

La nueva huida del derecho administrativo a través de la Inteligencia Artificial: los Sistemas de armas autónomos letales (SAAL)

Roser Martínez Quirante

Profesora Titular Derecho Administrativo UAB

Roser.martinez@uab.es

Palabras clave: arma, derecho, autonomía, delegación

Resumen (150 palabras) **Sumario** de la introducción, metodología, resultados y discusión.

Se trata de alertar sobre la aparición de los Sistemas Autónomos letales (SAAL) con capacidad de seleccionar y eliminar objetivos sin control humano. Se analizan las nuevas amenazas que provienen de la delegación de potestades letales a estos sistemas dotados de inteligencia artificial (IA) supuestamente más eficaces y eficientes que los seres humanos. Estas armas obvian la dimensión racional, cooperativa, intuitiva, moral y ética de las decisiones humanas, contradicen el DIH, el DIDH, las leyes de la guerra e, internamente, el derecho administrativo dado que el monopolio de la violencia legítima está en manos del Estado.

Los sistemas letales de armas autónomas deberían depender de la supervisión o juicio humano en, al menos, algunas de sus fases críticas (selección de objetivos o cancelación de la orden). El “control humano significativo” debe estar en la base de cualquier proceso autónomo que utilice la IA sobre todo cuando esta trate de acciones letales.

1. La IA y el derecho

El Estado de derecho en la nueva sociedad algorítmica tiene el reto de garantizar la interrelación entre el derecho a la seguridad humana y el derecho a la información. Para ello debe enfrentarse a una profunda transformación, reordenando jurídicamente sus estructuras y sus conceptos clásicos, como

consecuencia del desarrollo y avance de la inteligencia artificial en todos los campos de nuestra vida (Skinner, 2018).

Sabemos que la Inteligencia Artificial (IA) es una combinación de algoritmos creados con el propósito de crear sistemas computacionales que alcancen capacidades humanas (independientes al humano) a través de un proceso que incluye aprendizaje (adquisición de información, datos y reglas), razonamiento, autocorrección y, en último término, automejora (Tegmark, 2018).

Pero para el Derecho hay dos tipos básicos de IA. En primer lugar, la que utiliza la lógica a través de **algoritmos** que usan los principios racionales del pensamiento humano. En segundo lugar, la que se basa en **redes neuronales artificiales** que también funcionan con algoritmos pero que están diseñadas para que el sistema aprenda recursivamente (*Deep learning* o aprendizaje profundo) y pueda tomar decisiones que evolucionen en el tiempo.

Actualmente, la mayoría de proyectos de IA que utiliza la Administración pública para llevar a cabo sus fines en el orden social se basan en algoritmos del primer tipo, destinados a resolver problemas a través del procesamiento de capas y capas de información que son objeto de tratamiento más veloz y más eficiente que el que podría llevar a cabo un ser humano. La Comisión de la UE así lo aconseja en su Comunicación sobre IA de 2018 según la cual es necesario poner “el poder de la IA al servicio del progreso humano” (COM (2018) 237 final, de 25.04.2918).

Sin embargo, aunque al parecer la IA profunda (la que puede evolucionar con el tiempo y automejorarse recursivamente) no está implementada aún de forma extensa, sí empezamos a ver ámbitos públicos donde se está probando su eficacia.

Ciertamente, los poderes públicos utilizan cada día más este tipo de operaciones ciberneticas complejas para tomar decisiones regladas y automáticas, equivalentes, en lo sustancial, a la manera en que los seres humanos procesamos la realidad y la transformamos (Ramió, 2019).

El Derecho debe contemplar este nuevo escenario y asumir las nuevas implicaciones y consecuencias “del empleo de estas capacidades de inteligencia artificial para delegar en ella la toma de decisiones en su integridad a medida que se vayan mostrando mejores, más eficientes y más capaces que las tomadas por inteligencias humanas tradicionales” (Boix, 2020). Deberemos estar alerta porque se va a tratar de una nueva “huida”, en este caso no del derecho administrativo, sino de la decisión humana en pro de la aplicación de un algoritmo o de la utilización de redes neuronales artificiales.

El problema sobre todo surge cuando las decisiones que toman los poderes públicos no son regladas o automáticas sino que requieren la intervención de reconocimientos de voz, reconocimientos faciales, de patrones, etc. y pueden derivar en consecuencias irreversibles. Nos referimos a la posibilidad de que el Gobierno, investido de potestades discrecionales para la defensa nacional, delegue poderes exorbitantes como la de acabar con vidas humanas, a sistemas artificialmente inteligentes “huyendo” de sus competencias inherentes que son irrenunciables e indelegables. Se trata de la introducción de nuevos agentes inteligentes a través del uso de los SAAL (o LAWS, *Lethal autonomous weapons systems*, por sus siglas en inglés) máquinas de matar que pueden llegar a actuar sin control humano significativo (Martínez y Rodríguez, 2018).

2. La aparición de los SAAL sin control humano significativo

La definición de SAAL no está internacionalmente acordada ni regulada¹, pero vendría a ser toda aquella *arma que, con IA, selecciona, ataca y destruye objetivos que cumplen con unos parámetros predefinidos, y que una vez activada la orden no existe posibilidad humana de detenerla*.

Los sistemas autónomos letales desarrollados hasta el momento, en forma de artefacto volador (dron), terrestre (robot) o acuático, deberían depender de la supervisión o del juicio humano previo en, al menos, alguna de sus fases críticas (selección de objetivos, ataque o anulación de la orden de eliminar el objetivo). Los humanos deberían desempeñar sin lugar a dudas un papel crucial en el proceso de selección de objetivos y toma de decisiones cuando se utiliza un arma, por mucha inteligencia (artificial) que se le atribuya.

Sin embargo, escudándose en un uso militar de defensa nacional, se investigan y se desarrollan sistemas letales con autonomía total que podrían llegar a crear un panorama *Black Mirror*, puesto que ya no se trata de analizar el uso de armas automáticas o de sistemas comandados a través de control remoto, sino que se trata de armas que podrán decidir en base a los datos que en cada momento manejen.

Actualmente existe una situación de práctica *arregulación* de estas armas autónomas robotizadas lo que permite, por inactividad de los Estados anómicos (Waldman, 2006), una suerte de carrera competitiva sin ley entre

¹ El Convenio de la ONU sobre *certas armas convencionales* en vigor desde 1983 (CCW) tiene el objetivo de restringir o prohibir diferentes clases de armas que puedan considerarse excesivamente nocivas o que puedan producir efectos indiscriminados, y a pesar de la oposición de las grandes potencias armamentísticas, sería el marco regulatorio más adecuado para delimitar los usos y alcance de este tipo de tecnología.

potencias frente a la que los juristas no podemos permanecer como meros espectadores esperando un desastre anunciado.

Este desarrollo tecnológico en armas inteligentes sin control humano significativo no solo puede poner en peligro el equilibrio ya de por si inestable de la seguridad global, en un contexto en el que el derecho aún no se ha pronunciado de forma vinculante para la comunidad internacional, sino que puede amenazar a la misma seguridad interna de un país, como veremos más adelante (Scharre, 2018).

3. ¿Porqué la utilización de los SAAL es tan atractiva para los poderes públicos?

Los SAAL con control humano significativo pueden tener evidentes ventajas militares puesto que analizan capas de datos con mayor rapidez, y su respuesta es más precisa e instantánea (p.ej. los sistemas de defensa Phalanx, Mantis o Iron Drome). Los sistemas letales no tripulados pueden ampliar el campo de batalla sin arriesgar la vida de los propios soldados, sin que les afecten los factores físicos, emocionales o morales y además siendo más rápidos y menos costosos económicamente que los despliegues militares convencionales (p.ej. dron Loyal Wingman). Es decir, la distancia espacial entre el campo de batalla y los comandantes con competencias operacionales reduce costes personales y materiales y, además, la ausencia de sentimientos de venganza, odio o miedo de una máquina puede ofrecer una ventaja táctica en el combate.

Sin embargo, estos sistemas inteligentes letales están demostrando que son propensos al error porque los entornos en los que operan son complejos (no siempre se requieren acciones automáticas) y ya hemos visto en Yemen, Irak y Pakistán cómo de trágica puede llegar a ser la decisión de un algoritmo en un proceso automático si no interviene el juicio de un humano.

En este sentido, los mitos de la IA que describe claramente Rodríguez Álvarez, deben ser tenidos en cuenta para no atribuir más ventajas que las que realmente tiene esta tecnología (Rodríguez, 2018). Entre las falacias que el autor destaca está la de que IA puede tomar decisiones de forma más efectiva, más fiable, más ecuánime y más justa que un humano. Es totalmente erróneo, porque la IA reproduce por emulación los sesgos cognitivos de sus creadores, es decir tiende a reproducir la falta de imparcialidad que, de forma inconsciente, trasmite la persona que construye el sistema computacional algorítmico. La experta Cathy O’Neil en su obra *Armas de destrucción matemática* alertó en su día sobre que el creer en la infalibilidad de los algoritmos puede llevarnos a resultados dramáticos (O’Neil, 2017). Además, se ha hablado incluso de los riesgos de desarrollar en quienes manejan estas armas una mentalidad *Playstation* cuando se trata de proceder al ataque, que deshumaniza el conflicto y convierte la muerte en algo banal (Gutiérrez y Cervell, 2013).

Por otro lado, la facilidad para cometer abusos ha llevado a ejecuciones selectivas extrajudiciales, sumarias o arbitrarias por parte de algunos ejércitos de grandes potencias mundiales a través de esta tecnología. Y más aún lo van a ser en una sociedad donde los SAAL sean de uso dominante por parte de nuestros gobiernos a través de sus políticas militares.

Hasta ahora los sistemas armados ya disponen de hardware y software específico para funcionar. Su automatización proviene de algoritmos bien desarrollados que permiten operar de forma más rápida que un humano porque procesan la información mucho más rápidamente.

Pero el nuevo escenario se da cuando estas armas, debido un avance en la IA, van a disponer de redes neuronales complejas y van a tener la capacidad de

aprender de los datos que obtienen “en la nube” y automejorarse. El autoaprendizaje de los sistemas armados comporta que emulen el funcionamiento del cerebro y alimentándose continuamente con nuevos datos creando nuevas conexiones.

Por tanto, los SAAL se están comenzando a considerar “agentes racionales” dotados de inteligencia artificial que les permite adaptarse a contextos cambiantes, aprender durante el ciclo de su existencia y realizar tareas con diferentes grados de autonomía (Russel, S.*et al.*, 2010).

En este caso, a estos “agentes inteligentes” ¿se les puede atribuir la capacidad de decidir en entornos complejos cuando la potestad es consustancialmente discrecional?

Para contestar a esa pregunta, hemos de ver si hay verdaderamente un acto administrativo discrecional de decisión de acabar con la vida de otro en un conflicto bélico o el control humano significativo se refiere a un conjunto de decisiones conexas con diferente naturaleza jurídica según sea un comandante, un político o un simple soldado el responsable de ello.

4. Control humano significativo versus decisión autónoma de atacar

El Grupo de expertos gubernamentales de la Convención de armas convencionales (CCW) que nos reunimos desde 2014 en la sede de Naciones Unidas en Ginebra para aprobar un instrumento vinculante que limite esta tecnología, no hemos podido encontrar aún un significado exacto al concepto de “control humano significativo” que gira en torno a la acción directa que puede ejercer el operador sobre el arma en el momento de una ejecución

bélica.

Hemos de tener en cuenta que aparecen dos términos importantes en este complejo ámbito decisorio: el mando y el control.

El mando hace referencia a la autoridad de las Fuerzas Armadas de dirigir, coordinar, asignación de tareas y control de las fuerzas militares. El mando despliega decisiones estratégicas globales.

Por otro lado, **el control** se refiere a la autoridad de un oficial sobre una parte de las actividades militares y aquí se suelen desplegar las decisiones operacionales y tácticas. En este nivel, el oficial tiene el objetivo de transformar los objetivos establecidos por el mando superior en operaciones concretas. Posteriormente, el nivel táctico concreta el uso específico de las fuerzas militares.

En todos esos niveles se puede establecer la decisión de desplegar un sistema autónomo letal de un modo u otro, e incluso podemos decir que el control humano significativo debería exigirse en los tres niveles, estratégico, operacional y táctico mediante el factor de comunicación bidireccional. Sin embargo, esa misma complejidad a la hora de establecer cuándo y cómo, está haciendo difícil el acuerdo no sólo internacional, sino una regulación interna en los diferentes estados.

5. La decisión de disparar a partir de un algoritmo: la relación entre la administración civil y la militar

Como hemos visto, en las fases operacionales y tácticas, se pueden producir situaciones complejas en las que el humano no puede actuar con la rapidez

que se requiere. Por ello, se acude a sistemas armados automatizados a veces, semiautónomos, autónomos supervisados y autónomos completamente otras veces, que gracias a algoritmos procesan la información a una velocidad imposible para el hombre. El Comité Internacional de la Cruz Roja, en el último censo realizado en 2018, contabilizó unos 130 sistemas de armamento autónomo en el mundo (Roff, 2016), aunque otros recuentos los aproximan a los 300 (Rodríguez, J. *et al.*, 2019). Entre estos, se incluirían sistemas armamentísticos semi-autónomos (puesto que estarían sometidos a supervisión humana en algunas fases clave como la selección de objetivos) como podrían ser los misiles *Patriot*, o drones como el modelo *Reaper*, pero muchos de ellos con posibilidades de llegar a tener autonomía total.

Así pues, según Scharre el rol del militar puede ser más o menos importante según sea:

1. el operador esencial en un sistema armado: la máquina no puede operar sin el humano. La detección, persecución y fijación del objetivo está automatizada por los algoritmos, pero el comandante es quien toma la decisión de verificar el objetivo y autoriza el ataque en un tiempo determinado (p.ej. C-RAM para la protección de bases aéreas o el sistema antiaéreo MANTIS que detecta drones y misiles)
2. un *fail-safe*: el hombre fija las amenazas que deberían atacarse y las que deberían ignorarse, supervisa la operación y la puede cancelar, pero la máquina actúa de forma autónoma gracias a los algoritmos desde la detección e identificación del objetivo, hasta el ataque (p.ej. *Phalanx*, sistema de armamento de proximidad contra misiles antibuque; el *Thophy* que detecta misiles y granadas propulsadas o el *Arena* contra misiles antiblindaje).
3. un agente moral: el hombre sólo toma una decisión general de activar el

sistema y este decide a través del análisis de los datos que va adquiriendo de forma autónoma puesto que no hay comunicación entre ellos (p.ej. el dron israelí Harpy, con munición merodeadora para atacar defensas aéreas, es capaz de identificar personas y atacarlas o no si las considera combatientes y sus sensores evalúan los daños posteriormente; también el robot centinela Samsung SGR-A1 puede detectar soldados enemigos y les puede disparar sin intervención humana) (Scharre, 2018).

4. un mero espectador: cuando el sistema armado está compuesto por algoritmos y redes neuronales artificiales que lo capacitan para ser completamente autónomo para decidir cuándo, cómo y dónde está el objetivo a abatir. Un sistema así podrá crear o/y modificar sus objetivos, incorporando la observación de su entorno y la comunicación con otros agentes. Sus capacidades son: autodeterminación, autocompromiso, comunicación autónoma con otros sistemas, auto-modificación de objetivos basada en información adquirida de fuentes autónomas, planificación de objetivos, aprendizaje y adaptación constante.

En los casos segundo y tercero, si la decisión no es automática, sino que existe un componente discrecional y además ésta se delega a un sistema autónomo, pudiendo escoger el sistema entre varias soluciones posibles y válidas, entonces debemos advertir un problema para el derecho que no está preparado para regular las consecuencias de dichas acciones. Ahora bien, en el supuesto cuarto, el problema ya no se centra sólo en el derecho, sino que amenaza la propia supervivencia de nuestra sociedad.

Se trata de la posible quiebra de los principios más básicos de las potestades públicas, porque se estaría atribuyendo a un ser computacional artificial, sin humanidad, la potestad de decidir, en un conflicto a quién abatir en base al

“interés general”. Se le estaría atribuyendo una *licencia para matar* que en caso de conflicto bélico es exclusiva del poder público y para la defensa nacional en circunstancias excepcionales.

Así, ¿se hace necesaria una revisión del derecho de la administración civil o de la militar?

6. Relación entre la Administración civil y la militar en la atribución de potestades exorbitantes discrecionales

Hasta ahora, el uso de SAAL se justifica en la potestad de defensa nacional que tiene atribuida el Estado. Desde esta perspectiva la Administración militar, como rama específica de la Administración pública, bebe del Derecho constitucional en cuanto a sus principios programáticos y del Derecho administrativo en cuanto a su organización y caracteres entre los cuales está la “especial discrecionalidad de sus potestades” (Guaita, 1960).

En este sentido, la administración civil realiza actos de gestión de los intereses públicos y la militar, actos de fuerza, de imperio por “hombres de armas” concepto clásico reconocido por la doctrina. Respecto de estos últimos hay que analizar su alcance y hasta qué punto deben quedar en manos humanas sus decisiones.

El especialista en derecho militar, Guaita, destacó en su día que “la Administración militar (auxiliar de la civil) (...) debe asegurar y hacer posible que se presten y se satisfagan las necesidades generales con las garantías constitucionales (...) Las decisiones más graves, que se salen del marco de los simples actos administrativos para constituir actos políticos, como la

declaración de guerra, la firma de paz, etc. se confían al Parlamento. (...) Pero cuando hay una crisis jurídico-política, (...) y se declara en el país el estado de guerra, todo lo concerniente al orden público y a la seguridad interna pasa a ser (en virtud de un fenómeno de ósmosis jurídica) de la competencia de la administración militar, a la que queda sometida en esta materia y por completo la administración civil” (Guaita, 1960).

Entre la doctrina administrativa es unánime hasta ahora la idea de que “la defensa nacional requiere una maquinaria humana”, porque las complejas y exorbitantes potestades discrecionales que se le atribuyen requieren de una rendición de cuentas muy jerarquizada.

En palabras de Guaita “entendemos nosotros, con López Rodó que la potestad discrecional no implica, como decía Merkl, la ejecución de una norma superior, (...) sino un criterio de valoración, de oportunidad, en suma, de discreción, que nadie como la Administración, perito del interés público, puede tener. (...) Esta discrecionalidad, que de hecho es consustancial a la idea misma de Administración, está mucho más acentuada en la militar que en la civil. (...) La Administración militar es el *reducto irreducible* donde la discrecionalidad continúa teniendo un gran campo de aplicación” (Guaita, 1960).

Los Ejércitos son institutos armados, con lo cual, el principio jerárquico es un principio orgánico insoslayable para la Administración militar. La desobediencia o insubordinación de un funcionario militar hacia su superior puede revestir de una gran gravedad, aspectos a tener en cuenta cuando aparece en escena un nuevo elemento distorsionador: un sistema armado autónomo que puede llegar a tener la posibilidad de decidir cómo actuar, y en el fondo a autodeterminarse (Frischmann y Selinger, 2018).

7. La muerte a manos de un dron autónomo con IA no es una muerte digna

En los conflictos armados, *derecho a la vida* significa derecho a no ser asesinado de forma arbitraria o caprichosa, inexplicable o inhumana o como daño colateral, es decir, se protege el derecho a la dignidad humana. Se puede decir que la dignidad humana es un derecho más importante que el derecho a la vida, porque este puede perderse o soslayarse más fácilmente: en una sociedad civilizada puede haber ejecuciones legales, pero estas deben ser humanas y dignas (Lin, P. *et al.*, 2017).

Alemania posee una de las concepciones legales más desarrolladas al respecto de la dignidad humana: el artículo 1 de su Constitución establece que la dignidad humana es inviolable e incluso más importante que el derecho a la vida incluido en el artículo 2, que puede perderse bajo ciertas condiciones. Según la Carta Magna alemana, la dignidad humana es intangible, por lo que respetarla y protegerla es obligación de todo poder público. En 2005, la ley de Seguridad del Transporte Aéreo de ese país autorizó a sus fuerzas armadas a derribar aviones comerciales sospechosos de haber sido secuestrados por terroristas. Si un avión parecía estar rumbo a colisionar con un edificio, derribar el avión y sacrificar a los pasajeros se consideraba un mal menor con respecto a permitir que el vuelo continuara y matara a miles de personas en tierra. Pero en 2006 el Tribunal Constitucional Federal anuló esta ley por inconstitucional, ya que trataba a las personas como objetos; como parte del propio avión y no como individuos que merecen respeto y consideración. Tal y como apunta Lin, «la ley también habría tratado a las personas como números o estadísticas, predeterminando sus muertes por la posibilidad de

salvar un mayor número de vidas no identificadas. Y hay algo malo —algo irrespetuoso y deshumanizador— en hacer ética solo por números» (Lin (2017). Los sistemas autónomos letales no reconocen a las personas como seres humanos, sino simplemente como objetos o, peor, como *bytes* de información; y el tribunal alemán consideró de forma contundente que la sociedad civilizada no podía tratarlas del mismo modo.

El temor a un futuro distópico parece una razón legítima para una prohibición total o una moratoria de las armas autónomas letales mediante la aplicación del principio de precaución, pero a fin de defender esa posición deben fortalecerse previamente la noción de dignidad humana y la cláusula Martens, así como los conceptos relacionados con el control humano significativo (Lin, 2017). Recordemos que la aplicación de la Cláusula Martens comporta que las tecnologías emergentes sean juzgadas según los principios de la humanidad y los dictados de la conciencia pública cuando no estén ya cubiertas por otras disposiciones de los tratados.

Aunque esta tecnología pudiera llegar a ofrecer resultados mejores en base a un cálculo de coste-beneficio o llegara a conseguir comportamientos humanos, debe prohibirse por razones éticas y legales. Heyns, que tiene la misma opinión, la fundamenta en la concepción de Kant de la dignidad humana, según la cual las personas tienen derecho inherente a ser tratadas como seres humanos únicos y completos especialmente cuando sus vidas están en juego. Esa dignidad humana quedaría negada si las víctimas que quisieran apelar a la humanidad de su verdugo no pudieran por tratarse de un ser artificial. El poder ejecutivo debe ofrecer el debido respeto a la dignidad de la persona considerando su caso concreto y haciendo valoraciones y ajustes constantes. Y nada de ese *law enforcement* con las características de las

capacidades humanas pueden asegurarlo las armas autónomas, ya que faltaría el juicio humano adecuado en su actuación.

8. Estado de excepción algorítmico: La solución pasa por la regulación y el control de los algoritmos

McQuillan advierte con acierto de que el análisis, la vigilancia y la acumulación masiva y pormenorizada de datos a través de sistemas inteligentes están conduciendo a cambios en la gobernanza y daños en el núcleo de la sociedad civil. Se refiere a ello como «la implantación de estados de excepción algorítmicos» (Heyns, 2016).

Vivimos, y no nos damos cuenta, en un Estado policial artificialmente inteligente. Hasta el mismo Mark Zuckerberg, CEO de Facebook, reconoció implícitamente ante el Congreso estadounidense que estamos ante un Estado anómico y que es preciso uno regulador que no lo fíe todo al libre mercado: «*That federal regulation of Facebook and other Internet companies is inevitable*» (Kang y Roose, 2018). Será a través de esta legislación federal americana cuando se produzca una proyección internacional y, al final, una globalización, ya que podría tener efectos extraterritoriales sobre otros países, como ha ocurrido con la *Foreign Corrupt Practices Act* de 1977 (FCPA).

Sin embargo, hasta el momento no existen instrumentos internacionales legalmente vinculantes o incluso leyes nacionales que prohíban el desarrollo, producción y uso de los llamados *robots asesinos*.

Los fallos del mercado hicieron que el derecho interviera, y esa necesidad vuelve a ser acuciante. La única garantía de progreso y sostenimiento de los derechos ciudadanos ante la inteligencia artificial en los drones autónomos es la regulación.

Pero la regulación no sólo de las funciones de los sistemas sino de los algoritmos y de sus limitaciones. Se ha defendido que los algoritmos empleados en los sistemas armados por parte de la Administración pública civil o militar (en sus funciones de garantizar la seguridad) para la adopción de decisiones letales, podrían considerarse reglamentos. En este sentido, los algoritmos, cumplirían “una función material estrictamente equivalente a la de las normas jurídicas, al reglar y predeterminar la actuación de los poderes públicos” (Boix, 2020). Pero si es así, deberían cumplir con las mismas garantías que cualquier otra norma de rango reglamentario, es decir, ser inalterables, ser públicos y ser recurribles. Y lo cierto es que no es así. Los sistemas dotados de inteligencia artificial evolucionan porque se enriquecen de los datos que ofrece la nube, tienen un efecto *black box*, resultan ser indescifrables y los destinatarios no pueden recurrirlos.

Los sistemas de computación actuales utilizados en el mundo del Derecho van a afectar a las bases de la misma ordenación jurídica. Boix Palop ya alertó de los riesgos derivados de utilizar algoritmos en la toma de decisiones públicas bajo el argumento de una mayor eficiencia (Boix, 2007).

Solo podremos regular los SAAL si conocemos las implicaciones de estas nuevas posibilidades algorítmicas, su funcionamiento y sus consecuencias (deseadas y previstas... o no).

Boix ha alertado de la utilización de los algoritmos para la toma de decisiones de este tipo “dada la magnitud del riesgo en cuestión, que puede tener

consecuencias extintivas para la especie humana, y la irreversibilidad de la pérdida de control sobre estas inteligencias a partir de un determinado momento del proceso. Por ello quizás no solo haya que atender al principio de precaución en su declinación más tradicional, sino ir más allá e integrar la propia noción de precaución en una formulación más radical adaptada a todas las necesidades y exigencias del nuevo entorno” (Boix, 2020).

9. No cabe delegar la potestad discrecional de matar a los drones autónomos letales

Por todo lo expuesto, el derecho debe valorar la prohibición de los sistemas autónomos letales ya que una vez activados podrían seleccionar objetivos y acabar con la vida de las personas, de forma irrevocable y sin intervención humana, lo que supondría el otorgamiento de una facultad administrativa contraria al ordenamiento jurídico nacional e internacional.

La supresión de una vida humana solo se puede justificar legal o moralmente si no es arbitraria. Pero para no ser considerado arbitrario, el acto letal del agente debe basarse en una decisión informada y en un *juicio humano* cognitivo, pues solo la toma humana de decisiones garantiza el pleno reconocimiento del valor de la vida individual y de la importancia de su pérdida. Solo en ella entran en juego todos los complejos estándares modernos del derecho humanitario: proporcionalidad, compasión, utilización de métodos menos gravosos o menos restrictivos, cuidado constante, caballerosidad... y del derecho administrativo. En consecuencia, las acciones de los SAAL no son legítimas ni moralmente justificables y deberían estar prohibidas con arreglo al principio de dignidad humana y al *ius cogens*, que

como norma imperativa contiene las normas fundamentales del derecho humanitario (Buttha, *et al.*, 2016). Pero, además, el derecho administrativo en su concepción más amplia no debería permitir su utilización como armamento convencional de su personal civil o militar porque no cabe delegar a un algoritmo que tenga una aplicación discrecional, la potestad de acabar con la vida de seres humanos.

10. Los drones carecen de intuición

El derecho, por otro lado, no pone como requisito un elemento cognitivo fundamental, **la intuición humana**, cuando regula las facultades públicas discretionales en la toma de decisiones, quizás porque da por hecho que son los seres humanos quienes las llevan a cabo. Pero los SAAL pueden ser destinatarios hipotéticos de dichas atribuciones, por lo que debe analizarse si disponen de las capacidades que el derecho reclama y si por lo tanto pueden ejercer tales potestades cuando está en juego la vida de una persona.

Sea como sea, la intuición forma parte de nuestra misma esencia como humanos y de la de todas nuestras acciones, y ha jugado siempre un papel fundamental en la resolución de los conflictos violentos. Y los drones autónomos pueden estar dotados de mecanismos de imitación e incorporar procesos de tipo integrativo y cognitivo, pero no fenomenológicos. No podrán ser nunca intuitivos ni sentir emociones, sino solo replicarlas (Howard, Zhang y Horvitz, 2017). Como señala el neurocientífico G. Rizzolatti, descubridor de las neuronas espejo, «los robots pueden imitar, no sentir» (Rodella, 2018). Y si es así, si los algoritmos incluidos en drones, o en otro tipo de tecnología robótica, no pueden alcanzar las características humanas necesarias para

tomar decisiones discrecionales trascendentales referidas al ejercicio de la fuerza legítima contra las personas, no debería aceptarse el traspaso de tales competencias a los sistemas autónomos.

La potestad no ya de defender al Estado que los ha creado de ataques exteriores de otras naciones, sino de decidir que el enemigo está dentro del mismo Estado y que debe combatirlo segando vidas, reviste tanto *imperium* que no puede ser otorgada a seres artificiales sin emociones, sin intuición, humanas.

11. La pérdida de humanidad en las funciones de defensa de la seguridad: Drones autónomos letales

Es también necesario profundizar en nuevas formas de convivencia teniendo en cuenta que la deshumanización provocada ya por los sistemas autónomos *con control humano* en los conflictos bélicos deja en papel mojado todo lo que se había aprendido en la primera guerra mundial acerca de la cooperación y la dignidad humana (Axelrod, 1984), sobre comunicación no verbal y sobre la relación humana entre combatientes. El progreso en la comunicación no verbal humanitaria se detiene y hasta retrocede cuando se utilizan SAALs en tiempo de paz (preventivamente) o de guerra. En palabras de Sparrow, «debemos mantener una relación *interpersonal* con otros seres humanos incluso durante la guerra» o de lo contrario no estaremos respetando los mismos fundamentos del derecho (Sparrow, 2016). El avance enorme que, por ejemplo, supuso la *Tregua Dei* a partir del siglo XI (una convención según la cual los sacerdotes, las mujeres y la población de menor edad no debían recibir muerte bajo ninguna circunstancia) solo se podía conseguir entre

humanos. Las *zonas de seguridad reservadas* tampoco se habrían conseguido sin los componentes de humanidad y reciprocidad (Martínez, 2003).

Los defensores de estos nuevos sistemas de armamento, obviando la necesidad de este componente de humanidad, les atribuyen numerosas ventajas: reducción de los costes de los operativos, potencial único para desarrollar determinadas tareas más rápidamente que los humanos, capacidad de impactar sobre un objetivo incluso cuando los enlaces de comunicación se ven afectados... Arkin señala de ellos que «pueden ser diseñados para aceptar los riesgos más altos; pueden tener los mejores sensores; no se tambalearán ante emociones como el miedo o la cólera, que pueden incitar a seres humanos a actuar de manera inmoral; no padecerán prejuicios cognitivos que afectan al ser humano [...] e incluso pueden distinguir legítima y fiablemente los blancos legítimos de los ilegítimos» (Sparrow, 2016).

Se trata de ventajas ciertas y que no deben ser despreciadas, pero tampoco deben serlo, llevados de un utilitarismo exacerbado, los problemas serios que todo esto supone a nivel tanto teórico como práctico, y sobre todo legal; ni el hecho de que en numerosas ocasiones ha sido justamente el factor humano, la emoción humana (Fischer, 2008), la negociación (Uri, 2005), lo que ha impedido procesos de escalada bélica: existen numerosos ejemplos de hombres y mujeres de toda clase y condición que en un momento dado se negaron a apretar el botón que hubiera desencadenado uno (Rodríguez-Arana, 1993). Las guerras pudieron ir siendo más humanas porque la comunicación no verbal de la guerra de trincheras permitía momentos de tregua y baja letalidad sin que los soldados contendientes hubieran recibido ninguna orden en ese sentido (Axelrod, 1984).

En este sentido, los SAAL están contribuyendo a tener una visión deshumanizada de los conflictos bélicos, pero además, por ser tecnología de doble uso, también pueden deshumanizar la resolución de conflictos en tiempos de paz cuando se utilizan para la seguridad interna. Al igual que se ha reconocido que los ejércitos no deben utilizarse como herramienta común para solucionar conflictos ciudadanos, los SAAL no deberían utilizarse en ningún caso por parte de la Administración como recurso para substituir la decisión de las autoridades públicas de actuar preventivamente.

12. La segunda enmienda: el peligro de la expansión de los SAAL

En los Estados Unidos se garantiza constitucionalmente la autodefensa armada de sus ciudadanos y la milicia (Guardia Nacional de cada estado) a través de la Segunda Enmienda de la Constitución federal según la cual “el derecho del pueblo a tener y llevar armas no puede ser infringido”.

Las sentencias de la Corte Suprema Distrito de Columbia v. Heller (2008) y McDonald v. Chicago (2010) establecieron una jurisprudencia según la cual a los ciudadanos estadounidenses no solo se les reconoce el derecho fundamental a poseer, en su hogar, cualquier tipo de arma (incluso automática o militar) que se considere necesaria para su seguridad. Y es que se dice que, según el significado original, la segunda enmienda garantiza al ciudadano el poder de combatir tanto la criminalidad común como la posible tiranía del poder público, por lo que ni la Federación ni los estados podrían restringir ese derecho excepto ante delincuentes, enfermos mentales o en ciertos recintos como escuelas y recintos gubernamentales. El autor del fallo de Columbia v. Heller fue el juez Scallia según el cual esta enmienda recoge un derecho

individual y no colectivo por lo que el ciudadano no tiene porqué pertenece a una milicia para portar armas.

Por esa razón, cabe la posibilidad de una *vis expansiva* de la interpretación de las armas que pueden estar en poder de los ciudadanos para su defensa personal considerando la posibilidad de disponer de SAAL, de robots o drones autónomos para la defensa personal de los ciudadanos americanos y de sus Guardias Nacionales formadas al 50% por voluntarios. Ciertamente, podrían reclamar la posesión y tenencia de un dron o un minidron autónomo con inteligencia artificial para su propia seguridad lo cual podría conllevar una nueva epidemia de destrucción y muertes indiscriminadas en tiempo de paz y sin declaración de guerra, ni aplicación del DIH ni del *ius belli*.

Las armas más peligrosas van a ser las armas con IA más pequeñas debido a su poder de ocultación, letalidad y descontrol por parte de los poderes públicos. Y cuanto más imperceptibles para el ojo humano sean con la más sofisticada tecnología, será más difícil repeler o proteger a las víctimas inocentes contra su ataque. Los mini SAAL domésticos serán una nueva amenaza si su posesión se generaliza entre la población civil en forma de drones armados. Pero además, esta situación puede tener un denominado efecto imitación sobre los ciudadanos de otros países que, sin tener derecho constitucional a armarse pueden pretender equiparse con esta tecnología para garantizar su derecho constitucional a la vida y a su legítima defensa.

Como señala Kreps, “la revolución de los drones ya ha cambiado la guerra y pronto se convertirá en una herramienta común también en el contexto civil. Está claro que la tecnología de los drones ha llegado para quedarse” (Kreps y Fuhrmann, 2016).

Ahora bien, a pesar del minimalismo judicial estadounidense y de la alegalidad imperante respecto a los SAAL en todo el mundo, lo cierto es que existe un activo movimiento ciudadano a favor de la prohibición de dicha tecnología militar que está intentando transformar el horizonte de la seguridad global.

Sin embargo, dichos movimientos solo tienen el objetivo de que los Estados prohíban los SAAL para fines militares. No advierten que el terror no sólo puede provenir de estados irresponsables, corruptos o totalitarios, o de grupos terroristas, o de hackers que se vendan al mejor postor o a su ego, sino de ciudadanos libres de cualquier sospecha, pero que, teniendo derecho a armarse, como los estadounidenses, deciden hacerlo con sistemas autónomos de este tipo, que pueden provocar una auténtica escalada de accidentes, muertes y terror descontrolado.

Si las armas tradicionales de autodefensa (automáticas y semiautomáticas) hoy suman **diez veces más víctimas en sociedades armadas como Estados Unidos que en Europa**, imagínense lo que pasaría si estos ciudadanos, amparados por su Constitución, fueran equipados con armas letales autónomas con IA para su seguridad privada si no hay un claro pronunciamiento judicial en este sentido.

Esta polémica ya existe con las armas automáticas, y el fallo del juez federal del distrito de Massachusetts del 5 de abril de 2018, deja claro que las armas de asalto AR-15 no están comprendidas en la garantía de la segunda enmienda y pueden ser prohibidas, con lo que le da la razón a la fiscal general del Estado, Mauren Healey, que defendió la legalidad de las políticas estatales

restrictivas(<https://www.mass.gov/enforcing-the-massachusetts-assault-weapons-ban>).

Para proponer argumentos definitivos hay que tener en cuenta la cifra de 30.000 víctimas al año por armas de fuego (entre homicidios, asesinatos, accidentes y suicidios) en Estados Unidos a causa de la Segunda Enmienda y su amplia interpretación. Las masacres o las víctimas anónimas diarias por disparos de armas convencionales en las manos de los ciudadanos en tiempos de paz son una tragedia diaria desde los años 80, con casi 100 muertos al día. Y son armas que están físicamente vinculadas a una persona y son responsabilidad de quien aprieta el gatillo. No disparan si no hay un humano que lo haga. Sin embargo, la llegada de mini SAAL podría llevar a una proliferación de armas de autodefensa con IA que pueden decidir de forma independiente a quién disparar.

En lugar de los 300 millones de armas cortas y largas, automáticas y semiautomáticas que circula hoy en día en los Estados Unidos, podríamos enfrentarnos a 300 millones de drones letales con IA para la defensa personal de sus dueños, pero descontrolados, alienados físicamente de ellos (puesto que su actividad será independiente de los movimientos de sus “dueños”) y con una multiplicación exponencial de víctimas inocentes.

Por otro lado, deberemos tener en cuenta la influencia de la **“Doctrina del Castillo”** en los Estados Unidos (Martínez, 2003). Es una línea interpretativa según la cual se designa el domicilio como un lugar en el que la persona tiene protecciones e inmunidades que le permiten, en determinadas circunstancias, utilizar la fuerza letal para defenderse de un intruso y queda libre de enjuiciamiento legal por las consecuencias de la fuerza empleada. Su vestigio permanece de manera sobresaliente como un

conjunto de principios que se incorporan en diferentes grados tanto a través de la ley como de la jurisprudencia. Si a esto le sumamos la interpretación amplia de la Segunda Enmienda, la regulación de los SAAL es claramente urgente por la proyección internacional que puede tener su uso en manos de privados una vez sean asequibles y de tamaño doméstico. De lo contrario podrían convertirse en Armas de destrucción masiva individualizada (ADMI). Los ciudadanos americanos (y por emulación en el resto de países) van a sentirse empujados ya no solo a armarse con un arma convencional, una pistola o un arma de asalto automática como la AK46, sino a utilizar una **pocket-dron gun** para su seguridad. Los autores alertan de que pasaremos de una amenaza militar a una amenaza de armas de destrucción masiva o selectiva en manos de individuos libres de toda sospecha inicial.

El derecho no puede huir de sus responsabilidades. La regulación debe impedir la creación de estas ADMI antes de que sea demasiado tarde y debe forzar a que aflore la información que está en poder de los diferentes agentes involucrados (investigadores, empresas, gobiernos, etc.) sobre este nuevo peligro a nivel mundial para que su I+D+I no sea dar un cheque en blanco a la tecnología letal con Inteligencia artificial. Sería un atractivo negocio que la industria armamentística en combinación con la industria tecnológica especializada en IA no va a dejar de aprovechar. Pero ello comportaría un descontrol inadvertido aún, indescriptible e impredecible.

Antes de que sea demasiado tarde no deleguemos potestades indelegables a la tecnología letal o la humanidad va a sufrir un duro golpe de forma inesperada.

- Asaro (2012) “On banning autonomous weapon systems: human rights, automation and the dehumanization of lethal decision-making”, *International Review of the Red Cross*, 94, pp 687-709.
- Axelrod (1984), *The evolution of co-operation*, Nueva York: Basic.
- Boix Palop, A., (2007), “De McDonald's a Google: la ley ante la tercera revolución productiva”, *Teoría y Derecho*, núm.1, p.141-145.
- Boix Palop, A., (2020) “Los algoritmos son reglamentos: la necesidad de extender las garantías propias de las normas reglamentarias a los programas empleados por las administraciones para la adopción de decisiones”, *Revista de Derecho Público*, vol.1, p.225.
- Fischer (2008), *Las emociones en la negociación*, Granica.
- Frischmann, B., y Selinger, E., (2018), *Re-engineering humanity*, Cambridge University Press.
- Guaita, (1960), *Derecho administrativo especial*, tomo I.
- Gutiérrez, C., Cervell, M.J. (2013), “Sistemas de armas autónomas, drones y derecho internacional”, *Revista del Instituto Español de Estudios Estratégicos*, núm.2, p.30-31.
- Heyns, C. (2016) “Autonomous weapons systems: living a dignified life and dying a dignified death”, en Butha, *et al.*(eds.) *Autonomous weapons systems: law, ethics, policy*. Cambridge University Press.
- Howard, Zhang y Horvitz (2017), “Addressing bias in machine learning algorithms: a pilot study on emotion recognition for intelligent systems” en *IEEE Workshop on Advanced robotics and its social impacts*, Austin.
- Kang y Roose (2018), “Zuckerberg faces hostile Congress as calls for regulation mount”, *The New York Times*, 11 de abril de 2018.
- Kreps, S. y Fuhrmann, M. (2016), “Separating fact from fiction in the debate over drone proliferation”, *International Security*, 41(2), p.146.
- Lin, P. *et al.* (2017), *Robot ethics 2.0: from autonomous cars to artificial intelligence*, Oxford University Press.
- Martínez, R. (2003), *Armas: ¿libertad americana o prevención europea?*, Ariel.
- Martínez, R. y Rodríguez, J., (2018) *Inteligencia artificial y armas letales autónomas*, Trea.
- McQuillan (2015), “Algorithmic states of exception”, *European Journal of cultural studies*, 18, pág.564-576.
- O’Neil, (2017), *Weapons of math destruction: how big data increases inequality and threatens democracy*, New York: Broadway.
- Ramió, C. (2019), *Inteligencia artificial y administración pública: robots y humanos compartiendo el servicio público*, Catarata.
- Rodella (2018), “Los robots pueden imitar, no sentir”, *El País*, 28 abril.
- Rodríguez-Arana (1993), *Principios de ética pública: ¿Corrupción o servicio?*, Madrid: Montecorvo.
- Rodríguez, J. (2018), “Social Challenges of Artificial Intelligence: The case of Lethal Autonomous systems”. *International Journal of Law*, 244-267.
- Rodríguez, J. *et al.* (2019), “Nuevas Armas contra la ética y las personas”. *Informe 29 Centre Delàs d'estudis per la PAU*.
- Roff, H., (2016), “Artificial intelligence, Autonomous weapons, and meaningful human control”, de Global Security Initiative Arizona State University, FLI-RFP-AI1 program Future of Life Institute (<https://globalsecurity.asu.edu/robotics-autonomy>).
- Russel, S., *et al.*, (2010), *Artificial Intelligence. A modern approach*, New York:Pearson, 3^a.ed.
- Scharre, (2018), *Army of None: autonomous weapons and the future of war*, Norton and Company, pág. 322.

- Skinner, C. (2018), *Digital Human. The four revolution of humanity includes everyone*, Wiley.
- Sparrow (2016), “Robots and respect: assessing the case against autonomous weapon systems”, *Ethics and International Affairs*, 30, p.93-116.
- Tegmark, M., (2018), *Vida 3.0. Ser humano en la era de la inteligencia artificial*, Taurus.
- Uri (2005), *Alcanzar la paz: resolución de conflictos y mediación en la familia, el trabajo y el mundo*, Paidós.
- Waldman, P. (2006), *El Estado anómico: derecho, seguridad pública y vida cotidiana en América Latina*, Madrid: Editorial Iberoamericana.