. Partiendo de la base de que las armas autónomas letales serán aplicadas principalmente en un

contexto de conflicto armado, el hecho ilícito internacional ocurrirá en el marco del derecho internacional humanitario. A estos efectos, las armas autónomas letales plantean problemas principalmente con los principios de distinción entre personas civiles, militares y no combatientes, y proporcionalidad, según la fuerza del ataque. El incumplimiento de cualquiera de estos principios constituiría un hecho ilícito internacional, que debe serle atribuible al Estado para que genere responsabilidad.

La responsabilidad del Estado por infracción del *due diligence* responde al riesgo derivado de la realización de actividades en principio no prohibidas, pero potencialmente generadoras de daños a terceros. Este régimen, aún en desarrollo, suele alegarse en contextos de derecho ambiental, pero está siendo reivindicado por parte de la doctrina ante las dificultades de ejecutar el tipo de responsabilidad por hechos ilícitos, asegurando la reparación de las víctimas.

Finalmente, se propone un régimen de responsabilidad civil que funcionaría de forma supletoria para los tipos de responsabilidad anteriores, a fin de asegurar el cumplimiento de la responsabilidad.

2. DEFINIENDO LAS ARMAS AUTÓNOMAS LETALES

2.1. AUTONOMÍA Y ARMAS AUTÓNOMAS LETALES

Una definición aceptada y extendida de las armas autónomas letales es que son «unos sistemas que, una vez activados, pueden seleccionar y atacar objetivos sin necesidad de intervención adicional de un operador humano» (Departamento de Defensa de EE. UU., 2012: 13-14. También HRW, 2018a: 2).

Las armas autónomas letales adquieren su autonomía a través de algoritmos de inteligencia artificial y *machine learning*, lo que les convierte en sistemas de *software* avanzado capaz de distinguir objetos con los datos de que dispone.

Actualmente, hay tecnología cada vez más sofisticada que permite la actuación a kilómetros de distancia, como los drones que pueden manejarse en remoto. La creciente automatización del armamento permite prever el surgimiento de armas autónomas, que se distinguirían de las armas automáticas y las semiautomáticas o automatizadas. Esta distinción depende de la sofisticación técnica y capacidad de los sistemas. La principal diferencia entre las armas autónomas y las semiautomáticas es que, si bien

ambas actúan conforme a patrones que les fueron configurados, las armas autónomas son capaces de corregir sus procesos internos y ajustarlos a sus acciones a través de los datos recibidos (Barela y Plaw, 2021: 75).

Marra y McNeil (2012: 1150-1154) consideran que la principal característica del sistema autónomo es su capacidad de actuar sin intervención externa. Proponen tres atributos de autonomía: la frecuencia de la interacción entre la máquina y el operador, la capacidad de la máquina para responder en situaciones complejas de incertidumbre y su nivel de asertividad en la toma de decisiones.

El primer atributo se refiere a la cantidad de veces que el operador necesita dictarle instrucciones a la máquina. A mayor autonomía, menor necesidad de contacto. El segundo atributo trata de medir la adaptabilidad de la máquina: si es capaz de predecir y evitar colisiones, de evitar distintos obstáculos fijos o móviles o ser capaz de desempeñar sus funciones en un entorno cambiante. A mayor capacidad de adaptación, mayor autonomía. Finalmente, el tercer atributo se refiere al nivel de asertividad de la máquina que se mide por su habilidad para completar una misión, siendo capaz de cambiar los medios necesarios para conseguir el fin encomendado, sin intervención de persona alguna.

Schaub y Kristoffersen (2017: 8) amplían estas descripciones. Las máquinas automáticas responden a estímulos externos concretos (*inputs*) con *outputs* prestablecidos, sin que la máquina posea un sistema operativo ni esté configurada por algoritmos. En esta categoría se encontrarían las minas o los cables trampa. Los sistemas automatizados se basarían en un conjunto de alternativas diseñadas en función del estímulo percibido por la máquina, funcionando con algoritmos, lo que permitiría predecir su resultado. Ejemplos de armas automatizadas serían Predator en el aire, SWORDS en tierra y Aegial en el agua (Marra y McNeil, 2012: 1160-1173). También la cúpula de hierro de Israel (Schmitt y Thurnher, 2013: 236) que responde mandando misiles para interceptar la amenaza y luego al lugar del que partió el misil atacante. Todas ellas responden a un grupo prefijado de estímulos sin necesidad de intervención humana. Finalmente, las autónomas serían una evolución de las automatizadas por la complejidad de sus elementos, y sería potencialmente impredecible y con un comportamiento orientado a la consecución de objetivos.

En el ámbito de la ingeniería, Sheridan (Parasuraman, Sheridan y Wickens, 2000: 287) desarrolló una escala de diez grados que ilustra con ejemplos cuál es el nivel de sofisticación técnica que se espera que alcancen las armas autónomas:

Grado de autonomía	Descripción de la autonomía
1	El ordenador no ofrece asistencia alguna, el humano debe hacerlo todo.
2	Ofrece un conjunto de decisiones o alternativas.
3	Disminuye la selección a unas pocas.
4	Sugiere una entre ellas.
5	Ejecuta su recomendación si el humano lo aprueba.
6	Otorga al humano un periodo de tiempo para cancelar la operación antes de la ejecución automática.
7	Ejecuta automáticamente, después informa al humano de los resultados.
8	Solo informa al humano del resultado si el humano le pregunta.
9	Solo se informa al humano si el ordenador decide hacerlo.
10	El ordenador puede decidirlo todo, ignorando completamente al humano.

Si las armas autónomas consiguieran alcanzar el último paso, el control humano quedaría completamente aislado del proceso de toma de decisiones, especialmente en la selección de objetivos. El ICRC (2019: 7) y Heyns (2013: 8) describen a estas máquinas como aquellas que se basan en el paradigma percibir-pensar-actuar. Este paradigma se basa en la herramienta OODA Loop desarrollada por John Boyd (Marra y McNeil, 2012: 1144-1145), por el que las máquinas reciben datos del entorno que les permite, mediante el procesamiento con algoritmos, conocer el tipo de situación en la que se encuentran; procesadores que, entre varias vías de acción posibles, adoptarían una, y mecanismos de ejecución de la decisión tomada. En el caso de las armas autónomas letales, se trataría de recopilar imágenes de personas, procesar si se trata de amenazas o no y, en el caso de que lo sean, ejecutar el disparo.

2.2. OTROS CONCEPTOS TÉCNICOS REFERIDOS A LAS ARMAS AUTÓNOMAS LETALES

Al hablar de armas autónomas, se manejan distintos conceptos del campo de la informática que, no obstante, los juristas necesitan conocer para entender los riesgos potenciales asociados a las armas autónomas letales y los desafíos jurídicos que plantean.

Los algoritmos son «secuencias de instrucciones de programación que, ejecutadas por un ordenador, realizan cálculos o resuelven problemas» (ICRC, 2019: 13). Los algoritmos suelen ser predecibles en sus respuestas, pero pueden variar dependiendo de las circunstancias. Algunos algoritmos

permiten a la máquina reaccionar al entorno, pero otros no, y el sistema tiene una capacidad de adaptación al entorno muy baja (ICRC, 2019: 13). Es entonces cuando entra en juego la inteligencia artificial.

La inteligencia artificial es un programa informático capaz de procesar enormes cantidades de datos y de generar soluciones a partir de ellos. Destacan dos características de la inteligencia artificial: la primera de ellas es que un sistema inteligente que, una vez «entrenado», es capaz de ajustar su funcionamiento con base en los nuevos datos recibidos sin la intervención directa del ingeniero que lo programó originariamente. La segunda característica, que genera confusión, es que la inteligencia artificial funciona a partir de modelos probabilísticos y no de procesos cognitivos humanos (Baena, 2018: 31-32). Actualmente, los sistemas son de inteligencia débil, pero aun así son capaces de realizar traducciones, leer los labios o reconocer objetos (Baena, 2018: 36), que es la principal función de las armas autónomas letales. Cabe destacar que, con el avance de la inteligencia artificial, sistemas que antes eran inteligentes ahora son considerados meramente automatizados (ICRC, 2019: 13).

Los sistemas de *machine learning* son sistemas de inteligencia artificial capaces de «aprender» de los datos que tienen. En lugar de seguir programas prefijados, los sistemas basados en *machine learning* construyen su propio modelo basado en los datos proporcionados y son capaces de realizar acciones, identificar patrones o hacer predicciones. El problema con estos sistemas es que, a diferencia de con los algoritmos normales, los programadores no pueden predecir con exactitud cómo va a realizar el sistema la asociación con los datos, por lo que puede ser difícil de controlar. Este es el tipo de sistema que utilizan las armas autónomas letales (ICRC, 2019: 15).

Las redes neuronales son un modelo computacional avanzado que imita el funcionamiento del cerebro humano. Forman parte del *machine learning* y permiten al sistema adquirir, almacenar y enviar información de forma muy eficiente, siguiendo el modelo de conexiones neuronales del cerebro. El cerebro cambia sus conexiones neuronales a medida que aprende y realiza tareas, en la llamada plasticidad cerebral. Las redes neuronales computacionales también cambian sus conexiones para adaptarse al nuevo aprendizaje, y permiten al sistema cambiar su programación para producir mejores resultados (Krishnan, 2021: 220).

El *deep learning* es un concepto bastante novedoso, y se refiere a la aplicación de las redes neuronales a grandes bases de datos. El proceso de aprendizaje con *deep learning* consiste en enseñar a los ordenadores a diferenciar objetos enseñando un gran número de imágenes sobre esos objetos

precedentes de las grandes bases de datos. El procedimiento suele ser supervisado por programadores que le expliquen al sistema si sus soluciones son buenas o no y le corrijan, lo que ayudará al sistema a mejorar. Cuantos más datos se obtengan, mayor será la precisión que alcance el sistema en sus soluciones, y podría superar al cerebro humano al no verse influenciado por los sesgos cognitivos típicos del cerebro –cuestión aparte merecen los sesgos de las propias máquinas, que se tratan en el siguiente epígrafe–. También se espera que el ordenador sea capaz de llegar a soluciones a las que los seres humanos no serían capaces, al realizar nuevas asociaciones que los humanos no se habían planteado (Krishnan, 2021: 221).

Dos elementos propios de los sistemas informáticos, que serán claves a la hora de tratar las armas autónomas letales, son la predictibilidad y la fiabilidad. El primero se refiere a «la habilidad de decir o estimar que una determinada cosa pasará en el futuro o que será consecuencia de algo» (ICRC, 2019: 10). En las armas autónomas, se refiere a la capacidad de saber cómo funcionará el arma una vez puesta en funcionamiento. La segunda, la fiabilidad, se describe como «la cualidad de ser de confianza o comportarse generalmente bien». En las armas autónomas, esta cualidad se refiere a el arma actuará como se espera que lo haga, sin fallos ni consecuencias inesperadas (ICRC, 2019:10).

2.3. LÍMITES TECNOLÓGICOS DE LAS ARMAS AUTÓNOMAS

El funcionamiento de los sistemas de las armas autónomas puede verse viciado por sesgos y por la *black box*, provocando que no sean sistemas del todo fiables ni predecibles.

Los sesgos se refieren al «resultado discriminatorio que provoca la aplicación de un algoritmo programado a partir de datos alejados de la realidad, o que reflejen los prejuicios del humano que los ha creado o analizado» (Baena, 2018: 82). Se conocen varios sesgos (ICRC, 2019:18):

- Sesgo en el entrenamiento con datos o sesgo estadístico. Como los algoritmos de *machine learning* utilizan datos para definir sus modelos, limitaciones en la cantidad, cualidad o naturaleza de estos datos puede introducir sesgos en el funcionamiento de los datos. Una crítica frecuente que reciben las armas autónomas es que la mayoría de los datos proceden de sociedades occidentales, dificultando el reconocimiento facial de personas no blancas.

- Sesgo en el objetivo. El algoritmo les otorga un peso distinto o inapropiado a los elementos de los datos e ignora otros, llegando a conclusiones no respaldadas por los datos. En las armas autónomas, el sistema podría darle más peso al color de la vestimenta que a portar armas.
- Sesgo en el procesamiento. El propio algoritmo introduce un sesgo al analizar los datos. Con frecuencia, los programadores introducen estos sesgos para detectar y eliminar otros sesgos.
- Sesgo emergente. Este sesgo puede provocar que un algoritmo funcione de formas impredecibles. Lo provoca el entorno en que opera, no el programador ni el operador.
- Sesgo por transferencia de contexto. El algoritmo se utiliza en contextos distintos de aquellos para los que fue diseñado, provocando impredecibilidad.
- Sesgo interpretativo. Un usuario, puede ser un humano u otro sistema, malinterpreta la conclusión extraída por el primer algoritmo. Si el arma autónoma que está operando utiliza los datos de los objetivos extraídos por otro sistema, atacará por error a los objetivos incorrectos.
- Sesgos humanos El sistema hereda el comportamiento de las personas que realizaron los datos (Baena, 2018: 83). En las armas autónomas, si en los datos seleccionados se ve que solían seleccionarse a hombres con barba, el arma autónoma tenderá a seleccionar hombres con barba.

Mención aparte merece el sesgo de complacencia, que es un sesgo del operador provocado por el sistema: los pilotos confían tanto en él que no perciben los fallos de este, llevando a consecuencias fatales. En el estudio realizado por Endsley y O’Kiris (Endsley, 1995: 382-383), en el que se examinan los comportamientos de pilotos con la automatización de las aeronaves, observaron que los pilotos eran más propensos a tener fallos que concluían en accidentes fatales cuando los aviones eran automáticos. La automatización de las aeronaves provocaba que los pilotos perdiesen la consciencia sobre el entorno, entendida como «la percepción de los elementos del entorno en un tiempo y espacio determinados, la comprensión del significado de estos elementos, y la proyección de su situación en un futuro cercano».

La *black box* (Krishnan, 2021: 221) se encuentra en los sistemas que funcionan con redes neuronales. Como las asociaciones no se hacen solo aplicando códigos, sino que también interactúan con las bases de datos

y se modifican conforme a ellas, por lo que se complica la trazabilidad de las conclusiones. Para los propios programadores es difícil seguir las asociaciones y conexiones realizadas por la red neuronal para llegar a un determinado resultado. La *black box* complica la labor de supervisión tanto de los programadores como de los operadores y vuelve el sistema potencialmente impredecible.

2.4. IN THE LOOP, ON THE LOOP, OUT OF THE LOOP

Una vez explicado el concepto de autonomía que se maneja al hablar de armas autónomas letales, es necesario definir la posición de la persona, del operador humano. El concepto de *loop* sigue el proceso ya mencionado de OODA Loop diseñado por John Boyd (Marra y McNeil, 2012: 1144-1145), y obtiene de él el nombre. En inglés se manejan tres términos en función del lugar que ocupe el operador en el proceso de toma de decisiones (el *loop*): *in the loop*, *on the loop* y *out of the loop*.

In the loop se refiere al uso de herramientas por parte de combatientes, una categoría tradicional que incluye el empleo de armas y espadas (Malik, 2018: 622). Human Rights Watch ofrece una definición más centrada en las armas actuales, y es que los robots solo podrían seleccionar objetivos y actuar bajo las instrucciones del operador, como los drones teledirigidos (HRW, 2018a: 2).

On the loop es un tipo de control que se ejerce sobre armas desarrolladas tecnológicamente, pero en las que el operador conserva un poder de decisión relevante. El papel desempeñado por el operador puede ser de supervisión e intervención, aportando al operador datos suficientes sobre el entorno (ICRC, 2019: 9), o un poder de veto (Schaub, y Kristoffersen, 2017: 10). Human Rights Watch incluye ambos aspectos como control *on the loop* de las armas autónomas letales.

Out of the loop es un tipo de control mínimo, pues la máquina habría alcanzado plena autonomía y no necesitaría la intervención del operador (HRW, 2018a: 2). Es el último nivel de los descritos por Sheridan, que se ilustraban previamente. Que la máquina opere *out of the loop* es el eje central de la polémica que suscitan las armas autónomas letales, desde una perspectiva ética y jurídica. La disolución de la responsabilidad es patente, pues es difícil imputar la responsabilidad por los fallos de una máquina que carecía de control alguno.

Con el fin de limitar la autonomía de las armas autónomas letales y sus consecuencias, parte de la doctrina está desarrollando el concepto de

«control humano significativo» o «MHC» por sus siglas en inglés. El control humano significativo buscaría imponer un estándar mínimo de control cuando se utilicen las armas autónomas letales. Encuadrándose en el tipo de control *on the loop*, establece varios requisitos para su utilización, entre ellos la responsabilidad.

El control *on the loop*, tanto en su vertiente de supervisión como en el ejercicio del poder de veto, permite que el operador aún pueda tomar decisiones, y por tanto sería posible imputarle responsabilidad individual. Resulta relevante para su uso en conflictos armados, pues estas armas están pensadas para utilizarse en conflictos armados y, por tanto, si se cometiese alguna infracción del derecho internacional mediante su uso, sería del derecho internacional humanitario.

3. MANTENIENDO EL CONTROL

Los principios de proporcionalidad y distinción del derecho internacional humanitario imponen unos estándares sobre las armas empleadas en conflictos armados para que el uso sea legal. Como bien señala Article 36¹ en su informe (2013: 3) la aplicación de la inteligencia artificial a las armas es posible y legal gracias a la interacción entre las armas y el operador humano. En el caso de las armas semiautónomas, que funcionan por sensores, su conformidad a la ley depende de que el operador humano conozca tanto el funcionamiento de los sensores como el contexto sobre el que actúa el arma. Reino Unido aseguró que «siempre debería haber control y autorización humana al disparar» (2015, párr. 8).

Los cinco miembros permanentes del Consejo de Seguridad de la ONU han declarado la importancia de mantener el control humano sobre los sistemas autónomos: China: «las discusiones sobre la interacción humano-máquina [...] deberían definir el modo y el grado de implicación humana» (2018: 1); Rusia: «no ponemos en duda la necesidad de mantener el control humano sobre la máquina» (2018: 2); Francia: «el uso de la fuerza forma parte de la responsabilidad inherente [al individuo] y requiere de mayor elaboración y aclaración» (2018: 2); Reino Unido: «El uso de nuestras armas siempre permanecerá bajo control humano como garantía absoluta de la supervisión y autoridad humanas, y la responsabilidad por el uso de armas» (2018: 2); y, Estados Unidos: «los sistemas de armas autónomas y

¹ Organización no gubernamental radicada en Reino Unido dedicada a reducir el uso de armas y analizar las implicaciones de su uso.

semiautónomas deben ser diseñadas de forma que permitan a los comandantes y operadores ejercer niveles apropiados de juicio humano sobre el uso de la fuerza» (2012: 2).

Planteado por primera vez por la organización Article 36 en 2013, el concepto de control humano significativo (MHC, por sus siglas en inglés) ha tenido gran acogida a nivel internacional, siendo objeto de debate en las reuniones de la CCW en 2014. El Parlamento Europeo, en una declaración conjunta (2018: 4-5), utiliza el mantenimiento del control humano significativo para distinguir las armas autónomas letales (que estarían prohibidas por el derecho internacional) de las armas autónomas, que sí serían conformes a derecho.

El MHC busca establecer una garantía de control humano en el empleo de sistemas de armas autónomas letales. A pesar de que aún no se ha alcanzado un consenso sobre su definición, gobiernos como el de India, Rusia y, especialmente, Estados Unidos, se muestran críticos con este concepto, proponiendo EE. UU. su propio concepto, los niveles apropiados de juicio humano sobre el uso de la fuerza (Departamento de Defensa de EE. UU., 2012: 2).

La adopción de cualquiera de estos estándares de control minimizaría los riesgos derivados de su funcionamiento y permitiría operar en conflictos armados acorde al derecho internacional humanitario. Asegurando el control humano se asegura también la responsabilidad individual, que podría facilitar la responsabilidad del Estado.

En este apartado se explicarán qué formas de control pueden ejercerse sobre las armas autónomas letales, y después se explicarán las propuestas para definir el control humano significativo por parte de distintos organismos, las diferencias con el nivel apropiado de juicio humano (la noción de EE. UU.) y las críticas a las que se enfrenta.

3.1. EL CONTROL HUMANO SIGNIFICATIVO (MHC)

No hay todavía una definición aprobada para el control humano significativo, y distintas organizaciones exponen los elementos mínimos que debería integrar para poder convertirse en un estándar de control. Así, la organización Article 36 decidía promover el concepto de *meaningful human control*, que es el que goza de mayor aceptación en la actualidad. En un principio, el informe de 2013 se centró en explicar la relevancia del control humano significativo y se refirió a tres aspectos: la información recibida por el operador, la forma de ataque y la responsabilidad por los posibles daños ocasionados (Article 36, 2013: 4). Ha sido en 2016 (Article 36,

2016: 1-4) cuando ha perfilado los elementos que integran el control humano significativo:

- Tecnología predecible, fiable y transparente. La tecnología es predecible cuando puede preverse el curso de acción de la tecnología; fiable, cuando es difícil que falle y está diseñada para limitar los daños en caso de error; por último, es transparente si los usuarios pueden entender cómo funciona.
- Información veraz sobre el resultado esperado, la tecnología en uso y el contexto en que opera. El control humano sobre la tecnología se basa en la capacidad de decidir efectuar un ataque con la información disponible. Es muy relevante para determinar la intención del comandante, claro elemento subjetivo de la responsabilidad. Luego la información sobre el resultado buscado es importante también para evaluar las circunstancias en que se tomó la decisión y las consecuencias que el comandante quería evitar. La información permitirá conocer cuáles eran los objetivos buscados —*target profile*— para determinar si algunos objetivos fueron determinados como tales intencionadamente o no.
- Article 36 introduce aquí el elemento de la predictibilidad como la conexión entre la intención del comandante y la probabilidad de que los resultados coincidan con la voluntad del comandante. La predictibilidad es también una característica de la interacción entre la tecnología y el entorno, para evitar graves consecuencias. El entorno variará según sea tierra, aire o agua, y es esencial para modular la actuación si hay civiles cerca.
- Comprender la tecnología que se está utilizando es necesario para que el comandante pueda evaluar los daños potenciales a civiles y los riesgos derivados de la actuación que se propone realizar.
- Juicio y acción humanos oportunos, y el tiempo disponible para poder intervenir. Este elemento permite tener un punto de conexión con la teoría de la responsabilidad por las acciones realizadas y los daños provocados. Debe tenerse en cuenta el tiempo de que se dispone para la toma de decisiones.
- Responsabilidad a un cierto estándar. Article 36 llama a establecer un sistema de responsabilidad que no incluya únicamente al comandante, sino también a fabricantes y programadores.

El régimen de control humano significativo propuesto por Article 36 solo podría aplicarse a ataques individuales, porque en los colectivos no puede ejercerse control humano significativo al dirigirse contra un colectivo genérico (2013: 1 y 3).

Posteriormente, en 2014, el ICRC (2014: 2) señaló tres condiciones que debe cumplir el control humano significativo:

- El operador debe tener un manejo perfecto del entorno para poder reaccionar a cambios en el entorno o situaciones imprevistas.
- Debe haber participación cognitiva durante el ataque y tiempo suficiente para deliberar sobre la selección de objetivos, en referencia a la necesidad del ataque y los riesgos potenciales.
- Posibilidad de abortar el ataque en cualquier momento.

En 2019 (ICRC, 2019: 13-14) publicó un nuevo informe que proponía nuevas condiciones:

- El arma debe tener un cortafuegos que permita su inutilización en caso de error informático.
- Debe servir como un «catalizador para la responsabilidad», asegurando que habrá consecuencias legales para las actuaciones ilegales, de conformidad al derecho internacional.
- Debe asegurar que las decisiones sobre la vida y la integridad física las toma un agente moral, una persona, en lugar del algoritmo.
- El Centre for a New American Security (CNAS) distinguió en 2015 tres componentes del control humano significativo (Scharre y Horowitz, 2015: 15):
- El operador toma decisiones informadas y conscientes sobre las armas.
- El operador tiene información suficiente sobre el objetivo, el arma y el entorno como para asegurar la legalidad de la acción que va a desarrollar.
- El arma ha sido probada adecuadamente, y el operador ha recibido formación específica que asegura su control efectivo sobre el arma.

Sobre la responsabilidad, Scharre y Horowitz no incluyen un punto específico sobre ello, pero consideran que el control humano significativo conlleva la capacidad de ejercer la responsabilidad porque es posible conectar a un comandante con la acción (Scharre y Horowitz, 2015: 15).

Por último, el ICRC (2016: 4-5) estableció cinco criterios básicos para el control humano significativo:

- Predictibilidad del arma en el contexto en que se pretende usar.
- Fiabilidad del arma en el contexto en que se pretende usar.
- Supervisión humana durante el desarrollo y uso del arma.
- Tener información suficiente sobre el funcionamiento del arma en su contexto.
- Responsabilidad por las consecuencias de su uso.

Las propuestas sobre el contenido del control humano significativo expuestas por las organizaciones internacionales son diversas y, como se ha visto, siguen modulándose. Pese al esfuerzo realizado por estas organizaciones, el control humano significativo sigue siendo objeto de controversia y recibe críticas por parte de los Estados y los propios académicos.

Scharre y Horowitz, del CNAS, criticaron algunas de las propuestas del ICRC de 2014, concretamente la información recibida del entorno para poder tomar decisiones y la capacidad de abortar el ataque en cualquier momento. Estos requisitos ofrecen una visión irreal del control humano y sientan una exigencia poco realista sobre la predictibilidad del curso de acción (Scharre y Horowitz, 2015: 9). Además, la exigencia de abortar en cualquier momento es imposible de realizar, pues no puede retraerse un proyectil después de que haya sido lanzado (2015: 9). Extienden esta crítica sobre la información suficiente a Article 36, porque su estándar es igualmente irreal (2015: 11). El ICRC ha eliminado ambas exigencias del informe de 2019, como se ha podido apreciar, pero Article 36 lo ha mantenido, matizando el requisito.

Por otro lado, Crootof (2016: 57-58) considera que estos requisitos no eliminan el problema de la imprecisión del concepto. Términos como «información suficiente», «decisión informada» y «entrenamiento apropiado» son conceptos igualmente ambiguos y fáciles de rebatir.

Algunos Estados se muestran disconformes con el concepto, sugiriendo que es demasiado vago e impreciso (Crootof, 2016: 57-58). Estados Unidos propuso su propio estándar en 2012 que se verá en el apartado siguiente, y no ha apoyado la adopción del control humano significativo, considerándolo subjetivo y difícil de entender, además de entorpecer la discusión sobre la interacción humano-máquina (Estados Unidos, 2016, párr. 7). India expresó su preocupación en la CCW de 2014 porque la adopción de este concepto ayudase a legitimar estas armas (HRW, 2020, párr. «India»). El concepto de control humano significativo presupone el desarrollo de estrategias políticas y diplomáticas sobre la guerra y la estra-

tegia, estrategias que llevan desarrollándose en países occidentales desde la Guerra Fría y que acrecentarían la brecha con los demás países, especialmente aquellos que salen de la colonización. La propia idea de «control significativo» conlleva sesgos cognitivos basados en la visión que se tiene de seguridad y estrategia desarrollados en occidente durante décadas (Williams, 2021: 6-7). Rusia aporta una crítica similar, afirmando que es un concepto puramente subjetivo «que, en las manos apropiadas, podría convertirse en un instrumento para imponer el interés nacional de otros Estados» (Nadibaidze, 2022: 417). Recientemente, en la reunión de 2021, afirmó que se trataba de un concepto «propuesto por algunos países con el único propósito de politizar el debate sobre armas autónomas letales, sin ningún interés jurídico» (Nadibaidze, 2022: 421).

Polonia (2015: 1), por su parte, propuso en 2015 tomar el control humano significativo para un control *estatal* significativo, en el sentido de que los estados deben ser considerados responsables por las acciones de sus armas, y que esta debería extenderse a la compra, almacenamiento y fabricación de estas.

Pese a las críticas recibidas y la falta aún de una definición oficial, en los requisitos señalados por las principales organizaciones que tratan el tema del control humano significativo puede apreciarse un cierto consenso entre las propuestas: la importancia de la información sobre el entorno para poder tomar decisiones informadas, el operador ha recibido formación específica sobre el arma y la asunción de responsabilidad. También elementos como la predictibilidad y fiabilidad del arma son elementos recurrentes en la definición del control humano significativo.

3.2. NIVELES APROPIADOS DE JUICIO HUMANO SOBRE EL USO DE LA FUERZA (AHJ)

La principal alternativa al control humano significativo es el propuesto por Estados Unidos en 2012 (2016: 2), los niveles apropiados de juicio humano sobre el uso de la fuerza. Este concepto se recogió en la directiva 3000.09 del Departamento de Defensa de Estados Unidos, y realiza una enumeración detallada de criterios para poder utilizar un arma autónoma y la responsabilidad y obligaciones de sus órganos administrativos.

De acuerdo con la directiva, antes de su utilización deberán reunir las siguientes condiciones:

- El sistema operativo:
 - o Funciona como se esperaba en operaciones de prueba en ambientes realistas.

- Completa las tareas encomendadas por el comandante u operador, de acuerdo con su voluntad o, en su defecto, busca la autorización de otro operador para proseguir.
- Es lo bastante sofisticado como para prevenir fallos tecnológicos o pérdida de control del sistema.
- Tanto el *hardware* como el *software* deben diseñarse con:
 - Cortafuegos, *anti-tamper* (*software* que evita que un sabotaje lo modifique) y aseguramiento de la información.
 - Interfaces y controles para la interacción humano-máquina.
- Para que los operadores puedan tomar decisiones informadas y apropiadas (Departamento de Defensa de EE. UU., 2012: 2) sobre la selección de objetivos, deberán:
 - Ser fáciles de comprender por operadores formados.
 - Proveer un *feedback* sobre el estado del sistema.
 - Establecer procedimientos sencillos para activar y desactivar las funciones del sistema.

El operador debe manejar el arma no solo teniendo en cuenta las normas de derecho humanitario, sino también tratados aplicables, las reglas de seguridad de los sistemas de armas y las reglas de compromiso del Ejército. Según Saxon (2014: 102), esta flexibilidad comportaría que el operador se encuentre constreñido por los intereses políticos que precisen de la aplicación de un determinado juicio humano u otro.

Sobre la selección de objetivos por el sistema, la propia directiva impide que las armas autónomas letales puedan emplearse para seleccionar objetivos humanos, pero sí podrán seleccionarse otro tipo de objetivos si se está en posición de defensa de ataques enemigos, como las señales electromagnéticas emitidas por el radar enemigo (Estados Unidos, 2018: 3).

La directiva concluye detallando las tareas de supervisión que deben desarrollar varios órganos administrativos del departamento de defensa, así como el comandante a cargo de la operación, lo que permite seguir la cadena de mando de cara a la exigencia de responsabilidad. El informe presentado en 2018 (Estados Unidos, 2018: 8). añade que no puede darse un vacío en la responsabilidad porque las máquinas no son agentes morales, y por tanto «los seres humanos no eluden la responsabilidad por usar armas con funciones autónomas». En efecto, los comandantes deben realizar el juicio siguiendo los principios de proporcionalidad y distinción del derecho internacional humanitario, y son responsables por cualquier consecuencia que genere el arma, aunque esta sea autónoma.

El concepto estadounidense, pese a estar más detallado, se modula como un estándar más flexible, por ello se denomina apropiado y no significativo, pues «no puede haber un nivel de juicio humano fijo y que se pueda aplicar en todas las situaciones» (Estados Unidos, 2018: 2), y puede diferir sobre los tipos de armas, dominios y tipos de guerra²(Estados Unidos, 2016b, párr. 19), los contextos en los que se opera e incluso dentro de las distintas funciones de la propia arma (Estados Unidos, 2018: 2).

El concepto de «control humano» enturbia el diálogo sobre las armas autónomas letales porque se trata de un término sujeto a múltiples interpretaciones. Se ha entendido el «control» como la forma de garantizar el cumplimiento con el derecho internacional humanitario, pero no debería «ser un fin en sí mismo». La sofisticación técnica de un arma autónoma no implica necesariamente menor implicación humana en el proceso de toma de decisiones; de hecho, el uso de sensores y ordenadores permite a los operadores obtener parámetros sobre el uso de la fuerza que las armas tradicionales no han podido proporcionar, por lo que en verdad se incrementa el nivel de control sobre el arma (Estados Unidos, 2018: 7-8).

El concepto de los niveles apropiados de juicio humano sobre el uso de la fuerza es bastante similar al control humano significativo, coincidiendo además con requisitos del ICRAC, CNAS y Article 36; la obtención de información del entorno, la formación del operador y la asunción de responsabilidad son los puntos coincidentes principales. Los niveles apropiados de juicio humano están más desarrollados que el control humano significativo, especialmente en cuanto a requisitos técnicos, probablemente porque es una propuesta de un único país. Cabe mencionar que también exige la existencia de cortafuegos y un *software* que garantice –en la medida de lo posible– que no van a producirse fallos tecnológicos, medida que únicamente el ICRAC había mencionado. Entre sus requisitos también se expone la predictibilidad y fiabilidad del arma, aunque no se diga expresamente, y en ambos casos se apuesta por la responsabilidad del comando. También ambos coinciden en que presionar un botón no puede entenderse

² En su informe de 2016, Estados Unidos aportó ejemplos de situaciones en las que los niveles apropiados de juicio humano podían variar según el contexto: «In particular, the level of human judgment over the use of force that is appropriate will vary depending on factors, including, the type of functions performed by the weapon system; the interaction between the operator and the weapon system, including the weapon's control measures; particular aspects of the weapon system's operating environment (for example, accounting for the proximity of civilians), the expected fluidity of or changes to the weapon system's operational parameters, the type of risk incurred, and the weapon system's particular mission objective».

como ejercicio de control ni de juicio humano (Estados Unidos, 2018: 2 y Article 36, 2013: 2).

Los niveles apropiados de juicio humano adolecen del mismo problema que el control humano significativo: se utiliza terminología abierta e imprecisa, pero también más flexible que en el caso del control humano significativo.

Las principales diferencias son dos: la forma de entender el control y la selección de objetivos. En el primer caso, en el control humano significativo se entiende el control desde una perspectiva más tradicional, en la que el ser humano controla los movimientos de la máquina, mientras que, en el segundo, se sigue una perspectiva más tecnológica, de sofisticación técnica y desarrollo tecnológico. En cuanto a la selección de objetivos por el arma, solo es posible en el caso de los niveles apropiados de juicio humano, que no obstante impiden la selección de objetivos humanos. Esta última distinción entre objetivos humanos y materiales no ha sido abordada por ninguna de las organizaciones que trabaja en el control humano significativo.

Ambos estándares de control humano son muy similares, estableciendo los mismos requisitos mínimos y con el mismo fin, evitar la disolución de la responsabilidad y el incumplimiento del derecho internacional humanitario, por lo que es probable que se termine llegando a un consenso en el que se recojan elementos de ambos estándares.

4. LAS ARMAS AUTÓNOMAS LETALES Y EL DERECHO INTERNACIONAL HUMANITARIO

Los sistemas de armas autónomos letales carecen de momento de una regulación internacional específica por su novedad, pero, dado su potencial en los conflictos armados, suelen colocarse bajo el ámbito del derecho internacional humanitario y el derecho de los conflictos armados.

Las armas autónomas generan multitud de incógnitas sobre su conformidad al derecho internacional humanitario: es necesario un control efectivo sobre las armas para poder asegurar su conformidad al derecho internacional humanitario³ y es necesario un régimen de responsabilidad que garantice su cumplimiento.

³ Los principales instrumentos del derecho internacional humanitario son los cuatro Convenios de Ginebra de 1949 y sus protocolos adicionales de 1997. Convenios anteriores como el de San Petersburgo de 1868 forman parte del derecho consuetudinario. Crawford, E. y Pert, A. (2020). *International Humanitarian Law*. 2ª Ed. Cambridge: Cambridge University Press, pp. 5-19.

En las próximas páginas se mostrará la interacción de las armas autónomas letales con los principios del derecho internacional humanitario, especialmente los principios de distinción y proporcionalidad, donde la autonomía de las armas podría infringir estos principios y generar responsabilidad por parte del Estado en caso de que se produzca una conducta ilícita.

4.1. LAS ARMAS AUTÓNOMAS Y LOS PRINCIPIOS DE DERECHO INTERNACIONAL HUMANITARIO

El derecho internacional humanitario se asienta sobre seis principios básicos (Crawford y Pert, 2020: 43). La infracción de cualquiera de estos principios generaría responsabilidad del Estado. Sin embargo, los problemas con las armas autónomas letales se centran en los principios de distinción y proporcionalidad, que son los que se van a analizar.

4.1.1 El principio de distinción

El principio de distinción aparece recogido en el artículo 48 del Protocolo Adicional I a los Convenios de Ginebra de 1949. Forma parte también del derecho internacional consuetudinario y de los «principios cardinales» del derecho internacional humanitario (Crawford y Pert, 2020: 43-44).

En consecuencia, el principio de distinción exige separar entre civiles y combatientes, siendo los combatientes los únicos objetivos lícitos, extendiéndose la protección a los bienes civiles como escuelas o iglesias. La dificultad se encuentra en distinguir en todo momento según avanza el conflicto, y por ello los combatientes deben portar uniforme y señas que los identifiquen como tal (Crawford y Pert, 2020: 43-44). El artículo 51.7 AP I prohíbe además que se use a la población o los bienes civiles para favorecer o cubrir las operaciones militares. El mismo artículo 51, en sus apartados 4 y 5, establece la prohibición de ataques indiscriminados, siendo aquellos que no se dirigen contra objetivos militares o emplean medios de guerra prohibidos.

El desafío sobre las armas autónomas letales se centra en la capacidad de las armas para, por un lado, ser capaces de distinguir entre objetivos lícitos e ilícitos en un entorno cambiante (los combatientes que se rindan ya no son objetivos lícitos) y, por otro, que no produzcan ataques indiscriminados en caso de fallo. Por eso cobran especial relevancia el poder

de veto y los mecanismos de cortafuegos y salvaguarda en caso de error informático.

a) La selección de objetivos

La selección de objetivos se rige por el llamado derecho de selección de objetivos o Targeting Law. Hay una distinción entre objetivos materiales y personales. Primero se explicará la distinción para los bienes y luego para las personas.

Para asegurar el cumplimiento del artículo 48 AP I, es necesario definir cuáles son los objetivos militares y cuáles son los objetivos civiles. El artículo 52.2 AP I establece que los ataques deberán dirigirse únicamente contra objetivos militares.

La naturaleza cambiante de los métodos de guerra actuales hace imposible hacer un listado exhaustivo de los elementos que se consideran objetivos militares (Crawford y Pert, 2020: 188), pero el artículo 52.2 AP I sí fija cuatro criterios generales para distinguir los objetivos militares de los civiles: naturaleza, ubicación, finalidad y utilización.

La *naturaleza* se refiere a las cualidades intrínsecas del objeto, como un buque de guerra o un tanque. Estos dos ejemplos son objetivos militares sin necesidad de mayor consideración, pero puede haber objetos que en principio no tienen un uso militar, pero acaben teniéndolo, como una fábrica de coches que durante la guerra pasa a fabricar vehículos militares. La naturaleza de la fábrica ha cambiado, y pasa de ser un objetivo civil a uno militar (Crawford y Pert, 2020: 190).

Con *ubicación* se hace referencia a la importancia militar específica del objeto, como puede ser un aeropuerto o un puente. La toma de estos lugares o que se encuentren en una zona controlada por el Ejército rival lo convierten en un objetivo militar (Crawford y Pert, 2020: 190).

La *finalidad* del objeto se refiere al uso que se espera o se le puede dar. En el caso anterior de la fábrica de coches, su fin era civil (fabricar vehículos) pero pasa a tener un fin militar al fabricar vehículos de guerra, por lo que su carácter de objetivo civil cambia por su finalidad (Crawford y Pert, 2020: 190).

La *utilización* de un objeto civil como uno militar puede cambiar su carácter civil, incluso si este tipo de objetos están expresamente protegidos por el derecho. Es el caso de escuelas y hospitales, que pierden su inmunidad cuando los militares los utilizan como base. A estos efectos, el artículo 52.3 AP I declara que «en caso de duda acerca de si un bien que normalmente se dedica a fines civiles, tal como un lugar de culto, una casa u otra vivienda o una escuela, se utiliza para contribuir eficazmente a la acción

militar, se presumirá que no se utiliza con tal fin», por lo que se convierte en un objetivo militar.

Además de los criterios generales, algunos objetos civiles están expresamente protegidos por normas internacionales, entre ellos los bienes culturales, edificios de uso médico y con emblemas de Cruz Roja, Luna Roja o Cristal Rojo, y bienes necesarios para la supervivencia de los civiles (Crawford y Pert, 2020: 145-155 y 201-213).

El principio de distinción exige también distinguir entre personal con distinto estatus jurídico. Principalmente, entre civiles, combatientes, civiles que tomen parte en las hostilidades y *hors de combat*.

- Distinción entre civiles y combatientes

El artículo 48 AP I establece que solo los combatientes pueden ser objetivos válidos en conflicto armado y protege a los civiles. El Convenio de Ginebra para aliviar la suerte que corren los heridos y los enfermos de las Fuerzas Armadas en campaña, de 1949, extiende la protección de los civiles a los heridos, el personal religioso y el personal médico militar (artículos 24 y 25), mientras ejerzan funciones sanitarias y no combatan (Soni y Dominic, 2020: 31).

El artículo 43 AP I define a los combatientes como miembros de las Fuerzas Armadas de las partes. El artículo 50 AP I define a los civiles en sentido negativo, siendo todos aquellos que no son combatientes ni participan en las hostilidades.

En general, las personas que no estén participando en el conflicto ni porten armas o pertenezcan a una fuerza militar, serán consideradas civiles y gozarán del régimen más alto de protección y, ante la duda, serán consideradas civiles.

- Distinción entre civiles y civiles que participen en las hostilidades

Los civiles que participen en las hostilidades no están protegidos por el régimen de los civiles mientras dure su participación, como señala el artículo 51 AP I.

El protocolo no aporta mayor información sobre los civiles participantes en el conflicto para poder distinguirlos de la población civil, pero Cruz Roja sí aporta una distinción. El protocolo busca una participación directa, lo que implica un acto individual y en las hostilidades, que es un acto colectivo. Por tanto, los civiles partícipes realizarán actos individuales con relevancia en un acto colectivo (ICRC, 2010: 43). Para que un acto sea

considerado como participación directa en las hostilidades, debe cumplir tres requisitos:

1. El acto debe tener, o probabilidades de tener, efectos sobre la capacidad militar de una de las partes, o bien causar daños a personas y bienes protegidos,
2. Debe haber un vínculo causal directo entre este acto y el daño producido,
3. El propósito del acto debe ser provocar un daño en detrimento de una parte y beneficio de otra (ICRC, 2010: 43).

- Distinción entre militares y *hors de combat*

Los objetivos militares serán lícitos siempre y cuando conserven su estatus de combatientes. El artículo 41 API I define los *hors de combat*:

La justificación para esta distinción es que las personas en *hors de combat* se encuentran indefensas y precisan de protección por parte del derecho internacional (Soni y Dominic, 2020: 34).

- a) Las armas autónomas letales en el principio de distinción

Una vez vista la teoría sobre el principio de distinción, es necesario ponerla en relación con las armas autónomas letales. Pese a la incertidumbre asociada a las armas autónomas, especialmente a aquellas que funcionan mediante redes neuronales, no puede decirse que estén directamente prohibidas (Soni y Dominic, 2020: 30-35 y Heyns, 2013: 15-16).

La distinción entre objetos y personas es posible gracias a los sensores que incorporen las armas. Las armas autónomas estarán provistas de radares, láseres y sensores de infrarrojos que permitan a la máquina ajustar su posición y orientación en el entorno. Esto genera una visión (*machine vision*) que permitiría al arma obtener un mayor número de patrones con las imágenes para distinguir objetivos, aplicando inteligencia artificial (Sehrawat, 2017: 6). Sehrawat se limita a hablar de la inteligencia artificial, pero las redes neuronales son seguramente el elemento tecnológico más prometedor para las armas autónomas letales.

Las redes neuronales tienen mecanismos de procesamiento de la información más complejos y, sobre todo, un modo de «aprendizaje» muy superior al de la sola inteligencia artificial. Con el desarrollo del *deep learning* se puede llegar a reconocimientos mucho más exhaustivos. La fusión de sensores (Sehrawat, 2017: 7) produce una recopilación de información de todos los sensores de que dispone la máquina para obtener una visión

tridimensional de los datos. Pues bien, con la correcta interpretación de los datos podría llegarse a una mejor diferenciación de los objetivos: reconocimientos de puentes, hospitales... también de los combatientes para identificar si un bien ha cambiado de su uso o naturaleza, como el mencionado ejemplo de la fábrica de vehículos. Incluso en el caso del reconocimiento de personas, los sensores podrían diferenciar casos claros de civiles, al menos por su alta probabilidad: niños, por su estatura; personas mayores, con pelo o barba canos; heridos, con vendas, muletas o sangre.

La identificación mediante emblemas juega aquí un papel fundamental, como en los casos de protección de bienes culturales o los símbolos de Cruz Roja, Medialuna Roja o Cristal Rojo que se referenciaban en el epígrafe anterior. Si las máquinas ya realizan reconocimiento de objetos y lectura de labios (Baena, 2018: 36), pueden distinguir los emblemas reconocidos por el derecho internacional. Otros signos no reconocidos pero que podrían enseñarse a las máquinas es el reconocimiento de uniformes y emblemas como la «Z» que identifica los tanques rusos en la actual invasión de Ucrania.

También habría que enseñar a las redes neuronales las señales que identifiquen la rendición de los combatientes, como abandonar un buque, deponer las armas o enseñar una tela blanca (Scholz y Galliot, 2021: 65).

La distinción en estos términos sería, en principio, posible (Meza, 2018: 271). Pero en los conflictos armados se dan situaciones confusas que incluso a los humanos les resulta difícil evaluar (Sehrawat, 2017: 10): En los conflictos armados más recientes, especialmente aquellos que incluyen a actores no estatales, los insurgentes se hacen pasar por civiles (Sehrawat, 2017: 10), los conflictos se han trasladado a ciudades con alta densidad de población y los campos de batalla son cada vez más atípicos, e incluso el uso de uniformes se ha vuelto más infrecuente (Soni y Dominic, 2020: 32). Todo ello dificulta que las armas autónomas letales cumplan con el principio de distinción, pero es una dificultad a que también se enfrentan los soldados, operadores y jefes militares.

Es preciso recordar que, conforme a las exigencias tanto del control humano significativo como de los niveles apropiados de juicio humano, hay que mantener el control humano para que el uso de las armas autónomas sea conforme al derecho internacional humanitario, y que la única forma de control permitida es la *on the loop*. El tipo de control *on the loop* permite tanto la selección de posibles objetivos seleccionados por la máquina como la orden expresa de atacar, así como un poder de veto. En los casos dudosos, es necesaria la intervención del operador para garantizar la correcta distinción de los objetivos.

Soni y Dominic (2020: 34) proponen la instauración de un umbral de duda que, basado en porcentajes, impida la actuación del arma y obligue al operador a decidir. Meza Rivas propone también enseñar a las redes neuronales a consultar con el operador en caso de duda (Meza, 2018: 275).

En cualquier caso, la conformidad de las armas autónomas letales con el principio de distinción dependerá de la interacción con el operador y las acciones de este, y con la tecnología actual podría enseñárseles a distinguir emblemas y signos. Irónicamente, las redes neuronales son el mejor mecanismo para la labor de distinción, pero también son el principal factor de impredecibilidad. Las redes neuronales, mediante el procesamiento de datos, permiten a las máquinas perfeccionar la distinción entre objetivos y la identificación de señales y emblemas; de hecho, si la identificación es incorrecta, le permiten rectificar su asociación. Pero mientras que las redes neuronales permiten a la máquina refinar sus asociaciones, también aumenta la impredecibilidad, pues aumenta el riesgo de que las redes neuronales realicen conexiones aleatorias y no esperadas por el programador que las diseñó.

4.1.2. El principio de proporcionalidad

El principio de proporcionalidad exige valorar las medidas tomadas en la guerra de forma que sean proporcionadas, es decir, que la ventaja militar obtenida supere el daño provocado a los civiles y sus bienes (Crawford y Pert, 2020: 46). Se recoge fundamentalmente en el artículo 51.5.b. AP I, en el que se establece que no cumplen con el principio de proporcionalidad los ataques que causen daños excesivos en relación con la ventaja militar obtenida.

a) El alcance del principio de proporcionalidad

La ponderación que requiere el principio de proporcionalidad no consiste, simplemente, en contar las bajas civiles y compararlas con el número de bajas militares enemigas, sino en una verdadera ponderación que examine, caso por caso, la razonabilidad del ataque teniendo en cuenta las circunstancias (Schmitt y Thurnher, 2013: 254). Factores como el tipo de adversario, tipo de ataque o el tiempo son necesarios para realizar la ponderación (Van den Boogaard, 2015: 14). Hay tres componentes en la valoración: el principio de distinción, la ventaja militar esperada y el comandante razonable.

- El principio de distinción. Como se decía en el epígrafe anterior, las partes deben ser capaces de distinguir entre objetivos civiles y militares, incluso en los casos en los que la naturaleza o el uso de los objetos cambia, como la escuela que ha sido convertida en base militar.
- La ventaja militar esperada. La ventaja militar puede definirse como el resultado de «las consecuencias del ataque que directamente mejora las propias operaciones militares o entorpece las del enemigo» (Harvard College, 2010: 45), y debe referirse a los beneficios que resultan del ataque, no como un examen conjunto de las partes del ataque (Harvard College, 2010: 45). Es el conjunto de las tácticas y técnicas militares el que se debe examinar, pues el resultado puede no derivar de un ataque contra un objetivo específico, sino de un conjunto cumulativo de ellos (Van den Boogaard, 2015: 18).
- El jefe militar razonable. La decisión final del ataque recae en el comandante o jefe de operaciones. Por eso, en casos dudosos, se debe acudir al estándar del jefe militar razonable (Van den Boogaard, 2015: 20). En el Estatuto de Roma, artículo 8.2.b.iv, como el artículo 51.5.b AP I, se utiliza como parámetro la palabra «excesivo», y «claramente excesivo» en el primer caso. El Tribunal Penal Internacional para la ex-Yugoslavia sentenció que para saber si un ataque había sido proporcionado o no, era necesario examinar «si una persona razonable e informada, en las circunstancias del autor, haciendo un uso razonable de la información de la que disponía, habría esperado tantas víctimas civiles como consecuencia del ataque» (STPIY de 5 de diciembre de 2003, *caso Galić*, párrafo 58).

b) Las armas autónomas letales en el principio de proporcionalidad

El principio de proporcionalidad requiere realizar una ponderación que es discutible que las máquinas sean capaces de realizar. Schmitt y Thurnher (2013: 254-255) valoran la inclusión de la «metodología estimativa de daño colateral» (*collateral damage estimate methodology* o CDEM) en la programación de las armas autónomas letales para evitar ataques desproporcionados. También podrían incluirse algoritmos que determinasen casos claros positivos y negativos que permitiesen a la máquina evaluar la proporcionalidad del ataque, pero sigue habiendo casos difusos, y en cualquier caso sería el jefe militar quien realizase la ponderación (Schmitt y Thurnher, 2013: 256). Es posible que la máquina sea capaz de realizar la

ponderación en algunos casos claros, pero incapaz en los difusos (Heyns, 2013: 15).

Para que las armas autónomas pudiesen utilizarse de conformidad al principio de proporcionalidad, deben monitorizarse los efectos de las armas durante el ataque, y su capacidad para hacer daño. Los sensores incorporados a las armas autónomas permitirían calcular con mayor precisión el número de posibles bajas civiles en el contexto de la ponderación (Gannon y Khalaf, 2022: 1068).

Sin embargo, el carácter cambiante de los conflictos actuales, en el que los bienes civiles pasan rápidamente a ser militares y hay grandes movimientos de población, es complicado que el arma sea capaz de calcular el daño colateral, pese a las configuraciones informáticas que se le hayan hecho al respecto (Gannon y Khalaf, 2022: 1068).

Teniendo en cuenta los desafíos que plantea y siguiendo el razonamiento de Schmitt y Thurnher (2013: 254-257), la incorporación del CDEM y los sensores que puedan calcular con mayor precisión el número de bajas civiles no son sino fuentes de información para el operador o incluso el jefe militar. Recordando las exigencias del control *on the loop*, no es la máquina la que realiza la ponderación del ataque, ni siquiera decide cuándo atacar: la máquina propone posibles objetivos, y es tarea del operador y del jefe militar realizar el cálculo del daño colateral y realizar la ponderación con la ventaja militar esperada, ponderación que es inherentemente subjetiva (Meza, 2018: 337) y que la máquina no podrá realizar (Heyns, 2013: 15).

Incluso en el caso de que se permita a la máquina seleccionar objetivos tanto personales como patrimoniales —algo no permitido por ninguno de los estándares de control discutidos, MHC y AHJ— la decisión de atacar sigue en manos del operador, por lo que la máquina no realiza la ponderación. En caso de que la máquina tomase esa decisión y se delegase en ella la labor de ponderar, estaríamos ante un control *out of the loop* que no cumpliría los estándares de control que están discutiendo los países.

5. LA RESPONSABILIDAD DEL ESTADO EN LAS ARMAS AUTÓNOMAS LETALES

La petición de responsabilidad por las consecuencias de las armas autónomas letales plantea un desafío en este ámbito. La peculiaridad de las armas autónomas letales obliga a plantearse si los medios tradicionales de responsabilidad son adecuados, y por ello inclina a gran parte de la doctrina a inclinarse por la responsabilidad individual y no por la estatal, pese

a no ser excluyentes. Que la responsabilidad individual y la estatal sean complementarias aumenta la probabilidad de que se consiga efectivamente una reparación para las víctimas.

Como se decía en la introducción, exigir responsabilidad al Estado conlleva ciertas ventajas sobre la responsabilidad individual: puede provocar que se desincentive el uso de las armas autónomas por parte de los Estados y, además, los Estados tienen fondos suficientes como para satisfacer las reparaciones que se les exijan.

Aquí se van a proponer tres regímenes de responsabilidad por las que el Estado podría ser responsable de las armas autónomas letales, y se destinará un cuarto epígrafe a comentar la realidad procesal del caso.

Es preciso tener en cuenta que aquí únicamente se está hablando de «responsabilidad», que es el término castellano para referirse a los términos *responsibility*, *liability* y *accountability* del inglés. Se hará referencia al inglés en los casos en que sea preciso para facilitar su comprensión.

5.1. RESPONSABILIDAD INTERNACIONAL DEL ESTADO POR HECHOS ILÍCITOS

La responsabilidad del Estado se considera derecho consuetudinario y ha sido tratada por la Comisión de Derecho Internacional, ILC (Meza, 2018: 370)⁴, que articuló el Proyecto de Artículos sobre la Responsabilidad de los Estados por Hechos Ilícitos (ARSIWA, por sus siglas en inglés). El artículo 1 establece que «Todo hecho internacionalmente ilícito del Estado genera su responsabilidad internacional». Esta frase puede dividirse en los dos partes: por un lado, «hecho internacionalmente ilícito» intenta cubrir todas las conductas del Estado, ya sea una acción u omisión o ausencia de acción (Crawford, 2013:49). La comisión del ilícito dependerá de las condiciones de la obligación que se ha roto y las condiciones jurídicas de dicho acto (ILC, 2001: 2). Este artículo, que recoge un principio del derecho internacional, ha sido aplicado por la Corte Internacional de Justicia en casos como el Canal de Corfú o «Rainbow Warrior», cuando el tribunal arbitral afirmó que «cualquier violación de cualquier obligación por parte de un Estado, de cualquier origen, conllevará responsabilidad del Estado» (ILC, 2001: 32-33).

Por otro lado, el hecho «internacionalmente ilícito» comporta responsabilidad del Estado. Se refiere a las nuevas relaciones que emanan del

⁴ Meza Rivas, «el desarrollo y el uso de los sistemas de armas autónomas letales en los conflictos armados internacionales», 370.

derecho internacional a causa de la conducta del Estado (ILC, 2001: 32-33). Los artículos 20-25 se refieren a supuestos de no acción y vicios que no generan responsabilidad alguna.

El artículo 2 establece que la conducta que ocasione el ilícito internacional puede ser tanto una acción como una omisión, y presenta dos requisitos:

- La conducta debe ser atribuible al Estado.
- Constituye una violación de una obligación internacional del Estado.

Los dos anteriores constituyen los elementos subjetivo y objetivo de la responsabilidad del Estado, y aunque son términos que el texto evita (ILC, 2001: 35), son usuales en el lenguaje jurídico. Los artículos 4 a 11 establecen las conductas que le son atribuibles al Estado. También el artículo 91 AP I establece que los estados son responsables de todos los actos realizados por sus órganos, ya sea obedeciendo o desobedeciendo sus instrucciones. El artículo 4 ARSIWA afirma que «1. Se considerará hecho del Estado según el derecho internacional el comportamiento de todo órgano del Estado, ya sea que ejerza funciones legislativas, ejecutivas, judiciales o de otra índole, cualquiera que sea su posición en la organización del Estado y tanto si pertenece al Gobierno central como a una división territorial del Estado», y que se entenderá que el órgano «incluye toda persona o entidad que tenga esa condición según el derecho interno del Estado».

Es decir, que si el arma autónoma violase alguno de los principios del derecho internacional explicados en el capítulo anterior, al operar como parte del Ejército del Estado, su conducta sería directamente atribuible al Estado (Meza, 2018: 371). Al margen de la responsabilidad individual por el uso inadecuado de las armas autónomas, sería posible exigir responsabilidad al Estado (Malik, 2018: 631. También en Meza, 2018: 372, Marauhn, 2014: 2 y Chengeta, 2016: 48).

Luego, si un arma autónoma es desplegada en el marco de un conflicto armado (como se decía, no pretendía analizarse aquí la adecuación de las armas autónomas al derecho internacional humanitario, por lo que se supondrá que se están utilizando legalmente) y falla al distinguir un bien militar de un bien civil, o realiza un ataque manifiestamente desproporcionado que el operador no evitó, el Estado sería responsable por la violación de los principios del derecho internacional humanitario, que además de formar parte del derecho consuetudinario, están recogidos en convenciones y adquieren por ello fuerza legal.

El régimen de responsabilidad del Estado genera varios desafíos que hacen pensar en la adecuación de otras formas de responsabilidad del Estado. Primero, hay controversia sobre la procedencia de aplicar unas normas convencionales a las armas autónomas. Geiss (2015: 22) señala que no puede solo invocarse el artículo 91 AP I, porque este se refiere únicamente a actos «cometidos por las personas que formen parte de sus fuerzas armadas», y por tanto no podría aplicarse a las máquinas porque no son personas. La aplicabilidad de la norma dependería de su interpretación, y Geiss está adoptando aquí una interpretación muy literal. Siguiendo su punto de vista, es cierto que podría generarse aquí un vacío en la responsabilidad, pero podría imputársele al operador que ordenó el ataque.

Segundo, los habilitados para interponer denuncias ante los tribunales son los propios Estados, lo que plantea un problema para los individuos que hayan sido damnificados por las armas autónomas letales pero su Estado no quiera ejercer la acción de petición de responsabilidad al Estado responsable, por no hablar de los apátridas. El derecho internacional reconoce con carácter general un derecho de los individuos a buscar directamente la reparación por parte del Estado infractor, y este derecho se recoge en el artículo 33.2 ARSIWA: «[...] de cualquier derecho que la responsabilidad internacional del Estado pueda generar directamente en beneficio de una persona o de una entidad distinta de un Estado». En la práctica es difícil que los individuos puedan presentar acciones, por barreras legales y económicas (Malik, 2018: 639). Sin embargo, es un desafío que resulta independiente de las armas autónomas letales.

Tercero, Geiss (2015: 22) apunta que si se le exige una conducta dolosa o negligente al Estado, se generaría un vacío de responsabilidad. Sin embargo, los artículos dedicados a la atribución de la responsabilidad del Estado, en el capítulo II, no exigen dolo ni negligencia, sino un acto proveniente de los órganos del Estado. De hecho, si los órganos —las Fuerzas Armadas— se extralimitan en sus funciones, el Estado será responsable por las consecuencias de este exceso (ILC, 2001: 39) o incluso son contrarios a sus instrucciones (ILC, 2001: 39). Los comentarios al ARSIWA señalan sobre este punto que «en la ausencia de un requerimiento específico de elemento mental en términos de obligación primaria, en este caso solo el acto del Estado cuenta, independientemente de la intención» (ILC, 2001: 36).

Por último, queda la dificultad para invocar la responsabilidad ante los tribunales y cuáles son competentes, como se analizará en el último epígrafe de este capítulo.

5.2. RESPONSABILIDAD DEL ESTADO POR INFRACCIÓN DEL DUE DILIGENCE

El Estado tiene la obligación de proteger a sus ciudadanos y por ello debe prevenir, investigar y castigar estas conductas y e incluso procesar a los culpables (Meza, 2019: 371), aunque estos sean rebeldes (Chengeta, 2016: 48). Algunos tratados le imputan responsabilidad al Estado por no haber velado por el cumplimiento de la obligación⁵. Se considera una infracción del *due diligence*, y se recoge en el artículo 86 AP I.

Dos son los motivos principales para esta regla de responsabilidad: primero, por la falta de la diligencia debida por parte del Estado, y segundo, porque estaría discriminando a las víctimas de los agentes no estatales. Cobra especial relevancia en los contextos de posguerra contra líderes rebeldes o terroristas, para facilitar la reconciliación entre la población (Chengeta, 2016: 49).

Un ejemplo de esta forma de omisión del Estado se encuentra en el caso *Corfú Channel*, cuando Albania sabía, o debería haber sabido, de la existencia de minas en sus aguas territoriales pero no avisó a terceros Estados, ocasionando la responsabilidad del Estado albanés (ILC, 2001: 35). En palabras de la Corte, «es obligación de los Estados no permitir que su territorio sea utilizado para hacer actos contrarios a los derechos de otros Estados» (SCIJ de 9 de abril de 1949, *Reino Unido c. Albania*: 22). En las armas autónomas letales, los estados deben asegurar que grupos no estatales se hagan con ellas, o que productores fabriquen armas autónomas sin los debidos cortafuegos y medidas de control que establezca la ley.

Marauhn (2014: 2 y 5) considera que la infracción del *due diligence* es más relevante en el caso de que las armas autónomas estén operando *out of the loop*, porque permitiría atribuirle responsabilidad al Estado en cualquier caso. Sin embargo, el control *out of the loop* es poco probable por haber un cierto consenso en el manejo *on the loop* de las armas autónomas entre los Estados y las organizaciones dedicadas al estudio de las armas autónomas letales. Geiss (2016, párr. 7), por su parte, considera que es necesario concretar las obligaciones específicas que deben contemplar los Estados en relación con las armas autónomas, estableciendo un estándar de medidas a acordar según el riesgo potencial.

⁵ Entre otros, la Convención sobre la prohibición del desarrollo, la producción y el almacenamiento de armas bacteriológicas (biológicas) y tóxicas y sobre su destrucción y el Tratado sobre el comercio de armas de 2013.

5.3. RESPONSABILIDAD CIVIL POR LOS DAÑOS OCASIONADOS (STRICT LIABILITY)

Es un régimen de indemnización por los daños ocasionados. Numerosos autores están proponiendo este tipo de régimen para las armas autónomas letales, por el que los estados serían encontrados responsables por los daños ocasionados derivados del riesgo de utilizar armas autónomas letales (Hammond, 2015: 670).

La razón es evitar compensar el riesgo transferido que se produce sobre las víctimas, pues la acción del infractor —en el caso, del Estado al emplear armas potencialmente impredecibles— «impone un riesgo unilateral sobre la víctima en situaciones donde la víctima no impone el mismo riesgo sobre el infractor» (Hammond, 2015: 670).

Fuzaylova (2019: 1360) distingue cuatro requisitos establecidos por los tribunales —a nivel nacional— para reconocerlo:

- La actividad —el uso de armas autónomas— genera un riesgo potencial y previsible de provocar daños físicos.
- El riesgo es «muy significativo».
- El riesgo perdura incluso si todos los actores implicados realizan un uso razonable.
- La actividad no es de uso común.

El riesgo potencial se incrementaría con un control *out of the loop*, siendo inferior en un tipo de control *on the loop*.

El régimen de responsabilidad podría basarse tanto en un sistema que requiera probar el ilícito (*strict liability*) como de carga de la prueba (*presumed liability*) (Geiss, 2016, párr. 13). Ya hay algunos precedentes en el ámbito internacional, como la responsabilidad establecida en el Convenio sobre la responsabilidad internacional por daños causados por objetos espaciales de 1972, concretamente en su artículo 8, que dice así: «Un Estado que haya sufrido daños, o cuyas personas físicas o morales hayan sufrido daños, podrá presentar a un Estado de lanzamiento una reclamación por tales daños». Ya en el caso *Rainbow warrior* se estableció que había responsabilidad del Estado ya fuera un ilícito de origen contractual o por daños, y por tanto se genera la misma responsabilidad que si se hubiese violado un tratado (Crawford, 2013: 51-52).

Crootof (2016: 1386-1387) propone el régimen de compensación de derecho de daños como forma de evitar el vacío de responsabilidad derivado de la falta de intencionalidad en el tipo de responsabilidad del Estado por hechos ilícitos y la responsabilidad individual. Podría ser un régimen

supletorio de indemnización para los dos casos anteriores, porque el *strict liability* se le impone a un actor por los daños que haya ocasionado con su conducta (caso de responsabilidad por hechos ilícitos) (Crootof, 2016: 1394) y el *negligence liability* se utiliza para los daños ocasionados por su falta de cuidado (responsabilidad por infracción del *due care*) (Crootof, 2016: 1394). Además, el artículo 35 ARSIWA establece que incluso en caso de ausencia de violación internacional los estados estarán obligados a indemnizar a sus víctimas, por lo que puede convertirse en un buen régimen de responsabilidad adicional.

Sobre la implementación, Crootof propone la firma de tratados específicos sobre responsabilidad por daños que la Corte Penal Internacional examine (Crootof, 2016: 1397), la creación de un tribunal específico para este tipo de demandas (Crootof, 2016: 1398), o incorporar el régimen de responsabilidad internacional al derecho nacional (Crootof, 2016: 1398).

En definitiva, proponen la implementación de un régimen de responsabilidad civil que sea subsidiario a los dos regímenes anteriores por su mayor facilidad para imputar la responsabilidad al no requerir intencionalidad y existir ya previamente en otros ámbitos como el derecho espacial.

5.4. DESAFÍOS PROCESALES

No puede hablarse de responsabilidad del Estado sin mencionar la cuestión procesal. Tanto a nivel internacional como nacional, el enjuiciamiento de la responsabilidad del Estado se convierte en un desafío para los perjudicados.

5.4.1. La Corte Internacional de Justicia

Hammond identifica la Corte Internacional de Justicia (Hammond, 2015: 678-679) como primer tribunal en tener la jurisdicción para conocer de la responsabilidad generada por el uso de armas autónomas, pues previsiblemente se infringirán principios del derecho internacional humanitario. Sin embargo, la Corte solo tiene jurisdicción sobre Estados que hayan reconocido expresamente su jurisdicción para ese supuesto, y además solo admite demandas que provengan de los estados (Hammond, 2015: 678-679), lo que deja fuera a los individuos, como se alertaba en el primer epígrafe de este capítulo. Los estados tienen dos formas de consentir aceptar la jurisdicción de la corte:

Primero, ambos estados pueden firmar un acuerdo internacional que contenga una cláusula de sumisión a la Corte, según figura en el artículo 37 del Estatuto de la Corte. Sin embargo, ese mecanismo sería ineficiente porque los tratados ya firmados sobre derecho internacional humanitario no disponen de dicha cláusula (Hammond, 2015: 679).

Segundo, un Estado puede aceptar formalmente la jurisdicción de la Corte sobre una disputa con otro Estado miembro de la ONU. Solo tendrían que firmar una declaración de consentimiento ante la Corte, dirigidas al Greffier (CIJ, s.f., párr. 1-3). Sin embargo, apenas un tercio de los países reconoce la jurisdicción de la Corte (Hammond, 2015: 679), incluyendo a Estados Unidos, Rusia y China. Por tanto, la jurisdicción de la Corte también en este caso sería muy limitada. Además, en el caso de las armas autónomas, el Estado tendría pocos incentivos para someterse su jurisdicción, pues además tendría que enfrentar más costes (Hammond, 2015: 680).

Además de los problemas de jurisdicción, el problema principal que tiene la Corte es que carece de un medio coercitivo para hacer ejecutar sus fallos (Malik, 2018:639). En principio, el Consejo de Seguridad de Naciones Unidas podría ejecutar el fallo de la Corte, pero en la práctica no ha ejercido nunca este poder (Malik, 2018: 639). Huelga recordar que los cinco miembros permanentes del Consejo están desarrollando armas autónomas letales, por lo que es poco probable que ejerzan este poder contra ellos mismos.

5.4.2. Tribunales nacionales

Las víctimas de las armas autónomas podrían buscar el enjuiciamiento en los tribunales nacionales. Pero esta opción tampoco resulta fácil. Si acudiesen a tribunales de otros países para ejercitar la acción de responsabilidad contra su propio Estado o algún otro, muy pocos estados admiten demandas de responsabilidad civil por actos cometidos en otro Estado (Hammond, 2015: 680). Algunos estados permiten incluir una acción civil a la acción penal por un crimen que ocurrió en el extranjero, pero los estados no están sujetos a las leyes penales de los otros estados, por lo que esta cuestión quedaría irresoluta (Hammond, 2015: 680).

Otra opción sería acudir a sus propios tribunales nacionales. Para ello, primero, el Estado tiene que reconocer la acción que se pretende ejercer (Hammond, 2015: 682). Segundo, es necesario que el crimen internacional –lesiones, muerte...– se encuentre entre los supuestos que permiten enjuiciar al Estado y evitar la inmunidad del Estado (Hammond, 2015: 682).

La cuestión procesal supone un reto mucho más importante que la propia acción de responsabilidad internacional del Estado. Como se ha visto, los tres regímenes cubren las conductas activas y pasivas de los estados con relación a las armas autónomas, e incluso se está impulsando un tercer régimen supletorio para el caso de que no se reconozca ninguna de las anteriores. Sin embargo, ninguno de estos tres regímenes de responsabilidad sirve si no es posible encontrar tribunales que enjuicien estas acciones. Aun así, estas complicaciones procedimentales son anteriores al inicio de las armas autónomas letales y sus consecuencias se extienden a varios ámbitos del derecho, aunque principalmente es el derecho penal. Las víctimas de genocidio, trata de seres humanos, terrorismo o cualquier otra violación que emane del derecho internacional humanitario se encuentra ante las mismas complicaciones que las potenciales víctimas de las armas autónomas letales, por lo que el vacío de responsabilidad no es mayor que en los demás casos del derecho internacional.

6. CONCLUSIONES

- I. De los tres tipos de control que se barajan para las armas autónomas letales, *out of the loop* supone que las decisiones sobre el ataque las realiza la máquina. En cambio, *on the loop* solo permite que la máquina ofrezca una selección de objetivos sobre la que el operador debe decidir, o un poder de veto que evite violaciones de derecho internacional humanitario.
- II. *On the loop* es el tipo de control que cumpliría los dos estándares de control propuestos, el control humano significativo y los niveles apropiados de juicio humano, y que además permitiría asegurar la conformidad al derecho internacional humanitario y la imputación de responsabilidad al Estado.
- III. Es posible encontrar un consenso entre las organizaciones que desarrollan el control humano significativo, estableciendo como requisitos la decisión informada, que la orden de ataque la ejerza el operador, que el operador haya recibido formación para manejar la máquina y que se depure responsabilidad.
- IV. Los niveles apropiados de juicio humano, pese a ser el otro estándar propuesto, coincide con el anterior en la toma de decisiones informadas, la formación del operador y la petición de responsabilidad. Incorpora –aunque coincide aquí con el ICRAC– la exi-

gencia de mecanismos de cortafuegos en caso de error y permite expresamente seleccionar bienes como objetivos.

- V. Teniendo en cuenta la similitud entre ambos estándares, es probable que los dos terminen unificándose en uno solo.
- VI. En este trabajo se ha visto cómo interactúan las armas autónomas letales con los principios del derecho internacional humanitario y, aunque no están prohibidas, solo el tipo de control *on the loop* podría garantizar su cumplimiento por evitar en mayor medida las infracciones del DIH.
- VII. La distinción entre objetivos civiles y militares es posible que pueda ser realizada por la máquina gracias a sus sensores y los procedimientos de *machine learning*, pero es necesario que sea el operador el que autorice el ataque para garantizar el cumplimiento del principio.
- VIII. Las redes neuronales son el mejor mecanismo para que la máquina aprenda a distinguir emblemas y edificios porque permite a la máquina realizar asociaciones más complejas y corregirlas si son erróneas. El problema que plantean las redes neuronales es que en ocasiones estas mismas asociaciones entre imágenes son difíciles de predecir incluso para los programadores, lo que aumenta su impredecibilidad.
- IX. Para que las armas autónomas puedan ser conformes al principio de proporcionalidad, la ponderación entre la ventaja militar adquirida y el daño a la población civil debe ser realizada siempre por el jefe de operaciones pues, aunque la máquina tenga la posibilidad de calcular en términos porcentuales el daño potencial y hacer estimaciones matemáticas, la ponderación sigue siendo una labor exclusivamente humana.
- X. La otra crítica a este respecto que reciben las armas autónomas es la existencia del *moral buffer*, cuando los operadores delegan la carga moral de sus decisiones en las máquinas. Con el control *on the loop* se evita este sesgo, porque siguen siendo los operadores los que toman la decisión final.
- XI. Se proponen tres regímenes de responsabilidad del Estado para las armas autónomas letales: la responsabilidad internacional por hechos ilícitos, la responsabilidad por infracción del *due diligence* y el *strict liability*.
- XII. El problema principal que plantea la responsabilidad por hechos ilícitos es que los artículos sobre responsabilidad del Estado no

- son sino un proyecto, y carecen de fuerza legal, aunque sus normas adquieren cierto respaldo en el derecho consuetudinario.
- XIII. El problema se plantea hacia los individuos, pues este sistema de responsabilidad está pensado fundamentalmente para las reclamaciones entre Estados, y se deja fuera a individuos apátridas, víctimas de actores no estatales y víctimas cuyos Estados no hayan querido ejercer la reclamación de responsabilidad. Aunque se les reconoce el derecho a dirigir estas reclamaciones por sí mismos, las barreras legales y económicas lo dificultan en la práctica. De hecho, la propia CIJ no admite reclamaciones de individuos.
- XIV. La responsabilidad por infracción del *due diligence* facilita imputar responsabilidad al Estado en el caso de que la violación la hayan cometido agentes no estatales, como terroristas o corporaciones, porque consiste en una omisión del deber de cuidado. Adquiere especial relevancia en el manejo *out of the loop*, porque permitiría atribuirle responsabilidad al Estado.
- XV. Por último, el *strict liability regime* es una acción civil derivada de la comisión de un ilícito internacional. Autores especializados en las armas autónomas letales como Crootof o Geiss consideran que este es el mejor régimen de responsabilidad porque evitaría los problemas de atribución de la conducta al Estado, así como un juicio penal. Se erige como un régimen supletorio a los dos anteriores que busca cubrir posibles lagunas legales.
- XVI. Con los tres regímenes de responsabilidad propuestos es posible exigirle responsabilidad al Estado, tanto en un plano más teórico como práctico, y en el orden civil además del penal. Con el uso de estos tres tipos de responsabilidad, se evitaría por completo el famoso «vacío de responsabilidad» del que alertan las organizaciones especializadas en armas autónomas.
- XVII. La víctima de armas autónomas letales se enfrenta a desafíos procesales para poder pedir responsabilidad al Estado, y los propios Estados también. Además de no reconocer las demandas individuales, la CIJ también plantea problemas de jurisdicción y ejecución.
- XVIII. Apenas un tercio de los países del mundo reconoce la jurisdicción de la CIJ, lo que complica la exigencia de responsabilidad, y entre ellos EE. UU., Rusia y China, todos ellos actores principales en las armas autónomas letales.
- XIX. La CIJ carece de mecanismos coercitivos que le permitan hacer ejecutar sus fallos. En teoría, el Consejo de Seguridad de la ONU

podría ejercer ese poder, pero teniendo en cuenta que nunca se ha hecho con anterioridad y que los cinco miembros permanentes están desarrollando armas autónomas letales, parece bastante improbable que vayan a empezar a ejecutar los fallos en esta circunstancia.

- XX. Más que vacío de responsabilidad, se trata de un vacío de responsabilidad procesal. Pero este problema se extiende a las víctimas de otras violaciones de derechos humanos y DIH, y no es una facultad exclusiva de las armas autónomas letales, por lo que no puede hablarse de vacío de responsabilidad en las armas autónomas letales cuando no es diferente del vacío que se genera en otros ilícitos internacionales.

7. BIBLIOGRAFÍA

Libros y artículos de revista

- Baena, M. (2018). Empleo de la inteligencia artificial en la resolución de contenciosos internacionales, públicos y privados. Experiencias pioneras. En: *Cuadernos de la Escuela Diplomática. Selección de Memorias del curso selectivo de funcionarios de la Carrera Diplomática 2019-2020*. Escuela Diplomática de España, editor. Madrid, Escuela Diplomática de España, pp.19-96.
- Baker, D. P. (2021). The Robot Dogs of War. En: Gaillot, J., MacIntosh, D. y Ohlin, J. D. (eds.). *Lethal Autonomous Weapons*. Nueva York, Oxford University Press, pp. 25-39.
- Barela, S. J. y Plaw, A. (2021). Programming Decision? Requiring Robust Transparency for AWS. En: Gaillot, J., MacIntosh, D. y Ohlin, J. D. (eds.). *Lethal Autonomous Weapons*. Nueva York, Oxford University Press, pp. 73-88.
- Chengeta, T. (2016). Accountability Gap: Autonomous Weapon Systems and Modes of Responsibility in International Law. *Denver Journal of International Law & Policy*. Vol. 45, n.º 1, pp. 1-51.
- Crawford, E. y Pert, A. (2020). *International Humanitarian Law*. 2ª Ed. Cambridge, Cambridge University Press.
- Crawford, J. (2013). *State Responsibility*. Cambridge, Cambridge University Press.

- Crootof, R. (2016a). A meaningful floor for ‘meaningful human control’. *Temp. Int’l & Comp. L.J.* Vol. 30, pp. 53-62.
- Crootof, R. (2016b). War Torts: Accountability for Autonomous Weapons. *University of Pennsylvania Law Review*. Vol. 164, n.º 6, pp. 1347-1402.
- Endsley, M. (1995). The Out-of-the-Loop Performance and Level of Control in Automation. *Human factors*. Vol. 37, n.º 2, pp. 381-394.
- Fuzaylova, E. (2019). War Torts, autonomous weapon systems and liability: why a limited strict liability tort regime should be implemented. *Cardozo Law Review*. Vol. 40, pp. 1327-1366.
- Ganoon, F. y Khalaf, H. (2022). Autonomous weapons systems under humanitarian law. *Journal of Positive Psychology & Wellbeing*. Vol. 6, n.º 1, pp. 1063-1073.
- Hammond, D. (2015). Autonomous Weapons and the Problem of State Accountability. *Chicago Journal of International Law*. Vol. 15, n.º 2, pp. 652-687.
- Krishnan, A. (2021). Enforced transparency: A Solution to Autonomous Weapons as Potentially Uncontrollable Weapons Similar to Bioweapons. En: Gaillot, J., MacIntosh, D. y Ohlin, J. D. (eds.). *Lethal Autonomous Weapons*. Nueva York, Oxford University Press, pp. 219-235.
- Malik, S. (2018). Autonomous weapon systems: the possibility and probability of accountability. *Wisconsin International Law Journal*. Vol. 35, n.º 3, pp. 609-642.
- Marra, W. C. y McNeil, S. K. (2012). Understanding the ‘loop’: regulating the next generation of war machines. *Harvard Journal of Law and Public Policy*. Vol. 36, no. 3, pp. 1140-1185.
- Meza Rivas, M. J. (2018). El desarrollo y el uso de los sistemas de armas autónomas letales en los conflictos armados internacionales. Tesis doctoral inédita. Universitat de Barcelona. Disponible en: <https://www.tdx.cat/handle/10803/668711>
- Nadibaidze, A. (2022). Great power identity in Russia’s position on autonomous weapons systems. *Contemporary Security Policy*. Vol. 43, n.º 3, pp. 407-435.
- Parasuraman, R., Sheridan, T. y Wickens, C. (2000). A model for types and levels of human interaction with automation *IEEE transactions on systems, man and cybernetics. Part A: systems and humans*. Vol. 30, n.º 3, pp. 286-297.
- Saxon, D. (2014). A Human Touch: Autonomous Weapons, Directive 3000.09, and the ‘Appropriate Levels of Human Judgment over the

- Use of Force'. *Georgetown Journal of International Affairs*. Vol.15, n.º 2, pp. 100-109.
- Schaub, G. y Kristoffersen, J. W. (2017). In, on or out of the Loop? Denmark and Autonomous Weapons Systems. *Centre for military studies*, pp. 1-39.
- Schmitt, M. y Thurnher, J. (2013). Out of the loop: autonomous weapon systems and the law of armed conflict. *Harvard National Security Journal*. Vol. 4, n.º 2, pp. 231-281.
- Scholz, J. y Galliot, J. (2021). The Humanitarian Imperative for Minimally-Just AI in weapons. En: Gaillot, J., MacIntosh, D. y Ohlin, J. D. (eds.). *Lethal Autonomous Weapons*. Nueva York, Oxford University Press, pp. 57-72.
- Sehrawat, V. (2017). Autonomous weapon system: Law of armed conflict (LOAC) and other legal challenges. *Computer Law & Security Review*. Vol. 33, pp. 1-35.
- Van den Boogaard, J. (2015). Proportionality and autonomous weapon systems. *Journal of International Humanitarian Legal Studies*. Vol. 6, n.º 2, pp. 1-36.
- Williams, J. (2021). Locating LAWS: Lethal Autonomous Weapons, Epistemic Space and 'Meaningful Human Control. *Journal of Global Security Studies*. Vol. 6, n.º 4, pp. 1-18.

Otras publicaciones, informes y recursos de internet

- Amoroso, D. y Tamburrini, G. ICRAC. (2019). What makes human control over weapons systems 'meaningful'? Presentado ante la *Convention on Certain Conventional Weapons (CCW) Meeting of Experts on Lethal Autonomous Weapons Systems*. Ginebra. Ginebra, Naciones Unidas. Disponible en:
http://www.icrac.net/wp-content/uploads/2019/08/Amoroso-Tamburrini_Human-Control_ICRAC-WP4.pdf
- Article 36. (2013). Killer robots: UK Government Policy on Fully Autonomous Weapons. Reino Unido, Article 36. Disponible en:
https://article36.org/wp-content/uploads/2013/04/Policy_Paper1.pdf
- Article 36. (2016). Key elements of meaningful human control. Presentado ante la *Convention on Certain Conventional Weapons (CCW) Meeting of Experts on Lethal Autonomous Weapons Systems*. Ginebra, pp. 11-15, abril de 2016. Ginebra, Naciones Unidas. Dis-

- ponible en: <https://www.article36.org/wp-content/uploads/2016/04/MHC-2016-FINAL.pdf>
- China. (2018). Position Paper. Presentado ante la *Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be deemed to Be Excessively Injurious or to Have Indiscriminate Effects*. Ginebra, pp. 9-13, abril de 2018. Ginebra, Naciones Unidas. Disponible en: [https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_\(2018\)/CCW_GGE.1_2018_WP.7.pdf](https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_(2018)/CCW_GGE.1_2018_WP.7.pdf)
- Corte Internacional de Justicia. (s.f.). *Declarations recognizing the Jurisdiction of the Court as Compulsory*. Publicado en *Corte Internacional de Justicia*. 15 de septiembre 2022: Disponible en: <https://www.icj-cij.org/en/declarations>
- Estados Unidos. (2018). Human-machine interaction in the development, deployment and use of emerging technologies in the area of lethal autonomous weapons systems. Presentado ante la *Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be deemed to Be Excessively Injurious or to Have Indiscriminate Effects*. Ginebra, pp. 27-31, agosto de 2018. Ginebra, Naciones Unidas. Disponible en: [https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_\(2018\)/2018_GGE%2BLAWS_August_Working%2BPaper_US.pdf](https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_(2018)/2018_GGE%2BLAWS_August_Working%2BPaper_US.pdf)
- Estados Unidos. (2016a). Opening Statement to the Convention on Certain Conventional Weapons (CCW) Informal Meeting of Experts on Lethal Autonomous Weapons Systems. *U.S. Mission to International Organizations in Geneva*. 15 de septiembre 2022. Disponible en: <https://geneva.usmission.gov/2016/04/11/laws/>
- Estados Unidos. (2016b). U. S. Delegation Statement on ‘Appropriate Levels of Human Judgment’. *U.S. Mission to International Organizations in Geneva*. 15 de septiembre 2022. Disponible en: <https://geneva.usmission.gov/2016/04/12/u-s-delegation-statement-on-appropriate-levels-of-human-judgment/#:~:text=The%20United%20States%20agrees%20with,raised%20when%20designing%20and%20developing>
- Francia. (2018). Human-Machine Interaction in the Development, Deployment and Use of Emerging Technologies in the Area of Lethal Autonomous Weapons Systems. Presentado ante la *Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be deemed to Be Excessively Injurious or to*

- Have Indiscriminate Effects*. Ginebra, pp. 27-31 agosto de 2018. Ginebra, Naciones Unidas. Disponible en: [https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_\(2018\)/2018_GGE%2BLAWS_August_Working%2BPaper_France.pdf](https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_(2018)/2018_GGE%2BLAWS_August_Working%2BPaper_France.pdf)
- Geiss, R. (2016). *Autonomous Weapon Systems: Risk Management and State Responsibility* presentado ante el *Informal Meeting of Experts on lethal Autonomous Weapons Systems*, Ginebra, 11-15 de abril de 2016. Ginebra: Naciones Unidas. Disponible en: [https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Informal_Meeting_of_Experts_\(2016\)/Geiss-CCW-Website.pdf](https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Informal_Meeting_of_Experts_(2016)/Geiss-CCW-Website.pdf)
- Geiss, R. (2015). *The International-Law Dimension of Autonomous Weapon Systems*. *Friedrich Ebert Stiftung* (Alemania: octubre 2015). 15 de septiembre 2022. Disponible en: <https://library.fes.de/pdf-files/id/ipa/11673.pdf>
- Harvard College. (2010). *Commentary on the HPCR Manual on International Law Applicable to Air and Missile Warfare*. *Program on Humanitarian Policy and Conflict Research at Harvard University*. Massachusetts. Disponible en: https://www.researchgate.net/profile/Claude-Bruderlein/publication/264036862_Manual_on_International_Law_Applicable_to_Air_and_Missile_Warfare/links/59a911d50f7e9b27900e2f0e/Manual-on-International-Law-Applicable-to-Air-and-Missile-Warfare.pdf
- Heyns, C. (2013). Informe del Relator Especial sobre las ejecuciones extrajudiciales, sumarias o arbitrarias. Presentado ante el *Consejo de Derechos Humanos de la Asamblea General de Naciones Unidas*, 23.º período de sesiones. Ginebra, 9 de abril de 2013. Ginebra, Naciones Unidas. Disponible en: https://www.ohchr.org/sites/default/files/Documents/HRBodies/HRCouncil/RegularSession/Session23/A-HRC-23-47_en.pdf
- Human Rights Watch (2018a). *Losing Humanity: The Case Against Killer Robots*. Estados Unidos, Human Rights Watch, agosto de 2018. Disponible en: <https://www.hrw.org/sites/default/files/reports/Losing%20Humanity%20Executive%20Summary.pdf>
- Human Rights Watch. (2020). *Stopping Killer Robots: Country Positions on Banning Fully Autonomous and Retaining Human Control*. [Consulta: 21 de septiembre]. Disponible en: https://www.hrw.org/report/2020/08/10/stopping-killer-robots/country-positions-banning-fully-autonomous-weapons-and#_ftn130

- ICRAC. (2014). Statement on technical issues to the 2014 UN CCW Expert Meeting. Publicado el 14 de mayo de 2014. [Consulta: 21 de septiembre]. Disponible en: <https://www.icrac.net/icrac-statement-on-technical-issues-to-the-2014-un-ccw-expert-meeting/>
- ICRC (2019). Autonomy, artificial intelligence and robotics: technical aspects of human control. Ginebra, agosto de 2019. [Consulta: 21 de septiembre]. Disponible en: <https://www.icrc.org/en/document/autonomy-artificial-intelligence-and-robotics-technical-aspects-human-control>
- ICRC. (2010). Guía para interpretar la noción de participación directa en las hostilidades según el derecho internacional humanitario. *International Red Cross Committee*, Ginebra, diciembre 2010. [Consulta: 22 de septiembre]. Disponible en: <https://www.icrc.org/es/publication/guia-participacion-directa-hostilidades-derecho-internacional-humanitario-dih>
- ICRC. (2016). Views of the International Committee of the Red Cross (ICRC) on autonomous weapon system. Presentado en la *Convention on Certain Conventional Weapons (CCW) Meeting of Experts on Lethal Autonomous Weapons Systems*. Ginebra, 11 de abril de 2016. Ginebra, Naciones Unidas. Disponible en: <https://www.icrc.org/en/document/views-icrc-autonomous-weapon-system>
- International Law Commission, (2001). Draft articles on Responsibility of States for Internationally Wrongful Acts, with commentaries. *Yearbook of the International Law Commission*. Vol. 2, n.º 2, pp. 1-114. Disponible en: https://legal.un.org/ilc/texts/instruments/english/commentaries/9_6_2001.pdf
- Marauhn, T. (2014). An Analysis of the Potential Impact of Lethal Autonomous Weapons systems on Responsibility and Accountability for Violations of International Law. Presentado ante la *Convention on Certain Conventional Weapons (CCW) Meeting of Experts on Lethal Autonomous Weapons Systems*. Ginebra, pp. 13-16 de mayo de 2014. Ginebra, Naciones Unidas. Disponible en: [https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Informal_Meeting_of_Experts_\(2014\)/Marauhn_MX_Laws_SpeakingNotes_2014.pdf](https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Informal_Meeting_of_Experts_(2014)/Marauhn_MX_Laws_SpeakingNotes_2014.pdf)
- Polonia. (2015). Meaningful Human Control as a form of state control over LAWS. Presentado en la *Convention on Certain Conventional Weapons (CCW) Meeting of Experts on Lethal Autonomous Weapons Systems*. Ginebra, pp. 13-17 de abril de 2015. Ginebra, Naciones Unidas. Disponible en: https://docs-library.unoda.org/Convention_

- on_Certain_Conventional_Weapons_-_Informal_Meeting_of_Experts_(2015)/2015_LAWS_MX_Poland_characteristics.pdf
- Reino Unido. (2018). Human Machine Touchpoints: The United Kingdom's perspective on human control over weapon development and targeting cycles. Presentado ante la *Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be deemed to Be Excessively Injurious or to Have Indiscriminate Effects*. Ginebra, pp. 27-31 agosto de 2018) Ginebra, Naciones Unidas. Disponible en: [https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_\(2018\)/2018_GGE%2BLAWS_August_Working%2BPaper_UK.pdf](https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_(2018)/2018_GGE%2BLAWS_August_Working%2BPaper_UK.pdf)
- Reino Unido. (2015). Statement. Presentado ante el Informal Meeting of Experts on lethal Autonomous Weapons Systems Ginebra, pp. 13-17 de abril de 2015. Ginebra, Naciones Unidas. Disponible en: https://www.reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2015/meeting-experts-laws/statements/15April_UK.pdf
- Rusia. (2018). Russia's Approaches to the Elaboration of a Working Definition and Basic Functions of Lethal Autonomous Weapons Systems in the Context of the Purposes and Objectives of the Convention. Presentado ante la *Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be deemed to Be Excessively Injurious or to Have Indiscriminate Effects*. Ginebra, pp. 9-13 de abril de 2018. Ginebra, Naciones Unidas. Disponible en: [https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_\(2018\)/CCW_GGE.1_2018_WP.6_E.pdf](https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Group_of_Governmental_Experts_(2018)/CCW_GGE.1_2018_WP.6_E.pdf)
- Scharre, P. y Horowitz, M. (2015). Meaningful human control in weapon systems: A Primer. *Center for a New American Security*. [Consulta: 17 de septiembre]. Disponible en: https://s3.us-east-1.amazonaws.com/files.cnas.org/documents/Ethical_Autonomy_Working_Paper_031315.pdf?mtime=20160906082316&focal=none
- Soni, A. y Dominic, E. (2020). Legal and Policy Implications of Autonomous Weapons Systems. *The Centre for Internet & Society*. India. Disponible en: <https://cis-india.org/internet-governance/legal-and-policy-implications-of-autonomous-weapons-systems>

Legislación y jurisprudencia

Convención sobre la prohibición del desarrollo, la producción y el almacenamiento de armas bacteriológicas (biológicas) y tóxicas y sobre su destrucción. 10 de abril de 1972. Naciones Unidas, Convención sobre las Armas Biológicas.

Directiva n.º 3000.9: Autonomy in Weapon Systems. 21 de noviembre de 2012. Estados Unidos, Departamento de Defensa.

Estatuto de la Corte Internacional de Justicia. Naciones Unidas, Corte Internacional de Justicia.

<https://www.icj-cij.org/public/files/statute-of-the-court/statute-of-the-court-es.pdf>

Joint Motion for a Resolution pursuant to Rule 123(2) and (4) of the Rules of the Procedure on Autonomous weapon systems. 10 de septiembre de 2018. Unión Europea, Parlamento Europeo.

Protocolo I adicional a los Convenios de Ginebra de 1949 relativo a la protección de las víctimas de los conflictos armados internacionales. 8 de junio de 1977. Naciones Unidas, Convención de Ginebra.

Responsibility of States for Internationally Wrongful Acts. Diciembre de 2001. Nueva York, Asamblea General de Naciones Unidas.

Tercer Convenio de Ginebra relativo al trato debido a los prisioneros de guerra. 1949. Naciones Unidas, Asamblea General.

Tratado sobre el Comercio de Armas. 2013. Naciones Unidas, Asamblea General.

SCIJ de 9 de abril de 1949, *Reino Unido c. Albania.*

STPIY de 5 de diciembre de 2003, *caso Galić.*