



# SISTEMAS DE ARMAS AUTÓNOMAS:

el dilema de la rendición de cuentas

JIMENA SOFÍA VIVEROS ÁLVAREZ

SERIE

OPINIONES TÉCNICAS SOBRE TEMAS DE RELEVANCIA NACIONAL

42

# INSTITUTO DE INVESTIGACIONES JURÍDICAS

OPINIONES TÉCNICAS SOBRE TEMAS DE RELEVANCIA NACIONAL, núm. 42

Dra. Nuria González Martín  
*Coordinadora de la serie*

Lic. Mariana Ávalos Jiménez  
*Asistente de la serie*

## COORDINACIÓN EDITORIAL

Lic. Raúl Márquez Romero  
*Secretario Técnico*

Mtra. Wendy Vanesa Rocha Cacho  
*Jefa del Departamento de Publicaciones*

Jaime García Díaz  
*Cuidado de la edición*

José Antonio Bautista Sánchez  
*Formación en computadora*

Edith Aguilar Gálvez  
*Diseño de cubierta e interiores*



# **SISTEMAS DE ARMAS AUTÓNOMAS:**

el dilema de la rendición de cuentas

JIMENA SOFÍA VIVEROS ÁLVAREZ

Esta edición y sus características son propiedad de la Universidad  
Nacional Autónoma de México.

Prohibida la reproducción total o parcial por cualquier medio  
sin la autorización escrita del titular de los derechos patrimoniales.

Primera edición: 25 de mayo de 2021

DR © 2021. Universidad Nacional Autónoma de México

INSTITUTO DE INVESTIGACIONES JURÍDICAS

Circuito Maestro Mario de la Cueva s/n  
Ciudad de la Investigación en Humanidades  
Ciudad Universitaria, Coyoacán, 04510 Ciudad de México

Impreso y hecho en México

ISBN Serie Opiniones Técnicas sobre Temas de Relevancia Nacional: 978-607-30-1256-0

# Contenido

7

Prefacio

11

Introducción

13

Contextualización

31

Consideraciones de derecho internacional  
humanitario

53

El dilema de la responsabilidad

## Prefacio

En la actualidad, transitamos de una etapa en la que se ha tenido que convencer de la eficacia e idoneidad del uso de las tecnologías de la información y comunicación, a otra etapa donde parece que las plataformas quedan obsoletas, los algoritmos y el aprendizaje automatizado nos superan, y en donde se buscan, en definitiva, más avances tecnológicos con un desarrollo correcto y eficiente del ser humano, quien es el último recipiente de las bondades que esperamos generen dichos avances tecnológicos.

Aquí es donde la inteligencia artificial (IA) entra en juego, auxiliando en la toma de decisiones, agilizando la productividad, transformando nuestra manera de comunicarnos, ya no sólo entre seres humanos, sino en nuestra interacción con asistentes virtuales, bots y demás expresiones tecnológicas, revolucionando la manera de hacer negocios a partir de optimizar complejos procesos logísticos y predecir necesidades particulares de consumo, innovando en la automatización de toda clase de vehículos terrestres, aéreos y marítimos, entre tantas otras expresiones de una tecnología que ha adquirido la capacidad de aprender por sí sola, a partir de recopilar y procesar datos bajo lineamientos específicos. Todo lo anterior nos impone a los estudiosos del derecho la muy trascendente premisa de ceñir los alcances de estas tecnologías al manto protector de la ética y la responsabilidad, colocando al ser humano en el centro de sus intereses y finalidades.

En una sociedad futura, a corto o mediano plazo, a las bondades que acompañan a la IA, debemos incorporar las ventajas que presenta la complementariedad con la inteligencia humana (IH), y así crear tecnologías que se caractericen por una simbiosis correcta que aprovechen todas las capacidades. Las personas —a través de su IH— y la IA aportan diferentes habilidades y fortalezas. La verdadera pregunta es: ¿cómo puede la IH trabajar con la IA para producir inteligencia aumentada?

No hablamos de competencia, sino de balances (IA/IH) como metas primigenias a la que deberíamos enfocarnos, si realmente queremos plantear una sociedad futura con niveles de progreso correctos, satisfactorios y, por ende, felices.

Desafortunadamente, la historia de la humanidad da cuenta de innumerables ejemplos en los que el aprovechamiento (e incluso el desarrollo) de diversas tecnologías ha sido motivado por fines ajenos a la virtud y el florecimiento de nuestra especie. En la medida en que los avances científicos pongan a nuestro alcance herramientas cada vez más poderosas, resulta imperativo entender su magnitud y efectos en caso de que se les pretenda dar un mal uso. Actualizar y reformar de manera robusta y seria el andamiaje legal, para impedir que la enorme potencialidad de la inteligencia artificial se destine a la destrucción y aniquilamiento del ser humano, es un imperativo de la más alta trascendencia.

Nuestra autora, Jimena Sofía Viveros Álvarez, no sólo plantea estas premisas —contundentes por su validez— sino que va más allá, al exponer en la presente contribución, *El dilema de la rendición de cuentas respecto a los sistemas de armas autónomas*, la tecnología más avanzada y al mismo tiempo la más peligrosa para la especie humana.

No nos presenta una visión apocalíptica; al contrario, plantea una visión positiva que contribuye al desarrollo con algoritmos eficaces, sólo si el ser humano, que siempre estará detrás, tiene las habilidades para no instruir y repetir patrones inadecuados.

Pablo PRUNEDA GROSS\*  
Nuria GONZÁLEZ MARTÍN\*\*

---

\* Coordinador de la Línea de Investigación en Derecho e Inteligencia Artificial en el Instituto de Investigaciones Jurídicas de la Universidad Nacional Autónoma de México.

\*\* Investigadora titular "C" en el Instituto de Investigaciones Jurídicas de la Universidad Nacional Autónoma de México, y miembro del Sistema Nacional de Investigadores, nivel III.



## Introducción\*

[...] la primer máquina ultrainteligente es el último invento que el hombre necesita hacer, siempre que ésta sea lo suficientemente dócil para decirnos cómo mantenerla bajo control.<sup>1</sup>

A lo largo de la historia, las etapas del desarrollo de la humanidad han estado definidas por las herramientas y tecnología empleada por nuestra especie: desde el descubrimiento del fuego a la electricidad, de la rueda a los cohetes espaciales, de la escritura a la programación de computadoras, de la lanza a las armas nucleares.

En nuestro panorama moderno, es común que los debates legales versen sobre la legalidad de los desarrollos tecnológicos emergentes.

Sin embargo, actualmente estos debates están tomando un giro completamente diferente, ya que ahora discutimos situaciones en las que el criterio humano pudiera ser reemplazado o superado por decisiones derivadas de la programación de una computadora.

---

\* La autora es abogada criminalista mexicana, especializada en derecho penal internacional.  
Nota: todas las traducciones fueron realizadas por quien suscribe este artículo analítico.

<sup>1</sup> Good, Irving John, *Speculations Concerning the First Ultra-Intelligent Machine*, Nueva York, Academic Press Inc., 1965, p. 33.

Por lo tanto, nos encontramos ante una encrucijada, enfrentando la más significativa, y quizás *la última* invención de la humanidad: la inteligencia artificial (IA).

La naturaleza de estas discusiones es significativamente diferente, ya que requiere un mayor nivel de comprensión técnica especializada, conciencia ética y conocimientos legales para poder comprender las ramificaciones de esta nueva tecnología. Por ende, con toda razón este es un tema multifacético, interdisciplinario y sumamente controversial.

Más aún, el debate se torna particularmente relevante en el contexto de los conflictos armados, ya que el derecho internacional humanitario (DIH) acepta que durante los tiempos de guerra habrá muertes legítimas, destrucción y despliegue de armas a gran escala.

Por lo tanto, es imperativo que como sociedad entremos al análisis y discusión de las implicaciones fácticas y legales del creciente desarrollo de las tecnologías de IA con fines bélicos, en particular los sistemas de armas autónomas (SAA).<sup>2</sup>

El análisis que presento en este artículo tiene como objetivo evidenciar el vacío legal respecto al uso de dichas tecnologías en el ordenamiento jurídico actual, desde el enfoque del DIH.

El propósito de lo anterior es hacer un llamado a todos los actores relevantes y partes interesadas, para participar activamente en los debates que conduzcan a las reformas normativas y conceptuales necesarias para cerrar la insostenible brecha de responsabilidad y rendición de cuentas que genera esta laguna jurídica.

---

<sup>2</sup> También conocidas como sistemas de armas autónomas letales.

## Contextualización

En este apartado, la autora proporcionará una contextualización del debate en torno a la IA y los SAA, ofreciendo una guía conceptual de los términos y tecnologías que conforman el objeto de estudio en este artículo. Cabe precisar que este es un artículo *legal* que no pretende autoridad alguna en los aspectos científicos o meramente técnicos de dichos conceptos, ya que la mayoría de ellos ni siquiera están definidos de manera conclusiva por los principales expertos en sus campos, y porque la velocidad exponencial con la que evolucionan hace que sean de la naturaleza más fluida.

Comencemos con un examen semántico de las nociones relevantes para este estudio:

La palabra *artificial*, derivada del latín *artificialis*, es equivalente a los términos “facticio”, “sintético”, “falso”, “antinatural”; una cosa que es artificial está hecha por el hombre o es producida por humanos, generalmente para que se parezca a algo que sí es natural.<sup>3</sup>

Por otro lado, la palabra *inteligencia* es más difícil de definir. La inteligencia se explica en términos generales como la “capacidad de comprender”, “la capacidad de resolver

---

<sup>3</sup> Voz: “Artificial, adj.”, *Diccionario de la lengua española (DLE)*, Real Academia Española (RAE), disponible en <https://dle.rae.es/artificial?m=form>, ingresado el 2 de julio de 2020.

problemas” o simplemente como “conocimiento o el acto de entender”.<sup>4</sup> Sin embargo, esta noción todavía es cuestionada entre los psicólogos, ya que algunos de ellos la relacionan con el intelecto humano y, por lo tanto, se limitan al cerebro cognitivo. Así pues, esta noción estaría tradicionalmente ligada a la condición *humana*.

Para Stephen Hawking, “la inteligencia es la capacidad de adaptarse al cambio”.<sup>5</sup> Esta cita nos sirve para ligar el próximo concepto *compuesto*, la IA.

No hay consenso sobre una definición universal de este concepto, ya que los principales expertos en la materia la afinan y desafían constantemente.<sup>6</sup> El único aspecto definitivo y sobre el cual existe consenso es que la IA es una tecnología *muy* disruptiva.

Como campo de investigación, el nombre de IA se decidió durante un taller en la Universidad de Dartmouth en 1956, donde un grupo de científicos destacados se reunieron durante ocho semanas, llevando a cabo una lluvia de ideas sobre la concepción sobre el concepto de “máquinas que piensan”.<sup>7</sup>

---

<sup>4</sup> Voz: “Inteligencia”, *DLE*, RAE, disponible en <https://dle.rae.es/inteligencia?m=form>, ingresado el 2 de julio de 2020.

<sup>5</sup> Stephen Hawking declaró esto en su graduación de la Universidad de Oxford, “Professor Stephen Hawking: 13 of his most Inspirational Quotes”, *The Telegraph*, Londres, 8 de enero de 2016, disponible en [www.telegraph.co.uk/news/science/stephen-hawking/12088816/Professor-Stephen-Hawking-13-of-hismostinspirational-quotes.html](http://www.telegraph.co.uk/news/science/stephen-hawking/12088816/Professor-Stephen-Hawking-13-of-hismostinspirational-quotes.html), ingresado el 2 de julio de 2020.

<sup>6</sup> Scherer, Matthew U., “Regulating Artificial Intelligence Systems, Risks, Challenges, Competencies and Strategies”, *Harv. J. L. & Tech*, vol. 29, núm. 2, pp. 353, 359-362, 2016 (en adelante: Scherer, “Regulating AI”).

<sup>7</sup> Sileno, Giovanni, *History of AI, Current Trends, Prospective Trajectories*, Asser Institute, Winter Academy on Artificial Intelligence and International Law, 2021 (en adelante: Sileno, *History of AI*); Sileno menciona el grupo de 20 científicos e ingenieros notables, que estaban en el taller de Dartmouth en 1956, entre ellos: John McCarty (lenguaje LISP, cálculo de situaciones, lógicas no monótonas) Marvin Minsky (marcos, perceptrón, sociedad de mentes), Herbert Simon (teórico de la lógica, solucionador de problemas generales, racionalidad limitada), Allen Newell (teórico de la lógica, solucionador de problemas generales, el nivel de conocimiento), Ray Solomonoff

A grandes rasgos, puede entenderse como el uso de sistemas informáticos para realizar tareas que antes requerían de la inteligencia, cognición o razonamiento humanos.<sup>8</sup> Es una categoría de investigación destinada a desarrollar sistemas que sean capaces de resolver problemas o alcanzar metas en diferentes grados de dificultad mediante el razonamiento. Es decir, imitando las habilidades humanas de resolución de problemas, en algunos casos incluyendo la capacidad de aprender de la experiencia, y por lo tanto, mejorar las capacidades de la máquina sin ninguna intervención humana,<sup>9</sup> y que está diseñada para actuar como un agente racional.<sup>10</sup>

En ese sentido, la Organización para la Cooperación y el Desarrollo Económicos (OCDE) ha dictado cinco principios básicos para la regulación de la IA —*lato sensu*— en un documento de acuerdo general adoptado por 42 Estados miembros.<sup>11</sup>

- La IA debe estar al servicio de las personas y del planeta, impulsando un crecimiento inclusivo, el desarrollo sostenible y el bienestar.

---

(padre de la probabilidad algorítmica, teoría de la información algorítmica), Arthur Lee Samuel (primer algoritmo de aprendizaje automático para damas), W. Ross Ashby (pionero en cibernética, ley de la variedad requerida), Claude Shannon (padre de la teoría de la información) y John Nash (padre de la teoría de juegos).

<sup>8</sup> Voz: "Inteligencia (artificial)", *DLE*, DRAE, disponible en <https://dle.rae.es/inteligencia#2DxmhCT>, ingresado el 2 de julio de 2020.

<sup>9</sup> Kononenko, Igor y Kukar, Matjaz, *Machine Learning and Data Mining: Introduction to Principles and Algorithms*, 2007, p. 38; véase también Scherer, "Regulating AI", *supra* nota 6, p. 361.

<sup>10</sup> Russell, Stuart y Norvig, Peter, *Artificial Intelligence - A Modern Approach*, 3a. ed., 2010, pp. 4 y 5 (en adelante: Russell y Norvig, *AI-A Modern Approach*).

<sup>11</sup> Organización para la Cooperación y el Desarrollo Económicos, *Principios de la OCDE sobre IA*, disponible en <https://www.oecd.org/centrodemexico/medios/cuarentaydospaísesadoptanlosprincipiosdelaoocdesobreinteligenciaartificial.htm>, ingresado el 5 de julio de 2020.

- Los sistemas de IA deben diseñarse de manera que respeten el Estado de derecho, los derechos humanos, los valores democráticos y la diversidad, e incorporar salvaguardas adecuadas que permitan la intervención humana.
- Los sistemas de IA deben estar presididos por la transparencia y una divulgación responsable, a fin de garantizar que las personas sepan cuándo están interactuando con ellos y puedan oponerse a los resultados de esa interacción.
- Los sistemas de IA han de funcionar con robustez, de manera fiable y segura durante toda su vida útil, y los potenciales riesgos deberán evaluarse y gestionarse en todo momento.
- Las organizaciones y las personas que desarrollen, desplieguen o gestionen sistemas de IA deberán responder de su correcto funcionamiento en consonancia con los principios precedentes.

No obstante, el potencial de esta tecnología tiene un alcance muy amplio, por lo que es necesario hacer una distinción entre los diferentes tipos de IA:<sup>12</sup>

- IA acotada, que tiene un rango limitado de habilidades y es la IA que más prevalece en nuestro mundo actual. Está programada para realizar una tarea específica extremadamente bien.
- IA general, que está a la par con las capacidades humanas. Esta tecnología no se ha logrado aún. El propósito de la IA general es pensar, comprender y actuar de una manera indistinguible de la de un ser humano en una situación determinada. Cabe destacar que el objetivo de muchos proyectos relacionados con la IA es que estos sistemas puedan adaptarse a diferentes situaciones y

---

<sup>12</sup> O'Carroll, Brodie, *What are the 3 types of AI? A Guide to Narrow, General, and Super Artificial Intelligence*, 2017, disponible en <https://codebots.com/artificial-intelligence/the-3-types-of-ai-is-the-third-even-possible>, ingresado el 5 de julio de 2020.

funcionar sin control humano. Lo que los científicos no han logrado aún, es una forma de hacer que las máquinas sean *conscientes*, mediante la programación de un conjunto completo de habilidades cognitivas que hasta ahora sólo son conocidas en los humanos.

- Súper IA, que es incluso más capaz que un humano. Está destinada a superar nuestras capacidades y así superar las limitaciones de nuestra especie. Es superpoderosa y consciente de sí misma, más allá del sentido humano. Esto simboliza la máxima evolución de este campo, y es en lo que se basa la teoría de la *singularidad*, la cual dice que las máquinas algún día serán lo suficientemente inteligentes para programarse y mejorarse a sí mismas hasta independizarse de sus creadores humanos.

Por un lado, algunos tecnoescépticos creen que este escenario es improbable, por el otro, el inventor de Google, Ray Kurzweil, predice que esta “singularidad” ocurrirá en 2045 con la creación de una IA consciente de sí misma, que será millones de veces más poderosa que todos los cerebros humanos.<sup>13</sup>

La vía para lograr esto es a través de una tecnología llamada aprendizaje automático (en inglés *machine learning*) un concepto que fue inicialmente introducido en la década de 1940 por el matemático Alan Turing y desarrollada a través de su *juego de imitación*, lo que definió un estándar operativo para la inteligencia, conocido como el “Test de Turing”,<sup>14</sup> el cual se convirtió en la base de la IA que conocemos actualmente.<sup>15</sup>

---

<sup>13</sup> Business, AI, *Ray Kurzweil Predicts that the Singularity will take Place in 2045*, (2017), disponible en [https://aibusiness.com/document.asp?doc\\_id=760200](https://aibusiness.com/document.asp?doc_id=760200), ingresado el 5 de julio de 2020.

<sup>14</sup> Russell y Norvig, *AI-A Modern Approach*, *supra* nota 10, pp. 16 y 17.

<sup>15</sup> *Ibidem*, pp. 16 y 17.

En esencia, el aprendizaje automático es un proceso que permite que los sistemas artificiales mejoren a partir de la experiencia,<sup>16</sup> permite que las máquinas se adapten a nuevos entornos y actúen de una manera que les permita lograr el objetivo asignado independientemente de obstáculos imprevistos y sin una dirección explícita de su programador.<sup>17</sup>

Idealmente, el aprendizaje automático se convertiría en una solución para abordar de manera más eficiente, efectiva y precisa la imprevisibilidad, sin recibir órdenes del programador.<sup>18</sup>

Por otro lado, el aprendizaje profundo (en inglés *deep learning*) es una subárea del aprendizaje automático que se ocupa de algoritmos inspirados en la estructura y función del cerebro, llamadas redes neuronales artificiales.<sup>19</sup> Se basa en una jerarquía de aprendizaje de representación, que produce diferentes niveles de abstracción.<sup>20</sup> Básicamente, es una expansión del aprendizaje automático en capas multiplicadas, asimilando así una cantidad exponencial de datos.

Estas tecnologías tienen una ventaja asimétrica notable sobre los humanos, dado que pueden acumular conocimiento de bases de datos potencialmente infinitas y, a su vez, nos dejan con un conocimiento muy limitado de sus capacidades.

---

<sup>16</sup> Sileno, *History of AI*, *supra* nota 7.

<sup>17</sup> Russell y Norvig, *AI-A Modern Approach*, *supra* nota 10, p. 693; Coglianese, Carry y Lehr, David, "Regulating by Robot: Administrative Decision Making in the Machine-Learning Era", *Georgetown Law Journal*, SSRN, 2017, p. 1156 (en adelante: Coglianese y Lehr, "Regulating by Robot"); Rich, Michael L., "Machine Learning, Automated Suspicion Algorithms, and the Fourth Amendment", *U. PA. L. Rev.*, vol. 164, 2015-2016, pp. 871, 875 (en adelante: Rich, "Machine Learning").

<sup>18</sup> Russell y Norvig, *AI-A Modern Approach*, *supra* nota 10, p. 693; Coglianese y Lehr, "Regulating by Robot", *supra* nota 17, p. 1156; Rich, "Machine Learning", *supra* nota 17, p. 875.

<sup>19</sup> Brownlee, Jason, *What is Deep Learning?*, 2019, disponible en <https://machinelearningmastery.com/what-is-deep-learning/>, ingresado el 4 de febrero de 2021.

<sup>20</sup> Goodfellow, Ian *et al.*, *Deep Learning*, MIT Press, 2016.

Por consiguiente, existe una circunstancia que se debe tener en cuenta, un gran *caveat* derivado tanto del aprendizaje automático como del aprendizaje profundo, al cual podríamos referirnos como una “incógnita conocida”. Esta incógnita comprendería el proceso de *toma de decisiones*, también conocido como “la caja negra”. La razón de esto es que, a diferencia de los procesos de toma de decisiones de los humanos que son lineales, los procesos de toma de decisiones artificiales son demasiado complejos para que los entendamos nosotros, porque la propia máquina crea sus algoritmos, y es así como estos procesos conforman la caja negra.<sup>21</sup>

Esto significa que, independientemente de su “configuración” original, es el propio programa el que decide el valor adecuado que se le asigna a cada elemento que percibe.<sup>22</sup>

Además, el programador no sabe qué regla o incluso qué características específicas utilizó la máquina para llegar a una determinada conclusión, ni tampoco puede deconstruir las inferencias o rastrear los procesos de decisión que se aplicaron.<sup>23</sup>

A manera de ejemplo, una función de utilidad que esté programada para mitigar o evitar el sufrimiento humano, podría decidir matar en lugar de herir a una persona, ya que las personas no sufren cuando están muertas.<sup>24</sup>

---

<sup>21</sup> Nicholson Price II, W., “Black-Box Medicine”, *Harvard Journal of Law & Technology*, vol. 28, 2014-2015, pp. 419, 432-434 (en adelante: Nicholson, “Black-Box Medicine”); Rich, “Machine Learning”, *supra* nota 17, p. 886.

<sup>22</sup> Coglianese y Lehr, “Regulating by Robot”, *supra* nota 17, p. 1156. “No podemos saber realmente en qué características precisas se basa cualquier algoritmo de aprendizaje automático”.

<sup>23</sup> *Idem*.

<sup>24</sup> Para más ejemplos, véase Calo, Ryan, “Robotics and the Lessons of Cyberlaw”, *California Law Review*, vol. 103, 2015, pp. 542 y 543 (en adelante: Calo, “Robotics and the Lessons of Cyberlaw”).

En otras palabras, el programador controla los datos introducidos al programa en su fase de aprendizaje, proporciona pautas de optimización para la interpretación de estos datos (lo que se conoce como función de utilidad) y está al tanto del resultado que extrapola el programa, pero para todos los demás efectos, la entidad artificial se considera una caja negra que no ofrece una explicación intuitiva ni causal de sus acciones.<sup>25</sup>

En resumen, el uso de programas de aprendizaje automático debe hacerse siempre con la plena conciencia que conlleva el riesgo inherente de que no hay forma de predecir, comprender o auditar una decisión específica realizada por la IA en términos comprensibles para los humanos.<sup>26</sup>

Para asimilar estos conceptos, viene a la mente la siguiente cita: “Como sabemos, hay conocimientos conocidos; hay cosas que sabemos que sabemos. También sabemos que existen incógnitas conocidas; es decir, sabemos que hay algunas cosas que no sabemos. Pero también hay incógnitas desconocidas, las que no conocemos, no las conocemos”.<sup>27</sup>

---

<sup>25</sup> Shilo, Liron, *When Turing Met Grotius AI, Indeterminism, and Responsibility*, SSRN, 2018, p. 14 (en adelante: Shilo, *When Turing Met Grotius*).

<sup>26</sup> *Ibidem*, pp. 11 y 12, 18 y 19. Para más información sobre la caja negra, véase Nicholson, “Black-Box Medicine”, *supra* nota 21, pp. 432-437; 442-467. Él explica los pros y contras de las cajas negras en el contexto médico; Rich, “Machine Learning”, *supra* nota 17, pp. 886, 923 y 924. Describe el intercambio entre la utilización de algoritmos que tienen las cajas negras, pero que son de gran precisión en sus predicciones, y concluye que, a pesar de la exactitud favorable en la predicción, las cajas negras deberían ser más transparentes en aras de analizar la implicación de la cuarta enmienda por el uso de tales algoritmos.

<sup>27</sup> Respuesta de Donald Rumsfeld a una pregunta del Departamento de Defensa de Estados Unidos, durante una conferencia de prensa en febrero de 2002, disponible en <https://archive.defense.gov/Transcripts/Transcript.aspx?TranscriptID=2636>, ingresado el 7 de julio de 2020.

## Sistemas de armas autónomas

Es destacable que la tecnología subyacente de la IA tiene el potencial de adaptarse tanto a usos civiles como militares. Este estudio se centra en el segundo dominio, cuyas aplicaciones podrían incluir inteligencia, vigilancia y reconocimiento, navegación, comando y control multidominio, defensa antimisiles, defensa cibernética, manipulación de información, reconocimiento de objetivos y desarrollo de armas.<sup>28</sup>

Los sistemas de armas más recientes son completamente *sui generis* y dan origen a una categoría propia *de jure* y *de facto*.

Estos sistemas son *entidades* altamente sofisticadas capaces de imitar las habilidades humanas de toma de decisiones para ejecutar una variedad de tareas *sin* intervención humana alguna.<sup>29</sup> También conocidos como sistemas de armas autónomas letales, estas entidades están diseñadas para *iniciar activamente* y *tomar decisiones letales* en lugar de actuar simplemente como sistemas defensivos y/o reactivos.<sup>30</sup>

Como tal, esta categoría de entidades no está definida actualmente en nuestro ordenamiento jurídico, dada la predominante y novedosa característica de contar con un carácter

---

<sup>28</sup> Nasu, Hitoshi, *Artificial Intelligence and the Obligation to Respect and to Ensure Respect for International Humanitarian Law*, Exeter Centre for International Law, 2019, p. 5.

<sup>29</sup> Véase Schmitt, Michael N., "Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics", *Harvard National Security Journal*, 2013, p. 4 (en adelante: Schmitt, "AWS and IHL"); Anderson, Kenneth *et al.*, "Adapting the Law of Armed Conflict to Autonomous Weapon Systems", *International Legal Studies*, vol. 90, 2014, p. 386; véase Alston, Philip, *Interim report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions*, U.N. Human R. Comm., U.N. Doc. A/65/321, 23 de agosto de 2010, p. 18; Kerr, Ian y Szilagyi, Katie, *Asleep at the Switch? How Killer Robots Become a Force Multiplier of Military Necessity*, 2016, p. 333; Ben-Naftali, Orna y Triger, Zvi, "The Human Conditioning: International Law and Science Fiction", *Law, Culture and the Humanities*, vol. 14, 2016, p. 38.

<sup>30</sup> Shilo, *When Turing Met Grotius*, *supra* nota 25, p. 2.

autónomo a nivel cognitivo y de toma de decisiones. Por ello, es primordial recordar que, aunque pueden usarse de manera armamentista, no deben examinarse o definirse simplemente como armas, ya que son mucho más que eso.

Tras una observación minuciosa, uno se da cuenta de que estas entidades no son armas, plataformas convencionales ni agentes morales equivalentes a los humanos para efectos legales. Dicho esto, a menudo se les refiere como la primera, en ocasiones como lo segundo, y con frecuencia se les trata como el tercero.<sup>31</sup>

Naturalmente, este es otro concepto que carece de consenso en su definición, y cuyo margen de discrecionalidad varía enormemente. En un extremo del espectro, un SAA se considera un componente automatizado de un arma existente y, en el otro, como una plataforma que es capaz de detectar, aprender y lanzar ataques.<sup>32</sup>

De manera relacionada, y en un esfuerzo por simplificar —pero en efecto, ampliando— el concepto, académicos de Harvard han introducido el término “algoritmos de guerra”, definido como cualquier algoritmo que se expresa en código de computadora, que se efectúa a través de un sistema construido y que es capaz de operar en el contexto de un conflicto armado. Estos sistemas incluyen arquitecturas de autoaprendizaje que están en el centro de los debates más importantes sobre el reemplazo del criterio humano con elecciones derivadas de algoritmos.<sup>33</sup>

Como se explicó anteriormente, el hecho de que estas elecciones puedan ser difíciles de anticipar o desentrañar para los humanos, *vis-à-vis* la posibilidad de que sean autó-

---

<sup>31</sup> *Ibidem*, p. 15.

<sup>32</sup> Lewis, Dustin A. *et al.*, *War-Algorithm Accountability*, HLS PILAC, 2016, p. 5 (en adelante: Lewis *et al.*, *War-Algorithm*).

<sup>33</sup> *Ibidem*, p. 10.

nomos y capaces de actuar físicamente en el mundo, definitivamente confronta las nociones jurídicas tal y como las conocemos hoy en día.<sup>34</sup>

En consecuencia, los SAA desafían las nociones fundamentales e interrelacionadas del derecho internacional público, el DIH, el derecho penal internacional (DPI) y los esquemas de responsabilidad relacionados. Estos conceptos incluyen atribución, control, previsibilidad y capacidad de reconstrucción.<sup>35</sup>

Es comprensible que el científico estadounidense Max Tegmark se refiera a la transición hacia la autonomía como “la tercera revolución de las armas”, después de la invención de la pólvora en el siglo trece y la de las armas nucleares en el siglo veinte.<sup>36</sup>

## Autonomía

Ahora bien, es importante abordar dos aspectos claves de esta nueva categoría: la autonomía y la toma de decisiones.

Dado que nuestras sociedades están cada día más interconectadas, es relativamente sencillo llevar a cabo no sólo los ciberataques tradicionales, sino los ciberataques cinéticos. Estos ataques son asaltos virtuales que tienen consecuencias tangibles en el mundo físico causando daños, lesiones o incluso la muerte, únicamente a través de la explotación de sistemas y procesos de información vulnerables. Esto es, utilizar el ciberespacio para infligir daños físicos en plantas de energía nuclear, instalaciones de agua, oleoductos, fábricas, hospitales,

---

<sup>34</sup> Calo, “Robotics and the Lessons of Cyberlaw”, *supra* nota 24, p. 542.

<sup>35</sup> Lewis *et al.*, *War-Algorithm*, *supra* nota 32, p. 77.

<sup>36</sup> Tegmark, Max, *Life 3.0: Being Human in the Age of Artificial Intelligence*, Nueva York, 2017.

bancos, sistemas de tránsito y estructuras de apartamentos.<sup>37</sup> El atacante puede estar ubicado en un lugar seguro y alejado, afectando de forma remota las vidas humanas y desestabilizando gobiernos nacionales o extranjeros al atacar su infraestructura crítica, pues todo lo que se requiere es una conexión a internet.

A pesar de lo peligrosa que es esta noción, a simple vista, es digno de mención que se trata de un tipo de guerra controlada, operada y dirigida por humanos... *en principio*.

Sin embargo, lo que actualmente concebimos como *espacio de batalla*<sup>38</sup> se está volviendo gradualmente, pero cada vez más, libre de humanos, tanto en la práctica como en la teoría. Es fácil entender la razón de esto, ya que las máquinas tienen una ventaja sobre los humanos, dado que pueden ejecutar algunas tareas de manera más rápida, precisa y económica que si las hiciéramos nosotros.

Teóricamente, los seres humanos todavía estamos a cargo de *cuándo* comenzar una guerra, *contra quién* se peleará (*ius ad bellum*), qué *armas, medios y métodos* se utilizarían y qué *objetivos* se deben alcanzar (*ius in bello*). Sin embargo, a nivel táctico, las máquinas se están convirtiendo en los “micro-gestores” y los ejecutores de cómo lograr estos objetivos. Decidirán cada vez más a quién, qué y cuándo atacar para hacerlo. Esto es lo que se conoce como automatización del espacio de batalla.<sup>39</sup>

La pregunta que debemos hacernos en este punto es, ¿podemos permitir que estas decisiones se deleguen de manera legal (y ética) a las máquinas?

---

<sup>37</sup> Goud, Naveen, *What is a Cyber Kinetic Attack?*, disponible en <https://www.cybersecurity-insiders.com/what-is-a-cyber-kinetic-attack/>, ingresado el 20 de julio de 2020.

<sup>38</sup> Previamente conocido como “campo de batalla”, como se explicará *infra* en el epígrafe “Consideraciones del derecho internacional humanitario”.

<sup>39</sup> Shilo, *When Turing Met Grotius*, *supra* nota 25, p. 1.

Además, es sólo cuestión de tiempo antes de que esta automatización del espacio de batalla llegue al siguiente nivel, el volverse autónoma.

El Comité Internacional de la Cruz Roja (CICR) es la organización humanitaria encargada de salvaguardar los Convenios de Ginebra, y ha declarado que no se opone a las nuevas tecnologías de la guerra *per se*.<sup>40</sup>

En todos los casos, como requisito mínimo, cualquier tecnología de guerra nueva debe utilizarse, y debe poder utilizarse, de conformidad con las normas vigentes del DIH. Sin embargo, las características únicas de las nuevas tecnologías de guerra, las circunstancias previstas y esperadas de su uso y sus consecuencias humanitarias pronosticables, plantean dudas respecto a si las reglas existentes son suficientes o si deben aclararse o complementarse, a la luz de su presumible impacto.<sup>41</sup>

Por supuesto, ciertas tecnologías militares —como las que permiten una mayor precisión en los ataques— pueden ayudar a las partes en conflicto a minimizar las consecuencias humanitarias de la guerra. No obstante, como ocurre con cualquier nueva tecnología de guerra, su legalidad depende de la forma en que se utilicen en la práctica.

Para el CICR, una aplicación importante es el uso de herramientas de IA y aprendizaje automático para controlar *hardware* militar físico, en particular, el creciente número de sistemas robóticos no tripulados —en el aire, en tierra y en el mar— con una amplia gama de tamaños y funciones. La IA y el aprendizaje automático pueden permitir una mayor autonomía en estas

---

<sup>40</sup> CICR, Artificial Intelligence and Machine Learning in Armed Conflict: A Human-Centered Approach, Ginebra, 6 de junio de 2019, p. 1 (en adelante: CICR, AI and Machine Learning in Armed Conflict).

<sup>41</sup> CICR, "International Humanitarian Law and the Challenges of Contemporary Armed Conflicts", 32nd International Conference of the Red Cross and Red Crescent, Ginebra, octubre de 2015, pp. 38-47.

plataformas robóticas, ya sean armadas o desarmadas, y controlar todo el sistema o funciones específicas, como vuelo, navegación, vigilancia o selección de objetivos.<sup>42</sup>

La organización también está interesada en la aplicación de la IA y el aprendizaje automático para el desarrollo de armas cibernéticas como “armas digitales autónomas”, ya que se espera que cambien la naturaleza tanto de las capacidades de ciberdefensa y ciberataques, aumentando así la escala y cambiando la naturaleza y gravedad de los ataques.<sup>43</sup>

Naturalmente, una preocupación recurrente entre los académicos es que su naturaleza autónoma crea una brecha de responsabilidad<sup>44</sup> —la caja negra— que niega cualquier culpabilidad moral de los actores humanos que participaron en su creación, uso o despliegue. Está claro que las cuestiones relativas a la atribución de responsabilidad son importantes, porque el hecho de que alguien pueda ser considerado responsable de desviarse de las reglas acordadas, es una condición *sine qua non* para librar una guerra justa.<sup>45</sup>

---

<sup>42</sup> CICR, AI and Machine Learning in Armed Conflict, *supra* nota 40, p. 3.

<sup>43</sup> Brundage, M. *et al.*, *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*, 2018; United Nations Institute for Disarmament Research, *The Weaponization of Increasingly Autonomous Technologies: Autonomous Weapon Systems and Cyber Operations*, UNIDIR, 2017.

<sup>44</sup> Davison, N., “Autonomous Weapon Systems under International Humanitarian Law”, *UNODA Occasional Papers. Perspectives on Lethal Autonomous Weapon Systems*, núm. 30, noviembre de 2017 (en adelante: Davison, “AWS under IHL”), disponible en <https://www.icrc.org/en/document/autonomous-weapon-systems-under-international-humanitarian-law>, ingresado el 2 de agosto de 2020.

<sup>45</sup> Véase Sparrow, Robert, “Killer Robots”, *Journal of Applied Philosophy*, vol. 24, 2007, pp. 62, 67 (en adelante: Sparrow, “Killer Robots”).

## ¿Tomadores de decisiones?

Quizás la aplicación más amplia y de mayor alcance es el uso de la IA y el aprendizaje automático para la toma de decisiones, permitiendo la recopilación y el análisis de datos para identificar personas u objetos, evaluar patrones de vida o comportamiento, hacer recomendaciones para estrategias militares u operaciones, o hacer predicciones sobre acciones o situaciones futuras.<sup>46</sup>

Estos sistemas automatizados de toma de decisiones son efectivamente una expansión de las herramientas de inteligencia, vigilancia y reconocimiento, que utilizan IA y aprendizaje automático para automatizar el análisis de grandes conjuntos de datos y proporcionar “consejos” a los humanos en la toma de decisiones particulares y, cada vez más, para automatizar tanto el análisis y la posterior decisión y/o acción por parte del sistema.<sup>47</sup> Las aplicaciones relevantes de la IA y del aprendizaje automático incluyen el reconocimiento de patrones, procesamiento del lenguaje natural, reconocimiento de imágenes, reconocimiento facial y reconocimiento de comportamiento. La preocupación del CICR gira en torno al hecho de que el posible uso de estos sistemas es extremadamente amplio y puede incluir decisiones sobre a quién o qué atacar y cuándo,<sup>48</sup> decisiones sobre a quién detener y durante cuánto tiempo,<sup>49</sup> decisio-

---

<sup>46</sup> CICR, AI and Machine Learning in Armed Conflict, *supra* nota 40, p. 4.

<sup>47</sup> *Idem*.

<sup>48</sup> Estados Unidos, “Implementing International Humanitarian Law in the Use of Autonomy in Weapon Systems”, Documento de Trabajo, Convención sobre Prohibiciones o Restricciones del Empleo de Ciertas Armas Convencionales que Puedan Considerarse Excesivamente Nocivas o de Efectos Indiscriminados (CAC) Grupo de Expertos Gubernamentales, marzo de 2019.

<sup>49</sup> Deeks, Ashley, “Predicting Enemies”, *Virginia Public Law and Legal Theory Research Paper*, núm. 21, 2018, pp. 1549-54.

nes sobre estrategia militar —incluso sobre el uso de armas nucleares—<sup>50</sup> y las relacionadas con operaciones específicas, como los intentos de predecir o adelantarse a los adversarios.<sup>51</sup>

Es precisamente esta preocupación la que ha motivado discusiones agitadas sobre el papel de la toma de decisiones en la guerra, y sobre quién está mejor preparado para tomar decisiones de vida o muerte: los humanos o las máquinas. Sin embargo, esta es una cuestión casi teológica, ya que no está técnicamente claro que una máquina pueda cumplir en todo momento con el DIH o las reglas para las que fue programada, y por otro lado, un ser humano puede actuar deliberadamente en violación al DIH.<sup>52</sup>

En todo caso, también existe un desacuerdo significativo sobre el análisis costo-beneficio que podría resultar de distanciar a los combatientes humanos del campo de batalla y si los beneficios potenciales de salvar vidas de los SAA son superados por los riesgos inherentes al hecho de que la guerra también se vuelve, en un sentido práctico, más fácil de llevar a cabo,<sup>53</sup> lo cual tiene como resultado una mayor vulnerabilidad de las poblaciones civiles.

Los defensores de estas tecnologías argumentan que la IA y los sistemas de apoyo en la toma de decisiones, basados en el aprendizaje automático, pueden permitir a los humanos tomar mejores decisiones en la forma en que se conducen las hostilidades, y que éstas

---

<sup>50</sup> Boulanin, V. (ed.), *The Impact of Artificial Intelligence on Strategic Stability and Nuclear Risk*, vol. 1: *Euro-Atlantic Perspectives*, Stockholm International Peace Research Institute, 2019.

<sup>51</sup> Hill, S. y Marsan, N., "Artificial Intelligence and Accountability: A Multinational Legal Perspective", *Big Data and Artificial Intelligence for Military Decision Making, Meeting Proceedings STO-MP-IST-160*, NATO, 2018.

<sup>52</sup> Sassóli, Marco, "Autonomous Weapons and International Humanitarian Law: Advantages, Open Technical Questions and Legal Issues to be Clarified", *International Law Studies*, vol. 90, U.S. Naval War College, 2014, p. 310 (en adelante: Sassóli, "Autonomous Weapons and IHL").

<sup>53</sup> Lewis et al., *War-Algorithm*, *supra* nota 32, p. 8.

estarán apegadas al DIH. Afirman que esto minimizará los riesgos para los civiles, al facilitar una recopilación y un análisis de información más rápida y generalizada.

No obstante, el CICR correctamente reitera el problema de la caja negra, dado que esos mismos análisis algorítmicos también podrían facilitar peores decisiones, violaciones del DIH y exacerbar los riesgos para los civiles. Lo anterior, especialmente debido a las limitaciones actuales de la tecnología, tales como la imprevisibilidad, la falta de explicación y los prejuicios discriminatorios. Desde una perspectiva humanitaria, esta es una preocupación fundamental, ya que son riesgos de lesiones o muerte para las personas o de destrucción de objetos civiles, y además porque estas *decisiones* se rigen por la *lex specialis* del DIH.<sup>54</sup>

---

<sup>54</sup> CICR, AI and Machine Learning in Armed Conflict, *supra* nota 40, p. 7.

## Consideraciones de derecho internacional humanitario

El DIH, también llamado derecho de los conflictos armados o leyes de la guerra, es la rama del derecho internacional que regula la conducta en un conflicto armado. De conformidad con el principio de igualdad entre los beligerantes, el DIH vincula a todas las partes en un conflicto armado, incluidos los grupos armados no estatales.<sup>55</sup>

El DIH, también conocido como *ius in bello*, se aplica independientemente de cualquier consideración sobre la legalidad del uso de la fuerza, *ius ad bellum*. La separación entre la aplicación de estos dos regímenes legales es necesaria por razones tanto humanitarias como prácticas. Es una característica recurrente, en los conflictos armados, que al menos una de las partes difiera sobre quién recurrió ilegalmente a la fuerza o sobre respecto a quién la causa es justa. Sin embargo, independientemente de qué parte “tenga la razón”, hay civiles en todos los lados del conflicto a los que se les deben de garantizar los mismos niveles de protección. Por lo tanto, el DIH se esfuerza por limitar los efectos de los conflictos armados: por un lado, proporciona protección a quienes no participan (o ya no participan) en las hostilidades, como civiles, heridos, enfermos y náufragos, o prisioneros de guerra y, por otro lado, el DIH limita los medios y métodos de la guerra, es decir, las armas y tácticas que pueden utilizarse.

---

<sup>55</sup> Kleffner, Jan K., “The Applicability of International Humanitarian Law to Organized Armed Groups”, *International Review of the Red Cross*, vol. 93, 2011, p. 443.

Además del derecho consuetudinario, el principal *corpus iuris* que rige la conducción de las hostilidades son los cuatro convenios de Ginebra y sus tres protocolos adicionales.

En este punto, es importante reconocer la transformación dramática en la naturaleza de los conflictos armados, desde los tiempos en que estos convenios fueron redactados en 1949, justo después de la Segunda Guerra Mundial, donde los actores, combatientes, campos de batalla, armas y tácticas empleadas eran significativamente diferentes a las que prevalecen hoy.

Aunque estos convenios salvaguardan principios básicos muy importantes que siguen siendo extremadamente relevantes para la conducción de las hostilidades en la actualidad (como la proporcionalidad, la distinción, la necesidad militar y la humanidad), es innegable que éstos fueron redactados contemplando escenarios muy diferentes. Hoy en día, los conflictos armados son en su mayoría poco convencionales, no internacionales, irregulares y asimétricos.

El mayor cambio se ha producido en la noción del “campo de batalla”, que es un término utilizado para describir un lugar o terreno en el que se libró o se libra una guerra.<sup>56</sup> En la actualidad, la evolución de la conducción de las hostilidades ha llevado a que ahora se le denomine “espacio de batalla”.

En las guerras “antiguas”, los campos de batalla sólo se daban en dos dominios, tierra y mar. El tercer dominio, el aire, entró en juego después de la invención del avión por los hermanos Wright, y su ventajoso uso militar durante la Primera Guerra Mundial, y el cuarto, el espacio exterior, se volvió relevante tras el lanzamiento del primer satélite, Sputnik I, y del

---

<sup>56</sup> Voz: “Campo (de batalla)”, *DLE*, RAE, disponible en <https://dle.rae.es/campo>, ingresado el 20 de agosto de 2020.

primer misil balístico intercontinental, el R-7 Semyorka: ambos logrados por la Unión Soviética en 1957 durante la Guerra Fría.<sup>57</sup>

En 2010, *The Economist* declaró que “la guerra ha entrado en el quinto dominio: el *ciberespacio*”,<sup>58</sup> y un año después, el Departamento de Defensa de Estados Unidos incorporó oficialmente el nuevo dominio a su planificación, doctrina, recursos y operaciones, a su vez dando pie a que la OTAN reconociera también el ciberespacio como un dominio operacional en 2016.<sup>59</sup> Posteriormente, en los últimos años se ha hablado de un sexto dominio, algunos proponentes plantean que podría ser el espectro electromagnético o, aún más aterrador, la mente humana.<sup>60</sup>

Está claro que estos avances de tecnología hubieran sido imposibles de prever o predecir en 1949 por los redactores de los Convenios de Ginebra. En consecuencia, se puede conceder que debiera ser apropiado introducir nuevos conceptos en las leyes de la guerra, especialmente a la luz de tecnologías tan disruptivas como la IA y los SAA. A todas luces, es necesario hacerlo a fin de regular adecuadamente la gama contemporánea de conflictos armados, y cumplir con su propósito de asegurar una protección efectiva a los civiles contra el sufrimiento innecesario.

---

<sup>57</sup> The Strategist, *Why the Fifth Domain is Different*, Australian Strategic Policy Institute, 2019, disponible en <https://www.aspistrategist.org.au/why-the-fifth-domain-is-different/>, ingresado el 20 de agosto de 2020 (en adelante: The Strategist, *Why the Fifth Domain is Different*).

<sup>58</sup> The Economist, *War in the Fifth Domain*, 1 de julio de 2010, disponible en <https://www.economist.com/briefing/2010/07/01/war-in-the-fifth-domain>, ingresado el 20 de agosto de 2020.

<sup>59</sup> Estados Unidos, *United States Department of Defense Strategy for Operating in Cyberspace*, julio de 2011; véase también Schmitt, Michael N., *Tallinn Manual on the International Law applicable to Cyber Warfare*, NATO, Cooperative Cyber Defence Centre of Excellence, Cambridge University Press, 2013.

<sup>60</sup> The Strategist, *Why the Fifth Domain is Different*, *supra* nota 57.

## Sistemas de armas autónomas bajo el derecho internacional humanitario

Primero que nada, como marco legal, los Estados Parte de los Convenios de Ginebra de 1949 tienen la obligación general de respetar y garantizar el cumplimiento de los mismos, así como tomar las medidas necesarias para la represión de todos los actos violatorios a las disposiciones de aquéllos.<sup>61</sup> Esta obligación representa el núcleo de un sistema de responsabilidad colectiva,<sup>62</sup> pues se cuenta entre los medios disponibles para asegurar el cumplimiento de los convenios.<sup>63</sup>

Lo anterior es una obligación general no recíproca, aplicable en toda circunstancia —por lo cual se extiende incluso en tiempos de paz—<sup>64</sup> y que abarca tanto las obligaciones negativas de abstenerse de participar de cualquier forma en la violación de una norma de DIH,<sup>65</sup> como las obligaciones positivas de tomar todas las medidas factibles para cumplir con el régimen.<sup>66</sup> El alcance de estas obligaciones cubre las medidas de cumplimiento de todas las

---

<sup>61</sup> Convenios de Ginebra de 1949: Convenio de Ginebra I, sobre los Heridos y Enfermos de las Fuerzas Armadas en Campaña, artículo 49 (en adelante: CG I); Convenio de Ginebra II, sobre los Heridos, Enfermos y Náufragos de las Fuerzas Armadas en el Mar, artículo 50 (en adelante: CG II); Convenio de Ginebra III, sobre los Prisioneros de Guerra, artículo 129 (en adelante: CG III); y Convenio de Ginebra IV, sobre los Civiles, artículo 146 (en adelante: CG IV).

<sup>62</sup> Boisson de Chazournes, Laurence y Condorelli, Luigi, "Common Article 1 of the Geneva Conventions Revisited: Protecting Collective Interests", *IRRC*, vol. 82, 2000, pp. 67 y 68.

<sup>63</sup> CICR, *Commentary on the First Geneva Convention: Convention (I) For the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field*, 2a. ed., 2016, p. 121 (en adelante: CICR, *Commentary on the 1st GC*).

<sup>64</sup> *Ibidem*, pp. 127-129.

<sup>65</sup> CIJ, *Actividades Militares y Paramilitares contra el Gobierno de Nicaragua*, fondo, sentencia, 1986, pp. 220, 255.

<sup>66</sup> Véase Geiß, Robin, "The Obligation to Respect and Ensure Respect for the Conventions", en Clapham, Andrew *et al.* (eds.), *The 1949 Geneva Conventions: A Commentary*, Oxford University Press, 2015, pp.117-120, 130-132.

personas bajo su autoridad y jurisdicción, incluyendo a la población en su conjunto, así como otros actores estatales y no estatales.<sup>67</sup>

### *Naturaleza de los sistemas de armas autónomas*

Para cualquier experto legal, el primer paso para realizar un análisis de DIH es identificar la regla relevante que se aplicará en un contexto particular.

Una cuestión crucial para empezar es discernir si en el DIH los SAA se consideran un arma (o un sistema de armas) o si deben clasificarse de otra manera, como podría ser un combatiente. La opinión más común se inclina a clasificarlos como *armas*; <sup>68</sup> sin embargo, la perspectiva de que un SAA pueda ser considerado un combatiente, cuando, por ejemplo, se habla de la capacidad de toma de decisiones del sistema, también se ha discutido seriamente entre los expertos.<sup>69</sup>

El Grupo de Trabajo sobre el Derecho de la Guerra, del Departamento de Defensa de Estados Unidos, adopta este segundo enfoque al hacer una distinción entre los términos “armas” y “sistemas de armas”.<sup>70</sup> El primero se refiere a “todas las armas, municiones, material,

---

<sup>67</sup> Comparar CICR, Commentary on the 1st GC, *supra* nota 63, pp.155-156, con Kalshoven, Frits, “The Undertaking to Respect and Ensure Respect in All Circumstances: From Tiny seed to Ripening Fruit”, *Yearbook of International Humanitarian Law*, vol. 2, núm. 3, 1999.

<sup>68</sup> Sobre la combinación entre armas y “medios y métodos de guerra”, al menos en el contexto del artículo 36 del Protocolo adicional I a los Convenios de Ginebra de 1949 (PAI) sobre el examen jurídico de armas, véase Liu, Hin-Yan, “Categorization and Legality of Autonomous and Remote Weapons Systems”, *International Review of the Red Cross*, vol. 94, 2012, pp. 627, 636 (en adelante: Liu, “Categorization and Legality of AWS”).

<sup>69</sup> Shilo, *When Turing Met Grotius*, *supra* nota 25, pp. 8, 20.

<sup>70</sup> Liu, “Categorization and Legality of AWS”, *supra* nota 68, p. 635.

instrumentos, mecanismos o dispositivos que tengan el efecto previsto de herir, dañar, destruir o incapacitar personal o propiedades”, mientras que el segundo se concibe más ampliamente para incluir “el arma en sí misma y aquellos componentes necesarios para su funcionamiento, incluidas tecnologías nuevas, avanzadas o emergentes”.<sup>71</sup>

No obstante, para algunos académicos, “la capacidad de toma de decisiones autónoma empuja a estos sistemas tecnológicamente avanzados al límite de la noción de combatiente”. El manual militar alemán establece que “los combatientes son personas que pueden participar directamente en las hostilidades, *i.e.*, participar en una función indispensable del uso de un arma o un sistema de armas”. Este texto plantea aún otra pregunta sobre la manera y las implicaciones de diferenciar un “arma” de un “sistema de armas”.<sup>72</sup>

En esta línea de ideas, no se puede argumentar que los SAA son meras armas (objetos) e intentar aplicarles los estándares de proporcionalidad y distinción, ya que éstas sirven para examinar el raciocinio de las acciones llevadas a cabo por combatientes y comandantes (humanos) en el espacio de batalla.

A menos que se acepte que estas entidades están en la misma categoría que los agentes morales o los humanos, estos principios no son aplicables a los SAA, ya que no son aplicables a otras armas, medios o métodos de guerra.

En esencia, el corazón del debate radica en el hecho de que, por un lado, su *propósito* es participar en el combate y matar, y como tal, es congruente con la definición de arma, pero, por otro lado, su *naturaleza*, como un tomador de decisiones artificial que puede o no estar equipado con armas convencionales, no es congruente con las normas legales que

---

<sup>71</sup> *Ibidem*, pp. 627, 635 y 636.

<sup>72</sup> *Idem*.

conocemos hoy en día. Por lo tanto, efectivamente no existe un umbral para su clasificación y, por lo tanto, no está claro cuál sería la norma aplicable para ellos.

En mi opinión, la aproximación más precisa de su naturaleza *sui generis* es que deben seguir considerándose como un arma en el plano teórico, pero una especie de “humanoide” en la práctica.<sup>73</sup>

### *Legalidad de los sistemas de armas autónomas per se*

Suponiendo sin conceder, que fueran considerados como un arma, esto conduciría a un análisis de su legalidad de conformidad con los artículos 35 y 36 del Protocolo Adicional I a los Convenios de Ginebra (PAI).<sup>74</sup>

El artículo 35 establece las limitaciones básicas sobre la elección y el empleo de armas, y los medios y métodos de guerra, asimismo, el artículo 36 exige que los países realicen un examen jurídico de cualquier arma, medio o método nuevo que pretendan utilizar.

Respecto al primero, en términos generales, según el DIH, un arma o su uso pueden considerarse ilícitos en dos circunstancias.<sup>75</sup>

En primer lugar, el arma puede considerarse ilegal *per se*, es decir, en sí misma, ya sea porque el arma ha sido expresamente prohibida en el derecho internacional aplicable o

---

<sup>73</sup> Shilo, *When Turing Met Grotius*, *supra* nota 25, p. 18.

<sup>74</sup> Protocolo Adicional I a los Convenios de Ginebra de 1949, relativo a la Protección de Víctimas de los Conflictos Armados Internacionales, artículos 35 y 36, 1977 (en adelante: PAI); véase también su comentario en: Sandoz, Yves *et al.* (eds.), *The Additional Protocols of 8 June 1977 to the Geneva Conventions of 12 August 1949*, CICR, 1987 (en adelante: Sandoz *et al.*, *The AP of 1977*).

<sup>75</sup> Véase Boothby, William H., “Prohibited Weapons”, *Max Planck Encyclopedia of Public International Law*, 2015.

porque el arma no puede utilizarse de una manera compatible con el DIH. Hay que destacar que el debate sobre si los SAA (de cualquier manera en que se definan) deberían ser objeto de una prohibición preventiva sigue abierto y sin resolverse hasta el momento. Algunos defensores de esta prohibición preventiva han señalado el desarrollo del Protocolo sobre Láseres Cegadores (Protocolo IV de la CAC) como un precedente relevante que también podría ser aplicable a estas tecnologías.<sup>76</sup>

En la misma línea de pensamiento, un hecho histórico, que servirá como un precedente muy importante, es el Tratado sobre la Prohibición de Armas Nucleares (TPAN),<sup>77</sup> ya que el 24 de octubre de 2020 Honduras se convirtió en el país número 50 en ratificar este instrumento, cumpliendo así las condiciones para su entrada en vigor el 22 de enero de 2021.<sup>78</sup> Las discusiones sobre la aprobación y la oposición de los Estados a este Tratado se asemejan y se pueden relacionar con aquéllas sobre el trato de los SAA. Los efectos en cadena de la entrada en vigor del TPAN aún están por verse en los próximos años; sin embargo, definitivamente es un paso crucial y muy esperado en el camino hacia el desarme nuclear, junto con todas las demás armas de destrucción masiva.

---

<sup>76</sup> Para una visión ampliada de esta línea de razonamiento, véanse los principios fundamentales invocados por el preámbulo de la CAC, Ginebra, (1980); véase también el Protocolo IV de la CAC, "Protocolo sobre armas láser cegadoras", Doc. CCW/CONF.I/16, Parte I, 1995.

<sup>77</sup> Asamblea General de las Naciones Unidas, Convención sobre la Prohibición de la Utilización de Armas Nucleares, 11 de enero de 2006, A/RES/60/88 (en adelante: AGNU, Prohibición de la Utilización de Armas Nucleares).

<sup>78</sup> Naciones Unidas, UN Secretary-General's Spokesman - on the Occasion of the 50th Ratification of the Treaty on the Prohibition of Nuclear Weapons, 24 de octubre de 2020 (en adelante: ONU, UN SG Spokesman), disponible en <https://www.un.org/sg/en/content/sg/statement/2020-10-24/un-secretary-generals-spokesman-the-occasion-of-the-50th-ratification-of-the-treaty-the-prohibition-of-nuclear-weapons>, ingresado el 25 de octubre de 2020.

## Legalidad del uso de los sistemas de armas autónomas

La segunda consideración para determinar si un arma es ilegal se basa en un uso en particular; en otras palabras, sólo aquel uso prohibido, mas no el arma *per se*, sería ilegal.<sup>79</sup> Esto significa que se puede prohibir el uso de un arma determinada en la medida en que sus efectos no puedan controlarse y, por tanto, sean indiscriminados.<sup>80</sup>

La Corte Internacional de Justicia (CIJ) sentó un precedente notorio en su opinión consultiva sobre armas nucleares, ya que cambió el paradigma de evaluar las armas por su legalidad *per se*, al evaluarlas por la legalidad de su *uso*. En esta opinión consultiva no vinculante, la CIJ concluyó que las armas nucleares no son ilegales *per se* debido a que existe la posibilidad, aunque remota, de un uso legal de las mismas, aunque remoto, en situaciones extremas que amenacen la existencia del Estado.<sup>81</sup> Por lo tanto, si la legalidad de los SAA no se aborda frontalmente en función de los méritos de su naturaleza *sui generis*, esta opinión podría servir como un argumento potencialmente peligroso para la legalidad *per se* de todas las armas nuevas, tales como éstas, permitiendo un examen posterior de su *uso*.

En cuanto al “uso legal”, a efecto de cumplir con el DIH hay tres pasos posteriores a la determinación de la legalidad del arma. La primera es determinar si el objetivo contra el que se usa es legal; en segundo lugar, incluso si es un objetivo legal, se debe realizar una evaluación de proporcionalidad, y en tercer lugar, se deben tomar medidas de precaución destinadas a reducir los daños civiles incidentales.

---

<sup>79</sup> De conformidad con el artículo 35(2) del PAI.

<sup>80</sup> Sandoz *et al.*, *The AP of 1977, supra* nota 74, p. 623. Los comentarios al artículo 51(4)(c) no son claros, parecen aplicar tanto a armas cuya naturaleza es indiscriminada y a armas cuyos efectos no se pueden limitar. Parece ser que se busca evitar el uso indiscriminado y/o efectos de las armas en todas las etapas de su despliegue.

<sup>81</sup> Véase CIJ, Legalidad de la Amenaza o el Empleo de las Armas Nucleares, Opinión Consultiva, julio de 1996, p. 226 (en adelante: CIJ, Legalidad de Armas Nucleares).

Para cumplir con estos requisitos, los manuales militares han establecido que, con el propósito de que los beligerantes empleen legalmente un SAA, deben cumplir con los tres principios más relevantes del DIH: distinción, proporcionalidad y precaución.<sup>82</sup> Los cuales tienen su fundamento en el derecho convencional y consuetudinario.

- Distinción: implica distinguir entre objetivos militares y civiles o bienes de carácter civil y, en caso de duda, presumir la condición de civil de ambos.
- Proporcionalidad: requiere evaluar si el daño incidental que probablemente se inflija a la población civil o a bienes civiles sería excesivo en relación con la ventaja militar concreta y directa que se prevé de ese ataque en particular.
- Precaución: obliga a tomar todas las precauciones posibles para evitar, y en cualquier caso minimizar, daños incidentales a civiles y daños a bienes de carácter civil; y cancelar o suspender el ataque si resulta evidente que el objetivo no es un objetivo militar, o si se puede esperar que el ataque resulte en daños colaterales excesivos.

A la luz de lo anterior, al examinar la aplicabilidad de estos principios, uno se da cuenta de que un *arma*, en el sentido tradicional, no puede ser considerada responsable por no respetarlos. Como se mencionó anteriormente, es una paradoja definir a los SAA como *armas*, y luego debatir si pueden o no ser combatientes obedientes. Sin embargo, como se opinó anteriormente, la autora acepta que esta entidad no es *de facto* sólo un objeto, sino un agente más sofisticado que debe evaluarse mediante una métrica diferente.

---

<sup>82</sup> Suiza, Towards a “Compliance-Based” Approach to LAWS. Informal Meeting of Experts on Lethal Autonomous Weapons Systems 1 (Informal working paper), Ginebra, 2016, p. 13 (en adelante: Suiza, Towards a “Compliance-Based” Approach), disponible en [http://www.unog.ch/80256EDD006B8954/\(httpAssets\)/D2D66A9C427958D6C1257F8700415473/\\$file/2016\\_LAWS+MX\\_CountryPaper+Switzerland.pdf](http://www.unog.ch/80256EDD006B8954/(httpAssets)/D2D66A9C427958D6C1257F8700415473/$file/2016_LAWS+MX_CountryPaper+Switzerland.pdf), ingresado el 5 de agosto de 2020.

Para ampliar el argumento anterior, a partir de una simple lectura queda claro que el principio de distinción, codificado en el artículo 48 del PAI y desarrollado en los artículos 51(2), (3), (4)(a) y 52 del mismo, no se aplican al arma en sí, sino a su operador.<sup>83</sup> No obstante, cuando se trata de SAA, no hay una separación clara entre el arma y el operador.<sup>84</sup>

Aparte del principio de distinción, el principio de proporcionalidad es aún más revelador y también se ha utilizado de manera controvertida para examinar la legalidad de los SAA. La pregunta es si los sistemas de armas autónomas son capaces o no de realizar cálculos de proporcionalidad.<sup>85</sup> De manera disonante, debe tenerse en cuenta que esta característica no es de ninguna manera parte de un examen jurídico de armas.<sup>86</sup>

---

<sup>83</sup> Véase Schmitt, Michael N. y Thurnher, Jeffrey S., "Out of the Loop: Autonomous Weapon Systems and the Law of Armed Conflict", *Harvard National Security Journal*, vol. 4, 2013, pp. 231, 250-252 (en adelante: Schmitt y Thurnher, "Out of the Loop").

<sup>84</sup> Véase Schmitt, "AWS and IHL", *supra* nota 29, pp. 2, 8-10. Schmitt responde a HRW y otros críticos, e insiste en la distinción que debe hacerse al declarar un arma como indiscriminada *per se* y usar un arma de forma indiscriminada.

<sup>85</sup> Schmitt y Thurnher, "Out of the Loop", *supra* nota 83, p. 254.

<sup>86</sup> PAI, artículo 36; menciona que los Estados Parte al Protocolo tienen la obligación de llevar a cabo un examen jurídico durante el desarrollo y adquisición de un arma nueva, y previo a su uso en un conflicto armado. Para otros Estados, los exámenes jurídicos son una medida de sentido común para asegurar que las fuerzas armadas puedan llevar a cabo las hostilidades de conformidad con sus obligaciones internacionales.

## Aplicabilidad de la Cláusula Martens

Ante la falta de consenso internacional sobre este tema, la autora opina que, mientras no se llegue a un acuerdo, todavía podemos confiar en la Cláusula Martens,<sup>87</sup> piedra angular del DIH, como una forma legítima de protección alternativa.<sup>88</sup>

La cláusula se introdujo por primera vez en el Preámbulo de la Convención de La Haya de 1899 como resultado de los debates en las Conferencias de Paz de La Haya.<sup>89</sup> Desde entonces ha sido objeto de múltiples interpretaciones, algunos han sostenido que, especialmente los términos “leyes de la humanidad” y las “exigencias de la conciencia pública”,

<sup>87</sup> La cláusula fue introducida por primera vez en el preámbulo de la Convención de la Haya (II) de 1899, relativa a las leyes y costumbres de la guerra terrestre, y dice: “Mientras que se forma un Código más completo de las leyes de la guerra, las Altas Partes Contratantes juzgan oportuno declarar que, en los casos no comprendidos en las disposiciones reglamentarias adoptadas por ellas, las poblaciones y los beligerantes permanecen bajo la garantía y el régimen de los principios del Derecho de Gentes preconizados por los usos establecidos entre las naciones civilizadas, por las leyes de la humanidad y por las exigencias de la conciencia pública”.

<sup>88</sup> Véase Tribunal Militar de Estados Unidos en Núremberg, Estados Unidos vs. Alfred Krupp *et al.*, “El Caso Krupp”, Sentencia (1948), Digesto Anual y Reportes de Casos de Derecho Internacional Público, 1948, p. 622; como lo cita el juez Shahabuddeen en su opinión disidente, en la Opinión Consultiva sobre la legalidad de las armas nucleares: “El preámbulo [de la Convención de la Haya no. IV de 1907] es mucho más que una declaración piadosa. Es una cláusula general, que convierte los usos establecidos entre las naciones civilizadas, las leyes de la humanidad y las exigencias de la conciencia pública en el criterio legal que se aplicará siempre y cuando las disposiciones específicas de la Convención y las Regulaciones anexadas a la misma no cubran determinados casos que ocurren en la guerra, o concomitantes a ésta”.

<sup>89</sup> En estos debates, las delegaciones debían encontrar una manera de darle fuerza legalmente vinculante a las reglas contenidas en la Declaración de Bruselas de 1874 relativas al estatus de civiles que se levantan en armas en contra de una fuerza ocupante, en particular, los derechos del poder ocupante y el derecho de los pueblos a la resistencia. Para evitar el mismo fracaso, la delegación de Bélgica sugirió que se mantuviera sin regulación; no obstante, el delegado de Rusia Fyodor Fyodorovich Martens dijo que la intención de la Declaración no era abolir el derecho de los pueblos a defender sus países, sino darles el derecho a hacerlo bajo ciertas condiciones y con más garantías de las que tenían antes, y procedió a leer una declaración que se insertó en el *procès-verbal*; véase CICR, *Commentary on the 1st GC*, *supra* nota 63, p. 3286.

tienen un valor normativo autónomo bajo el derecho internacional, ya que el primer término se ha asociado con la noción de “consideraciones elementales de humanidad”,<sup>90</sup> mientras que el segundo ha sido identificado como la motivación de los Estados, organizaciones o individuos que llevaron a la adopción de tratados de DIH.<sup>91</sup>

Otros argumentan que opera dentro del alcance del artículo 38 del Estatuto de la Corte Internacional de Justicia, sugiriendo que la cláusula podría acelerar la creación del DIH consuetudinario, reduciendo la necesidad de la práctica de los Estados cuando una norma consuetudinaria potencial está respaldada por las “leyes de la humanidad” o las “exigencias de la conciencia pública”, como expresiones imperativas de *opinio iuris*.<sup>92</sup> Asimismo, se ha propuesto que sirve como guía en la interpretación del DIH, como una aclaración de los “principios generales del derecho reconocidos por las naciones civilizadas” y como un recordatorio de la validez continua del derecho internacional consuetudinario.<sup>93</sup>

De hecho, la Comisión de Derecho Internacional ha manifestado que la cláusula “prevé que incluso en casos no cubiertos por acuerdos internacionales específicos, los civiles y combatientes permanecen bajo la protección y autoridad de los principios del derecho internacional derivados de la costumbre establecida, de los principios de humanidad y de las exigencias de la conciencia pública”.<sup>94</sup>

<sup>90</sup> *Ibidem*, p. 3291.

<sup>91</sup> *Idem*.

<sup>92</sup> Véase, por ejemplo, Salas Extraordinarias de las Cortes de Camboya, Caso núm. 001/18-07-2007-ECCCIOCIJ (PTC 02), *Amicus Curiae* Brief of Professor Antonio Cassese and Members of the Journal of International Justice on Joint Criminal Enterprise Doctrine, 27 de octubre de 2008, p. 35.

<sup>93</sup> CICR, Commentary on the 1st GC, *supra* nota 63, pp. 3295 y 3296.

<sup>94</sup> International Law Commission, Report of the International Law Commission on the Work of its Forty-sixth Session, 2 May- 22 July 1994, Official Records of the General Assembly of the United Nations, Doc. A/49/10, *Yearbook of the International Law Commission*, vol. II, núm. 2, 1994, p. 131.

En conclusión, la cláusula ha adquirido el rango de norma consuetudinaria y subraya que en los casos no cubiertos por los tratados de DIH, las personas afectadas por conflictos armados nunca se verán completamente privadas de protección. Al contrario, la conducta de los beligerantes permanece regulada, como mínimo, por los principios del derecho de gentes, las leyes de la humanidad y por los dictados de la conciencia pública.<sup>95</sup>

Como prueba de lo anterior, la cláusula ha sido incluida en varios tratados de DIH, a lo largo de los años, como en la Convención de La Haya de 1907,<sup>96</sup> los Convenios de Ginebra de 1949 (CG I-IV)<sup>97</sup> y sus protocolos adicionales (PA I-II),<sup>98</sup> la Convención sobre Prohibiciones o Restricciones en el Empleo de Ciertas Armas Convencionales que puedan Considerarse Excesivamente Nocivas o de Efectos Indiscriminados de 1980 (CAC),<sup>99</sup> la Convención sobre Municiones en Racimo de 2008,<sup>100</sup> y sus elementos pueden encontrarse en otros tratados,

---

<sup>95</sup> Véase CICR, "La Cláusula de Martens y el derecho de los conflictos armados", disponible en <https://www.icrc.org/es/doc/resources/documents/misc/5tdlcy.htm>, ingresado el 20 de noviembre de 2020.

<sup>96</sup> El preámbulo de la Convención de la Haya de 1907, incluyó la Cláusula Martens.

<sup>97</sup> Los cuatro convenios de Ginebra tienen un artículo idéntico en cuanto a la denuncia, el cual menciona que la denuncia: "No surtirá efecto alguno sobre las obligaciones que las Partes en conflicto hayan de cumplir en virtud de los principios del derecho de gentes, tal como resultan de los usos establecidos entre naciones civilizadas, de las leyes de humanidad y de las exigencias de la conciencia pública"; véase CG I, artículo 63; CG II, artículo 62; CG III, artículo 142; CG IV, artículo 158.

<sup>98</sup> Véase PAI, artículo 1(2); Protocolo Adicional II a los Convenios de Ginebra de 1949, relativo a la Protección de Víctimas en Conflictos Armados No Internacionales, Preámbulo, 1977 (en adelante: PAII).

<sup>99</sup> El preámbulo menciona que, en casos no previstos por la convención, la población civil y los combatientes permanecerán en todo momento bajo la protección de los principios del derecho internacional.

<sup>100</sup> El preámbulo reconoce la protección por los principios del derecho internacional en casos no previstos por la convención.

como la Convención sobre la Prohibición de Minas Antipersonales de 1997,<sup>101</sup> e incluso el Estatuto de Roma de la Corte Penal Internacional de 1998.<sup>102</sup>

Por lo tanto, la autora comparte la opinión del Gobierno suizo, en relación con los SAA, en el sentido de que la Cláusula Martens ofrece “una protección alternativa importante en la medida en que se debe referir a las ‘leyes de la humanidad y las exigencias de la conciencia pública’ en caso de que el DIH no sea lo suficientemente preciso o riguroso”.<sup>103</sup> En esta línea de pensamiento, se desprende que no puede decirse que lo que no está explícitamente prohibido está permitido si contraviene los principios enunciados en la cláusula, los cuales pueden entenderse como obligaciones positivas para casos de operaciones militares, donde pudieran existir consecuencias humanitarias insostenibles.<sup>104</sup>

Es digno de mención que la CIJ, en su Opinión Consultiva de 1996 sobre la Legalidad de la Amenaza o el Empleo de Armas Nucleares, tuvo la previsión de afirmar que la cláusula “demostró ser un medio eficaz de abordar la rápida evolución de la tecnología militar”,<sup>105</sup> y que el hecho de que algunas armas no hayan sido mencionadas en los protocolos adicionales

---

<sup>101</sup> El preámbulo menciona: “Poniendo de relieve el papel que desempeña la conciencia pública en el fomento de los principios humanitarios [...]”.

<sup>102</sup> El preámbulo se refiere a la “conciencia de la humanidad”, pues menciona: “Teniendo presente que, en este siglo, millones de niños, mujeres y hombres han sido víctimas de atrocidades que desafían la imaginación y conmueven profundamente la conciencia de la humanidad”.

<sup>103</sup> Suiza, Towards a “Compliance-Based” Approach, *supra* nota 82, p. 4, en donde se cita el preámbulo de la CAC y el artículo 1(2) del PAI.

<sup>104</sup> *Ibidem*, p. 3, citando el artículo 57(2)(a) del PAI y los cuatro convenios de Ginebra en sus artículos 49, 50, 129 y 146, respectivamente.

<sup>105</sup> CIJ, Legalidad de Armas Nucleares, *supra* nota 81, p. 78.

a los Convenios de Ginebra de 1977 (u otros tratados) no permite sacar conclusiones legales relativas a su uso.<sup>106</sup>

Además, la Corte consideró que el no aplicar las normas básicas del derecho humanitario, como esta cláusula, sería incompatible con los principios del DIH, que son aplicables a todos los tipos de armas. Por lo tanto, se rechazó la tesis de que los principios del DIH no se aplican a las armas más nuevas.<sup>107</sup>

Más aún, se afirmó que la continuidad de la existencia y la aplicabilidad de la Cláusula Martens es indudablemente una afirmación de que los principios y normas del derecho humanitario son aplicables para las armas nucleares,<sup>108</sup> y la autora sostiene firmemente que el mismo razonamiento debe replicarse respecto a los SAA.

El juez Shahabuddeen argumentó que su inclusión en la Convención de La Haya de 1899 tenía la intención de llenar los vacíos que dejaba el derecho convencional.<sup>109</sup> Afirmó que tiene autoridad propia, autosuficiente y concluyente para tratar los principios de humanidad como principios del derecho internacional, que pueden por sí mismos ejercer fuerza jurídica para regir la conducta militar en los casos en que no exista una norma relevante en el derecho de los tratados, dejando el contenido preciso de la norma o estándar insinuado a una determinación a la luz de condiciones cambiantes.<sup>110</sup> Así, cuando se trata del uso de un arma en

---

<sup>106</sup> *Ibidem*, p. 84.

<sup>107</sup> *Ibidem*, pp. 85 y 86.

<sup>108</sup> *Ibidem*, p. 87.

<sup>109</sup> "Opinión disidente del juez Shahabuddeen", en CIJ, Legalidad de Armas Nucleares, *supra* nota 81, p. 406. Él considera que la palabra "permanecen" en la cláusula (en inglés, *remain*) indica que, para el tiempo de su adopción, ya había ciertos principios en el derecho internacional que proveían protección a los beligerantes y los civiles, por lo que la cláusula no se puede limitar a principios que se crearán en el futuro.

<sup>110</sup> *Idem*. En apoyo a esto, el juez Shahabuddeen cita a mr. Sean McBride: "las declaraciones en las Convenciones de la Haya [...] por virtud de la Cláusula Martens, importaron al derecho humanitario principios que fueron mucho

particular, la opinión de los Estados sólo es relevante por su valor al indicar el estado de la conciencia pública, no en términos de *opinio iuris* sobre la legalidad de dicha arma.<sup>111</sup>

Por otro lado, en el *Caso de las Inmunidades Jurisdiccionales del Estado*, el juez Cançado Trindade emitió una opinión disidente en contra de una orden, analizando más profundamente la cláusula, afirmando que tiene más de un siglo de validez continua, y sigue siendo una advertencia contra el supuesto de que se permitiría todo lo que no esté expresamente prohibido por los convenios sobre DIH, además de que los principios que consagra seguirían siendo aplicables, independientemente de la aparición de nuevas situaciones.<sup>112</sup>

Consideró que al entrelazar los principios de humanidad y las exigencias de la conciencia pública, la cláusula establece una “interdependencia orgánica” de protección legal legítima en beneficio de la humanidad.<sup>113</sup> Además, la autora comparte enfáticamente la opinión del juez de que los “principios de humanidad” y las “exigencias de la conciencia pública” pertenecen al dominio de *ius cogens*, y considera la cláusula como “una expresión de la *raison d’humanité* que impone límites a la *raison d’État*”.<sup>114</sup>

Más aún, en el *Caso sobre las Obligaciones Relativas a las Negociaciones sobre el Cese de la Carrera de Armas Nucleares y el Desarme Nuclear*, el juez Cançado Trindade emitió

---

más allá de lo que se escribió en la convención; dotándolas de una dimensión dinámica que no estaba limitada por el tiempo”.

<sup>111</sup> *Ibidem*, p. 410.

<sup>112</sup> CIJ, *Alemania v. Italia: con Intervención de Grecia. Inmunidades Jurisdiccionales del Estado*, Orden de 6 de julio de 2010, Opinión disidente del juez Cançado Trindade, 2010, pp. 137-139 (en adelante: CIJ, *Inmunidades del Estado*); véase también Corte Interamericana de Derechos Humanos, *Case of Barrios Altos v. Perú*, Sentencia del 14 de marzo de 2001, Fondo, Opinión concurrente del juez Cançado Trindade, 2001, pp. 22-24.

<sup>113</sup> CIJ, *Inmunidades del Estado*, *supra* nota 112, p. 139.

<sup>114</sup> *Idem*. Véase también *infra* nota 115; el juez Cançado repite este argumento en su opinión disidente en el caso entre las Islas Marshall y Pakistán.

otra opinión disidente en la que consideró que los “principios de humanidad” y las “exigencias de la conciencia pública”, evocados por la cláusula, permean no sólo el derecho de los conflictos armados, sino todo el derecho internacional.<sup>115</sup> Para él, la Cláusula Martens salvaguarda la integridad del derecho invocando los principios del derecho de gentes, las “leyes de la humanidad” y las “exigencias de la conciencia pública”. Por lo tanto, la ausencia de una norma convencional no es concluyente, y la ausencia de una disposición convencional que prohíba expresamente las armas nucleares no significa que éstas sean legales o legítimas.<sup>116</sup>

En consecuencia, tal y como se mencionó anteriormente,<sup>117</sup> el TPAN,<sup>118</sup> el cual entró en vigor el 22 de enero de 2021, ha confirmado decisivamente estas declaraciones jurisdiccionales en el ámbito convencional.<sup>119</sup>

En conclusión, todo lo anterior sirve para argumentar que a pesar de que en la actualidad existe una aparente laguna jurídica en forma de *lex specialis* respecto a la legalidad de los SAA *per se* y de su uso, éstos nunca podrían ser empleados de forma inconsistente con las exigencias de la conciencia pública o consideraciones básicas de humanidad. En la opinión de la autora, deshumanizar la conducta de un conflicto armado y desplegar armas que pueden funcionar de manera incontrolable o impredecible va en contra de estos mismos principios y, por lo tanto, en cualquier caso, es ilegal *ab initio*.

---

<sup>115</sup> CIJ, *Islas Marshall v. Pakistán, Obligaciones Respecto de las Negociaciones sobre la Cesación de la Carrera de Armamentos Nucleares y el Desarme Nuclear*, Sentencia sobre jurisdicción y admisibilidad, Opinión disidente del juez Cançado Trindade, 2016, p. 195.

<sup>116</sup> *Ibidem*, p. 196.

<sup>117</sup> Como se explicó *supra* en el epígrafe “Consideraciones del derecho internacional humanitario”.

<sup>118</sup> AGNU, Prohibición de la Utilización de Armas Nucleares, *supra* nota 77.

<sup>119</sup> ONU, UN SG Spokesman, *supra* nota 78.

## *Inquietudes de derecho internacional humanitario en concreto*

Como se ha dicho antes, para el CICR los SAA son una preocupación inmediata desde una perspectiva humanitaria, jurídica y ética, dado el riesgo de pérdida del control humano sobre las armas y el uso de la fuerza.<sup>120</sup> Esta pérdida de control plantea riesgos para la población civil, debido a consecuencias impredecibles;<sup>121</sup> cuestiones legales, porque los combatientes deben emitir juicios específicos del contexto al llevar a cabo ataques según el DIH, y cuestiones éticas,<sup>122</sup> porque la agencia humana en las decisiones de uso de la fuerza es necesaria para defender la responsabilidad moral y la dignidad humana. Por estas razones, el CICR ha instado a los Estados a identificar elementos prácticos de control humano determinante como base para los límites acordados internacionalmente a la autonomía en los sistemas de armas, centrándose en lo siguiente:<sup>123</sup>

- ¿Qué nivel de supervisión humana, intervención y capacidad de desactivación se requiere durante la operación de un arma que selecciona y ataca objetivos sin intervención humana?

---

<sup>120</sup> CICR, ICRC Statements to the CCW Group of Governmental Experts on Lethal Autonomous Weapons Systems, Ginebra, marzo de 2019, pp. 25-29, disponible en [https://www.unog.ch/80256ee600585943.nsf/\(httpPages\)/5c00ff8e35b6466dc125839b003b62a1?OpenDocument&ExpandSection=7#Section7](https://www.unog.ch/80256ee600585943.nsf/(httpPages)/5c00ff8e35b6466dc125839b003b62a1?OpenDocument&ExpandSection=7#Section7), ingresado el 5 de agosto de 2020.

<sup>121</sup> Davison, “AWS under IHL”, *supra* nota 44, pp. 13-15.

<sup>122</sup> CICR, Ethics and Autonomous Weapon Systems: An Ethical Basis for Human Control?, Report of an Expert Meeting, 3 de abril de 2018 (en adelante: CICR, Ethics and AWS), disponible en <https://www.CICR.org/en/document/ethics-and-autonomous-weapon-systems-ethical-basis-human-control>, ingresado el 5 de agosto de 2020.

<sup>123</sup> CICR, The Element of Human Control, Working Paper, CCW Meeting of High Contracting Parties, CCW/MSP/2018/WP.3, 20 de noviembre de 2018, disponible en [https://www.unog.ch/80256EDD006B8954/\(httpAssets\)/810B2543E1B5283BC125834A005EF8E3/\\$file/CCW\\_MSP\\_2018\\_WP3.pdf](https://www.unog.ch/80256EDD006B8954/(httpAssets)/810B2543E1B5283BC125834A005EF8E3/$file/CCW_MSP_2018_WP3.pdf), ingresado el 5 de agosto de 2020.

- ¿Qué nivel de previsibilidad, en términos de su funcionamiento y las consecuencias de su uso, y confiabilidad, en términos de probabilidad de falla o mal funcionamiento, se requiere?
- ¿Qué otras limitaciones operativas se requieren para el arma, en particular en las tareas, objetivos (e.g., material o personal), entorno de uso (e.g., áreas des-pobladas o pobladas), duración de la operación autónoma (i.e., limitaciones de tiempo) y alcance del movimiento (i.e., limitaciones de espacio)?

Además de estas preocupaciones, los expertos en el tema han señalado que actualmente no está claro si los SAA serán capaces de formular e implementar las siguientes decisiones de evaluación y juicios de valor apegados al DIH:<sup>124</sup>

- La presunción de la condición de civil en caso de duda;
- La evaluación del exceso del daño incidental esperado en relación con la ventaja militar anticipada;
- La traición a la confianza en el DIH en relación con la prohibición de la perfidia;<sup>125</sup>
- La prohibición de la destrucción de la propiedad civil excepto cuando lo exijan “imperativamente” las necesidades de la guerra; y,
- ¿Cómo se verían los posibles estándares operativos que incluyan un “control humano determinante” (incluso en las operaciones de selección de objetivos), “control estatal determinante” y “niveles apropiados de juicio humano”?

---

<sup>124</sup> Lewis *et al.*, *War-Algorithm*, *supra* nota 32, p. 103.

<sup>125</sup> Voz: “Perfidia”, que significa: “Deslealtad, traición o quebrantamiento de la fe debida”, *DLE*, RAE, disponible en <https://dle.rae.es/perfidia>, ingresado el 8 de agosto de 2020; véase también Henckaerts, J. M. y Doswald-Beck, Louise, “Customary International Humanitarian Law”, vol. I: “Rules”, *International Review of the Red Cross*, 3a. ed., Cambridge, 2009, p. 221 (en adelante: Henckaerts y Doswald, “Customary IHL”); la Regla 65 menciona que el matar, herir o capturar a un adversario, valiéndose de medios pérfidos, está prohibido.

## Observaciones finales

Es innegable que nos encontramos en una encrucijada en nuestra sociedad, donde nos enfrentamos a definir si nuestra especie humana es única en su esencia como agentes racionales y morales, o si podemos replicar esa condición nosotros mismos en las máquinas que creamos, aceptando que somos no tan únicos después de todo.

Aparte de este tipo de dilemas existenciales, lo que está claro es que la tecnología avanza a un ritmo exponencial, en particular los sistemas capaces de “autoaprendizaje”, cuyas “elecciones” pueden ser difíciles de prever o desentrañar para los humanos, y sus “decisiones” se consideran que “reemplazan” el criterio humano; además, es especialmente problemático cuando éstas operan en contextos de conflicto armado. Por lo tanto, estas tecnologías deben ser cuidadosamente reguladas para prevenir el sufrimiento humano y garantizar la rendición de cuentas por parte de los humanos.

Debido a lo anterior, es imperativo reconocer la singularidad de los SAA, tanto en el sentido literal como teórico de la palabra, así como la falta de normatividad existente que les pueda ser aplicable.

Por lo tanto, como se indicó en la introducción, el propósito de este artículo es incentivar a todos los actores relevantes a sumarse a las discusiones necesarias para crear un régimen legal que regule el desarrollo, uso y rendición de cuentas de una nueva categoría jurídica de entidades que puedan abarcar adecuadamente las características *sui generis* de los SAA.

## El dilema de la responsabilidad

### *Planteamiento del problema: el reconocimiento de las capacidades y riesgos de la inteligencia artificial*

En el epígrafe “Contextualización”, se han abordado algunas de las principales características de las tecnologías de IA y sobre su aplicación en los SAA.<sup>126</sup>

Sin embargo, para los propósitos de este apartado, es importante profundizar en ello analizando algunas de sus características más relevantes.<sup>127</sup>

Primero, estas tecnologías tienen el potencial de actuar de manera impredecible.<sup>128</sup> Se ha vuelto cada vez más normal que la IA dependa del aprendizaje automático, que en esencia es una forma de *software* que, posteriormente a su configuración inicial, continúa

---

<sup>126</sup> Como se mencionó *supra* en el apartado “Sistemas de armas autónomas”.

<sup>127</sup> Abbott, Ryan y Sarch, Alex, “Punishing Artificial Intelligence: Legal Fiction or Science Fiction”, *UC Davis Law Review*, vol. 53, 2019, pp. 330-332, (en adelante: Abbott y Sarch, “Punishing AI”).

<sup>128</sup> Véase Yasserí, Taha, “Never Mind Killer Robots — Even the Good Ones Are Scarily Unpredictable”, *Phys.Org*, agosto de 2017, disponible en <https://phys.org/news/2017-08-mind-killer-robots-good-scarily.html>, ingresado el 11 de febrero de 2021; Shane, Janelle, “Why Did the Neural Network Cross the Road?”, *Ai Weirdness*, 2018, disponible en <https://aiweirdness.com/post/174691534037/why-did-the-neural-network-cross-the-road>, ingresado el 11 de febrero de 2021.

desarrollándose en respuesta a los datos que adquiere sin ninguna programación explícita adicional.<sup>129</sup> Por lo tanto, la IA puede actuar de formas que sus programadores originales tal vez no pretendían o incluso ni habían previsto.<sup>130</sup> El emprendedor Elon Musk también ha expresado sus preocupaciones instando al establecimiento de una autoridad reguladora que supervise el desarrollo de la IA, advirtiendo que ésta podría ser la causa más probable de una Tercera Guerra Mundial.<sup>131</sup>

En segundo lugar, la IA tiene el potencial de actuar de forma inexplicable. Puede ser posible observar lo que ha hecho una IA, pero no cómo o por qué tomó ese curso de acción.<sup>132</sup> Esto es lo que comúnmente se conoce como la “caja negra”,<sup>133</sup> como ya se mencionó anteriormente.<sup>134</sup> Este es un resultado natural derivado del aprendizaje automático, dado que la IA habrá estado expuesta a miles de millones de datos,<sup>135</sup> por lo tanto es imposible rastrear qué datos en específico desencadenaron un resultado en particular. En una conferencia tecno-

---

<sup>129</sup> Véase Castelvechi, Davide, “Can We Open the Black Box of AI?”, *Nature*, 5 de octubre de 2016 (en adelante: Castelvechi, “Can We Open the Black Box”), disponible en <https://www.nature.com/news/can-we-open-the-black-box-of-ai-1.20731>, ingresado el 11 de febrero de 2021.

<sup>130</sup> Ha habido un enfoque reciente en las decisiones sesgadas por los algoritmos de aprendizaje automático, a veces debido al sesgo implícito de un programador, a veces debido a datos de entrenamiento sesgados; véase, e.g., DeBrusk, Chris, “The Risk of Machine-Learning Bias (and How to Prevent It)”, *MIT Sloan Management Review*, 2018, disponible en <https://sloanreview.mit.edu/article/the-risk-of-machine-learning-bias-and-how-to-prevent-it/>, ingresado el 11 de noviembre de 2020.

<sup>131</sup> Véase Osborne, S., “Elon Musk Calls for Urgent Laws on Robot as They Will Soon Be Risk to Public”, *Express*, 2017, disponible en <https://www.express.co.uk/news/science/885344/elon-musk-artificial-intelligence-robotics-regulation>, ingresado el 19 de octubre de 2020.

<sup>132</sup> Véase, e.g., Castelvechi, “Can We Open the Black Box of AI?”, *supra* nota 129.

<sup>133</sup> *Idem*.

<sup>134</sup> Como se vio *supra* en el epígrafe “Contextualización”.

<sup>135</sup> Castelvechi, “Can We Open the Black Box of AI?”, *supra* nota 129.

lógica llevada a cabo en Lisboa en 2017, el físico Stephen Hawking advirtió sobre los riesgos de la IA al afirmar que ésta puede ser el peor caso de inteligencia humana.<sup>136</sup>

En tercer lugar, la IA puede actuar de forma autónoma. Esto se debe a las dos características mencionadas anteriormente. En el contexto de este estudio, esto implicaría que una IA puede causar daño sin ser operada directamente por un individuo. La IA puede recibir información sensorial, establecer objetivos, evaluar los resultados en función de criterios, tomar decisiones y ajustar su comportamiento para aumentar la probabilidad de éxito —todo sin estar dirigida por comandos humanos—. <sup>137</sup> Incluso puede darse el caso de que el programador que pone en marcha una IA no sea capaz de recuperar el control de la misma —y posiblemente hasta sea intencionalmente así—. <sup>138</sup> Es un hecho ampliamente conocido que algunos de los principales ejércitos del mundo, incluida la Fuerza Aérea de Estados Unidos, ya utilizan tecnologías semi y totalmente autónomas, y que han invertido recursos significantes para continuar aumentando la autonomía de estos sistemas.<sup>139</sup>

---

<sup>136</sup> Véase Murphy, M., “Stephen Hawking: AI Could Be Best — or Worst — Thing in Human History”, *MarketWatch*, New York City, 2017, disponible en <https://www.marketwatch.com/story/stephen-hawking-ai-could-be-best-or-worst-thing-in-human-history-2017-11-06>, ingresado el 29 de octubre de 2020.

<sup>137</sup> Abbott y Sarch, “Punishing AI”, *supra* nota 127, p. 333.

<sup>138</sup> *Ibidem*, p. 331. Mencionan que la “DAO” fue el intento más notorio de crear una organización autónoma descentralizada, cuyo propósito era desplegar una entidad que no pudiera ser controlada por sus creadores, actuando sin dirección adicional; debía operar a través de pequeños contratos, o reglas preprogramadas, de conformidad con códigos disponibles públicamente e inalterables; sin embargo, poco después de su lanzamiento falló por errores de programación y por los *hackers*; véase Falkon, Samuel, “The Story of the DAO — Its History and Consequences”, *The Startup*, 2017, disponible en <https://medium.com/swlh/the-story-of-the-dao-its-history-and-consequences-71e6a8a551ee>, ingresado el 12 de diciembre de 2020.

<sup>139</sup> Palmer, A., “Autonomous UAS: a Partial Solution to America’s Future Airpower Needs”, A Research Paper in Partial Fulfillment of the Graduation Requirements, Air University, Alabama, 2010.

En cuarto lugar, la interacción entre la IA acotada, que está diseñada para realizar tareas “limitadas” o “específicas”,<sup>140</sup> y aquellas que podrían desarrollarse como las IA “generales”, capaces de realizar cualquier tarea a la par de las habilidades humanas,<sup>141</sup> o probablemente incluso mejor, como concluyó un estudio realizado por investigadores de las universidades de Oxford y Yale.<sup>142</sup> Sin embargo, no está claro cuándo y cómo esta tecnología se desarrollará y se desplegará por completo.<sup>143</sup>

Por supuesto, en un momento dado, cualquier máquina convencional también podría actuar de manera impredecible, inexplicable o autónoma. Sin embargo, es mucho más probable que la IA exhiba estas características y en mayor medida.<sup>144</sup> También se puede estar de acuerdo en que las implicaciones interdisciplinarias de estas tecnologías nos hacen cada vez menos aptos para comprender, y mucho menos para regular, su comportamiento, lo que a su vez exacerba su imprevisibilidad.

De hecho, las máquinas han causado daños desde la Antigüedad, y los robots han causado muertes desde al menos la década de 1970.<sup>145</sup> Sin embargo, con excepción de los casos en que las máquinas se han utilizado intencionalmente para causar daño, la mayoría

---

<sup>140</sup> Abbott y Sarch, “Punishing AI”, *supra* nota 127, pp. 331 y 332; véase Abbott, Ryan, “Everything is Obvious”, *UCLA L. Rev.*, vol. 66, 2019, p. 25.

<sup>141</sup> *Idem*.

<sup>142</sup> Grace, K. *et al.*, “Viewpoint: When Will AI Exceed Human Performance? Evidence from AI Experts”, *Journal of Artificial Intelligence Research*, vol. 62, 2018, pp. 729-754.

<sup>143</sup> Véase, de manera general, Müller, Vincent C. y Bostrom, Nick, “Future Progress in Artificial Intelligence: A Survey of Expert Opinion”, en Müller, Vincent C. (ed.), *Fundamental Issues of Artificial Intelligence*, 2016, p. 555. Describe una encuesta en donde los expertos piensan que la IA no será realidad sino hasta algunas décadas después.

<sup>144</sup> Abbott y Sarch, “Punishing AI”, *supra* nota 127, p. 332.

<sup>145</sup> *Idem*; véase Abbott, Ryan, “The Reasonable Computer: Disrupting the Paradigm of Tort Liability”, *The George Washington Law Review*, vol. 86, 2018, p. 8.

de estos eventos han sido calificados como accidentes, o como resultado de negligencia o imprudencia por parte del operador. No obstante, estos accidentes implican la aplicación de leyes de carácter penal,<sup>146</sup> las cuales no pueden aplicarse contra las propias máquinas que causaron daño.

El fondo del asunto es que la IA se diferencia de las máquinas convencionales en las formas esenciales antes mencionadas, lo cual ha desconcertado tanto a los académicos como a los profesionales del derecho con respecto a la aplicación de los conceptos tradicionales del derecho penal tal y como los conocemos.

En este sentido, la clave es considerar el elemento de reducibilidad, porque si una IA se involucra en una conducta que sería criminal si la cometiera una persona, y el acto es reducible, entonces normalmente habrá un ser humano que podría ser considerado penalmente responsable por ella. Por el contrario, si una conducta de AI no es efectivamente reducible, es posible que no haya otra parte involucrada a la que pueda sancionarse, en cuyo caso la actividad criminal ocurriría con impunidad.<sup>147</sup> Tal como lo afirmó el relator especial sobre Ejecuciones Extrajudiciales, Sumarias o Arbitrarias, si la naturaleza de un arma hace imposible la responsabilidad por sus consecuencias, su uso debe considerarse no ético e ilegal.<sup>148</sup>

Por lo tanto, es imperativo que todas las conductas de la IA puedan ser reducibles. Como actualmente se sabe, las tecnologías de IA, incluso cuando se comportan de forma au-

---

<sup>146</sup> Véase United States Department of Justice, “BP Exploration and Production Inc. Agrees to Plead Guilty to Felony Manslaughter, Environmental Crimes and Obstruction of Congress Surrounding Deepwater Horizon Incident”, U.S. Dept Justice, 2012, disponible en <https://www.justice.gov/opa/pr/bp-exploration-and-production-inc-agrees-plead-guilty-felony-manslaughter-environmental>, ingresado el 13 de enero de 2021.

<sup>147</sup> Abbott y Sarch, “Punishing AI”, *supra* nota 127, pp. 369-375.

<sup>148</sup> Heyns, Christof, Informe del Relator Especial sobre las Ejecuciones Extrajudiciales, Sumarias o Arbitrarias, U.N. Doc. A/HRC/23/47, 2013, p. 80 (en adelante: Heyns, Informe del Relator Especial).

tónoma, en la medida en que una persona las utiliza como herramienta para cometer un delito, y éstas funcionan de manera previsible, el delito se puede reducir a la conducta del individuo que causó el daño.<sup>149</sup> Incluso cuando la IA causa un daño imprevisible, aún puede ser reducible si, por ejemplo, un individuo crea una IA para robar información financiera, pero por un error de programación hace que la IA apague una red eléctrica alterando o dañando infraestructuras críticas.<sup>150</sup> Este es un problema familiar en derecho penal, por lo que se han desarrollado herramientas doctrinales mediante las cuales es posible imponer responsabilidad.<sup>151</sup>

Sin embargo, con las IA que se conocen actualmente, cada vez es más difícil reducir un delito causado por una IA a un individuo, debido a la autonomía, complejidad y falta de explicación del actuar de la IA. Además, debido a que una gran cantidad de personas pueden contribuir al desarrollo de una IA durante un largo periodo de tiempo,<sup>152</sup> y también debido a las funciones de aprendizaje automático, puede ser difícil atribuir la responsabilidad a un individuo específico por un proceso en el que la IA ha recopilado información sobre cómo comportarse basándose en el acceso a miles de millones de datos de fuentes heterogéneas.<sup>153</sup>

---

<sup>149</sup> Abbott y Sarch, "Punishing AI", *supra* nota 127, p. 334.

<sup>150</sup> *Idem*.

<sup>151</sup> *Idem*.

<sup>152</sup> En el 2017, por ejemplo, más de 4500 empleados de Microsoft contribuyeron con *software* de libre acceso almacenado en GitHub, una plataforma de desarrollo para almacenar código de libre acceso; véase Asay, Matt, "Who Really Contributes to Open Source", *Infoworld*, 7 de febrero de 2018, disponible en <https://www.infoworld.com/article/3253948/who-really-contributes-to-open-source.html>, ingresado el 2 de febrero de 2021; véase Lardinois, Frederic y Lunden, Ingrid, "Microsoft Has Acquired GitHub for \$7.5B in Stock", *Techcrunch*, 4 de junio de 2018, disponible en <https://techcrunch.com/2018/06/04/microsoft-has-acquired-github-for-7-5b-in-micro-soft-stock/>, ingresado el 2 de febrero de 2021.

<sup>153</sup> Determann Lothar y Perens, Bruce, "Open Cars", *Berkeley Tech. L.J.*, vol. 32, 2017, pp. 915, 988.

En nuestra coexistencia con agentes inteligentes, el pronóstico de la combinación de estos elementos retrata un escenario poco alentador que nos lleva a cuestionar hasta qué punto los seres humanos aún podemos controlar estas tecnologías, especialmente cuando tienen impactos cinéticos en el mundo físico. Intuitivamente, esto nos confronta con las incómodas preguntas que han preocupado a académicos e investigadores.<sup>154</sup>

¿Qué sucede si estas tecnologías provocan daños? ¿Quién será responsable?  
¿Nos enfrentamos a una seria brecha de responsabilidad?

Naturalmente, esta cuestión se vuelve aún más urgente ante las tecnologías de IA aplicadas en SAA con las que se pone en riesgo la vida de la población civil. Aquellos que argumentan que efectivamente habrá una brecha de responsabilidad<sup>155</sup> en los casos de civiles que sean asesinados ilegalmente mediante el uso de un SAA, usan este argumento para justificar la prohibición o restricción de los mismos.<sup>156</sup>

Este desacuerdo se muestra de una manera muy sencilla entre Estados. Los últimos informes del Grupo de Expertos Gubernamentales sobre Tecnologías Emergentes en el Área

---

<sup>154</sup> Véase, por ejemplo: Willick, M., "Artificial Intelligence: Some Legal Approaches and Implications", *AI Magazine*, vol. 4, núm. 2, 1983, p. 5, disponible en <https://aaai.org/ojs/index.php/aimagazine/article/view/392>, ingresado el 18 de noviembre de 2020; Karnow, Curties E. A., "Liability for Distributed Artificial Intelligences", *Berkeley Technology Law Journal*, vol. 11, núm. 1, 1996, p. 188, disponible en <https://lawcat.berkeley.edu/record/1115611?ln=en>, ingresado el 18 de noviembre de 2020.

<sup>155</sup> Chengeta, Thompson, "Accountability Gap: Autonomous Weapon Systems and Modes of Responsibility in International Law", *Denver Journal of International Law and Policy*, vol. 45, núm. 1, 2020, pp. 2, 4 (en adelante: Chengeta, "Accountability Gap"); Sparrow, "Killer Robots", *supra* nota 45, p. 62; CICR, *Autonomous Weapon Systems: Implications of Increasing Autonomy in the Critical Functions of Weapons*, 2016, p. 44; Weizmann, Nathalie *et al.*, "Autonomous Weapon Systems Under International Law", Geneva Academy of International Humanitarian Law, Academy Briefing, núm. 8, 2014, p. 24 (en adelante: Weizmann *et al.*, "AWS under International Law").

<sup>156</sup> McDougall, Carrie, "Autonomous Weapon Systems and Accountability: Putting the Cart Before the Horse", *Melbourne Journal of International Law*, vol. 20, 2019, pp. 7, 13 (en adelante: McDougall, "AWS and Accountability").

de los Sistemas de Armas Autónomas Letales (GEG)<sup>157</sup> registran el hecho de que los Estados están bastante divididos sobre el desarrollo y uso futuro de los SAA. Por un lado, los Estados prohibicionistas sostienen que los SAA deberían prohibirse en su totalidad o al menos colocar las armas bajo una moratoria.<sup>158</sup> Otro bloque está pidiendo negociaciones sobre un tratado que regule los SAA, argumentando que su potencial debe estar constreñido primordialmente por el requisito de que garanticen el control humano determinante.<sup>159</sup> Algunos otros

<sup>157</sup> El GEG se estableció por los Estados Parte a la CAC, y es el principal foro de discusiones entre los Estados para el desarrollo de estas armas; véase GEG, Working Paper by the Bolivarian Republic of Venezuela on Behalf of the Non-Aligned Movement (NAM) and Other States Parties to the Convention on Certain Conventional Weapons (CCW), CCW/GGE.1/2020/WP.5, 2020; también, GEG, United Kingdom Expert Paper: The Human Role in Autonomous Warfare, CCW/GGE.1/WP.6, 2020.

<sup>158</sup> Heyns, Informe del Relator Especial, *supra* nota 148; Human Rights Watch (HRW) & International Human Rights Clinic, "Mind the Gap: The Lack of Accountability for Killer Robots", Human Rights Watch, 2015 (en adelante: HRW, "Mind the Gap"), disponible en <https://www.hrw.org/report/2015/04/09/mind-gap/lack-accountability-kill-er-robots#>; Stewart, Darren M., "New Technology and the Law of Armed Conflict", *International Law Studies*, vol. 87, 2011, pp. 291-294 (en adelante: Stewart, "New Technology and the Law of Armed Conflict"); O'Connell, Mary Ellen, "Banning Autonomous Killing: The Legal and Ethical Requirement That Humans Make Near-Time Lethal Decisions", en Evangelista, Matthew y Shue, Henry (eds.), *The American Way of Bombing: Changing Ethical and Legal Norms, from Flying Fortresses to Drones*, Cornell University Press, 2014, pp. 224, 236.

<sup>159</sup> Véase, e.g., Margulies, Peter, "Making Autonomous Weapons Accountable: Command Responsibility for Computer-Guided Lethal Force in Armed Conflicts", en Ohlin, Jens David (ed.), *Research Handbook on Remote Warfare*, Edward Elgar Press, 2016, p. 19 (en adelante: Margulies, "Making Autonomous Weapons Accountable"); Chengeta, "Accountability Gap", *supra* nota 155, pp. 2, 4; Malik, Swati, "Autonomous Weapon Systems: The Possibility and Probability of Accountability", *Wisconsin International Law Journal*, vol. 35, núm. 3, 2018, pp. 621-625 (en adelante: Malik, "AWS"); Guiora, Amos N., "Accountability and Decision Making in Autonomous Warfare: Who Is Responsible?", *Utah Law Review*, vol. 2017, núm. 2, 2017, pp. 393-398 (en adelante: Guiora, "Accountability and Decision Making"). De hecho, Guiora sugiere que los estándares de responsabilidad sean más estrictos para SAA (p. 418). Hauptman, Allyson, "Autonomous Weapons and the Law of Armed Conflict", *Military Law Review*, vol. 218, invierno de 2013, pp. 170, 193; Beard, Jack M., "Autonomous Weapons and Human Responsibilities", *Georgetown Journal of International Law*, vol. 45, núm. 3, 2014, pp. 617, 674, 675, 681; Aaronson, Michael, "Robots don't Kill People, it's the Humans We Should Worry About", *The Conversation*, 31 de mayo de 2013, disponible en <https://theconversation.com/robots-dont-kill-people-its-the-humans-we-should->

argumentan que una declaración política sería suficiente, y otros siguen en contra de cualquier forma de regulación internacional más allá de las reglas existentes del derecho internacional.<sup>160</sup>

Al respecto, numerosas preocupaciones sobre los SAA se han planteado por una variedad de actores, incluido el CICR<sup>161</sup> y la Campaña para Detener a los Robots Asesinos (*Campaign to Stop Killer Robots*), la cual es una coalición de 106 organizaciones no gubernamentales, con actividad en 54 países.<sup>162</sup> Los académicos están igualmente divididos, con opiniones tanto a favor como en contra de los SAA.

Si bien se desconoce cómo continuarán evolucionando estas tecnologías y en qué circunstancias podrán implementarse, las cuestiones más importantes que deben abordarse en los debates son aquéllas relativas a la responsabilidad y rendición de cuentas en relación con éstas.

Se ha señalado que uno de los principales argumentos de quienes piden la prohibición, la moratoria o regulación de los SAA, es que el uso de tales armas resultará en una bre-

---

*worry-about-14779*, ingresado el 10 de marzo de 2021; CICR, *Ethics and AWS*, *supra* nota 122, pp. 1, 2, 11. Davison, "AWS under IHL", *supra* nota 44, pp. 5, 17 y 18; Weizmann *et al.*, "AWS under International Law", *supra* nota 155, p. 5.

<sup>160</sup> Grupo de Expertos Gubernamentales de las Altas Partes Contratantes en la Convención sobre Prohibiciones o Restricciones del Empleo de Ciertas Armas Convencionales que Puedan Considerarse Excesivamente Nocivas o de Efectos Indiscriminados, Informe del Periodo de Sesiones de 2018 del Grupo de Expertos Gubernamentales sobre las Tecnologías Emergentes en el Ámbito de los Sistemas de Armas Autónomas Letales, UN Doc. CCW/GGE.1/2018/3, 23 de octubre de 2018, p. 7, y anexo III, p. 23 (en adelante: GEG, Informe de Sesiones 2018).

<sup>161</sup> Véase generalmente CICR, "Autonomous Weapon Systems, Implications of the Increasing Autonomy in the Critical Functions of Weapons", Expert Meeting, Suiza (15-16 de marzo de 2016); para tener acceso a este documento y a otros, véase CICR, *New Technologies and IHL*, disponible en <https://www.icrc.org/en/war-and-law/weapons/ihl-and-new-technologies>, ingresado el 11 de marzo de 2021.

<sup>162</sup> Véase la página web con la campaña para detener a los robots asesinos (*Campaign to Stop Killer Robots*), disponible en <https://www.stopkillerrobots.org>, ingresado el 11 de marzo de 2021.

cha de responsabilidad, lo que significa que habrá una virtual impunidad por cualquier violación de la ley como resultado de su uso.

En relación con lo anterior, cabe mencionar que algunos académicos han argumentado que la penalización por el uso de SAA sería preferible a una prohibición de éstos por cuatro motivos:

En primer lugar, porque poner el énfasis directamente en las personas limitaría otros tipos de problemas relacionadas con la firma y ratificación de tratados internacionales, sobre la ejecución de éstos y sobre la responsabilidad de los Estados.<sup>163</sup> En segundo lugar, consideran que esto transmitiría el mensaje de que cualquier situación de impunidad derivada de los SAA sería de naturaleza jurídica y no a causa de la insuficiencia técnica. En tercer lugar, porque una prohibición prematura puede detener indebidamente el desarrollo de tecnologías autónomas que pueden tener aplicaciones legítimas en el ámbito civil. Y cuarto, porque el enfoque de criminalización puede rescindirse fácilmente,<sup>164</sup> en caso de que posteriormente se resolvieran las preocupaciones sobre el uso de los SAA.<sup>165</sup>

No obstante, como se explicará a continuación,<sup>166</sup> sería un paso en falso asumir que la responsabilidad penal individual por el uso de los SAA podría determinarse en ausencia de una prohibición previa o un acuerdo internacional al respecto.

---

<sup>163</sup> Liu, Hin-Yan, "Refining Responsibility: Differentiating Two Types of Responsibility Issues Raised by Autonomous Weapons Systems", en Bhuta, Nehal *et al.* (eds.), *Autonomous Weapons Systems: Law, Ethics, Policy*, Cambridge University Press, 2016, p. 344 (en adelante: Liu, "Refining Responsibility").

<sup>164</sup> Véase McDougall, "AWS and Accountability", *supra* nota 156, p. 25. Considera que la afirmación de que la criminalización se rescindirá más fácilmente que una prohibición, o al menos una moratoria, o que tendrá un efecto diferenciado en aplicaciones civiles, puede ser poco probable.

<sup>165</sup> *Idem.*

<sup>166</sup> Véase *infra* epígrafe "Responsabilidad a través del derecho penal internacional".

Sin embargo, la opinión de la autora es que si bien los análisis sobre la legalidad o la conveniencia del uso de estas armas —los cuales conducen a una posible regulación o prohibición— y aquéllos sobre la problemática derivada de la brecha de responsabilidad pueden estar entrelazados conceptualmente, estas discusiones tienen diferentes objetivos *in se y*, por lo tanto, deben analizarse por separado.

Para los propósitos de este apartado, sólo se explorará el debate sobre la brecha de responsabilidad en el sentido de rendición de cuentas.

### *¿Qué tipo de responsabilidad podemos esperar?*

Algunas de las primeras consultas sobre los desafíos de la responsabilidad legal por acciones de máquinas inteligentes surgieron hace más de veinte años, reconociendo que cuando una máquina alcanza cierto nivel de inteligencia, hasta el punto de tomar decisiones por sí misma, surgen dificultades para imputar responsabilidad.<sup>167</sup> En la actualidad, esas preocupaciones siguen siendo igual de válidas al enfrentarnos a la misma interrogante, ya que no importa cuánto aumente la autonomía de las máquinas ni cuán sofisticadas se vuelvan, siguen sin tener agencia moral.<sup>168</sup>

---

<sup>167</sup> Véase Perri 6, “Ethics, Regulation and the New Artificial Intelligence. Part II: Autonomy and Liability”, *Information, Communication and Society*, vol. 4, núm. 3, 2001, pp. 406-434, 414.

<sup>168</sup> Wagner, Markus, “Taking Humans Out of the Loop: Implications for International Humanitarian Law”, *Journal of Information, Law & Science*, vol. 21, 2011, pp. 5, 11 (en adelante: Wagner, “Taking Humans out of the Loop”); Asaro, Peter, “On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-Making”, *International Review of the Red Cross*, vol. 94, 2012, p. 693 (en adelante: Asaro, “On Banning AWS”).

En este punto, es fundamental detenerse y hacer dos aclaraciones conceptuales importantes a considerar durante el resto de esta discusión.

La primera es reconocer que los SAA pueden tener dos facetas, en la que pueden ser completamente autónomas y el ser humano estar “fuera del proceso”, y en la que son semiautónomas ya que operan de forma automática en conjunto con los humanos que están “dentro del proceso”.<sup>169</sup> Sin embargo, dado que hasta ahora la línea de la discusión internacional parece apuntar a asegurar un “control humano determinante”<sup>170</sup> como un intento de contrarrestar las objeciones contra los SAA, es de gran relevancia identificar cuál sería la dinámica que resultaría de su uso y qué implicaciones operativas tendría, dado que este término está lejos de ser definido de manera homogénea.<sup>171</sup>

En cualquier caso, ambos escenarios deben abordarse en términos de rendición de cuentas, ya que de otro modo resultarían igualmente en impunidad.

---

<sup>169</sup> Véase Bo, Marta, “The Human-Weapon Relationship in the Age of Autonomous Weapons and the Attribution of Criminal Responsibility for War Crimes”, Working Draft, 2019, p. 1; véase también Marra, William C. y McNeil, Sonia K., “Understanding the Loop: Regulating the Next Generation of War Machines”, *Harv. J. L. & Pub. Policy*, vol. 36, 2013, pp. 1139, 1150. Marra y McNeil afirman que la diferencia entre los términos de “automatización” y “autonomía” —que son similares— radica en que los sistemas automatizados no se dirigen a sí mismos y carecen de capacidad de toma de decisión, mientras que los autónomos son capaces de ser independientes en el establecimiento y persecución de sus propósitos.

<sup>170</sup> GEG, Informe de Sesiones 2018, *supra* nota 160, p. 22.

<sup>171</sup> Para explorar los significados de “control humano determinante”, véase Horowitz, Michael C. y Scharre, Paul, “Meaningful Human Control in Weapons Systems: A Primer”, Working Paper, Centre for New American Security, 2015, disponible en <https://www.cnas.org/publications/reports/meaningful-human-control-in-weapon-systems-a-primer>, ingresado el 15 de marzo de 2021; Ekelhof, Merel, “Autonomous Weapons: Operationalizing Meaningful Human Control”, *Humanitarian Law and Policy*, 2018, disponible en <https://blogs.icrc.org/law-and-policy/2018/08/15/autonomous-weapons-operationalizing-meaningful-human-control/>, ingresado el 15 de marzo de 2021.

La segunda es reconocer que, dadas las particularidades de la IA y los SAA, ambos han sido ampliamente sujetos a antropomorfismos. Sin embargo, esto es un error cognitivo, porque induce erróneamente a la expectativa de que la IA se adhiera a normas sociales o patrones (humanos) de comportamiento preestablecidos.<sup>172</sup> Tal como se mencionó anteriormente,<sup>173</sup> tras una observación minuciosa, uno se da cuenta de que estas entidades no son armas, plataformas convencionales ni agentes morales equivalentes a los humanos para fines legales. Sin embargo, a menudo se les refiere como lo primero, en ocasiones como lo segundo, y con frecuencia se les trata como lo tercero.<sup>174</sup>

Lo anterior es relevante para esta discusión, pues independientemente de cuán avanzada pueda llegar a ser la tecnología, la máquina nunca será un agente moral responsable<sup>175</sup> y, por lo tanto, la responsabilidad nunca se le podrá transferir a ella. Por lo menos en este aspecto, los Estados han llegado a un consenso.<sup>176</sup>

---

<sup>172</sup> Abbott y Sarch, "Punishing AI", *supra* nota 127, p. 333. "No vamos a intentar articular las diferencias no funcionales entre el razonamiento humano y algorítmico, un tema que ha fascinado y confundido a científicos desde los 1950s".

<sup>173</sup> Véase *supra* el apartado "Sistemas de armas autónomas".

<sup>174</sup> Shilo, *When Turing Met Grotius*, *supra* nota 25, p.15.

<sup>175</sup> Esto se afirma como un hecho axiomático por la mayoría de los autores en el tema. Véase, e.g., Malik, "AWS", *supra* nota 159; Heyns, Informe del Relator Especial, *supra* nota 148; Sparrow, "Killer Robots", *supra* nota 45, pp. 65-68, 71-73; Sharkey, Noel E., "The Evitability of Autonomous Robot Warfare", *International Review of the Red Cross*, vol. 94, núm. 886, 2012, pp. 787, 790. Para un punto de vista en contrario, véase Sullins, John P., "When is a Robot a Moral Agent?", *International Review of Information Ethics*, vol. 6, 2006, p. 23; Ohlin, Jens David, "The Combatant's Stance: Autonomous Weapons on the Battlefield", *International Law Studies*, vol. 92, 2016, pp. 1 y 2 (en adelante: Ohlin, "The Combatant's Stance").

<sup>176</sup> En medio de una escasa lista de 10 "posibles principios rectores", el GEG acordó que "la responsabilidad humana de las decisiones sobre el uso de sistemas de armas debe mantenerse, ya que la rendición de cuentas no puede transferirse a las máquinas". Esto debe tenerse en cuenta durante todo el ciclo de vida del sistema de armas. GEG, Informe de Sesiones 2018, *supra* nota 160, p. 4.

Además, como se explicará a detalle más adelante,<sup>177</sup> se ha establecido una analogía entre la relación del comandante humano *vis-à-vis* un subordinado humano, con la del comandante humano *vis-à-vis* un robot.<sup>178</sup> La reiterada referencia de la persona que despliega un SAA como *comandante* genera la impresión errónea de que los SAA son en cierto modo combatientes o soldados.<sup>179</sup>

Dicho esto, otro punto muy importante en el que se ha llegado a un consenso entre los Estados es que “la responsabilidad por el desarrollo, despliegue y uso de cualquier sistema de armas emergente en el marco de la CAC debe garantizarse de conformidad con el derecho internacional aplicable, incluso mediante el funcionamiento de tales sistemas dentro de una cadena humana de mando y control”.<sup>180</sup>

Además, los Estados también acordaron que “los seres humanos deben seguir siendo responsables en todo momento de las decisiones sobre el uso de la fuerza de conformidad con el derecho internacional aplicable”.<sup>181</sup>

Una visión operativa que demuestra la necesidad de responsabilidad humana detrás de la toma de decisiones de una IA es el Manual del Derecho de la Guerra del Departamento de Defensa de Estados Unidos,<sup>182</sup> que incluye la disposición 6.5.9.3 “Las obligaciones de distin-

---

<sup>177</sup> Véase *infra* el apartado “Usuarios: responsabilidad de superiores jerárquicos”.

<sup>178</sup> Chengeta, “Accountability Gap”, *supra* nota 155, p. 3.

<sup>179</sup> Docherty, Bonnie, “Losing Humanity: The Case Against Killer Robots”, *Human Rights Watch*, vol. 1, pp. 4, 33 y 34, 42 y 43 (en adelante: Docherty, “Losing Humanity”).

<sup>180</sup> GEG, Informe de Sesiones 2018, *supra* nota 160, p. 4.

<sup>181</sup> *Ibidem*, p. 23 (a).

<sup>182</sup> U.S. Department of Defense, Law of War Manual, 2016, disponible en <https://dod.defense.gov/Portals/1/Documents/pubs/DoD%20Law%20of%20War%20Manual%20%20June%202015%20Updated%20Dec%202016.pdf?ver=2016-12-13-172036-190>, ingresado el 10 de agosto de 2020.

ción y proporcionalidad del derecho de guerra se aplican a las personas y no a las armas mismas” y estipula que “el derecho de la guerra no le requiere a las armas tomar determinaciones legales, incluso si el arma (por ejemplo, a través de computadoras, *software* y sensores) puede caracterizarse como capaz de hacer determinaciones fácticas, como el disparar o seleccionar y atacar un objetivo”. Esto se ha considerado una señal de que las armas robóticas nunca serán sujetos de responsabilidad,<sup>183</sup> lo que nos incita a formular la inevitable pregunta: ¿qué prospectivas de responsabilidad realmente existen para los SAA?

Ahora bien, es necesario hacer una pausa en la nomenclatura “rendición de cuentas” para vislumbrar qué podría implicar realmente el alcance de este concepto en el contexto de este debate.

Se puede afirmar con seguridad que la rendición de cuentas es un concepto amplio que se ha utilizado como término general para describir diversas formas de responsabilidad, incluida la responsabilidad del Estado, procedimientos administrativos y disciplinarios emprendidos en respuesta a violaciones del DIH (incluso abarcando la justicia militar), responsabilidad civil y responsabilidad penal.<sup>184</sup> Todas estas modalidades son de carácter complementario entre sí y de ninguna manera son excluyentes unas de otras.<sup>185</sup> La rendición de cuentas es muy importante en el derecho internacional porque cuando existe una brecha de responsabilidad, el derecho de las víctimas a interponer un recurso y obtener reparaciones se ve afectado adversamente.<sup>186</sup>

---

<sup>183</sup> Bryson, J. J. *et al.*, “Of, for, and by the People: the Legal Lacuna of Synthetic Persons”, *Artificial Intelligence and Law*, vol. 25, 2017, pp. 273-291.

<sup>184</sup> McDougall, “AWS and Accountability”, *supra* nota 156, p. 7.

<sup>185</sup> Chengeta, “Accountability Gap”, *supra* nota 155, p. 3.

<sup>186</sup> Burke, Megan y Persi-Vicentic, Loren, “Remedies and Reparations”, in Casey-Malsen, S. (ed.), *Weapons Under International Human Rights Law*, 2014, pp. 542-589 (en adelante: Burke y Persi, “Remedies and Reparations”); Moffett, Luke, *Justice for Victims before the International Criminal Court*, 2014, p. 146.

También es digno de mención que el diseño, desarrollo y/o uso de SAA podría implicar principios generales y reglas que se encuentran en varios campos del derecho internacional público, tales como *ius ad bellum*, DIH, derecho internacional de derechos humanos, DPI, derecho espacial, entre otros.

En consecuencia, los recursos legales para la rendición de cuentas en cada uno de estos regímenes pueden variar significativamente en su alcance. Para los propósitos de este estudio, la rendición de cuentas se entiende como el deber de responsabilizarse por el ejercicio de autoridad sobre el diseño, desarrollo o uso (o una combinación de éstos) de los SAA, reconociendo que la autoridad puede ser ejercida por una amplia gama de actores.<sup>187</sup>

Con un enfoque de abajo hacia arriba, la primera y más intuitiva aproximación sería la responsabilidad penal individual bajo el derecho internacional por la comisión de crímenes internacionales, como lo son los crímenes de guerra que involucren el uso de los SAA. Para determinar esta responsabilidad penal, es necesario verificar la comisión de un crimen definido en el derecho internacional, establecer la jurisdicción competente sobre ese crimen, delimitar qué modalidad de responsabilidad aplica a la conducta de un individuo en particular, demostrar que se cumplan los elementos materiales (*actus reus*) y subjetivos (*mens rea*) de dicha modalidad de responsabilidad por el crimen del que se trate, en su caso valorar la existencia de una justificación jurídica, y finalmente en el evento de que se condene, imponer una sentencia.<sup>188</sup>

Por el contrario, en una aproximación de arriba hacia abajo, la segunda vía es invocar la responsabilidad del Estado derivada de actos u omisiones que impliquen el uso de SAA, cuando esos actos u omisiones supongan una violación de una norma de derecho internacio-

---

<sup>187</sup> Lewis et al., *War-Algorithm*, *supra* nota 32, p. 11.

<sup>188</sup> *Ibidem*, pp. 12, 77.

nal existente.<sup>189</sup> Para atribuir la responsabilidad del Estado, se requiere establecer la existencia de una norma consagrada en algún tratado o por alguna costumbre internacional, discernir la obligación jurídica derivada de dicha norma, identificar si hay una violación a la misma, y lo más importante, *atribuir* dicha violación a un Estado, después determinar si hay excluyentes aplicables, y si no las hay, imponer reparaciones.<sup>190</sup>

Por otro lado, un enfoque idealmente paralelo a los dos mencionados anteriormente sería aplicar un escrutinio de gobernanza. Si bien las consecuencias derivadas de esta opción pueden parecer “más laxas”, este enfoque debe incluirse considerando que, en el contexto geopolítico actual, podría convertirse en el recurso más accesible en la práctica. Esta opción contempla la medida en que una persona o entidad está y deba estar sujeta a, o deba ejercer formas de escrutinio, monitoreo o regulación interna o externa (o una combinación de ellas) en relación con el diseño, desarrollo o uso de un SAA. Algunos ejemplos de escrutinio de gobernanza incluyen el monitoreo independiente, el desarrollo regulatorio, la adopción de resoluciones y códigos de conducta no vinculantes, el diseño normativo de arquitecturas técnicas (incluida la maximización de la auditabilidad sobre los algoritmos) y la autorregulación comunitaria.<sup>191</sup>

La autora continuará explorando las alternativas legales más prevalentes a lo largo de la serie del estudio más amplio; sin embargo, para los propósitos de este apartado, la rendición de cuentas se relaciona únicamente con la responsabilidad penal individual.

Dado que se trata de una cuestión a nivel mundial que incumbe a toda la humanidad, comencemos por explorar la cuestión de la rendición de cuentas desde la perspectiva del DPI.

---

<sup>189</sup> *Ibidem*, pp. 83 y 84.

<sup>190</sup> *Ibidem*, pp. 54, 84.

<sup>191</sup> *Ibidem*, p. 91.

## Responsabilidad a través del derecho penal internacional

En este punto, es importante recordar la *raison d'être* y la relevancia de esta discusión. Los mecanismos de rendición de cuentas son esenciales para lograr la disuasión<sup>192</sup> mediante el cumplimiento de las obligaciones del DIH de enjuiciar infracciones graves y crímenes de guerra,<sup>193</sup> procurando la prevención que es fundamental para la protección de civiles,<sup>194</sup> y que también son fundamentales para garantizar el derecho a reparación de las víctimas.<sup>195</sup> En resumidas cuentas, la rendición de cuentas se ha denominado “el *quid* del derecho internacional”.<sup>196</sup>

Por lo tanto, como una cuestión de pragmatismo, el derecho internacional no puede limitarse a la fijación de estándares para los gobiernos, los actores no estatales y sus agentes, sino que debe englobar las consecuencias por el incumplimiento de los mismos.<sup>197</sup> Asimismo, las normas de DIH, algunas de las cuales son *ius cogens*, carecen de valor sin mecanismos

---

<sup>192</sup> Véase, e.g., Guiora, “Accountability and Decision Making”, *supra* nota 159, p. 398; Guiora menciona que decisiones de “matar/no matar” autorizadas por el Estado, donde los estándares de responsabilidad no son inherentes ni integrales, es similar a autorizar el “Viejo Oeste”; Stewart (“New Technology and the Law of Armed Conflict”, *supra* nota 158, p. 292) se refiere a “las cuestiones de política pública más amplias asociadas con la posibilidad de que las operaciones militares se lleven a cabo en un *entorno sin culpa*”.

<sup>193</sup> CG IV, artículo 146, requiere que las violaciones graves se criminalicen y se enjuicien. También hay una obligación a enjuiciar una gama más amplia de crímenes de guerra bajo el derecho consuetudinario.

<sup>194</sup> Véase, e.g., Heyns, Informe del Relator Especial, *supra* nota 148, p. 75.

<sup>195</sup> Véase, e.g., Chengeta, “Accountability Gap”, *supra* nota 155, p. 5.

<sup>196</sup> *Ibidem*, p. 49; véase también Malik, “AWS”, *supra* nota 159, p. 620; Guiora, “Accountability and Decision Making”, *supra* nota 159, p. 398; HRW, “Mind the Gap”, *supra* nota 158; Heyns, Informe del Relator Especial, *supra* nota 148, p. 75.

<sup>197</sup> Ratner, Steven *et al.*, *Accountability for Human Rights Atrocities in International Law: Beyond the Nuremberg Legacy*, 3a. ed., Oxford, 2009 (en adelante: Ratner *et al.*, *Accountability for Human Rights Atrocities*).

de rendición de cuentas por no respetarlas.<sup>198</sup> Sin duda, se puede argumentar que la brecha de responsabilidad por las violaciones del derecho internacional plantea efectivamente una amenaza global para el mantenimiento de la paz y la seguridad.<sup>199</sup>

La autora coincide con la afirmación de que, “después de todo, sin responsabilidad, el derecho internacional no es más que el proverbial *brutum fulmen* - un rayo inofensivo”.<sup>200</sup>

Asimismo, como se mencionó anteriormente, la responsabilidad también es fundamental, porque está intrínsecamente relacionada con el derecho a la reparación de las víctimas, tanto civiles como militares en casos del uso ilegal de un arma que ha sido declarada como ilegal o que tiene efectos indiscriminados como método de guerra, o por el uso de la fuerza desproporcionado o excesivo durante la aplicación de la ley para el mantenimiento del orden público. Lo anterior también se extiende al incumplimiento intencional o negligente de proteger a las víctimas de armas dañinas, ya que lo anterior ha sido reconocido como conducta ilegal equivalente a una violación de derechos humanos.<sup>201</sup> La autora opina que los SAA se encuadran perfectamente como armas dañinas, de las cuales todas las personas tienen derecho a protección según los estándares internacionales,<sup>202</sup> por lo que los desafíos en términos de rendición de cuentas que plantea su uso deben tomarse muy en serio.

---

<sup>198</sup> Chengeta, “Accountability Gap”, *supra* nota 155, p. 5; Seibert-Fohr, Anja, *Prosecuting Serious Human Rights Violations*, Oxford, 2009, pp. 292 y 293.

<sup>199</sup> Véase Jones, John R.W.D. y Powles, Steven, *International Criminal Practice*, 3a. ed, 2003, p. 2, (en adelante: Jones y Powles, *International Criminal Practice*).

<sup>200</sup> Chengeta, “Accountability Gap”, *supra* nota 155, p. 5.

<sup>201</sup> Burke y Persi, “Remedies and Reparations”, *supra* nota 186, p. 554.

<sup>202</sup> Véase *supra* en apartado “Aplicabilidad de la Cláusula Martens”. Chengeta comparte este punto de vista en: Chengeta, “Accountability Gap”, *supra* nota 155, p. 5.

El derecho a la reparación, entendido como un proceso que tiene por objeto brindar justicia a las víctimas, remover o reparar en la medida de lo posible el daño causado por los actos ilícitos mediante la prevención y la disuasión,<sup>203</sup> es relevante en el contexto de este apartado, ya que es el deber de los Estados hacer efectivos los derechos de las víctimas mediante la investigación de violaciones a los derechos humanos y el enjuiciamiento de los perpetradores.<sup>204</sup>

En este sentido, la rendición de cuentas puede caracterizarse como una “compleja amalgama de derecho y de un amplio espectro de procesos sancionadores que trasciende las divisiones ortodoxas de sujetos del derecho internacional”.<sup>205</sup>

Como resultado de lo anterior, la responsabilidad penal individual se ha convertido en parte del derecho internacional consuetudinario<sup>206</sup> con el fin de disuadir y/o sancionar actos ilícitos cometidos, no sólo durante conflictos armados internacionales y no internacionales,<sup>207</sup>

---

<sup>203</sup> Véase David, Roman y Choi Yuk-ping, Susanne, “Victims on Transitional Justice: Lessons from the Reparation of Human Rights Abuses in the Czech Republic”, *Human Rights Quarterly*, vol. 27, 2005, pp. 392-393; Pisillo Mazzeschi, Riccardo, “Reparation Claims by Individuals for State Breaches of Humanitarian Law and Human Rights: An Overview”, *Journal of International Criminal Justice*, vol. 1, 2003, pp. 339, 344.

<sup>204</sup> Comité de Derechos Humanos de las Naciones Unidas, Observación General núm. 31 [80], “Naturaleza de la obligación jurídica general impuesta a los Estados Parte en el Pacto”, CCPR/C/21/Rev.1/Add.13, 2004, pp. 2 y 3; véase Corte Europea de Derechos Humanos, *Aksoy v. Turkey*, Sentencia, 1996; véase Comisión Africana de Derechos Humanos y de los Pueblos, *Social and Economic Rights Action Centre And Centre for Economic and Social Rights v. Nigeria*, Comunicación núm. 155/96, 2001, pp. 44-48.

<sup>205</sup> Ratner *et al.*, *Accountability for Human Rights Atrocities*, *supra* nota 197, p. 3.

<sup>206</sup> Swart, Bert, “Modes of International Criminal Liability”, en Cassese, Antonio (ed.), *The Oxford Companion to International Criminal Justice*, 2009, pp. 82, 91 (en adelante: Swart, “Modes of International Criminal Liability”).

<sup>207</sup> Tribunal Internacional Penal para la exYugoslavia (TIPY), *Fiscal c. Tadić*, Decisión relativa a la Apelación de la Defensa concerniente a la Excepción Prejudicial de Incompetencia, 2 de octubre de 1995, p. 129.

sino también durante tiempos de paz.<sup>208</sup> Esta es la principal forma de rendición de cuentas analizada en este apartado.

## Los crímenes

El DPI es la rama del derecho internacional público que establece la responsabilidad penal individual por crímenes internacionales, es decir, crímenes de guerra, crímenes de lesa humanidad, genocidio y el crimen de agresión. Su propósito dentro del ordenamiento jurídico internacional es de carácter multifacético, ya que tiene como objetivo contribuir a la disuasión general y específica, la incapacitación, rehabilitación, reconciliación, justicia para las víctimas, retribución, esclarecimiento de la verdad, la promoción del orden internacional, el establecimiento y mantenimiento de una paz inclusiva y duradera, etcétera.<sup>209</sup>

Por lo tanto, para que el DPI sea aplicable, debe existir un crimen internacional.

En este sentido, viene a la mente la siguiente cita del difunto Cherif Bassiouni, uno de los padres del derecho internacional moderno: “Los crímenes internacionales se han desarrollado hasta la fecha, sin [...] una definición acordada de lo que constituye un crimen internacional, cuáles son los criterios para la criminalización internacional y cómo se distinguen los crímenes internacionales”.<sup>210</sup>

---

<sup>208</sup> Los crímenes de lesa humanidad y genocidio no requieren la existencia de un conflicto armado; véase Bouganim-Shaag, Efrat y Naggan, Yael, “Emerging Voices: Peace-Time Crimes Against Humanity and the ICC”, *Opinio Juris*, 2013, disponible en <http://opiniojuris.org/2013/08/30/emerging-voices-peace-time-crimes-humanity-icc/>, ingresado el 20 de febrero de 2021; véase también, artículo 1 de la Convención para la Prevención y Sanción del Delito del Genocidio, 1948.

<sup>209</sup> McDougall, “AWS and Accountability”, *supra* nota 156, p. 28.

<sup>210</sup> Bassiouni, M. Cherif, *Introduction to International Criminal Law*, Transnational Publishers, 2003, p. 111.

No obstante, de la jurisprudencia de las cortes y tribunales penales internacionales se podría concluir que sólo se ha atribuido la responsabilidad penal individual por conductas que ya están prohibidas por la costumbre o el derecho convencional bajo otra rama del derecho internacional como el DIH, el derecho internacional de los derechos humanos o el *ius ad bellum*.<sup>211</sup>

El Estatuto de Roma fue adoptado en 1998 con el fin de establecer una institución permanente y global, *i.e.*, la Corte Penal Internacional (CPI), para perseguir y enjuiciar a los responsables de los crímenes más aborrecibles conocidos por el hombre. Como tal, el Estatuto estaba destinado a proporcionar una orientación normativa de lo que esos crímenes podrían materializarse en el futuro, aunque ésta estuviera limitada a las referencias temporales de los conflictos del siglo XX —los cuales evidentemente no empleaban las tecnologías sofisticadas que hoy en día se empiezan a conocer—.

Prueba de ello es el hecho de que la mayoría de las discusiones giraron en torno a la cuestión de la costumbre, requiriendo pruebas de una prohibición ya ampliamente aceptada de la conducta en cuestión, o incluso de la preexistencia de un crimen según la costumbre internacional.<sup>212</sup> Por lo tanto, esto significó que las definiciones de los crímenes del Estatuto se adoptaron mirando hacia atrás, en el tiempo, no hacia adelante.<sup>213</sup>

<sup>211</sup> Dinstein, Yoram, "International Criminal Law", *Israel Law Review*, vol. 20, núm. 2-3, 1985, pp. 206, 221; Bantekas, Ilias y Nash, Susan, *International Criminal Law*, 2a. ed., Cavendish Publishing, 2003, p. 5; Cassese, Antonio, *International Criminal Law*, 2a. ed., Oxford University Press, 2008, pp. 11 y 12.

<sup>212</sup> Conferencia Diplomática de Plenipotenciarios de las Naciones Unidas sobre el establecimiento de una Corte Penal Internacional, Resumen de los Registros de las Reuniones Plenarias y de la Comisión Plenaria, UN Doc. A/CONF.183/13, vol. II, 15 de junio-17 de julio de 1998; en particular: 150 (Reino Unido), 151 (Eslovenia), 154 y 155 (Canadá), 155 (Israel), 158 (Siria), 160 (Nueva Zelanda), 160 (Grecia), 162 (Bélgica), 164 (Francia), 187 (Jordán), 270 (China), 277 (Suiza), 277 (Brasil), 278 (Corea), 285 (Bosnia y Herzegovina), 287 (Indonesia), 289 (Rusia).

<sup>213</sup> Hebel, Herman von y Robinson, Darryl, "Crimes within the Jurisdiction of the Court", en Lee, Roy S. (ed.), *The International Criminal Court: The Making of the Rome Statute*, Kluwer International, 1999, pp. 79, 104, 122-123;

Los redactores previeron el establecimiento de un Grupo de Trabajo sobre las Enmiendas, un órgano de la Asamblea de los Estados Parte de la CPI que considera la necesidad de realizar enmiendas tanto al Estatuto de Roma como a las Reglas de Procedimiento y Prueba de la Corte.<sup>214</sup>

Además, su mandato establece que, para proponer un nuevo crimen, se debe ponderar si éste pudiera considerarse como uno de los crímenes más graves y de trascendencia para la comunidad internacional en su conjunto, y si el crimen en cuestión está basado en una prohibición ya existente en el derecho internacional.<sup>215</sup>

Aquí debemos detenemos y recordar que el propósito de este estudio es señalar que existen lagunas legales considerables cuando se trata de los SAA, ya que tanto el DPI y el derecho penal, en general, ha sido ideado por humanos, para humanos. Es por ello que estas lagunas deben ser atendidas urgentemente para evitar que el vacío legal se entienda como una falta de prohibición.

No obstante, sería posible argumentar que los valores fundamentales que corren riesgo por el uso de SAA ya gozan de una protección por el ordenamiento legal internacional.

---

Schabas, William A., *An Introduction to the International Criminal Court*, Cambridge University Press, 2001, p. 23. Para una comparación de las definiciones de los crímenes en el Estatuto de Roma con las del derecho consuetudinario, véase Cassese, Antonio, "Genocide", en Cassese, Antonio *et al.* (eds.), *The Rome Statute of the International Criminal Court: A Commentary*, Oxford University Press, vol. 1, 2002, p. 335; Cassese, Antonio, "Crimes against Humanity", en Cassese, Antonio *et al.* (eds.), *The Rome Statute of the International Criminal Court: A Commentary*, Oxford University Press, vol. 1, 2002, p. 353; Bothe, Michael, "War Crimes", en Cassese, Antonio *et al.* (eds.), *The Rome Statute of the International Criminal Court: A Commentary*, Oxford University Press, vol. 1, 2002, p. 379.

<sup>214</sup> Corte Penal Internacional, Reglas de Procedimiento y Prueba, Doc. No. ICC-ASP/1/3, adoptadas el 9 de septiembre de 2002.

<sup>215</sup> Fortalecimiento de la Corte Penal Internacional y la Asamblea de Estados Parte, Doc. No. ICC-ASP/11/Res.8, Anexo II, Mandato del Grupo de Trabajo sobre las Enmiendas, 21 de noviembre de 2012, p. 9.

Por ejemplo, se podría hacer referencia a la Cláusula Martens,<sup>216</sup> al derecho a la vida, a la integridad personal, a la protección contra violaciones de derechos humanos, la seguridad jurídica, la necesidad de rendición de cuentas y reparaciones, al igual que la prohibición de realizar ataques de forma indiscriminada y desproporcionada, entre otros, concluyendo que ya existe una prohibición de violar cualquiera de estos valores legales fundamentales.

También se puede argumentar efectivamente que lo que resulta importante es la aplicabilidad de los Convenios de Ginebra, aunque éstos son en principio tecnológicamente neutros, ya que prohíben un resultado (e.g., matar civiles) y los medios usados son indiferentes.<sup>217</sup>

En otras palabras, el argumento es que la naturaleza *sui generis* de los SAA, *i.e.*, su imprevisibilidad, requiere penalización, no porque los SAA sean ilegales *per se*, sino porque el despliegue de SAA aumenta el riesgo de incumplimiento de las reglas de DIH destinadas a proteger a los no-combatientes.<sup>218</sup> Básicamente, lo que se pretende es prohibir toda conducta que pueda conducir a un resultado no deseado (y criminal) respecto a cuya comisión no se pueda determinar el individuo en quién recae la responsabilidad penal, volviéndola, *a priori*, virtualmente impune.

Existe la opinión de que, si hay crímenes de guerra reconocibles, debe haber criminales reconocibles.<sup>219</sup> Sin embargo, cuando se trata del DPI, tal como lo aplican actualmente las cortes y los tribunales internacionales, a menudo se necesita que estos valores fundamentales sean explícitos para establecer la responsabilidad penal individual de una determinada persona.

---

<sup>216</sup> Como se mencionó *supra* en “Aplicabilidad de la Cláusula Martens”. Véase Davison, “AWS under IHL”, *supra* nota 44, p. 8.

<sup>217</sup> McDougall, “AWS and Accountability”, *supra* nota 156, p. 27.

<sup>218</sup> *Ibidem*, p. 28.

<sup>219</sup> Véase, generalmente, Walzer, Michael, *Just and Unjust Wars: A Moral Argument with Historical Illustrations*, 2015.

Dada la naturaleza estricta del derecho penal en general, respetando la máxima *nullum crimen, nulla poena sine praevia lege*,<sup>220</sup> las conductas relacionadas con los SAA deben prohibirse expresamente antes de que se tipifiquen como crímenes. Por lo tanto, al final, esto no es una cuestión meramente legal, sino también una cuestión de política pública e internacional.

## El régimen del Estatuto de Roma

Al examinar los crímenes internacionales tal y como se definen en el Estatuto de Roma, podríamos anticipar que los SAA podrían resultar involucrados en la comisión de al menos los siguientes crímenes: genocidio, crímenes de lesa humanidad, crímenes de guerra y crimen de agresión.<sup>221</sup> Podría parecer que el crimen más obvio sería los crímenes de guerra, no obstante, la autora opina que se debe de dar énfasis al crimen de agresión, lo cual será materia de un estudio subsecuente. En todo caso, un despliegue completo de SAA podría usarse para perpetrar un genocidio o crímenes de lesa humanidad, especialmente debido a las funciones de selección de objetivos de SAA.

No obstante, esto sólo es un ejercicio hipotético, dada la naturaleza impredecible de estas tecnologías, y del hecho de que aún no conocemos todas sus capacidades.

---

<sup>220</sup> Véase, generalmente, Kreß, Claus, “*Nulla poena nullum crimen sine lege*”, *Max Planck Encyclopedia of Public International Law*, 2010, disponible en <https://www.legal-tools.org/doc/f9b453/pdf/>, ingresado el 10 de febrero de 2021.

<sup>221</sup> Estatuto de Roma de la Corte Penal Internacional, artículos 5-8 bis (en adelante: Estatuto de Roma).

La responsabilidad penal individual depende de dos factores básicos, la intención criminal (*mens rea*) y la acción delictiva (*actus reus*) por parte del perpetrador.<sup>222</sup> Este último comprende los elementos objetivos del crimen, tales como la conducta prohibida del autor (ya sea un acto o una omisión), sus consecuencias, el vínculo de causalidad entre el acto y la consecuencia y, en ocasiones, las circunstancias específicas relacionadas con el contexto, sujeto u objeto del crimen o sus modalidades.<sup>223</sup>

Con respecto al primer factor, la idea de castigar solamente a quienes tienen una mente culpable deriva de nociones del derecho natural (*ius naturale*) y derechos humanos<sup>224</sup> que datan desde 1819, cuando Bagshaw afirmó que la noción de que “ningún hombre debería ser castigado, excepto por su propia culpa” es un máxima del derecho natural.<sup>225</sup>

En consecuencia, es necesario analizar el *actus reus* y la *mens rea*, así como los elementos contextuales específicos de cada crimen de la CPI, para determinar si son compatibles o no con conductas que pudieran ser cometidas por o con el uso de SAA. Si bien, estos

---

<sup>222</sup> Véase Jones y Powles, *International Criminal Practice*, supra nota 199, pp. 414-424; Badar, Mohamed, *The Concept of Mens Rea in International Criminal Law: The Case for a Unified Approach*, 2013, pp. 234-252; Klip, Andri y Sluiter, Goran, *Annotated Leading Cases of International Criminal Tribunals: The International Criminal Tribunal for the Former Yugoslavia*, 2001, p. 321; Doría, José et al., *The Legal Regime of The International Criminal Court: Essays in Honour of Professor Igor Blishchenko*, 2009, p. 144; Marchuk, Iryna, *The Fundamental Concept of Crime in International Criminal Law: A Comparative Law Analysis*, 2013, p. 134; Bonaft, Beatrice, *The Relationship Between State and Individual Responsibility for International Crimes*, 2009, p. 247; “Trial of Bruno Tesch et al. (Zyklon B Case), UNWCC, Case Number 9, British Military Court (1946)”, *Law Reports of Trials of War Criminals*, 1949, pp. 93-104.

<sup>223</sup> Stahn, Carsten, *A Critical Introduction to International Criminal Law*, Cambridge University Press, 2019, p. 22.

<sup>224</sup> Véase Cobbett, William, *Cobbett's Parliamentary History of England: From the Norman Conquest, in 1066 to the Year 1803*, 1819, p. 1079.

<sup>225</sup> *Idem*.

crímenes pueden ser cometidos mediante diversas conductas o modalidades, los siguientes son algunos elementos generales requeridos para su comisión:<sup>226</sup>

Genocidio

*Actus reus:*

- La conducta específica debe estar dirigida contra una o más personas pertenecientes a un grupo determinado, ya sea nacional, étnico, racial o religioso (un grupo protegido).

*Mens rea:*

- El autor tiene la intención de cometer el acto, provocar sus efectos o tiene conocimiento de que ocurrirán en el curso ordinario de los hechos.
- El autor es consciente de que su conducta tuvo lugar en el contexto de un patrón manifiesto de conducta similar dirigida contra ese grupo protegido o fue una conducta que por sí misma podría producir tal destrucción.

*Dolus specialis:*

- Más específicamente, el autor tuvo la intención de destruir, total o parcialmente, ese grupo nacional, étnico, racial o religioso, como tal.

Crímenes de lesa humanidad

*Actus reus:*

- La conducta (*ataque*) debe estar dirigida contra una o más personas pertenecientes a la población civil.

<sup>226</sup> CPI, Elementos de los Crímenes, Doc. No. ICC-ASP/1/3 y Corr.1 (en adelante: CPI, Elementos de los Crímenes); véase CPI, Documentos Oficiales de la Conferencia de Revisión del Estatuto de Roma de la Corte Penal Internacional, Kampala, RC/11 (2010).

*Mens rea:*

- El autor tiene la intención de cometer el acto, provocar sus efectos o tiene conocimiento de que ocurrirá en el curso ordinario de los hechos.
- El autor tenía conocimiento de que la conducta era parte o tenía la intención de que la conducta fuera parte de un ataque generalizado o sistemático contra una población civil.

Elementos contextuales:

- La conducta se cometió como parte de un ataque generalizado o sistemático dirigido contra una población civil.

Crímenes de guerra

*Actus reus:*

- Artículo 8(2)(a): graves infracciones de los Convenios de Ginebra de 1949.
- Artículo 8(2)(b): otras graves violaciones de las leyes y costumbres aplicables en los conflictos armados internacionales.
- Artículo 8(2)(c): graves violaciones del artículo 3 común, *i.e.*, los actos específicos contra personas que no participan activamente en las hostilidades.
- Artículo 8(2)(e): violaciones de las leyes y costumbres aplicables en conflictos armados no internacionales.

*Mens rea:*

- El autor tiene la intención de cometer el acto, provocar sus efectos o tiene conocimiento de que ocurrirán en el curso ordinario de los hechos.

- El autor conoce las circunstancias de hecho que establecieron la existencia de un conflicto armado.<sup>227</sup>

Elementos contextuales:

- Para los crímenes en virtud del artículo 8(2)(a) y (b): la conducta tuvo lugar en el contexto de un conflicto armado internacional y estuvo relacionada con él.
- Para los crímenes previstos en el artículo 8(2)(c) y (e): la conducta tuvo lugar en el contexto de un conflicto armado no internacional y estuvo relacionada con él.

Crimen de agresión

*Actus reus:*

- El acto de agresión es el uso de la fuerza armada por parte de un Estado contra la soberanía, la integridad territorial o la independencia política de otro Estado, o de cualquier otra forma incompatible con la Carta de las Naciones Unidas.<sup>228</sup>

*Mens rea:*

- El autor tiene la intención de cometer el acto, provocar sus efectos o tiene conocimiento de que ocurrirá en el curso ordinario de los hechos.
- El autor conocía las circunstancias de hecho que determinaban que tal uso de la fuerza armada era incompatible con la Carta de las Naciones Unidas, así como una violación manifiesta a la misma.

---

<sup>227</sup> No se requiere que el autor haya hecho una evaluación sobre la existencia de un conflicto armado ni su carácter internacional o no internacional. Únicamente se exige conocimiento de las circunstancias de hecho que hayan determinado la existencia de un conflicto armado.

<sup>228</sup> El crimen de agresión no prevé diferentes modalidades para su comisión.

Cabe decir que, en términos de *mens rea*, el artículo 30 del Estatuto de Roma establece una disposición general aplicable de manera adicional a los elementos subjetivos de cada crimen. Establece que una persona puede ser considerada penalmente responsable y ser castigada, sólo si los elementos materiales del crimen se cometen con *intención* y *conocimiento*. Por un lado, define que hay “intención” cuando la persona busca participar en la conducta, causar la consecuencia o es consciente de que la consecuencia ocurrirá en el curso ordinario de los acontecimientos. Por otro lado, interpreta el “conocimiento”, al estar consciente de que existe una circunstancia o de que ocurrirá una consecuencia en el curso ordinario de los acontecimientos.<sup>229</sup>

El hecho de que la disposición general contenida en el artículo 30 efectivamente actualiza un requisito adicional<sup>230</sup> ha sido ampliamente cuestionado, ya que se ha reflejado como el establecimiento de un umbral más alto y, por lo tanto, de manera injusta para algunos autores, limita el espectro de personas que puedan considerarse responsables en virtud del Estatuto.

## Modos de responsabilidad

Después de analizar los elementos *ratione materiae* de los crímenes que posiblemente se cometan por o con el uso de SAA, debemos considerar también los elementos *ratione personae* necesarios para establecer la responsabilidad penal individual tal y como la conocemos.

---

<sup>229</sup> Estatuto de Roma, artículo 30; véase CPI, *Fiscal c. Thomas Lubanga Dyilo*, Decisión de Confirmación de Cargos, 7 de febrero de 2007, pp. 350-352.

<sup>230</sup> CPI, *Fiscal c. Thomas Lubanga Dyilo*, Sentencia conforme al artículo 74 del Estatuto, 5 de abril de 2012, pp. 1014-1018.

La persona que comete el crimen es el autor o perpetrador,<sup>231</sup> y puede haber diferentes (co)autores del mismo crimen, siempre que las acciones de cada persona satisfagan los elementos necesarios para su comisión.<sup>232</sup>

Por lo tanto, ahora debemos examinar las disposiciones aplicables del Estatuto de Roma, *i.e.*, los artículos 25, sobre los modos de responsabilidad, y 28, sobre la responsabilidad del superior jerárquico.

Comencemos desmenuzando el artículo 25, sobre la responsabilidad penal individual.<sup>233</sup> En primer lugar, debemos señalar que la competencia de la Corte está limitada a personas naturales, excluyendo *ab initio* la posibilidad de atribuir responsabilidad a la propia tecnología.

En segundo lugar, será penalmente responsable y podrá ser condenada la persona que:

- Cometa ese crimen por sí solo, con otro o por conducto de otro, sea éste o no penalmente responsable;
- Ordene, proponga o induzca la comisión de ese crimen, ya sea consumado o en grado de tentativa;

---

<sup>231</sup> Voz: "Perpetrador, adj.", *DLE*, RAE, disponible en <https://dle.rae.es/perpetrador>; véase también la voz: "Perpetrar", *DLE*, RAE, disponible en <https://dle.rae.es/perpetrar?m=form>, ambos ingresados el 10 de marzo de 2021.

<sup>232</sup> Véase TIPY, *Fiscal c. Kunarac et al.*, Sentencia, 22 de febrero de 2001, p. 390; Tribunal Internacional Penal para Ruanda, *Fiscal c. Kayishema & Ruzindana*, Sentencia, 1 de junio de 2001, pp. 187, 192; TIPY, *Fiscal c. Krstic*, Sentencia, 1 de agosto de 2001, p. 601.

<sup>233</sup> Estatuto de Roma, artículo 25; véase, generalmente, Ambos, Kai, "Article 25. Individual Criminal Responsibility", en Triffterer, Otto (ed.), *Commentary on the Rome Statute of the International Criminal Court: Observers' Notes, Article by Article*, 2a. ed., 2008.

- Con el propósito de facilitar la comisión del crimen, sea cómplice o encubridor o colabore de algún modo o instigue la comisión o tentativa de comisión de un crimen, incluso suministrando los medios para su comisión;
- Contribuya de cualquier otro modo en la comisión o tentativa de comisión de tal crimen por parte de un grupo de personas que tengan un fin común. Esta contribución debe ser intencional y se hará con el objetivo de promover la actividad o propósito delictivo del grupo, o a sabiendas de la intención del grupo de cometer el crimen;
- Con respecto al crimen de genocidio, incita directa y públicamente a otros a cometerlo;
- Intenta cometer un crimen de este tipo mediante actos que supongan un paso importante para su ejecución, aunque el crimen no se consume por circunstancias ajenas a su voluntad.

Cabe señalar que el crimen de agresión sólo puede ser cometido por una persona que esté efectivamente en condiciones de ejercer control sobre o dirigir la acción política o militar del Estado que cometió el acto de agresión.

Asimismo, el Estatuto establece claramente que ninguna disposición relativa a la responsabilidad penal individual afectará la responsabilidad de los Estados en virtud del derecho internacional.

Además de las formas de responsabilidad previstas en el artículo 25, el Estatuto prevé una modalidad de responsabilidad diferente en el artículo 28, que se conoce como responsabilidad del superior jerárquico, la cual se abordará a detalle más adelante.<sup>234</sup> Esta modalidad es aplicable a los jefes militares o personas que actúen efectivamente como jefe militar por

---

<sup>234</sup> Véase *infra* en “Usuarios: responsabilidad de superiores jerárquicos”.

crímenes cometidos por las fuerzas bajo su mando y control efectivo, o su autoridad y control efectivo, según sea el caso, como resultado de no haber ejercido control sobre sus fuerzas.

La distinción entre los diversos modos de responsabilidad es de suma importancia cuando se trata de dictar sentencia.<sup>235</sup>

### *En concreto: ¿quién podría ser responsable por hacer qué?*

Es un hecho que la naturaleza única de los SAA requiere la participación de diversos actores en las diferentes etapas de su desarrollo, evaluación y durante su uso hasta el despliegue final.<sup>236</sup>

En este sentido, la autora sostiene que estos actores podrían agruparse a grandes rasgos en tres categorías:

- Creadores (fabricantes, desarrolladores, ingenieros en robótica, programadores);
- Usuarios (comandantes, soldados u operadores civiles); y
- Autorizadores (líderes civiles y militares).

En primer lugar, es primordial enfatizar que, en congruencia con los principios de rendición de cuentas previstos por el derecho internacional, la responsabilidad penal de una persona no afecta la responsabilidad de otra.<sup>237</sup>

---

<sup>235</sup> Jones y Powles, *International Criminal Practice*, *supra* nota 199, pp. 414 y 415.

<sup>236</sup> Heyns, Informe del Relator Especial, *supra* nota 148, p. 79.

<sup>237</sup> Estatuto de Roma, artículo 25 (4); TIPY, *Fiscal c. Tadic*, Sentencia de Apelación, 15 de julio de 1999, pp. 227-229; la Sala describe los elementos que se necesitan para cooperar o incitar.

Por consecuencia, el hecho de que se le pueda atribuir responsabilidad penal a un fabricante no excluye que los usuarios finales puedan, a su vez, asumir un tipo diferente de responsabilidad penal,<sup>238</sup> sin embargo, no son necesariamente dependientes una de la otra.

Tradicionalmente, para los efectos de imponer responsabilidad penal a un combatiente, el DIH y el DPI no se preocupan por el fabricante del arma utilizada.<sup>239</sup> Lo que les es de interés es el portador del arma, el individuo que eligió usar esa arma en particular o quién ordenó o autorizó su despliegue.<sup>240</sup> El razonamiento detrás de esto es que el combatiente es quien efectivamente tiene el control del arma y también quien toma las decisiones con respecto a qué arma usar.

Para el usuario final (la persona que despliega el arma), la regla de oro y la más básica es que nunca debe usar un arma cuyos efectos no pueda controlar.<sup>241</sup> Los combatientes sólo deben usar aquellas armas que no contravengan sus responsabilidades bajo el derecho

---

<sup>238</sup> TIPY, *Fiscal c. Delalić*, Sentencia de Apelación, 20 de febrero de 2001, p. 182; véase también Grodzinsky, Frances *et al.*, "Moral Responsibility for Computing Artifacts, 'the Rules' and Issues of Trust", *Computer Science & Information Technology*, 2012, p. 16 (en adelante: Grodzinsky *et al.*, "Moral Responsibility for Computing Artifacts"); la regla 2 prevé: "La responsabilidad compartida de los artefactos informáticos no es un juego de suma cero. La responsabilidad de un individuo no se reduce simplemente porque más personas se involucran en el diseño, desarrollo, implementación o uso del artefacto".

<sup>239</sup> Véase Chengeta, "Accountability Gap", *supra* nota 155, p. 35.

<sup>240</sup> Véase PAI, artículo 75(4)(b); PAIL, artículo 6(2)(b); CG IV, artículo 33; Convención de la Haya (IV) de 1907, Relativa a las Leyes y Costumbres de la Guerra Terrestre y su Reglamento, artículo 50; Henckaerts y Doswald, "Customary IHL", *supra* nota 125, pp. 372 y 373; la regla 102 versa sobre la responsabilidad penal individual. Weizmann *et al.*, "AWS under International Law", *supra* nota 155, p. 25; nótese la crítica al enfoque de Heyn: "por violar el principio fundamental de que no se puede imponer ninguna pena a una persona por un acto del que no es responsable".

<sup>241</sup> Protocolo Adicional I, artículo 51(4).

internacional.<sup>242</sup> Estas mismas limitaciones también se aplican a los líderes que sean responsables de tomar las decisiones sobre la autorización de qué armas pueden ser empleadas por sus fuerzas armadas.

Por supuesto, de manera adicional, los fabricantes pueden ser coautores, coadyuvantes o cómplices del crimen si se cumplen las condiciones requeridas.

Además, aquéllos en posiciones de liderazgo que autorizan su desarrollo, uso y/o despliegue también deben asumir otra forma de responsabilidad penal por su participación catalizadora. Está claro que, para poder ser justos, cada uno de estos actores debe ser responsable en su propia capacidad.

## Usuarios: responsabilidad de superiores jerárquicos

Ya que los comandantes son los actores más cercanos para ejercer un “mando y control efectivo”, algunos académicos opinan que el artículo 28 del Estatuto de Roma sobre la responsabilidad del superior jerárquico es el más adecuado para tratar con los operadores de SAA.<sup>243</sup>

Sin embargo, la autora opina firmemente que ésta no es una solución adecuada ni deseable.

---

<sup>242</sup> *Idem.*

<sup>243</sup> Véase, generalmente, Margulies, “Making Autonomous Weapons Accountable”, *supra* nota 159.

La responsabilidad del superior jerárquico se ha convertido en parte del derecho internacional consuetudinario,<sup>244</sup> y es ampliamente considerada como una herramienta para reforzar la disuasión, así como prevenir la impunidad.<sup>245</sup> Esta modalidad permite que los comandantes sean responsables de las acciones de sus subordinados por no haber prevenido o castigado la comisión de crímenes por parte de éstos.<sup>246</sup>

Tanto en el DIH como en el DPI, se entiende que un comandante es una persona natural que ejerce autoridad sobre personas naturales dentro de una operación militar.<sup>247</sup> Asimismo, el hecho de que el artículo 28 del Estatuto de Roma utilice términos como “fuerzas” y “subordinados” que pueden ser sometidos a enjuiciamiento y condenas, denota claramente que se refiere a los seres *humanos* en el campo de batalla,<sup>248</sup> no a las máquinas.<sup>249</sup>

En consecuencia, la responsabilidad de superiores jerárquicos no es aplicable a los individuos que despliegan SAA, dado que ninguna analogía puede ser trazada entre la relación del comandante y sus subordinados *humanos*, y la interacción de un humano operando una

<sup>244</sup> TIPY, *Fiscal c. Delalić*, Sentencia, 16 de noviembre de 1998, pp. 330-343 (en adelante: TIPY, Sentencia de Delalić); Jones y Powles, *International Criminal Practice*, *supra* nota 199, pp. 432 y 433.

<sup>245</sup> Funk, T. Markus, *Victim's Rights and Advocacy in the International Criminal Court*, 2010, p. 16.

<sup>246</sup> Swart, “Modes of International Criminal Liability”, *supra* nota 206, p. 88; véase International Criminal Law Services, *Modes of Liability: Superior Responsibility*, Module 10 of Training Materials, 2018, pp. 3-7.

<sup>247</sup> Jones y Powles, *International Criminal Practice*, *supra* nota 199, p. 424; Smidt, Michael L., “Yamashita, Medina, and Beyond; Command Responsibility in Contemporary Military Operations”, *Military Law Review*, vol. 164, 2000, pp. 168-69, 176 (en adelante: Smidt, “Yamashita, Medina and Beyond”).

<sup>248</sup> Véase Chengeta, “Accountability Gap”, *supra* nota 155, p. 32. Chengeta compara la definición de “comandante” del Estatuto de Roma y aquella en el derecho consuetudinario internacional; véase también Smidt, “Yamashita, Medina and Beyond”, *supra* nota 247, p. 176.

<sup>249</sup> *Idem*. Véase también Mettraux, Guénaël, *The Law of Command Responsibility*, 2009, pp. 5-11 (en adelante: Mettraux, *The Law of Command Responsibility*).

máquina.<sup>250</sup> La reiterada referencia de la persona que despliega SAA como *comandante* alimenta la impresión errónea de que los SAA son en cierto modo combatientes,<sup>251</sup> lo cual abona a la distorsión conceptual antropomórfica explicada anteriormente.<sup>252</sup>

Además, bajo el régimen del Estatuto de Roma, para que un comandante sea considerado responsable de las acciones de sus subordinados, son necesarios seis elementos básicos que deben satisfacerse:<sup>253</sup>

- 1) Las fuerzas armadas cometieron crímenes bajo la jurisdicción de la Corte;
- 2) El acusado es un comandante militar *de jure* o *de facto*;
- 3) El acusado tenía control efectivo sobre las fuerzas que cometieron los crímenes;
- 4) El acusado hubiere sabido o, debido a las circunstancias, hubiere debido saber que las fuerzas estaban cometiendo o estaban a punto de cometer tales crímenes;
- 5) El acusado no tomó todas las medidas necesarias y razonables a su alcance para prevenir o reprimir la comisión de tales crímenes ni sometió el asunto a las autoridades competentes para su investigación y enjuiciamiento; y
- 6) Los crímenes cometidos por las fuerzas deben haber sido resultado de la falta de un ejercicio de control apropiado.

---

<sup>250</sup> *Ibidem*, p. 50.

<sup>251</sup> Docherty, "Losing Humanity", *supra* nota 179, pp. 4, 33 y 34, 42 y 43.

<sup>252</sup> Véase *supra* en "¿Qué tipo de responsabilidad podemos esperar?".

<sup>253</sup> Estatuto de Roma, artículo 28; véase también, Protocolo Adicional I, artículos 86(2), 87; CPI, *Fiscal c. Jean Pierre Bemba Gombo*, Sentencia conforme al artículo 74, 21 de marzo de 2016, p. 170 (en adelante: CPI, Sentencia de Bemba).

Los elementos anteriores son el resultado de la jurisprudencia que ha sido perfeccionada cuidadosamente por varios tribunales y cortes penales internacionales.<sup>254</sup> Sin embargo, cabe señalar que esta jurisprudencia sufrió un giro inesperado después de la controvertida decisión de la Sala de Apelaciones de la CPI en el *Caso del Fiscal vs. Bemba Gombo*,<sup>255</sup> en la que el imputado fue absuelto, al menos en parte, por ser un mando a distancia que operaba en un país extranjero.<sup>256</sup>

En particular, en la sentencia de apelación, la Sala de Apelaciones se centró en el quinto elemento, es decir, si Bemba *tomó todas las medidas necesarias y razonables* dentro de su poder para prevenir o reprimir la comisión de tales crímenes o someter el asunto a las

---

<sup>254</sup> Véase CPI, Sentencia de Bemba, *supra* nota 253, pp. 170-213; TIPY, Sentencia de Delalić, *supra* nota 244, pp. 338-340; TIPY, *Fiscal c. Galik*, Sentencia, 5 de diciembre de 2003, p. 173; véase Swart, "Modes of International Criminal Liability", *supra* nota 206, pp. 88 y 89.

<sup>255</sup> CPI, *Fiscal c. Bemba Gombo*, Sentencia de Apelación en contra de la sentencia conforme al artículo 74, 8 de junio de 2018, pp. 167-171. Bemba fue absuelto de las convicciones de crímenes de lesa humanidad y crímenes de guerra con base en el artículo 28. La mayoría de la Sala de Apelaciones sostuvo que la responsabilidad de "tomar todas las medidas necesarias y razonables" estaba intrínsecamente conectada a las posibilidades materiales del comandante para prevenir la comisión de los crímenes o referirlos a autoridades competentes. Consideraron que para saber si el comandante tomó estas medidas, se deben considerar los crímenes de los que el comandante tenía conocimiento, además de que, según la mayoría, el comandante no debe forzosamente tomar todas las medidas previstas, pues se deben considerar las realidades operativas, enfatizando que existe un riesgo real que debe evitarse en la evaluación de lo que un comandante debe hacer con el beneficio de la retrospectiva, y que la simple yuxtaposición del hecho de que ciertos delitos fueron cometidos por sus subordinados, con una lista de medidas de lo que hipotéticamente pudo haber hecho, no demuestra, por sí mismo, que el comandante actuó de manera irrazonable. Se añadió que debe identificarse específicamente lo que el comandante debió haber hecho *en concreto*.

<sup>256</sup> *Ibidem*, pp. 170 y 171. La mayoría consideró que en el fallo de primera instancia no se prestó suficiente atención a que las tropas de Bemba operaban en el extranjero, causando dificultades a Bemba como comandante remoto.

autoridades competentes para su investigación y enjuiciamiento, y a diferencia de la Sala de Primera Instancia, llegó a la conclusión de que sí lo hizo.<sup>257</sup>

De lo anterior, se desprende que el umbral del “control efectivo” requiere que el comandante tenga la capacidad material para prevenir o reprimir la comisión de crímenes o someter el asunto a las autoridades competentes.<sup>258</sup> Sin embargo, como se ha dicho anteriormente, las máquinas no tienen agencia moral<sup>259</sup> y, por lo tanto, no pueden ser sancionadas.<sup>260</sup>

En ese sentido, la naturaleza de la responsabilidad de superiores no permite que los comandantes abduquen sus obligaciones morales y legales de determinar si el uso de la fuerza es apropiado en una situación dada.<sup>261</sup> Cuando delegan obligaciones a un subordinado, aún conservan el deber de supervisar la conducta de ese agente humano responsable. En consecuencia, mientras que los SAA no sean agentes humanos responsables, los comandantes no pueden delegarles ninguna autoridad.<sup>262</sup>

Otro requisito para esta forma de responsabilidad es la evidencia de que el comandante hubiera sabido o, dadas las circunstancias del momento, hubiere debido saber que los crímenes estaban a punto de cometerse o se estaban cometiendo.<sup>263</sup> Si bien esto reduce el

---

<sup>257</sup> *Ibidem*, pp. 120-136, 184-194.

<sup>258</sup> *Ibidem*, p. 167; véase también, CPI, Sentencia de Bemba, *supra* nota 253, pp. 183 y 184.

<sup>259</sup> Véase *supra* el apartado “¿Qué tipo de responsabilidad podemos esperar?”.

<sup>260</sup> Chengeta, “Accountability Gap”, *supra* nota 155, p. 11; véase también Wagner, “Taking Humans out of the Loop”, *supra* nota 168, pp. 5, 11; Asaro, “On Banning AWS”, *supra* nota 168, p. 693; Himma, Kenneth Einar, “Artificial Agency, Consciousness, and the Criteria for Moral Agency: What Properties Must an Artificial Agent Have to be a Moral Agent?”, *Ethics & Information Technologies*, vol. 11, 2009, pp. 19-29.

<sup>261</sup> Asaro, “On Banning AWS”, *supra* nota 168, p. 701.

<sup>262</sup> *Idem*.

<sup>263</sup> Véase Estatuto de Roma, artículo 28; CPI, Sentencia de Bemba, *supra* nota 253, pp. 50-53, 170, 196.

requisito del elemento subjetivo hasta compararse con un umbral de negligencia, dada la naturaleza impredecible de los SAA, algunos autores argumentan que los posibles operadores podrían defenderse argumentando que, a su saber y entender, los SAA cumplirían con los principios del DIH para la selección de objetivos.

Sin embargo, la autora contraargumenta, al considerar que los usuarios deberían tener —como mínimo— una buena comprensión de las capacidades de la IA, y por lo tanto, del riesgo de impredecibilidad antes de usar el arma.

Contrario a aquellos argumentos defensivos, el resultado de la impredecibilidad de un SAA con niveles de autonomía altos o completos, funcionando en entornos no estructurados,<sup>264</sup> es que una vez desplegado, todas sus acciones son atribuibles al individuo que lo desplegó.<sup>265</sup> En este sentido, el despliegue de un SAA por sí solo es un ejercicio de control suficiente por parte del usuario.<sup>266</sup>

La opinión de la autora es que, si hay margen para cualquier impredecibilidad en relación con el despliegue de un arma como tal y como con los SAA, entonces es razonablemente predecible esperar el peor escenario, lo cual generaría la responsabilidad penal correspondiente por los crímenes cometidos.<sup>267</sup> En otras palabras, cuando un crimen se comete como resultado del despliegue de un SAA, la responsabilidad penal recaería sobre el individuo que lo desplegó.<sup>268</sup>

---

<sup>264</sup> Chengeta, “Accountability Gap”, *supra* nota 155, p. 34. Chengeta considera que al defender los SAA, Schmitt ignora el problema de la impredecibilidad; véase, generalmente, Schmitt, “AWS and IHL”, *supra* nota 29.

<sup>265</sup> Véase Schmitt, “AWS and IHL”, *supra* nota 29, pp. 16 y 17, 33.

<sup>266</sup> Véase Chengeta, “Accountability Gap”, *supra* nota 155, p. 34.

<sup>267</sup> Schmitt, “AWS and IHL”, *supra* nota 29, pp. 16-18, 33.

<sup>268</sup> Weizmann *et al.*, “AWS under International Law”, *supra* nota 155, pp. 24 y 25.

Por lo tanto, en la etapa de su desarrollo, no se les debe dar autonomía ni funciones que hagan que dejen de ser armas para convertirse, *de facto*, en robots combatientes. Más bien, los SAA siempre deben desarrollarse de manera que sigan siendo armas en manos de un combatiente que pueda ser responsable en los casos en que se cometan crímenes.<sup>269</sup>

Es comprensible que los conceptos legales a veces se puedan ajustar para abordar nuevas situaciones, sin embargo, respecto al intento de igualar ficticiamente a los SAA con combatientes, el concepto de responsabilidad de superiores jerárquicos no puede estirarse tanto sin perder su esencia.

En la opinión de la autora, el único caso en el que el tema de la responsabilidad del superior jerárquico es relevante, es cuando el comandante o el civil que supervisa al individuo que programa o que despliega un SAA sabía o debería haber sabido que su subordinado estaba programando o utilizando un SAA de forma ilegal, y no hizo nada para prevenir o detenerlo o sancionarlo después del hecho.<sup>270</sup> Lo anterior obedece la misma línea de razonamiento aplicable en relación con otras armas.

En consecuencia, la responsabilidad del superior no puede ser aplicable a una relación *humano-máquina* porque no existe un fundamento legal para asignar el estatus de combatiente a los SAA, éstos son armas y quienes las despliegan son los combatientes. En conclusión, y desde una perspectiva legal, el SAA no puede ni debe poder cometer crímenes. Como dijo Séneca, “una espada nunca es un asesino, es una herramienta en las manos del asesino”.<sup>271</sup>

---

<sup>269</sup> Sassóli, “Autonomous Weapons and IHL”, *supra* nota 52, pp. 308, 324.

<sup>270</sup> Mettraux, *The Law of Command Responsibility*, *supra* nota 249, p. 55; Schmitt, “AWS and IHL”, *supra* nota 29, pp. 33 y 34.

<sup>271</sup> Letters to Lucilius, 1st c., citado en Thomsett, Michael C. y Thomsett, Jean F. (eds.), *War and Conflict Quotes*, 1997, p. 158.

No obstante, a la luz de los obstáculos explicados anteriormente, un grupo diferente de académicos ha sugerido una revisión de la doctrina de la responsabilidad del superior jerárquico, a fin de facilitar la identificación del verdadero culpable. Esto requeriría bajar el estándar del elemento subjetivo.

En esta línea de pensamiento, se argumenta que, con “una modesta revisión de la doctrina” para extender su aplicación a la supervisión de máquinas,<sup>272</sup> la responsabilidad del mando se aplicaría si hubiera un requisito de “diligencia dinámica” por parte de los comandantes.<sup>273</sup>

Este enfoque requeriría por lo menos lo siguiente: una estructura de comando dedicada, pericia técnica, monitoreo humano en tiempo real (incluida una capacidad del SAA para solicitar una revisión), revisión periódica y frecuente del rendimiento, la entrada de parámetros dinámicos que rigen el uso de SAA con relación al tiempo, distancia y el máximo de daño colateral esperado, así como que las decisiones de selección de objetivos sean transparentes e interpretables para los humanos.<sup>274</sup>

Esta alternativa busca ofrecer una solución más adecuada al caracterizar normativamente al comandante como un perpetrador directo, sin necesidad de apoyarse en una doctrina retorcida de responsabilidad de mando que equipara una máquina autónoma con un subordinado.<sup>275</sup>

---

<sup>272</sup> Margulies, “Making Autonomous Weapons Accountable”, *supra* nota 159, p. 441.

<sup>273</sup> *Ibidem*, p. 406.

<sup>274</sup> *Ibidem*, pp. 431-440; véase Hauptman, Allyson, “Autonomous Weapons and the Law of Armed Conflict”, *Military Law Review*, vol. 218, invierno de 2013, pp. 194 y 195; CICR, Ethics and AWS, *supra* nota 122, pp. 1, 5; Stewart, “New Technology and the Law of Armed Conflict”, *supra* nota 158, pp. 291 y 292.

<sup>275</sup> Se llegó a conclusiones similares en: HRW, “Mind the Gap”, *supra* nota 158; Chengeta, “Accountability Gap”, *supra* nota 155, pp. 31-34; Roff, Heather M., “Killing in War: Responsibility, Liability, and Lethal Autonomous Robots”, in Allhoff, Fritz *et al.* (eds.), *Routledge Handbook of Ethics and War: Just War Theory in the Twenty-First*

En cualquier caso, considerando todo lo anterior, a pesar de que la responsabilidad del superior jerárquico podría ser vista por algunos como una solución atractiva dado que reconocería la autonomía de un SAA, en este estudio se sostiene que es una vía inadecuada para cerrar la brecha de rendición de cuentas, dados los desafíos legales actuales para cumplir con los criterios doctrinales de dicha figura.

### Creadores: comisión por cooperación

Por otro lado, algunos académicos han examinado la responsabilidad del fabricante, desarrollador o programador de un SAA, argumentando que éstos ejercerán un mayor control no sólo sobre la gama de acciones que el sistema de armas es capaz de realizar, sino también sobre las acciones específicas que efectivamente realizará después de su despliegue.<sup>276</sup>

La autora señala que es más probable que las acciones de este tipo de actores les conciernan a los tribunales nacionales, a menos que su conducta satisfaga todos los elementos constitutivos de un crimen dentro de la jurisdicción de la CPI.<sup>277</sup>

---

*Century*, Routledge, 2013, pp. 352, 358; Amoroso, Daniele y Tamburrini, Guglielmo, "Autonomous Weapon Systems and Meaningful Human Control: Ethical and Legal Issues", *Current Robotics Reports*, 2020, p. 19 (en adelante: Amoroso y Tamburrini, "AWS and Meaningful Human Control"); Crotoft, Rebecca, "War Torts: Accountability for Autonomous Weapons", *University of Pennsylvania Law Review*, 2016, pp. 1379-1381.

<sup>276</sup> McFarland, Tim y McCormack, Tim, "Mind the Gap: Can Developers of Autonomous Weapon Systems be Liable for War Crimes?", *International Law Studies*, vol. 90, 2014, p. 366 (en adelante: McFarland y McCormack, "Mind the Gap").

<sup>277</sup> Estatuto de Roma, artículo 5; véase CPI, Elementos de los Crímenes, *supra* nota 226.

En ese caso, se alega que el artículo 25(3)(c) del Estatuto de Roma, que se refiere a la cooperación,<sup>278</sup> sería el más adecuado para tratar con diseñadores y fabricantes de SAA.<sup>279</sup>

Esta disposición establece una forma de responsabilidad accesoria donde siempre se requiere la intención,<sup>280</sup> *i.e.*, el “propósito” de facilitar el crimen, ya que el mero conocimiento no es suficiente para imponer responsabilidad bajo este artículo.<sup>281</sup> Según la Corte, lo que se requiere es que la persona brinde asistencia para la comisión de un crimen y que, al incurrir en esta conducta, pretenda facilitar la comisión del mismo.<sup>282</sup> Dicha asistencia no necesita ser “sustancial”,<sup>283</sup> ya que la responsabilidad de los accesorios o cómplices, a los

---

<sup>278</sup> Cabe señalar que la cooperación e instigación se derivan del lenguaje usado en la versión en inglés del Estatuto “*aids, abets, or otherwise assists*”, mientras que la versión en español habla de “ser cómplice o encubridor o colabore de algún modo en la comisión [del crimen]”, no obstante, en ambos supuestos el artículo recoge la cooperación como forma de responsabilidad accesoria.

<sup>279</sup> McFarland y McCormack, “Mind the Gap”, *supra* nota 276, pp. 376-378.

<sup>280</sup> En algunos casos, “la intención” requerida por el artículo 25(3)(c) ha sido establecida por la Corte analizando el artículo 25(3)(d), que simplemente requiere “conocimiento” en oposición a *inter alia*, lo requerido por el artículo 25(3)(c); por ejemplo: CPI, *Fiscal c. Lubanga Dyilo*, Decisión de confirmación de cargos, 29 de enero de 2007, p. 337 (en adelante: CPI, Confirmación de Cargos de Lubanga); CPI, *Fiscal c. Mbarushimana*, Decisión relativa a la solicitud del fiscal de una orden de arresto en contra de Callixte Mbarushimana, 28 de septiembre de 2010, pp. 38-39; CPI, *Fiscal c. Mbarushimana*, Decisión de confirmación de cargos, 16 de diciembre de 2011, p. 289 (en adelante: CPI, Confirmación de Cargos de Mbarushimana).

<sup>281</sup> CPI, Confirmación de Cargos de Mbarushimana, *supra* nota 280, p. 274; véase también CPI, *Fiscal c. Ngudjolo Chui*, Sentencia conforme al artículo 74 del Estatuto, Opinión concurrente de la juez Christine Van den Wyngaert, 18 de diciembre de 2012, p. 25; analiza el hecho de que el Estatuto de Roma añade un elemento mental más estricto para ayudar o instigar (es decir, la intención o propósito), ya que el artículo 7(1) del Estatuto del TIPY, meramente requería “conocimiento”.

<sup>282</sup> CPI, *Fiscal c. Blé Goudé*, Decisión de confirmación de cargos, 11 de diciembre de 2014, p. 167 (en adelante: CPI, Confirmación de Cargos de Blé Goudé).

<sup>283</sup> CPI, *Fiscal c. Dominic Ongwen*, Decisión de confirmación de cargos, 23 de marzo de 2016, p. 43; CPI, *Fiscal c. Al Mahdi*, Versión pública de la Decisión de confirmación de cargos, 24 de marzo de 2016, p. 26.

que este artículo hace referencia, requiere una contribución menor que la de los autores principales.<sup>284</sup>

Concretamente, el cómplice debe “cooperar”, lo cual implica dar asistencia práctica o material, proporcionando así los medios para la comisión de un crimen. Por otro lado, también puede “instigar”, lo cual significa brindar asistencia moral o psicológica al perpetrador principal, alentando o mostrando simpatía por la comisión del crimen, la cual no necesita ser explícita.<sup>285</sup>

La Corte ha señalado que este tipo de asistencia debe tener un efecto en la comisión del crimen, sin embargo, la contribución no requiere un umbral específico y la participación del cómplice no tiene por qué ser una condición *sine qua non* para la comisión del crimen principal. Los únicos requisitos son que el individuo haya fomentado o facilitado la comisión de dicho crimen, antes, durante o después del hecho,<sup>286</sup> y que haya sido con el propósito de hacerlo.<sup>287</sup>

Es importante recordar que este modo de responsabilidad es accesorio, y como tal deriva de la conducta de un autor principal.<sup>288</sup> Esto significa que depende de la comisión, o al menos de la tentativa de comisión, de un crimen por parte del autor principal, aunque no se requiere que este último sea identificado, acusado o condenado.<sup>289</sup>

---

<sup>284</sup> CPI, *Fiscal c. Lubanga Dyilo*, Sentencia conforme al artículo 74 del Estatuto, 14 de marzo de 2012, pp. 997 y 998.

<sup>285</sup> CPI, Sentencia de Bemba, *supra* nota 253, pp. 88 y 89.

<sup>286</sup> CPI, *Fiscal c. Jean Pierre Bemba et al.*, Sentencia conforme al artículo 74 del Estatuto, 19 de octubre de 2016, p. 96 (en adelante: CPI, Sentencia de Bemba *et al.*).

<sup>287</sup> *Ibidem*, pp. 90-97.

<sup>288</sup> Véase CPI, Confirmación de cargos de Lubanga, *supra* nota 280, p. 337.

<sup>289</sup> CPI, Sentencia de Bemba *et al.*, *supra* nota 286, pp. 83-85.

Además, se argumenta que ni siquiera sería necesario demostrar la existencia de un plan común entre el fabricante y la persona que despliega el SAA. Según la jurisprudencia existente, debido a que el que coopera o instiga siempre será cómplice de un crimen perpetrado por otra persona,<sup>290</sup> no se requiere prueba de la existencia de un plan común y mucho menos de la preexistencia de tal plan.<sup>291</sup> Es posible que la persona que despliega el SAA (el autor principal) ni siquiera sepa sobre la contribución del cómplice (el fabricante o programador).<sup>292</sup>

No obstante, la autora identifica dos obstáculos principales para el enjuiciamiento de los creadores como cómplices.

El primero es el requisito de la *mens rea*, pues debe probarse que actúan con:<sup>293</sup>

- a) Conocimiento de la intención eventual del autor principal para cometer el crimen;
- b) Conocimiento de que su conducta ayudaría a la comisión del crimen; y
- c) El propósito de facilitar dicho crimen.

El segundo es que, si se intenta presentar cargos por crímenes de guerra, en la mayoría de los casos será difícil establecer el elemento contextual requerido, relativo a que la conducta del creador tuvo lugar en el contexto de un conflicto armado y estuvo asociada con éste, porque generalmente, su contribución se completaría en la fase de desarrollo del arma, lo que probablemente podría ocurrir antes del comienzo del conflicto armado en cuestión.<sup>294</sup>

---

<sup>290</sup> TIPY, *Fiscal c. Kordić & Erkez*, Sentencia, 26 de febrero de 2001, p. 399.

<sup>291</sup> Véase TIPY, *Fiscal c. Tadić*, Sentencia de apelación, 15 de julio de 1999, pp. 227-229.

<sup>292</sup> Chengeta, "Accountability Gap", *supra* nota 155, p. 22.

<sup>293</sup> McFarland y McCormack, "Mind the Gap", *supra* nota 276, p. 380; véase CPI, Confirmación de cargos de Blé Goudé, *supra* nota 282.

<sup>294</sup> *Ibidem*, pp. 372-374.

Por lo tanto, debe haber una revisión del elemento contextual, el cual debería enmendarse para incluir explícitamente, o interpretarse para incluir los actos de preparación previos al inicio del conflicto armado, siempre que la consumación del crimen se haya producido en el contexto pertinente.<sup>295</sup>

Entonces, es posible que la responsabilidad por complicidad, tal y como se encuentra regulada actualmente, no proporcione una solución satisfactoria, y por lo tanto, también debe estar sujeta a adaptaciones. Tiene que considerarse que un ser humano siempre debe decidir cómo programar el sistema, y claramente ese individuo debe también ser responsable por programarlo para llevar a cabo actos que equivalieran a crímenes de guerra.<sup>296</sup> Se ha señalado que es un deber del creador garantizar que los SAA sean seguros, lo más posible, tanto para los combatientes como para los no combatientes.<sup>297</sup>

### Consideraciones para reformas

Como se puede observar arriba, hay una falta de paralelos aplicables en las disposiciones del Estatuto de Roma. En consecuencia, para poder cerrar adecuadamente la brecha de impunidad, la comunidad internacional debe dar pie a adaptaciones legales que abarquen las nuevas realidades operativas, ya sea incorporando un crimen específico para los SAA o aceptando un modo de responsabilidad compatible con éstos.

---

<sup>295</sup> *Ibidem*, p. 384.

<sup>296</sup> Schmitt, "AWS and IHL", *supra* nota 29, p. 33.

<sup>297</sup> Arkin, Ronald C., *Governing Lethal Behavior: embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture*, 2011, p. 9.

Lo anterior puede suceder de dos maneras, ya sea mediante una ampliación al marco legal existente a través de interpretaciones jurisdiccionales o con la incorporación normativa de nuevos elementos.

Dado que lo primero sólo puede ocurrir *ex post facto*, es imperativo movilizar los esfuerzos necesarios para materializar lo segundo.

Para ello, el artículo 21(1)(c) del Estatuto reconoce que, en casos excepcionales, los principios generales del derecho, derivados de las leyes nacionales de los sistemas jurídicos del mundo, podrían ser leyes aplicables por la Corte.<sup>298</sup> La autora considera que la laguna legal que rodea a los SAA merecería directamente invocar esta disposición.

En consecuencia, debemos echar un vistazo a los modelos de responsabilidad penal que prevalecen en la mayoría de los sistemas legales nacionales para encontrar el esquema más adecuado para los SAA.

Es importante aclarar que los dos modelos que se explican a continuación no son alternativos entre sí, podrían aplicarse de manera coordinada y simultánea para dar pie a un régimen completo de responsabilidad penal en el contexto específico de participación de un sistema de IA.<sup>299</sup> Como resultado, cuando un SAA y seres humanos estén involucrados directa o indirectamente en la perpetración de un crimen específico, sería mucho más difícil evadir la responsabilidad penal.

---

<sup>298</sup> Si bien la Corte no ha tenido que resolver muchas cuestiones sobre el artículo, puede hacerlo como se demuestra en: CPI, *Fiscal c. Katanga*, Versión pública de la apelación relativa a la orden de la Sala II del 24 de marzo de 2017, titulada "Orden de reparaciones conforme al artículo 75 del Estatuto", 8 de marzo de 2018, p. 148; CPI, *Fiscal c. Lubanga*, Decisión relativa a las prácticas usadas para preparar y familiarizar a testigos, 1 de diciembre de 2007, pp. 40 y 41.

<sup>299</sup> Hallevy, Gabriel, "The Criminal Liability of Artificial Intelligence Entities — From Science Fiction to Legal Social Control", *Akron Law Journals*, 2016, p. 196 (en adelante: Hallevy, "The Criminal Liability of AI").

## Modos de responsabilidad

### Autoría mediata

En la mayoría de los sistemas legales,<sup>300</sup> cuando una persona causa que un agente inocente cometa un delito —i.e., cuando se da cauce a que un niño,<sup>301</sup> alguien con alguna discapacidad mental,<sup>302</sup> o una persona que carece de intención criminal participe en tal conducta—,<sup>303</sup> dicha persona es penalmente responsable como autor mediato.<sup>304</sup> En tales casos, el intermediario se considera un mero instrumento y el autor mediato es el verdadero autor.<sup>305</sup> Ese autor mediato es responsable de la conducta del agente inocente, y la responsabilidad se determina con base en la conducta<sup>306</sup> y a su estado mental.<sup>307</sup>

---

<sup>300</sup> Véase Fletcher, G. P., *Rethinking Criminal Law*, New York, Oxford University Press, 2000, p. 639; véase CPI, *Fiscal c. Germain Katanga and Mathieu Ngudjolo Chui*, Decisión de confirmación de cargos, 30 de septiembre de 2008, p. 495 (en adelante, CPI, Katanga y Chui).

<sup>301</sup> *Maxey v. United States*, 30 App. D.C. 63 (App.D.C.1907); *Commonwealth v. Hill*, 11 Mass. 136 (1814); *R v. Michael*, (1840) 2 Mod. 120, 169 E.R. 48.

<sup>302</sup> *Johnson v. State*, 142 Ala. 70, 38 So. 182 (1904); *People v. Monks*, 133 Cal. App. 440, 24 P.2d 508 (Cal. App.4Dist.1933).

<sup>303</sup> *United States v. Bryan*, 483 F.2d 88 (3rd Cir.1973); *Boushea v. United States*, 173 F.2d 131 (8th Cir.1949).

<sup>304</sup> *Morrisey v. State*, 620 A.2d 207 (Del.1993); *Conyers v. State*, 367 Md. 571, 790 A.2d 15 (2002); CPI, Katanga y Chui, *supra* nota 300, p. 495.

<sup>305</sup> Hallevy, "The Criminal Liability of AI", *supra* nota 299, p. 179.

<sup>306</sup> *Dusenbery v. Commonwealth*, 220 Va. 770, 263 S.E.2d 392 (1980).

<sup>307</sup> *United States v. Tobon-Builes*, 706 F.2d 1092 (11th Cir.1983); *United States v. Ruffin*, 613 F.2d 408 (2nd Cir.1979).

Sobre este punto, ha surgido un enfoque bastante controversial que establece una analogía entre el uso belicoso de un SAA y el reclutamiento y uso de niños soldados,<sup>308</sup> pues ambos no son agentes “conscientes” que cometen un delito, y por consiguiente están exentos de responsabilidad por su participación o comisión de crímenes internacionales.<sup>309</sup> El símil se da por el hecho de que, aunque los niños soldados son autónomos —quizás incluso más que un SAA— “carecen de total autonomía moral”,<sup>310</sup> lo cual vicia su comprensión de las complejas dimensiones morales de sus actos, convirtiéndolos en sujetos incapaces e inimputables,<sup>311</sup> y por lo tanto, inelegibles para el papel de combatiente, al igual que los SAA.<sup>312</sup>

Al respecto, el artículo 26 del Estatuto de Roma establece que la Corte no tendrá jurisdicción sobre ninguna persona que sea menor de 18 años en el momento de la presunta comisión de un crimen. Aunque, en opinión de la autora, es bastante arriesgado equiparar a niños humanos con SAA, ya que la tipificación del reclutamiento y la utilización de niños soldados como crimen tiene como fin la protección de los menores de edad,<sup>313</sup> no de aquéllos a quienes éstos podrían llegar a dañar. En este sentido, sería extremadamente inapropiado otorgar a los SAA la misma consideración de “agente inocente” que se otorga a los niños.

Además, los académicos que promueven esta postura afirman que el vacío que puede generar lo anterior, en términos de responsabilidad penal, se evita mediante la clara

---

<sup>308</sup> Sparrow, “Killer Robots”, *supra* nota 45, pp. 73 y 74; véase Henckaerts y Doswald, “Customary IHL”, *supra* nota 125, pp. 482-485; en lo referente a la regla 136, sobre el reclutamiento de niños soldados.

<sup>309</sup> Liu, “Refining Responsibility”, *supra* nota 163, pp. 343 y 344; Davison, “AWS under IHL”, *supra* nota 44, p. 17.

<sup>310</sup> Sparrow, “Killer Robots”, *supra* nota 45, p. 73.

<sup>311</sup> *Ibidem*, p. 73.

<sup>312</sup> Henckaerts y Doswald, “Customary IHL”, *supra* nota 125, pp. 482-485, Rule 136.

<sup>313</sup> Liu, “Refining Responsibility”, *supra* nota 163, p. 343.

prohibición de introducir niños soldados en cualquier conflicto armado.<sup>314</sup> Afirman que un individuo no sería responsable por los crímenes cometidos por los niños soldados, más bien sería responsable por haberlos introducido en el campo de batalla, siendo entidades irresponsables.<sup>315</sup> Esta línea de razonamiento podría servir para promover la propuesta de crimen específico sobre los SAA que se detalla más abajo,<sup>316</sup> o al menos una postura de prohibición *ab initio* de los SAA.

En cualquier caso, la verdadera pregunta es ¿quién es el autor mediato? Como se indicó anteriormente, la autora considera tres posibles grupos de personas, los creadores, los usuarios y los que autorizan.

Respecto a esas tres posibilidades, el *actus reus* del crimen ha sido realizado por el sistema de IA. El modelo de responsabilidad de autoría mediata considera la conducta cometida por el SAA como si fuera del programador, del usuario o de quién autoriza sobre la base de su uso instrumental como agente inocente,<sup>317</sup> ya que legalmente es simplemente una máquina.

Sin embargo, se podría argumentar que este modelo encuentra mayor complejidad cuando el sistema de IA no ha sido diseñado específicamente para cometer el crimen en cuestión, y lo cometió de acuerdo con sus capacidades de aprendizaje profundo, es decir, la experiencia o conocimiento que ha adquirido por sí mismo.

---

<sup>314</sup> *Idem.*

<sup>315</sup> *Ibidem*, pp. 343 y 344.

<sup>316</sup> Véase *infra* el apartado “Crimen específico de los sistemas de armas autónomas”.

<sup>317</sup> El sistema de IA se utiliza como instrumento y no como participante, aunque utiliza sus características de procesamiento de información; véase, e.g., Cross, George R. y Debessonnet, Cary G., “An Artificial Intelligence Application in the Law: COLIPS, a Computer Program that Processes Legal Information”, *High Tech. Law Journal*, vol. I, 1986, p. 329.

En este sentido, algunos académicos argumentarían que, considerando que este modelo requiere la intención de los programadores o de los usuarios de cometer un crimen a través del SAA utilizando algunas de sus capacidades de manera instrumental, si cupiera la posibilidad del que el sistema de IA sea considerado un “agente semiinocente” en razón a un cierto grado de autonomía, este modelo de responsabilidad podría ser cuestionado.<sup>318</sup>

## Responsabilidad por consecuencia natural probable

Para los casos en los que el modelo anterior no pueda proporcionar una solución adecuada, el modelo de responsabilidad por consecuencia natural probable podría entrar en juego. Este modelo también podría ser aplicable como respuesta al estado actual de los debates, en los que el argumento de la “impredecibilidad” se utiliza como excusa para la responsabilidad penal.

Este modelo de responsabilidad considera el escenario en el que efectivamente los creadores y/o usuarios se involucran en el funcionamiento de la IA, y/o los autorizadores en su uso y despliegue, pero éstos no pretenden cometer ningún delito con la máquina y no saben de la comisión del crimen hasta que ya ha ocurrido, sin planearlo ni participar en el mismo.

En concreto, en lo que se basa este modelo es en la capacidad de los creadores, usuarios y/o autorizadores de prever la potencial comisión del crimen, responsabilizándolos en la medida en que dicho crimen sea una consecuencia natural y probable de la conducta de

---

<sup>318</sup> Lacey, Nicola y Wells, Celia, *Reconstructing Criminal Law — Critical Perspectives on Crime and the Criminal Process*, 2a. ed., 1998, p. 53.

esa persona, *i.e.*, la creación, uso o autorización de un SAA. A grandes rasgos, este enfoque implica rebajar el elemento mental a algo parecido a la imprudencia o la negligencia.<sup>319</sup>

Tradicionalmente, esta doctrina se utiliza para imponer responsabilidad penal a los cómplices o perpetradores negligentes.<sup>320</sup> Con respecto al primero, la norma establecida por tribunales y comentaristas es que la responsabilidad del cómplice se extiende a los actos del autor que fueran consecuencia natural y probable<sup>321</sup> de un esquema delictivo al cual el cómplice hubiera ayudado o incitado.<sup>322</sup> Esto ha sido ampliamente aceptado en los estatutos y recodificaciones<sup>323</sup> de responsabilidad penal de los cómplices, incluidos los estatutos de la CPI y de los tribunales *ad hoc*.<sup>324</sup>

---

<sup>319</sup> Véase, *e.g.*, Ohlin, “The Combatant’s Stance”, *supra* nota 175, pp. 21-23; véase McDougall, “AWS and Accountability”, *supra* nota 156, p. 22. McDougall discute las ideas de Ohlin, al mencionar que los comandantes deberían ser enjuiciados con base en una perpetración indirecta. Bajo el artículo 25(3)(a) del Estatuto de Roma, los individuos son responsables penalmente por actuar a través de una organización que controlaban de tal forma, que sus órdenes resultaron en el crimen y fueron llevadas a cabo por la organización. Ergo, Ohlin considera que la doctrina puede ser reorientada para “cambiar el lenguaje metafórico de las máquinas a uno literal sobre responsabilidad de las máquinas”. McDougall, menciona además que Ohlin añade esta complicación, igualando un SAA a un soldado subordinado, para ver a la máquina como un agente culpable.

<sup>320</sup> Hallevey, “The Criminal Liability of AI”, *supra* nota 299, 2016, p. 184.

<sup>321</sup> *United States v. Powell*, 929 F.2d 724 (D.C.Cir.1991).

<sup>322</sup> Clark, William M. y Marshall, William L., *Law of Crimes*, 7a. ed., 1967, p. 529; Bowes Sayre, Francis, “Criminal Responsibility for the Acts of Another”, *Harv. L. Rev.*, vol. 43, 1930, p. 689; *People v. Prettyman*, 14 Cal.4th 248, 58 Cal.Rptr.2d 827, 926 P.2d 1013 (1996); *Chance v. State*, 685 A.2d 351 (Del.1996).

<sup>323</sup> Hallevey, “The Criminal Liability of AI”, *supra* nota 299, pp. 241-247; véase *State v. Kaiser*, 260 Kan. 235, 918 P.2d 629 (1996); *United States v. Andrews*, 75 F.3d 552 (9th Cir.1996).

<sup>324</sup> Bergsmo, Morten y Stahn, Carten (eds.), *Quality Control in Preliminary Examination: Volume 2*, Bruselas, TOAEP, 2018, p. 199. “[la cooperación e incitación] ha sido aceptada en el DPI y se ha incluido en los estatutos de todos los tribunales y cortes creados posteriormente a la guerra fría, [...] de hecho, [este modo de responsabilidad] ha criminalizado la conducta de llevar a cabo la guerra por proxis (cuando estos proxis cometen crímenes)”.

En relación con el segundo, el modelo de responsabilidad por consecuencia natural probable requiere que el perpetrador esté en un estado mental de negligencia, nada más.<sup>325</sup> Los creadores, usuarios y/o autorizadores no están obligados a saber sobre la comisión futura de un crimen como resultado de su actividad, ya que dicha comisión es una consecuencia natural probable de sus actos.<sup>326</sup>

Una persona negligente, en un contexto criminal, es una persona que no sabe sobre el crimen ni pretende ese resultado, pero está en una situación en la que una persona razonable podría haberlo sabido, ya que el crimen específico es una consecuencia natural probable de la conducta de esa persona.<sup>327</sup> La negligencia actualiza una omisión de conciencia o una omisión de conocimiento, no de actos.

El modelo de responsabilidad por consecuencia natural probable ofrecería un modelo de responsabilidad alternativo por crímenes cometidos por SAA, basados en negligencia, cuando los elementos de los crímenes subyacentes requieren una *mens rea* diferente.<sup>328</sup> La lógica detrás de esto es que un creador, usuario o autorizador razonable podría haber previsto la comisión del crimen y, por lo tanto, tuvo la oportunidad de prevenirlo en la etapa de origen.

Cuando el SAA lleva a cabo el *actus reus* del crimen, el individuo en cuestión puede ser considerado negligente, aún si ningún ilícito fuera planeado deliberadamente, o podría ser

---

<sup>325</sup> Hallevey, "The Criminal Liability of AI", *supra* nota 299, p. 183.

<sup>326</sup> *Idem*.

<sup>327</sup> Fine, Robert P. y Cohen, Gary M., "Is Criminal Negligence a Defensible Basis for Criminal Liability?", *Buff. L. Rev.*, vol. 16, 1966, p. 749; Hart, Herbert L. A., "Negligence, Mens Rea and Criminal Responsibility", *Oxford Essays in Jurisprudence*, 1961, p. 29; Stuart, Donald, "Mens Rea, Negligence and Attempts", *Crim. L. R.*, 647, 1968.

<sup>328</sup> American Law Institute, *The Model Penal Code — Official Draft and Explanatory Notes*, 1985, p. 312; *State v. Linscott*, 520 A.2d 1067 (Me.1987).

considerado plenamente responsable por un crimen en específico, si éste derivara de otro que sí fuera planeado deliberadamente, aunque el primero no formara parte del esquema criminal original.<sup>329</sup>

Sin embargo, este enfoque debe ser cuidadosamente modulado, ya que puede llegar a parecerse a lo que se conoce como *responsabilidad objetiva* (*strict liability*, en inglés). En algunos sistemas de derecho penal, existe responsabilidad objetiva cuando una persona es responsable por cometer una acción, independientemente de cuál fue su intención al cometer la acción.<sup>330</sup> La lógica detrás de esto es que la conciencia del perpetrador sobre lo que está haciendo no niega el hecho de que llevó a cabo la conducta en cuestión.<sup>331</sup>

Este concepto genera cierta controversia, algunos académicos se oponen a él por considerar injusto que una persona sea considerada responsable de actos más allá de sus intenciones, o incluso pese a la falta de ellas.<sup>332</sup> En cualquier caso, la responsabilidad objetiva generalmente resulta en castigos más indulgentes que otras *mentes reae*, lo cual podría mitigar los argumentos de la percibida injusticia.<sup>333</sup>

---

<sup>329</sup> Hallevey, "The Criminal Liability of AI", *supra* nota 299, pp. 184 y 185.

<sup>330</sup> Véase, por ejemplo, Legal Information Institute, *Strict liability*, Cornell Law School, disponible en [https://www.law.cornell.edu/wex/strict\\_liability](https://www.law.cornell.edu/wex/strict_liability), ingresado el 14 de marzo de 2021.

<sup>331</sup> En derecho penal, los crímenes por posesión y el abuso infantil son ejemplos de responsabilidad objetiva.

<sup>332</sup> Singer, Richard G., "The Resurgence of Mens Rea: The Rise and Fall of Strict Criminal Liability", *Boston College Law Review*, vol. 30, núm. 2, 1989, pp. 406-408; Duff, R. A., *The Realm of Criminal Law*, 2018, p. 19.

<sup>333</sup> *Ibidem*, p. 383; Levenson, Laurie L., "Good Faith Defenses: Reshaping Liability Crimes", *Cornell Law Review*, vol. 78, pp. 404, 433 y 434.

## Crimen específico de los sistemas de armas autónomas

Dados los desafíos asociados con las revisiones planteadas anteriormente en relación con la responsabilidad de superiores jerárquicos y la responsabilidad por complicidad, la autora se une a otros académicos en la contemplación de un crimen específico diseñado a la medida para los SAA, el cual se centra en un tercer grupo de perpetradores que son los líderes y tomadores de decisiones. Esta vía naturalmente sería complementaria a todo lo anterior.

Como ya se mencionó brevemente,<sup>334</sup> un tipo diferente de solución es criminalizar la introducción de SAA en el espacio de batalla con el argumento de que son agentes irresponsables.<sup>335</sup> De esta forma, se propone que la brecha de responsabilidad se pueda cerrar mediante el establecimiento de un crimen por la adquisición o autorización del uso de SAA.<sup>336</sup>

Aunque parecería que esta línea de razonamiento está más dirigida a fortalecer los argumentos para la prohibición del uso de estas tecnologías en lugar de buscar una compatibilidad con el marco del DPI actual, la autora opina que la creación de este crimen podría proporcionar una alternativa viable. Éste estaría más orientado como un crimen de liderazgo, en la medida en que los perpetradores necesitarían estar en una posición que les envista de la autoridad para tomar esas decisiones. En opinión de la autora, hay dos niveles consecutivos de responsabilidad derivados de este enfoque.

El primer nivel involucra a los líderes militares y/o civiles responsables de la adquisición y despliegue de estos sistemas de armas.<sup>337</sup> Estos perpetradores serían responsables

---

<sup>334</sup> Véase *supra* el apartado "Autoría mediata".

<sup>335</sup> Liu, "Refining Responsibility", *supra* nota 163, p. 344.

<sup>336</sup> *Idem*.

<sup>337</sup> Corn, Geoffrey S., "Autonomous Weapons Systems: Managing the Inevitability of 'Taking the Man out of the Loop'", en Bhuta, Nehal *et al.* (eds.), *Autonomous Weapons Systems: Law, Ethics, Policy*, Cambridge University

desde el momento que se autorice el uso de SAA, independientemente de las fallas objetivamente predecibles que pudieran presentarse, esto incluye el examen jurídico del arma y el proceso de validación de cumplimiento.<sup>338</sup>

El segundo más bien responsabilizaría a los comandantes que, posterior a que el primer grupo de líderes haya autorizado su uso, evalúen la capacidad de los SAA para realizar las tareas que se le asignan de conformidad con el DIH, y además otorguen la autorización para desplegarlos en una determinada operación.<sup>339</sup>

Lo anterior equivale a identificar un “autorizador”, una persona que sería penalmente responsable si hubiera debido ser consciente del riesgo sustancial e injustificado de un daño como resultado de la conducta del SAA, y esto se configuraría si, dadas sus circunstancias y conocimientos, el no anticipar este riesgo conllevara a una desviación grave del estándar de cuidado que se esperaría de una persona razonable en esa situación.<sup>340</sup>

Estos enfoques se derivan del modelo de responsabilidad por consecuencia natural probable, descrito anteriormente, y se basan en el argumento de que la persona que aprueba el uso belicoso del SAA, proporcionaría a la Corte más elementos para determinar la responsabilidad por las muertes de civiles causadas por un SAA impredecible, que los elementos que

---

Press, 2016 (en adelante: Com, “AWS”). Geoffrey Com dibuja paralelos entre los comandantes que adquieren los sistemas y la doctrina de responsabilidad de mando. McDougall, “AWS and Accountability”, *supra* nota 156, p. 22.

<sup>338</sup> *Ibidem*, p. 235.

<sup>339</sup> Jain, Neha, “Autonomous Weapons Systems: New Frameworks for Individual Responsibility”, en Bhuta, Nehal *et al.* (eds.), *Autonomous Weapons Systems: Law, Ethics, Policy*, Cambridge University Press, 2016, p. 314 (en adelante: Jain, “AWS”); véase también Ohlin, “The Combatant’s Stance”, *supra* nota 175, pp. 28 y 29. Ohlin menciona que para definir este crimen se requiere que el crimen sea “menos culposos” que los otros crímenes del DPI.

<sup>340</sup> *Ibidem*, p. 318.

puede proporcionar la persona a la que simplemente le ordenaron que activara el sistema o el programador que trabajó en su fase de desarrollo.<sup>341</sup>

Por otro lado, un grupo diferente de académicos ha argumentado que ambos enfoques condicionan la responsabilidad penal a la falta de cuidado adecuado en la decisión de adquirir o usar un SAA y, por lo tanto, no cubrirían el escenario en el que todas las evaluaciones relevantes sobre el cumplimiento del DIH son apropiadas, pero dada la impredecibilidad originada por “la caja negra” de los SAA hubiese civiles que murieran ilegalmente.

Al respecto, la autora señala que, según el artículo 49 del Protocolo Adicional I, el instante decisivo para el establecimiento de la responsabilidad penal es el lanzamiento del ataque.<sup>342</sup> Por lo tanto, en el contexto de los SAA, el momento decisivo es el momento en que un ser humano delega decisiones potencialmente letales al SAA.<sup>343</sup> Esta conclusión también se deriva del mandato de que los beligerantes sólo pueden elegir armas cuyos efectos puedan controlar.<sup>344</sup>

Si existe la posibilidad de que el SAA, debido su nivel de autonomía, actúe de manera impredecible, y en consecuencia pudiera cometer un crimen, entonces está claro que el combatiente no tiene un control significativo sobre el arma al no poder limitar sus efectos.<sup>345</sup>

---

<sup>341</sup> Hallevy, “The Criminal Liability of AI”, *supra* nota 299, pp. 181-185.

<sup>342</sup> Protocolo Adicional I, artículo 49(1), 1977; véase también CICR, *Commentary on the Additional Protocol I to the Geneva Conventions*, 1987, pp. 1879-1882.

<sup>343</sup> CICR, *Ethics and AWS*, *supra* nota 122, p. 21.

<sup>344</sup> Protocolo Adicional I, artículo 51(4)(c).

<sup>345</sup> Schmitt dice que “los SAA no son ilegales *per se*. Su autonomía no tiene relación directa con la probabilidad de que causen sufrimiento innecesario, ni evita que sean dirigidas contra combatientes y objetivos militares, y no necesariamente resultarán en efectos que el atacante no pueda controlar”; véase Schmitt, “AWS and IHL”, *supra* nota 29, p. 35.

Algunos académicos pretenden argumentar que, como resultado de la característica de impredecibilidad ya conocida de los SAA, es difícil —si no imposible— establecer la *mens rea*, y por lo tanto disminuye la culpabilidad del individuo que lo despliega.<sup>346</sup> Una objeción notable es que esto abriría demasiado la brecha de responsabilidad, lo que conduciría a circunstancias en las que el operador, el comandante o el programador podrían no ser verdaderamente culpables.<sup>347</sup>

La autora considera que esta visión es bastante conveniente para quienes pretenden desplegar estos sistemas, y que, por el contrario, la responsabilidad surge de este mismo hecho, ya que no puede haber justificación legítima para utilizar armas impredecibles y/o incontrolables desde un inicio.

Si prevaleciera la opinión de que “el problema de la impredecibilidad” es irresoluble, la única conclusión posible sería que los SAA deben prohibirse o restringirse hasta que sea posible garantizar un control humano determinante.<sup>348</sup>

En la reunión de expertos de la Convención sobre Armas Convencionales de 2014 sobre los SAA, la delegación de Estados Unidos sugirió que el “control humano determinante” comienza con la fabricación de diferentes componentes del SAA, continúa durante la programación del *software* y se extiende hasta el despliegue final de los sistemas de armas autóno-

---

<sup>346</sup> Véase United Kingdom Ministry of Defence-Development, Concepts, and Doctrine Centre, *The UK Approach to Unmanned Aircraft Systems*, JDN 2-11, 2011, p. 510.

<sup>347</sup> Chengeta, “Accountability Gap”, *supra* nota 155, p. 15; Malik, “AWS”, *supra* nota 159, p. 634; Liu, “Refining Responsibility”, *supra* nota 163, pp. 326 y 327; Jain “AWS”, *supra* nota 339, pp. 303, 320-322; Roff, *supra* nota 275, p. 355; Heyns, Informe del Relator Especial, *supra* nota 148, p. 80; Amoroso y Tamburrini, “AWS and Meaningful Human Control”, *supra* nota 275, pp. 20 y 21.

<sup>348</sup> Chengeta, “Accountability Gap”, *supra* nota 155, pp. 23-27, 50; Beard *supra* nota 159, p. 681.



mas.<sup>349</sup> Por lo tanto, se sugirió que al considerar lo que significa “control humano determinante” de los SAA, debería haber una “captura [de] la gama completa de actividad humana que tiene lugar en el desarrollo, adquisición, despliegue y uso de sistemas de armas; incluyendo el criterio de un comandante u operador respecto a emplear un arma en particular para lograr un efecto determinado en un campo de batalla específico”.<sup>350</sup>

La noción de control sobre el arma es fundamental para la responsabilidad de la persona que la usa y la despliega.<sup>351</sup> Para que haya un control determinante, la programación por sí sola no es suficiente. Existe la necesidad de alguna forma de supervisión después de la activación. Dicha supervisión debe ser en tiempo real. Las acciones de un SAA deben estar bajo el control de un combatiente humano que apruebe objetivos, prevenga o aborte misiones, siempre que la situación lo requiera.<sup>352</sup>

En cualquier caso, la autora está de acuerdo con los partidarios de estos enfoques, argumentando que —como mínimo— promovería una diligencia debida activa,<sup>353</sup> y de esta manera abordaría algunas de las preocupaciones de política relacionadas con la brecha de responsabilidad, independientemente de cuántos juicios pudieran eventualmente materializarse.

---

<sup>349</sup> Closing Statement U.S.; The Convention on Certain Conventional Weapons (CCW); Informal Meeting of Experts on Lethal Autonomous Weapons Systems (2014); U.S. Delegate Closing Statement, disponible en [http://www.unog.ch/80256EDD006B8954/%28httpAssets%29/6D6B35C716AD388CC1257CEE004871E3/\\$file/1019.MP3](http://www.unog.ch/80256EDD006B8954/%28httpAssets%29/6D6B35C716AD388CC1257CEE004871E3/$file/1019.MP3).

<sup>350</sup> *Idem*.

<sup>351</sup> Véase, e.g., Sassóli, “Autonomous Weapons and IHL”, *supra* nota 52, pp. 324-25; CICR, Ethics and AWS, *supra* nota 122, pp. 11-13.

<sup>352</sup> Sassóli, “Autonomous Weapons and IHL”, *supra* nota 52, pp. 323-325.

<sup>353</sup> Jain, “AWS”, *supra* nota 339, p. 319; Com, “AWS”, *supra* nota 337, pp. 235, 241. Com sugiere que esto puede ser el “talón de Aquiles operativo como resultado de la vacilación para perseguir [SAA]”.

## Observaciones finales

### ¿Existe realmente una brecha de responsabilidad?

Cabe destacar que hay algunos académicos que niegan por completo la existencia de tal brecha en el *statu quo*, principalmente debido a una comprensión conservadora (o posiblemente limitada) sobre los alcances de la tecnología de IA en los SAA. Sostienen que siempre habrá una conexión directa entre cualquier daño causado y un ser humano,<sup>354</sup> de acuerdo con dos suposiciones.

La primera es que es inverosímil que un SAA pueda llegar a ser realmente independiente de los humanos,<sup>355</sup> y la segunda, que incluso si esto fuera posible, sólo podrían implementarse si su uso pudiera cumplir con los criterios de legalidad<sup>356</sup> y, por lo tanto, se garantizaría que sus acciones sean atribuibles a un humano en todo momento.<sup>357</sup>

---

<sup>354</sup> Véase, e.g., Cass, Kelly, "Autonomous Weapons and Accountability: Seeking Solutions in the Law of War", *Loyola of Los Angeles Law Review*, vol. 28, núm. 3, 2015, pp. 1017, 1049-1053.

<sup>355</sup> Schmitt rechaza la posibilidad de que los SAA actúen de manera impredecible, y menciona que los robots "no se revelaran", que eso es un invento de Hollywood. Schmitt, "AWS and IHL", *supra* nota 29, p. 7; véase también Schmitt y Thurnher, "Out of the Loop", *supra* nota 83.

<sup>356</sup> McFarland y McCormack, "Mind the Gap", *supra* nota 276, p. 195. Consideran que usar un arma, cuando es imposible tomar precauciones suficientes, es un acto ilegal por sí mismo; y si las precauciones se toman, se puede cerrar la brecha de responsabilidad. Sassóli, "Autonomous Weapons and IHL", *supra* nota 52, pp. 324 y 325. "No creo que la posesión de capacidades de toma de decisión rompa la cadena de responsabilidad, pues siempre serán los humanos los que definan cómo funcionara la autonomía"; *cf.*, McDougall, "AWS and Accountability", *supra* nota 156, pp. 16 y 17. Ella critica a McFarland, pues nota que, pese a su conclusión, identificó varios escenarios donde la responsabilidad no se podía atribuir a una persona.

<sup>357</sup> Charles J. Dunlap Jr. esencialmente afirma que los SAA sólo pueden ser legalmente desplegados en escenarios que permitan atribuir responsabilidad, y menciona que pensar que no puede haber responsabilidad porque el

De acuerdo con la autora, la primera suposición carece de información o previsión sobre el crecimiento exponencial de los desarrollos tecnológicos,<sup>358</sup> y la segunda se relaciona más con la legalidad del arma *per se* que con la responsabilidad por las consecuencias de su uso. No hace falta decir que la legalidad de estas armas también es una cuestión crítica, sin embargo, su análisis involucra diferentes factores por lo que se aborda como el tema de estudio de la parte de DIH de este artículo.

En cualquier caso, después de considerar todo lo anterior, la opinión de la autora es que aparentemente es posible establecer la responsabilidad penal individual en situaciones en las que el ser humano está “en o dentro del proceso”. Esta expresión se refiere a la situación en la que el ser humano todavía mantiene el control, tanto de la operación como de la tecnología.<sup>359</sup>

Sin embargo, la brecha de rendición de cuentas se produce cuando el ser humano está “fuera del proceso”. Este escenario implica el despliegue de un sistema autónomo letal que puede operar en un entorno estructurado (como la identificación de objetivos sobre la base de criterios preprogramados) o en un entorno abierto y no estructurado (equipado con un grado de capacidad de aprendizaje).<sup>360</sup> A la luz de esta posibilidad, es claro que el DPI

---

SAA actuó de forma impredecible es incorrecto, pues si alguien lanza un ataque con efectos impredecibles incurre en un acto ilegal. Dunlap Jr., Charles J., “Accountability and Autonomous Weapons: Much Ado about Nothing”, *Temple International and Comparative Law Journal*, vol. 30, núm. 1, 2016, pp. 63, 71.

<sup>358</sup> Como ha escrito Tapson: “hasta que no veamos a una IA hacer algo totalmente impredecible, no nos daremos cuenta de nuestra visión limitada de las posibilidades”, véase Tapson, Jonathan, “Google’s Go Victory Shows AI Thinking Can Be Unpredictable, and That’s a Concern”, *The Conversation*, 18 de marzo de 2016, disponible en <https://theconversation.com/googles-go-victory-shows-ai-thinking-can-be-unpredictable-and-thats-a-concern-56209>, ingresado el 11 de marzo de 2021.

<sup>359</sup> Schmitt y Thurnher, “Out of the Loop”, *supra* nota 83, pp. 276 y 277.

<sup>360</sup> CICR, Ethics and AWS, *supra* nota 122, p. 9.

debe revisarse, preferiblemente *ex ante*, en el dominio normativo para que pueda surtir efectos preventivos, de lo contrario vendrá de forma *ex post facto*, a través de una interpretación jurisdiccional.

La opinión de la autora es que el primer paso para cerrar la brecha de responsabilidad es responsabilizar a los líderes que toman decisiones irresponsables sobre el desarrollo y despliegue de SAA. Esta sería la solución más eficaz, ya que cerraría *ab initio* la existencia de la brecha.

De hecho, responsabilizar a los líderes es un compromiso internacional sumamente arraigado, como puede observarse con la inclusión del artículo 27 del Estatuto de Roma relativo a la improcedencia del cargo oficial de un jefe de Estado, miembro de un Gobierno, o Parlamento, representante electo o cualquier tipo de funcionario gubernamental, quienes en ningún caso serán eximidos de responsabilidad penal en virtud de su cargo.

El mismo compromiso fue reiterado en 2000 por el Consejo de Seguridad de la ONU, en la Resolución 1329, que enfatizó el enjuiciamiento de figuras de liderazgo por crímenes de guerra en el contexto de los tribunales *ad hoc*.<sup>361</sup> Para la autora, los líderes sólo tienen dos opciones: desplegar SAA únicamente cuando se pueda garantizar por completo un control humano determinante o no autorizar su uso en lo absoluto. Sin duda alguna, cualquier otro resultado debería resultar en su responsabilidad penal.

Por lo tanto, la autora invita a todos los lectores a exigir que nuestros líderes tomen la decisión correcta de impedir un uso o despliegue ilícito o prematuro, evitando así que los crímenes se cometan en primer lugar.

Sin embargo, también es importante recordar que la responsabilidad penal individual se actualiza en varios niveles. La atribución de responsabilidad penal a líderes políticos y otras

---

<sup>361</sup> United Nations Security Council Resolution 1329, U.N. Doc. S/RES/1329, 30 de noviembre de 2000.

figuras de alto rango no excluye la responsabilidad de la persona o personas implicadas en el desarrollo y/o despliegue final del arma.<sup>362</sup> Aunque los tribunales y cortes internacionales se concentran en los “peces grandes”, como una cuestión de política interna, los “peces pequeños” también necesitan ser enjuiciados, aunque sea a nivel nacional.<sup>363</sup>

Como reflexión final, la autora comparte las cinco reglas que han sido desarrolladas por éticos prácticos y teóricos sociales que insisten en el principio de que los humanos no pueden ser excusados de su responsabilidad moral por el diseño, desarrollo o despliegue de artefactos informáticos.<sup>364</sup>

Las reglas establecen lo siguiente:

Regla 1: las personas que diseñan, desarrollan o despliegan un artefacto informático son moralmente responsables por ese artefacto y por los efectos previsibles de ese artefacto. Esta responsabilidad se comparte con otras personas que diseñan, desarrollan, despliegan o utilizan conscientemente el artefacto como parte de un sistema sociotécnico.

Regla 2: la responsabilidad compartida de los artefactos informáticos no es un juego de suma cero. La responsabilidad de un individuo no se reduce simplemente porque más personas se involucran en el diseño, desarrollo, implementación o uso del artefacto. En cambio,

---

<sup>362</sup> Véase TIPY, Sentencia de Delalić, *supra* nota 244, p. 1280. En el contexto de la CPI, si bien el principio de complementariedad evita que la Corte juzgue a todos los autores de los crímenes; en las jurisdicciones nacionales, el principio *aut dedere aut judicare* es aún aplicable en relación con los perpetradores de menor rango.

<sup>363</sup> Jones y Powles, *International Criminal Practice*, *supra* nota 199, pp. 412-414.

<sup>364</sup> Véase Grodzinsky *et al.*, “Moral Responsibility for Computing Artifacts”, *supra* nota 238; las reglas siguen una noción de responsabilidad objetiva, en la cual la responsabilidad se reconoce frente al despliegue de un SAA; véase Arkin, Ronald, “The Robot Didn’t Do It”, Position Paper for a Workshop on Anticipatory Ethics, Responsibility and Artificial Agents, 2013, disponible en <http://www.cc.gatech.edu/ai/robot-lab/online-publications/positionpaper3.pdf>, ingresado el 10 de marzo de 2021.

la persona es responsable por los comportamientos y efectos del artefacto después del despliegue, en la medida en que estos efectos sean razonablemente previsibles por esa persona.

Regla 3: las personas que de manera consciente utilizan un artefacto informático en particular son moralmente responsables de ese uso.

Regla 4: las personas que conscientemente diseñan, desarrollan, implementan o utilizan un artefacto informático, lo hacen responsablemente sólo cuando hacen un esfuerzo razonable para tener en cuenta los sistemas sociotécnicos a los cuales está incorporado el artefacto.

Regla 5: las personas que diseñan, desarrollan, implementan, promueven o evalúan un artefacto informático no deben engañar a los usuarios, explícita o implícitamente, sobre los efectos predecibles del artefacto o sobre los sistemas sociotécnicos a los cuales está incorporado el artefacto.

Si bien estas no son reglas vinculantes en el ámbito legal, la autora las considera relevantes, ya que incluyen todos los elementos necesarios para servir como pautas en la construcción de un marco jurídico que sí pueda serlo.