





ROSER MARTÍNEZ QUIRANTE  
JOAQUÍN RODRÍGUEZ ÁLVAREZ

Inteligencia artificial  
y armas letales autónomas  
*Un nuevo reto para Naciones Unidas*

Trea Ensayos

COLECCIÓN TREA ENSAYOS

Primera edición: septiembre del 2018

© del texto: Roser Martínez Quirante y Joaquín Rodríguez  
Álvarez, 2018

© de esta edición: Ediciones Trea  
Polígono de Somonte / María González la Pondala, 98, nave D  
33393 Somonte-Cenero. Gijón (Asturias)  
Tel.: 985 303 801 / Fax: 985 303 712  
trea@trea.es / www.trea.es

Corrección: Pablo Batalla Cueto  
Producción: Patricia Laxague Jordán  
Impresión: Gráficas Ápel  
Encuadernación: Encuastur

D. L.: AS 02571-2018  
ISBN: 978-84-17140-74-8

Impreso en España. *Printed in Spain*

Todos los derechos reservados. No se permite la reproducción total o parcial de este libro, ni su incorporación a un sistema informático, ni su transmisión en cualquier forma o por cualquier medio, sea este electrónico, mecánico, por fotocopia, por grabación u otros métodos, sin el permiso previo por escrito de Ediciones Trea, S. L.

La Editorial, a los efectos previstos en el artículo 32.1 párrafo segundo del vigente TRLPI, se opone expresamente a que cualquiera de las páginas de esta obra o partes de ella sean utilizadas para la realización de resúmenes de prensa.

Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra solo puede ser realizada con la autorización de sus titulares, salvo excepción prevista por la ley. Diríjase a CEDRO (Centro Español de Derechos Reprográficos) si necesita fotocopiar o escanear algún fragmento de esta obra ([www.conlicencia.com](http://www.conlicencia.com); 91 702 19 70 / 93 272 04 47).

# Índice

<b>Prólogo</b> .....	9
<b>Introducción</b> .....	15
<b>1. Contexto</b> .....	25
Cuerpos y tecnología como objeto de consumo.....	28
El sustrato tecnológico .....	36
Sistemas de armamento letal autónomo .....	51
<b>2. Tecnología y sociedad: implicaciones de la inteligencia artificial</b> .....	65
Humanidad y tecnología.....	65
Inteligencia artificial y percepción .....	72
Sistemas de armamento e inteligencia artificial .....	82
<b>3. El derecho ante la inteligencia artificial en los LAWS</b> ..	91
Introducción .....	91
Los LAWS como amenaza a los derechos a la dignidad humana y a la vida .....	94
El derecho frente al Estado de excepción algorítmico... ..	107
Política, opacidad y connivencia .....	119
Un nuevo autoritarismo .....	130
<b>Epílogo</b> .....	139
<b>Bibliografía</b> .....	149



ROSER: a mi hija Muriel, un regalo de la vida que no deja de sorprenderme con su inmenso cariño, creatividad y afán por saber, y que se integra, por suerte, en una generación que sueña con un nuevo mundo bajo la *intelligence du rêve* y sobre todo con la *puissance de la douceur*.

JOAQUÍN: a Diego, por ser capaz de hacerme ver más allá.





## Prólogo

En la actualidad, una de las mayores amenazas que se le presentan a la humanidad es lo que algunos llaman la *tercera revolución en la guerra*, cuyo producto final podría ser la automatización total del conflicto armado por medio de máquinas de matar capaces de operar por sí solas, sin un control humano significativo. Las principales potencias del planeta (Rusia, China y Estados Unidos), así como otras naciones (como el Reino Unido o Israel), están desarrollando ya tanques, barcos, aviones de combate, submarinos y otras armas que pueden rastrear, identificar y atacar objetivos sin la necesidad de un controlador humano.

Estos desarrollos tecnológicos en curso obligan claramente a un debate internacional sobre si se debe permitir que la decisión de matar a un ser humano sea delegada en sistemas de armamento autónomo. Y ese debate ha comenzado ya. La discusión ha abarcado desde las implicaciones morales y legales del asunto<sup>1</sup> hasta preocupaciones de orden técnico y operativo,<sup>2</sup> pasando por cuestiones relacionadas con la seguridad internacional: se ha expresado la preocupación de que los algoritmos de combate desconocidos que controlan las armas autónomas interactúen de formas impredecibles, lo que podría imposibilitar que las revisiones

<sup>1</sup> Asaro (2012); Heynes (2013). Ver también Heynes (2016), cap. 2.

<sup>2</sup> Ver Sharkey (2012) y Sharkey (2010).

de armas garanticen el cumplimiento del derecho internacional humanitario.<sup>3</sup>

Parece evidente que en el futuro previsible no podremos garantizar ese cumplimiento. Además de los problemas que revisten los principios de distinción y proporcionalidad en lo que respecta a determinar la legitimidad de los objetivos, los sistemas autónomos armados son por definición menos predecibles que otros. Aún no está claro cómo podría garantizarse en este caso la calidad de las revisiones de armas contempladas en el artículo 36 del Protocolo Adicional I de los Convenios de Ginebra.<sup>4</sup> Además, el Departamento de Defensa de Estados Unidos ha señalado una serie de problemas informáticos que el uso de sistemas armamentísticos autónomos puede comportar: error humano, fallos de interacción hombre-máquina, mal funcionamiento, degradación de comunicaciones, errores de codificación de *software*, ciberataques enemigos, infiltración en la cadena de suministro industrial, bloqueo, *spoofing*, señuelos y otras contramedidas o acciones enemigas y situaciones imprevistas en el campo de batalla.<sup>5</sup>

Algunos opinan que tales armas deberían poder usarse legalmente en ciertas circunstancias muy limitadas, mientras que otros sostienen que en algún momento en el futuro estos sistemas de armamento podrán cumplir íntegramente con el derecho internacional humanitario. Pero ambos conjuntos de estudiosos se refieren a una tecnología sujeta a tal derecho que nadie sabe cómo crear.

No hay nada de malo en la ambición tecnológica o en una agenda de investigación general pensada para los dominios

<sup>3</sup> Sharkey (2011).

<sup>4</sup> Convenciones de Ginebra, 12 de agosto de 1949, 1125 UNTS 3, artículo 36.

<sup>5</sup> Departamento de Defensa de los Estados Unidos (DoD): *Autonomy in weapon systems*, directiva 3000.09, 21.11.2012.

civiles, pero hay o debería haber menos lugar para la conjetura y la experimentación cuando se discuten tecnologías como los sistemas de armamento autónomo. Así, por ejemplo, hoy se experimenta con el *fútbol robótico*: un gran desafío investigador que se acomete con la convicción de que se trata de una oportunidad de probar la tecnología robótica y sus avances en aplicaciones reales. El objetivo final es desarrollar un equipo de robots humanoides autónomos capaces de derrotar a los campeones del mundo en 2050. Nadie sabe si tal cosa será posible, pero entretanto se van desarrollando avances que pueden aplicarse en otros lugares.<sup>6</sup> Para esta investigación, el éxito del objetivo final no es vital: si la empresa falla, podrá inventarse un nuevo deporte en el que humanos y robots jueguen juntos pero que disponga de reglas de enfrentamiento que otorguen a los robots la misma oportunidad de victoria.<sup>7</sup>

El asunto cobra otro cariz cuando hablamos de armas. Si al final resulta que cumplir escrupulosamente con el derecho internacional humanitario no es posible o tan fácil como se pensaba, ¿qué hacemos con este nuevo sistema de armamento automático? ¿Qué pasa si entretanto nos involucramos en conflictos serios? Es posible que haya que modificar lo que significa el cumplimiento del derecho y las reglas de enfrentamiento para otorgarle un lugar a las nuevas armas; algo que ya hubo que hacer en el pasado cuando los bombardeos aéreos y la guerra submarina transformaron radicalmente los conflictos bélicos.

Las limitaciones actuales en el uso de la tecnología para la guerra se deben en parte a que Estados tecnológicamente capaces, como el Reino Unido o Estados Unidos, han dejado

<sup>6</sup> Datteri y Tamburrini (2013).

<sup>7</sup> Bhuta, Beck, Geib, Liu y Krebs (2016).

claro que tiene que haber alguna forma de supervisión o juicio humano de las decisiones que comporten letalidad. En el Reino Unido, el subsecretario de Estado parlamentario para la Defensa, John Jacob Astor de Hever, dijo lo siguiente en 2013: «[El] Ministerio [de Defensa] no tiene actualmente la intención de desarrollar sistemas que operen sin intervención humana [...] la operación de los sistemas de armas siempre estará bajo control humano».<sup>8</sup> Por su parte, el Departamento de Defensa de Estados Unidos declaró lo siguiente cuando publicó su primer documento de política sobre armas autónomas: «Los sistemas de armas autónomas y semiautónomas se diseñarán de tal manera que permitan que los comandantes y los operadores ejerzan niveles apropiados de juicio humano sobre el uso de la fuerza».<sup>9</sup>

En ninguno de los dos países, sin embargo, se ha dejado completamente claro qué tipo de supervisión humana se empleará. La afirmación de que habrá un ser humano en el ciclo de control es demasiado vaga: puede significar simplemente que un humano programará el sistema para una misión o presionará el botón que lo active. Solo con suerte puede esto querer decir que existirá una deliberación humana sobre la legitimidad de un objetivo antes de iniciar un ataque sobre él.

Estamos ante un momento crítico en nuestra historia, en el cual ya se perpetran asesinatos automatizados que son un paso de gigante en la revolución industrial de la guerra. Sin embargo, hasta 2012 no se celebró ninguna discusión internacional al respecto. Solo en octubre de aquel año la sociedad civil dio un paso al frente en forma de una reunión de organizaciones no gubernamentales que se celebró en Nueva York coincidiendo con el aniversario de la prohibición de

<sup>8</sup> Cf. <<https://publications.parliament.uk/pa/ld201213/ldhansrd/text/130326-0001.htm#st>>.

<sup>9</sup> Saxon (2014).

las minas terrestres antipersonales. Como parte de la misma, yo fui invitado a impartir una charla sobre los peligros de los sistemas autónomos de armamento. Posteriormente, un grupo de siete oenegés (Iniciativa de Mujeres Nobel, Human Rights Watch, Artículo 36, Comité Internacional para el Control de Armas Robóticas, Mine Action Canada, Pugwash y PAX) decidió liderar una campaña internacional contra los robots asesinos y presentarla en el Parlamento del Reino Unido en abril de 2013.

En noviembre de 2013 se dio cumplimiento al mandato de celebrar una reunión de expertos de cuatro días en la ONU (en Ginebra), y más concretamente en el marco del Convenio sobre Ciertas Armas Convencionales (CCAC), consagrado a la protección de los civiles en los conflictos armados y también a evitar a los combatientes sufrimientos innecesarios. 121 Estados forman parte del CCAC, y todos tienen el mismo derecho a voto y a veto. Más tarde, hubo dos semanas más de reuniones de expertos en 2015 y 2016 antes de que el CCAC decidiera colectivamente pasar al siguiente nivel en 2017: trasladar el asunto a un grupo de expertos gubernamentales abierto a los mismos 121 Estados, pero con el encargo de discutir ya abiertamente y en firme en lugar de simplemente seguir las deliberaciones de los paneles de expertos y cuestionarlas.

Para el año 2018, la cuestión del control humano significativo de las armas se ha tratado ampliamente en declaraciones de más de ochenta naciones. La campaña ha crecido para incluir a más de setenta oenegés y la mayoría de los Estados aceptan ya que se necesita una regulación internacional de estas nuevas armas y su uso. 26 Estados-nación se han adherido concretamente a nuestro llamamiento a un nuevo protocolo internacional para prohibir los sistemas de armas autónomas. Los últimos han sido Austria y China.

El Comité Internacional para el Control de las Armas Robóticas, que presido, desempeña un papel principal en esta campaña y en su defensa en la ONU. Y estamos muy orgullosos de que los autores de este libro, Joaquín y Roser, sean dos de nuestros miembros.

NOEL SHARKEY, presidente de ICRC  
(International Committee for Robot Arm Control)

# Introducción

ROSER MARTÍNEZ Y JOAQUÍN RODRÍGUEZ<sup>10</sup>

Vivimos en un *mundo líquido* en el que se entremezclan fragmentos de sociedades complejas y disociadas, como fases de un sueño que apenas tiene sentido.<sup>11</sup> Habitamos un espacio y un tiempo en el que la tecnología se erige como frontera entre lo deseado y lo temido, prometiéndonos confort pero augurando nuevos conflictos; una encrucijada histórica en la que ha de determinarse el futuro no solo de las próximas generaciones, sino de la misma vida sobre el planeta.

Podemos afirmar que nuestro presente está definido en gran parte por la aceleración exponencial del sistema tecnocientífico, lo que propicia el surgimiento de nuevos riesgos sistémicos y transformaciones trascendentales que comprometen no solo la resiliencia del sistema, sino la de los paradigmas científicos sobre los que asentamos el entramado jurídico-institucional que le da forma.<sup>12</sup>

Podríamos afirmar también que la evolución del sistema tecnológico alberga el origen de sendas crisis jurídica y filosófica que se configuran como síntomas de un salto caótico en el que pasado, presente y futuro parecen coexistir en

<sup>10</sup> Queremos agradecer la colaboración de nuestro catedrático, el doctor Manuel Ballbé, en el proceso de elaboración de este libro, ya que el descubrimiento de muchas de las obras que hemos utilizado para defener los argumentos aquí planteados han sido fruto de su incansable perseverancia en la investigación y a su generosidad en compartirlas con sus discípulos.

<sup>11</sup> Rocca (2008).

<sup>12</sup> Giddens (1999).

un equilibrio inestable plagado de amenazas.<sup>13</sup> La crisis, tal y como afirmaba Gramsci, «consiste precisamente en el hecho de que lo viejo muere y lo nuevo no puede nacer: en este interregno se verifican los fenómenos morbosos más variados».<sup>14</sup>

Todos esos fenómenos, procesos y sistemas tecnológicos tienen hoy el potencial de amenazar seriamente no solo la estabilidad del sistema, sino la de la propia especie. Por ello, es necesario establecer marcos regulatorios en base a una nueva noción de *derecho anticipatorio* como forma jurídica inherente a la condición *post-humana*,<sup>15</sup> o lo que es lo mismo, un *derecho post-antropocéntrico* que nos guíe durante esta etapa de transición a partir de un imperativo ético consensuado socialmente.

Hoy, una novedosa e inquietante tecnología militar anuncia el surgimiento de una nueva carrera armamentística que puede alterar dramáticamente los frágiles equilibrios geopolíticos del planeta en un momento en que los paradigmas que regulan la escena internacional están en fase crítica de autodestrucción por las políticas de Donald Trump o las que representa el *Brexit* en el Reino Unido en tanto que debilitamiento de la Unión Europea; pero también de transformación y reforzamiento ante el auge de una China autoritaria o de la Rusia de Putin como nuevos actores clave en una gobernanza global cada vez más desequilibrada.

En este sentido, es innegable que hasta la fecha estábamos inmersos en un ambiguo proceso de globalización/americанизación que por un lado representó la conquista de nuevos derechos gracias al esfuerzo de los movimientos por los derechos civiles de los afroamericanos, de las mujeres, LGTBIQ,

<sup>13</sup> Consúltese la *teoría general de los sistemas evolutivos* de Csányi (1989).

<sup>14</sup> Gramsci (1999).

<sup>15</sup> Arendt (2015); Rodríguez (2016).



medioambientales, antimonopolios, sanitarios (*Obamacare*), etcétera, que culminaron en las presidencias de Clinton y Obama;<sup>16</sup> pero por otro un retroceso regulador acometido por las presidencias de Reagan, Bush y Trump y que nos ha llevado a crisis financieras como la del 2008, así como a graves crisis ambientales y en general a una merma de derechos y a un escenario futuro tenebroso. Ahora bien, como contraposición también ha habido un proceso paralelo de globalización/europeización competitivo que ha servido para recoger el modelo regulador progresista norteamericano pero corregir o advertir algunos de sus errores y disfunciones. En Estados Unidos se habla de un *efecto Unión Europea* que ha sido criticado contundentemente por medios como el *Wall Street Journal*, que lo definió como un *imperialismo regulatorio* (*regulatory imperialism*).<sup>17</sup> Son términos injustos. Lo que es realmente el *efecto Unión Europea* es un intento de asegurar un *Law's Empire* de los derechos humanos sobre el resto del mundo,<sup>18</sup> con regulaciones progresistas de todo tipo (sociales, sanitarias, alimentarias, laborales, etcétera), algunas de ellas promovidas en origen a través del activismo de grupos norteamericanos. Para poder comerciar con Europa —un mercado de quinientos millones de personas—, el resto de Estados ha tenido que adaptarse a dichas regulaciones, y ello ha supuesto un cierto freno al capitalismo salvaje, thatcheriano-reaganita, de las últimas décadas y una cierta vuelta a un capitalismo social y regulador como aquel cuyo máximo exponente fue el presidente demócrata Franklin D. Roosevelt.

<sup>16</sup> Ballbé y Martínez (2009); Sunstein (2016b). Vid. también Epp, C. (2014).

<sup>17</sup> «Regulatory imperialism», *The Wall Street Journal*, 26 de octubre de 2017 [en línea], <<https://www.wsj.com/articles/SB119334720539572002>>. [Consulta: 7-9-2018].

<sup>18</sup> Dworkin (1986). Véase también Álvarez (2009), citado en Ballbé y Martínez (2009), p. 209.

Actualmente enfrentamos el dilema, que ya planteaba Álvarez, de si el derecho internacional contemporáneo va a ser el *imperio de la ley* o la *ley del imperio*. Hay tendencias contrapuestas en ese sentido y una a regresar al unilateralismo norteamericano, como prueba que Donald Trump haya propuesto crear una Fuerza Militar Espacial (USSF) independiente que probablemente acabe albergando la tecnología más puntera en lo que respecta a los sistemas inteligentes que analizaremos a continuación. Se hace acuciante que en Naciones Unidas se alcance un consenso que frene dicho unilateralismo imperialista. En este sentido, el presidente de Francia, Emmanuel Macron, ha anunciado que impulsará la soberanía europea ante el avance del nacionalismo populista norteamericano y que procurará que la Unión Europea reclame su autonomía militar.<sup>19</sup>

En cuanto a otras grandes potencias, como Rusia o China, es cierto que su influencia puede hacer cambiar de estrategia a los actores democráticos tradicionales, pero cabe hacer notar que la articulación entre esa competición y la cooperación entre bloques está siendo canalizada satisfactoriamente a través de Naciones Unidas, que ha creado un inmenso corpus de *soft law* pero también de *hard law* de cuyo éxito y alcance no somos del todo conscientes. La ONU fomenta un encomiable proceso deliberativo entre los Estados y los grupos a fin de lograr una mayor seguridad humana,<sup>20</sup> reconoce el activismo de grupos que en todos los campos presionan y condicionan a los Estados y aporta una *expertise* de alto contenido científico y jurídico en beneficio de los derechos humanos. Es decir: en la actividad de esta organización no solo hay competición entre bloques, sino también búsqueda de consensos a través

<sup>19</sup> M. Bassets: «Macron impulsa la soberanía militar europea ante Estados Unidos», *El País*, 27 de agosto de 2018.

<sup>20</sup> Alemán (2016) y Gómez Hinojosa (2018).

de la cooperación para llegar a soluciones en las que todos se sientan participantes, protagonistas y ganadores. A través de la interacción cotidiana entre los reguladores y funcionarios especializados de las diferentes agencias administrativas estatales, de los distintos movimientos ciudadanos y grupos de interés, etcétera, se lleva a cabo un verdadero *law-making*.<sup>21</sup> A los protagonistas de este proceso, Slaughter los ha denominado *los nuevos diplomáticos*.<sup>22</sup>

Este papel de *law-maker* de Naciones Unidas cobra especial importancia ante el desarrollo de hibridaciones de inteligencia artificial en el campo armamentístico, porque estos sistemas podrían suponer una amenaza mayor que la nuclear. Nos referimos a los LAWS (*lethal weapons autonomous system*): una nueva tipología de armas que pueden llegar a ser totalmente autónomas, es decir, carentes de control humano significativo en fases críticas de su uso.<sup>23</sup> Si a estos sistemas se les suman los últimos avances en *inteligencia artificial fuerte* o *general*, podríamos vernos abocados a un mecanismo imparabile de destrucción no solo genocida, sino selectiva e individualizada, que tendrá como autor y testigo únicamente a un ente sintético. Uno de los objetivos de este trabajo es alertar de que la nueva amenaza letal no puede conservar la denominación de *autónoma*: en realidad, estas LAWS con inteligencia artificial general serán independientes, y por lo tanto proponemos denominarlas LIWS (*lethal independent weapons systems*). De otro modo, estaremos relativizando la gravedad de la amenaza.<sup>24</sup>

<sup>21</sup> Teoría defendida por Álvarez (2005).

<sup>22</sup> Slaughter (2004), citado en Ballbé y Martínez (2009).

<sup>23</sup> Unesco (2017), p. 25.

<sup>24</sup> El ICRC (*International Committee for Robot Arms Control*) se ha convertido en el centro interuniversitario e interinstitucional (compuesto por académicos especialistas en nanotecnología, derecho, etcétera) más innovador para poner orden en este nuevo sistema de armas autónomas de destrucción masiva.

Los LAWS sin control humano significativo existen, se conocen y ya están perpetrando dramáticos errores como el perpetrado en agosto de 2018 en Yemen por el SAQR1, un verdadero LAWS. Se trata de un dron saudí, parece que de patente americana, que tiene la capacidad de transportar misiles y bombas guiadas por láser y puede volar a más de 2500 kilómetros de distancia y a una altura de 25 000 pies. La venta de este tipo de armas se prohibió durante la presidencia de Obama, pero tal medida fue revocada por Trump, lo cual puede estar posibilitando la transferencia de dicha tecnología a otros Estados y empresas privadas. El 9 de agosto de 2018, la coalición liderada por los saudíes en la guerra del Yemen provocó con ella la muerte inadmisible de decenas de niños que viajaban en un autobús y se dirigían hacia un campamento de verano de Unicef, lo cual fue considerado por Naciones Unidas un crimen de guerra. Consciente o inconscientemente, países autoritarios como el saudí quieren fulminar del escenario de guerra a quienes controlan la aplicación del derecho humanitario en los conflictos, pero en este caso el tiro salió por la culata: las destrozadas mochilas azules de Unicef fueron profusamente fotografiadas y dieron la vuelta al mundo.

A partir de estas constataciones, el objetivo principal de esta obra es profundizar en el necesario debate legal, administrativo y científico/tecnológico que ha de concitar esta nueva amenaza que se percibe como una de las más graves y preocupantes, no solo para los equilibrios geopolíticos que mantienen esta paz frágil en la que vivimos hoy en día, sino para la supervivencia de nuestra especie. Estas nuevas formas de armamento suponen una ruptura del contrato social y constitucional global en que se basa la protección de los derechos humanos. Hemos creado armas automáticas y armas autónomas y al final llegaremos a desarrollar armas to-

talmente independientes, incluso basadas en la nanotecnología; y se hace evidente la urgencia de una regulación de estos nuevos sistemas de armamento independientes del control humano inicialmente ejercido sobre los mismos. Los futuros LIWS deberían ser prohibidos de forma anticipada y preventiva; y debería serlo también la mera investigación conducente a ese objetivo. Hablamos de unas armas que no solo representan un desafío tecnológico, sino también uno jurídico y otro ético,<sup>25</sup> pues ponen en cuestión los tratados internacionales surgidos en el período de entreguerras y consolidados en las décadas posteriores a la segunda guerra mundial.

Elon Musk, cofundador de SpaceX y Tesla, ha alertado de que «en la era de la inteligencia artificial podríamos crear un dictador inmortal del que nunca escaparíamos», así como que «la competición por el desarrollo de la inteligencia artificial se ha convertido en el mayor riesgo para una tercera guerra mundial, ya que el país que lidere la investigación en inteligencia artificial llegará a dominar los asuntos globales».<sup>26</sup>

En este contexto, y como ya se ha apuntado, organismos internacionales como Naciones Unidas cobran una relevancia especial a la hora de aunar los esfuerzos para generar instrumentos jurídicos articuladores de la estrategia preventiva. Los LAWS/LIWS están llamados a ser una de las piezas claves de un nuevo complejo industrial-militar que ve en ellos la próxima gran revolución del sector, representando un nuevo estadio post-nuclear y post-bioquímico;<sup>27</sup> y el *prin-*

<sup>25</sup> Tasioulas (2018).

<sup>26</sup> «Elon Musk teme que un robot dictador inmortal acabe dominando el mundo», *Europa Press*, 6 de abril de 2018 [en línea], <<http://www.europapress.es/ciencia/laboratorio/noticia-elon-musk-teme-robot-dictador-inmortal-acabe-dominando-mundo-20180406123847.html>>.

<sup>27</sup> Famosa advertencia del general y presidente estadounidense Dwight D. Eisenhower en su discurso de despedida en 1961: «Nuestro trabajo, los recursos y los medios de subsistencia son todo lo que tenemos; así es la estructura mis-

*cipio jurídico de precaución* (de tradición europea) puede convertirse en un elemento decisivo para frenar esta carrera armamentística.<sup>28</sup>

Se trata nada menos que de prevenir genocidios o crímenes contra la humanidad, como ha señalado una de las más autorizadas autoras en esta materia, Samantha Power, embajadora de la ONU durante la presidencia Obama, profesora de la Harvard Kennedy School of Government y Premio Pulitzer por *Problema infernal: Estados Unidos en la era del genocidio*. En esa obra, la autora analiza cómo ha reaccionado el gobierno de Estados Unidos ante los distintos genocidios del siglo xx y denuncia su negativa a actuar o a emplear la palabra *genocidio* para referirse a algunos de ellos y que se haya refugiado en la ausencia de intereses estadounidenses en los países involucrados. Siguiéndola, se constata fácilmente que la inactividad estadounidense frente a la amenaza de los LIWS no es más que el mantenimiento y la continuación de esos desentendimientos.

En este libro, a lo largo del primer capítulo describiremos el contexto y el estado de la cuestión que envuelve el surgimiento de esta nueva generación de armas que apunta a erigirse como la nueva frontera de la carrera armamentística, fundamentalmente derivada del uso extensivo de la robótica con carencias jurídicas y reguladoras evidentes. Además,

---

ma de nuestra sociedad. En los consejos de gobierno, debemos evitar la compra de influencias injustificadas, ya sea buscadas o no, por el complejo industrial-militar. Existe el riesgo de un desastroso desarrollo de un poder usurpado y [ese riesgo] se mantendrá. No debemos permitir nunca que el peso de esta conjunción ponga en peligro nuestras libertades o los procesos democráticos».

<sup>28</sup> Los riesgos tecnológicos no pueden ser calculados de acuerdo con modelos tecnocráticos tradicionales como si fueran una función estadísticamente previsible de la probabilidad y de sus efectos. Regular tecnologías nuevas es un reto para ley debido a los problemas de incertidumbre y los conocimientos limitados en la valoración y la dirección de los riesgos tecnológicos aparejados. Cf. Weimer y Marin (2016) y Sunstein (2005).

¿2005a o  
2005b?

abordaremos los contextos sociopolítico y tecnocientífico que rodean el alumbramiento de estos sistemas de armamento autónomo.

El siguiente capítulo estará consagrado a analizar las relaciones existentes entre sociedad y tecnología con el objetivo de iluminar las interdependencias existentes entre ambas esferas, así como a ilustrar al lector acerca de cómo la tecnología puede adquirir valores de índole determinista e introducir nuevas relaciones éticas y estéticas, prestando especial atención a aquellos dilemas relativos al surgimiento de nuevas formas de inteligencia y los retos que representan para lo humano.

En el capítulo previo al epílogo nos ocuparemos de los dilemas jurídicos a que dan lugar los LAWS subrayando la importancia que tiene el derecho administrativo a la hora de generar instrumentos vinculantes para su regulación, sin dejar de lado los enormes retos que supone para el derecho el surgimiento de la inteligencia artificial y su aplicación. Analizaremos los parámetros que obligan al desarrollo de nuevas normas jurídicas y regulaciones preventivas que deberían tener por objeto incluso el control sobre investigaciones que se lleven a cabo en un contexto de innovación científica y tecnológica en este ámbito. Además, abordaremos asimismo el agotamiento o erosión del derecho internacional militar y humanitario a la hora de lidiar con unas tipologías de armamento que eran imposibles de concebir en el momento en que se redactaron los Convenios de Ginebra.

Finalmente, en el epílogo trataremos aquellas vías de actuación de que puede echarse mano para asegurar un progreso de la inteligencia artificial que vaya aparejado a un desarrollo humano sostenible en virtud de una concepción ampliada de la dignidad y la seguridad humanas.





# Contexto

ROSER MARTÍNEZ Y JOAQUÍN RODRÍGUEZ

La transformación del sistema al calor del proceso de globalización, así como de la incorporación masiva de nuevas tecnologías, ha permitido definir una nueva comprensión integral de la realidad para una nueva clase global cuya nueva religión se asienta en la fe en el progreso tecnocientífico (sin controles, sin límites, sin regulación), condicionando tanto la investigación como la política. Y ello sucede en un momento en que se alumbra un marco tecnológico que tiene la capacidad inherente de transformar la forma en que nos comunicamos y comprendemos tanto nuestro contexto como la otredad y a nosotros mismos.

Desde Nueva York hasta Bombay pasando por París, Dubái, Moscú o Pekín, la clase global (con unos patrones de consumo, ocio y valores estandarizados) comparte una nueva fe en las promesas de la tecnología, mientras que otros —aquellos que están vinculados a la tierra, al contexto natural— luchan para sobrevivir en un momento en que la tecnología amenaza directamente su estilo de vida y sus capacidades de supervivencia.

Podríamos afirmar que nos encontramos ante una encrucijada de futuros, de diferentes posibilidades, en la que la utopía y la distopía se confunden; ante un momento histórico en el que tenemos la posibilidad no solo de definir nuestro tiempo, sino también el de aquellos que están por venir. La revolución digital, en colaboración con nuestras capacidades científicas

y tecnológicas, nos sitúa frente a un complejo mapa de decisiones en el que intereses opuestos intentan definir un marco que pueda determinar el nuevo tiempo. Y es precisamente en la niebla de la complejidad y la incertidumbre donde pueden surgir los monstruos; monstruos que en algunas ocasiones adquieren formas extremadamente inocentes mientras crecen en las sombras, incluso al margen de la propia intención de sus creadores, como primeras materializaciones concretas del futuro por venir.

La tecnología, y especialmente la inteligencia artificial y sus aplicaciones potenciales, configura una nueva experiencia holística de la vida después de la cuarta revolución industrial.<sup>29</sup> En ella, los espacios y los tiempos convergen de una manera que erosiona la separación tradicional entre la realidad material y la digital. La consolidación de la comprensión de lo *post-humano* nos transporta a un *mundo feliz* huxleyano en el que los deseos, las aspiraciones e incluso la felicidad se trasladan de la tierra de la construcción social a la de la reproducción digital; y *sets* tecnológicos como el *Internet de las cosas* (concepto que se refiere a la red de objetos cotidianos interconectados)<sup>30</sup> tienen un papel muy importante que jugar al ser capaces de compartir una experiencia conectada entre lo *real* y lo *digital*.

Tegmark señala, siguiendo la misma línea que Putnam, que «la fructífera colaboración entre humanos y máquinas parece prometedora en muchas áreas, incluida la ciencia, donde la inteligencia artificial podría ayudarnos a los humanos a al-

<sup>29</sup> Precisamente la Escuela de Prevención y Seguridad Integral (EPSI) fundada por el catedrático de derecho Manuel Ballbé hace veinte años tenía esta intención holística respecto del estudio de la seguridad (seguridad pública, privada, alimentaria, informática, ambiental, etcétera). Cf. la página *web* de la EPSI: <<http://www.uab.cat/web/escola-de-prevencio-i-seguretat-integral-1345721289258.html>>.

<sup>30</sup> Rejón (2016), p. 183.

canzar una comprensión más profunda y hacer realidad todo nuestro potencial». <sup>31</sup> Pero hemos de tener en cuenta que, si bien la industria puede impulsar muchos desarrollos en esta materia, la academia debe desempeñar el papel esencial de proporcionar nuevas ideas técnicas y reunir a investigadores de todas las disciplinas (ciencias sociales y jurídicas, ciencias cognitivas y humanidades, ciencias de la computación y de la estadística, etcétera). <sup>32</sup> De hecho, en palabras de Jordan, se está creando una nueva rama del saber que combina todo ello, por lo que «tenemos una oportunidad real de concebir algo históricamente nuevo: una disciplina de ingeniería centrada en el ser humano». <sup>33</sup>

Así pues, nos encaminamos a un *mundo programable* en el que aquellos que tenemos el privilegio de cabalgar esta ola de *progreso* tenemos también la responsabilidad de desarrollar un sistema que pueda otorgar a cada uno de nosotros la posibilidad de vivir sin miedo, garantizando la seguridad integral y que la dignidad humana sea recuperada como derecho y valor fundamental de progreso social y ecológico. <sup>34</sup> De dibujar un nuevo horizonte «para un mundo en el que seamos socialmente iguales, humanamente diferentes y totalmente libres», como decía Rosa Luxemburgo. Sin embargo, el camino no es fácil en absoluto, y el primer paso es una mejor comprensión de nuestro contexto y sustrato tecnológico.

<sup>31</sup> Tegmark (2018) y Putnam y Feldstein (2003).

<sup>32</sup> Wladawsky-Berger (2018).

<sup>33</sup> Cit. *ibídem*.

<sup>34</sup> Vid. el trabajo recopilatorio sobre la dignidad como derecho y como valor en Barak (2015). Cf. también Frischmann y Selinger (2018).

## Cuerpos y tecnología como objeto de consumo

Los cambios sistémicos desatados a lo largo de las últimas décadas han arrojado como consecuencia una reconfiguración profunda de los pilares sobre los que se asentaban los viejos paradigmas científicos, sociales, jurídicos, etcétera, erosionados por nuevos caos, contradicciones, complejidades e incertidumbres sobre los cuales nos ilustró mejor que nadie Ilya Prigogine, Premio Nobel de Física cuyas investigaciones sentaron las bases de la *teoría del caos* afirmando que «el caos posibilita la vida y la inteligencia. El cerebro ha sido seleccionado para volverse tan inestable que el menor efecto puede conducir a la formación de orden».<sup>35</sup>

En los últimos tiempos, la vida es a menudo entendida como un objeto de consumo y un mecanismo de producción. Valores intrínsecos al sujeto, como los propios de lo que entendemos por *dignidad humana*, no son ya simplemente obviados, sino, con mucha frecuencia, ignorados o degradados. La vida se cuantifica a través de algoritmos, el cuerpo se constituye en unidad vital de consumo dentro del ciclo de producción y la muerte es asumida como *colateral* en un ámbito que se extiende de lo bélico a lo productivo. Además, la salvaguarda de los derechos de los sujetos queda supeditada a los balances fiscales de las grandes corporaciones, que determinan sus condiciones de trabajo no en base a criterios éticos, ni tan siquiera legales, sino simplemente económicos. Es cierto, por otro lado, que esta situación ha provocado paralelamente el surgimiento de movimientos ciudadanos a favor de colectivos desfavorecidos, que combaten para frenar esta tendencia. Como resultado de tales

<sup>35</sup> Prigogine y Stengers (1984). Cf. también su obra *El fin de las certidumbres* (1996) y Sardar (2010).

presiones, de la litigación, del afloramiento de la información, etcétera, poco a poco ha ido reconociéndose derechos a las minorías y las empresas van viéndose obligadas a adaptar sus políticas económicas a la llamada *responsabilidad social corporativa*.<sup>36</sup>

Es decir: nos encontramos ante una fase de desarrollo del capitalismo globalizado consistente entre otras cosas en la transición desde la *biopolítica* de Foucault hasta la *necropolítica* de Mbembe.<sup>37</sup> En tal contexto, los esfuerzos de los movimientos ciudadanos y de las organizaciones internacionales son necesarios para frenar ese ejercicio del poder consistente en el control de los cuerpos desde una perspectiva tanto material como utilitarista. Y ello nos transporta hasta la fase actual de expansión del sistema, impregnada del dominio de los simulacros de la hiperrealidad.<sup>38</sup>

Las estructuras del *capitalismo avanzado* que desentraña Jameson en su obra *El posmodernismo o la lógica cultural del capitalismo avanzado* y el *capitalismo casino* y también señala Ballbé,<sup>39</sup> casi más caóticas que en los orígenes del mercado anómico,<sup>40</sup> ya no se limitan a ejercer un control represivo sobre el sujeto de una forma tanto física (material) como ideológica (subjetiva), sino que objetivizan el dominio sobre la vida y la muerte. Pero a la vez se produce una reacción contrapuesta que pugna por transitar de la desregulación a la *re-regulación* (ambiental, financiera, laboral, etcétera). Si tal cambio no se produce, se alumbrará una transmutación del gobierno tradicional en uno privado, indirecto, que no

<sup>36</sup> Ballbé y Martínez (2009).

<sup>37</sup> Foucault (1997); Foucault y Varela (1978); Mbembe (2012).

<sup>38</sup> Baudrillard (1994).

<sup>39</sup> Ballbé y Cabedo (2013).

<sup>40</sup> Durkheim teorizó en el siglo XIX sobre el concepto de *anomia* en la división del trabajo, es decir: normas que hacen inestables las relaciones del grupo, impidiendo su integración cordial. Cf. Ballbé (2006) y Waldman (2006).

destruirá al Estado, pero que traspasará el ejercicio del poder coactivo a élites paraestatales que se organizarán al margen del bien público y del interés general. Y de ahí se pasará a la instauración de un orden necropolítico, basado en el control y el uso económico del poder de dar muerte; a un sistema cuyo mantenimiento requerirá de nuevos y sofisticados sistemas de control social.

En esta fase, el sistema ya no solo busca «vigilar y castigar»<sup>41</sup> a través de un complejo sistema jurídico-institucional, sino que, como resultado del crecimiento exponencial del conocimiento científico, el Estado y la comunidad son más conscientes de los procesos vitales asociados a ciclos como el consumo. Se ha configurado un presente en el que el sistema *sabe* que opta por decidir quién vive y quién muere y cómo se producirá dicha muerte,<sup>42</sup> generando toda una compleja estrategia escatológica que puede materializarse en un amplio menú de finales posibles: violencia, guerra, enfermedad, intoxicación, extenuación, carestía, etcétera. Y todo ello en un mundo que, pese a haber amenguado notablemente las distancias físicas gracias a las tecnologías de la comunicación y el transporte, todavía reproduce modelos de explotación pretéritos, con prácticas que se basan en una extracción masiva de recursos de los países periféricos y en la restricción de la libertad de tránsito de las personas, condenando a muchos sujetos a la explotación y a la muerte mientras la extenuación ecológica a la que se ha sometido a ciertos territorios incrementa la necesidad de esa libertad. Regiones enteras del globo sufren una desertización y un empobrecimiento hídrico que impiden el acceso a agua a comunidades enteras que se ven en consecuencia forzadas a abandonar

<sup>41</sup> Foucault (1990).

<sup>42</sup> Strawser (2017).

sus tierras ante la imposibilidad de mantener las actividades tradicionales de cultivo. Existe toda una *emigración ecológica* que echa anualmente a miles de personas en los brazos de las mafias de tráfico de personas,<sup>43</sup> generando sucesivas crisis humanitarias mientras los países desarrollados aplican políticas activas de control de fronteras y limitaciones de entrada, tales como la reciente aprobación de la construcción *voluntaria* de centros de control de inmigrantes en los países miembros de la Unión Europea.<sup>44</sup>

Nos hallamos ante una nueva servidumbre de la gleba. Occidente sustenta sus dinámicas de consumo en una versión refinada de la idea de *espacio vital* que consiste en que ya no sea necesario ejercer el control efectivo de un territorio, sino simplemente el de su economía. Se garantiza la extracción de materiales mientras se restringe la traslación de individuos. Millones de hombres y mujeres son condenados a la pobreza, su dignidad humana es obviada y desatendida y nuevas formas de comercialización del riesgo, como los derechos de emisión, alumbran sofisticadas erosiones sociales y ambientales; dinámica esta que contrasta con aquellas propias de las clases sociales de los países occidentales y de las élites del resto del planeta, que se benefician de un complejo sistema de privilegios. Representan una fracción de la población mundial que no oye, no ve y no escucha la crisis humanitaria, ecológica y social que atravesamos como especie y como planeta. Tal vez ello ocurra porque los elementos de referencia alejan la tragedia diaria del mundo y nos impiden identificarnos como especie más allá de las barreras de raza,

<sup>43</sup> Beine y Parsons (2015); Janashvili (2018; en prensa).

<sup>44</sup> «La UE acuerda la creación voluntaria de centros “controlados” para migrantes en los países miembros», *El Diario*, 29 de junio de 2018 [en línea], <[https://www.eldiario.es/desalambre/UE-acuerda-creacion-voluntaria-inmigrantes\\_o\\_787421311.html](https://www.eldiario.es/desalambre/UE-acuerda-creacion-voluntaria-inmigrantes_o_787421311.html)>.

nación y clase (construidas para fracturar una hipotética unidad de acción orientada a garantizar un desarrollo sostenible y responsable). El modelo así configurado nos conduce a una extenuación social y ecológica y requiere de nuevos instrumentos de control y de manufacturación del consentimiento en base a la reconfiguración del humano como consumidor despojado de toda dignidad intrínseca.

Nuestro tiempo se configura en torno a la *deregulation* del 2000 y con vistas a un capitalismo caótico y desregulado que convierte la propia vida, los cuerpos, en objetos de consumo; de usar y tirar. Los procesos productivos se deslocalizan a lugares en los que la regulación ecológica o laboral son prácticamente inexistentes y se permite que los trabajadores no solo se expongan a riesgos inasumibles en Europa, sino también a situaciones extremas de explotación laboral rayana en la esclavitud, dándose incluso casos lacerantes de utilización de mano de obra infantil. En ese contexto, la inteligencia artificial mal aplicada puede jugar un papel clave como herramienta de cristalización de la desigualdad e invisibilización de las crisis humanitarias. Ejemplos como el de Cambridge Analytics y su papel en las elecciones estadounidenses de 2016 nos dan pistas acerca de un modelo a caballo entre Huxley y Orwell, en el que el *soma* puede combinarse con altas dosis de represión dependiendo del eslabón de la cadena en el que participen los individuos. La vida en sí es un artefacto de consumo y su valor depende exclusivamente de la oferta y la demanda.

Es necesario matizar en este punto que cuando nos referimos a vida y cuerpos, no lo hacemos únicamente a los humanos o antropomorfos, sino a un todo que tenga en cuenta las complejas relaciones ecosistémicas de interdependencia entre especies. Vida «es un proceso capaz de preservar su complejidad y de replicarse. Pero lo que se replica no es la



materia (hecha de átomos), sino la información (compuesta por *bits*) que especifica cómo están dispuestos los átomos». <sup>45</sup> Dicha concepción nos obliga a plantearnos una suerte de *seguridad post-humana* que adquiere una nueva dimensión con la aparición de la inteligencia artificial y el *wetware* (esto es, la interacción entre *software* y tejido orgánico) en unas complejas relaciones de interdependencia interespecies. De ahí la importancia de estudios transversales que centren su interés en la *seguridad integral* (concepto acuñado por Naciones Unidas) <sup>46</sup> y superen así la fase actual, que se sostiene en el consumo constante de cuerpos y unidades vitales, ya sea con fines médicos, científicos, laborales, etcétera.

El desarrollo tecnológico comenzará probablemente, en próximas fases de expansión, a ocupar capas del sistema que hasta ahora estaban reservadas a los humanos merced al desarrollo de diferentes formas de inteligencia artificial, acelerando la disolución de lo humano en favor de lo no humano o quizás, con algo de suerte, de lo *post-humano*, encaminándonos a una *vida 3.0*. <sup>47</sup> Ya se está estructurando un campo de juego en el que los avances en robótica y nanotecnología, y más especialmente los relativos al llamado *machine learning*, suponen un reto para la paz global. <sup>48</sup> Ya se está generando un nuevo contexto de impunidad y falta de control democrático en el que el conflicto y la guerra se desarrollan en base a una lógica que no tiene nada que ver con aquella que inspiró los

<sup>45</sup> Tegmark (2017), p. 40.

<sup>46</sup> Fernández Pereira (2006).

<sup>47</sup> Se considera vida 1.0 a la primera, a la simple, a la biológica, a la que es resultado de la evolución. La vida 2.0 es la que conocemos actualmente, la vida cultural, en la que los humanos podemos aprender nuevas habilidades complejas y modificar sus objetivos. La vida 3.0 es la vida tecnológica que aún no existe, pero podrá rediseñar su software y su hardware y no esperar a evolucionar a través de distintas generaciones. Tegmark, (2017, p. 40.)

<sup>48</sup> Geib y Lahmann (2017).

tratados fundamentales que regulan la guerra y su desarrollo, tales como el Convenio de La Haya (1899 y 1907).<sup>49</sup>

Fue Albert Einstein quien en una entrevista realizada por Alfred Wener en *Liberal Judaism* afirmó lo siguiente: «No sé con qué armas se libraré la tercera guerra mundial, pero sí que en la cuarta usarán palos y piedras». Hoy en día, parece que nos encontramos más cerca de ofrecer una respuesta a la primera incógnita: si no hacemos nada para remediarlo, la tercera guerra mundial se luchará con sistemas de armamento autónomo/independiente. Y sus consecuencias, tal y como anticipó Einstein, pueden ser desastrosas para la especie.

Necesitamos urgentemente la intervención de Naciones Unidas para conseguir una nueva aproximación que supere los marcos tradicionales del humanismo en tanto que antropo- y androcéntrico y ponga la mirada en un tiempo en que la armonía de las relaciones interespecies se desvele como clave en el mantenimiento de los equilibrios globales y la biodiversidad que sostiene la vida en el planeta. Debemos abogar por un sistema que abrace las singularidades cognitivas como algo inherente a nuestro tiempo y que al mismo tiempo se despoje del supremacismo biológico que ha provocado la extenuación ecológica del planeta a manos de nuestra especie.<sup>50</sup> Ello pasaría, entre otras cosas, por hacer de la inteligencia artificial lo que pretenden iniciativas como AI For Good, de la Unión Internacional de Telecomunicaciones (UIT, un organismo vinculado a Naciones Unidas), y al mismo tiempo restringir su aplicación en el ámbito militar. Precisamos una nueva comprensión de lo humano como un sistema de interdependencias complejas en el que no solo lo natural, sino también lo artificial se encuentre regulado y

<sup>49</sup> Sunstein (2016a).

<sup>50</sup> Sloterdijk (2003).

donde la noción de *control humano significativo* sea completamente desarrollada y regulada a través de instrumentos legales adecuados. El *posthumanismo* puede ser la alternativa a un mundo que vive bajo el dominio de la hiperrealidad, en el que será cada vez más difícil distinguir entre lo natural y lo artificial, así como entre lo material y lo digital y donde los patrones básicos de consumo cultural propios del humanismo se diluirán ante una avalancha de nuevas formas de expresión y comunicación.

A fin de cuentas, la tecnología *per se* no es el peligro: este subyace en los usos que los humanos pueden hacer de la misma y en su desregulación. Cabe recordar que, muy a menudo, lo que elegimos creer resulta transformarse en una profecía autocumplida. En Occidente hemos creído que los robots vendrán a desplazarnos; en Japón, sin embargo, han sido vistos como serviciales y amistosos; y en ambos casos el desarrollo de la tecnología se ha sujetado a esas narrativas.

Actualmente, hay demasiados elementos que inducen a pensar que en los próximos años tendrá lugar una penetración masiva de la inteligencia artificial en el ámbito militar. Existe el miedo a que la tecnología termine sustituyendo a los humanos en todos los ámbitos, motivo por el cual el *posthumanismo* debe ser capaz de desarrollar un pensamiento ecológico que, en el sentido más amplio de esta palabra, tenga en cuenta no solo el entorno natural sino también el tecnológico y regule claramente las limitaciones de ciertas aplicaciones de la inteligencia artificial (y obligue a que incluyan componentes de humanidad), precisamente para que tales presagios no ocurran.

Ahora bien, antes de adentrarnos en el análisis de estas aplicaciones, es necesario abordar el sustrato tecnológico de esta cuarta revolución industrial.

## El sustrato tecnológico

La tecnología, no importa cuán bien diseñada esté, es solo un amplificador de la intención y la capacidad humanas.

Kentaro Toyama<sup>51</sup>

La inteligencia artificial representa uno de los mayores desafíos de los patrones actuales de desarrollo de la tecnología.<sup>52</sup> Esto se debe fundamentalmente a que su desarrollo viene acompañado de nuevas nociones de inteligencia al margen de la conciencia y a que abre la puerta a procesos de decisión al margen de un control humano significativo o, lo que es lo mismo, sin supervisión humana en fases críticas.

Así, un artículo publicado por *The Atlantic* titulado «Cómo los algoritmos pueden reducir los puntajes de crédito de las minorías» revela cómo el uso masivo de algoritmos de inteligencia artificial por parte de las entidades financieras puede cristalizar en dinámicas de marginación sobre las minorías, haciendo necesario incluir controladores humanos que puedan corregir dichos sesgos.<sup>53</sup> De la misma forma, *The Guardian* alertaba sobre la aparición de «prejuicios» relativos a género y raza en inteligencias artificiales debido al procesamiento del lenguaje natural en fuentes abiertas que altera la teórica neutralidad de la inteligencia artificial,<sup>54</sup> algo que fue

<sup>51</sup> Toyama (2015).

<sup>52</sup> Scherer (2017). A menudo, la inteligencia artificial en robots se quiere presentar como un paso más en el Estado del bienestar de la sociedad, como las abejas-robot autónomas que suplen la falta de poblaciones de abejas polinizadoras, patente que Walmart acaba de registrar y que pretende desarrollar con este fin. Se trata de robots que, usando sensores y cámaras, volarían autónomamente polinizando cultivos siguiendo el resultado algorítmico establecido.

<sup>53</sup> Waddell (2016).

<sup>54</sup> Devlin (2017).

reconocido por compañías como Facebook, que prometieron aumentar las fases bajo control humano.<sup>55</sup> Esta dinámica nos permite observar que buena parte de los riesgos asociados al *machine learning* y a la inteligencia artificial radica directamente en los patrones referenciales de aprendizaje; y ello nos obliga a preguntarnos qué podemos enseñar como humanidad a estas nuevas inteligencias teniendo en cuenta que tenemos sociedades misóginas, racistas, clasistas, etcétera, y cómo podemos eliminar dichos sesgos en pos de la creación de una sociedad más justa y equitativa. Ahora bien: antes de avanzar sobre los riesgos éticos vinculados a estos *sets* tecnológicos es necesario explorar la propia conceptualización de la tecnología.

La inteligencia artificial, como concepto, ha estado impregnada, prácticamente desde su primera formulación, de nuestras mayores fantasías, entretejida en cualquier tipo de escenario imaginable: desde futuros de riqueza, confort y prosperidad hasta la aniquilación de nuestra propia especie,<sup>56</sup> algo a lo que pueden haber contribuido en gran medida la literatura y el cine.<sup>57</sup>

Pero, ¿qué es realmente la inteligencia artificial? Según Eric Horvitz, «no es realmente una sola cosa, sino un rico conjunto de subdisciplinas y métodos; visión, percepción, discurso y diálogo; decisiones, planificación, robótica, etcétera, siendo necesario considerar todas estas diferentes disciplinas y métodos para buscar soluciones verdaderas en la generación de valor para los seres humanos y las organizaciones».<sup>58</sup> Es decir, estamos ante un concepto amplio que representa no

<sup>55</sup> Makridakis (2017).

<sup>56</sup> Barrat (2013).

<sup>57</sup> Se ha dicho que gracias a los sueños, a la imaginación, a la creatividad se ha desarrollado y se desarrollará la mejor tecnología. Maderer (2017).

<sup>58</sup> ITU y XPRIZE (2017).

solo las aspiraciones de los sectores tecnocientíficos, sino el advenimiento de un sistema completamente nuevo tanto en el plano económico como en el sociológico, el antropológico, el filosófico y el legal. Para comprenderlo, es necesario establecer una distinción clara entre su estado actual de desarrollo y sus potencialidades. Hacen falta elementos descriptivos claros que nos permitan comprender su significado y alcance y explorar sus definiciones e implicaciones.

Así, nuestro primer enfoque puede ser el proporcionado por la Enciclopedia Británica, que define la inteligencia artificial como

La capacidad de una computadora digital o un robot controlado por computadora para realizar tareas comúnmente asociadas con seres inteligentes. El término se aplica con frecuencia al proyecto para desarrollar sistemas dotados de los procesos intelectuales característicos de los humanos, como la capacidad de razonar, descubrir significado, generalizar o aprender de la experiencia pasada. Desde el desarrollo de la computadora digital en la década de 1940, se ha demostrado que las computadoras se pueden programar para llevar a cabo tareas muy complejas, como por ejemplo descubrir pruebas de teoremas matemáticos o jugar al ajedrez, con gran habilidad. Aun así, a pesar de los continuos avances en la velocidad de procesamiento de la computadora y la capacidad de la memoria, todavía no hay programas que puedan igualar la flexibilidad humana en dominios más amplios o en tareas que requieren mucho conocimiento cotidiano. Por otro lado, algunos programas han alcanzado los niveles de desempeño de expertos y profesionales humanos en la realización de ciertas tareas específicas, por lo que la inteligencia artificial en este sentido limitado se encuentra en aplicaciones tan diversas como el diagnóstico médico, motores de búsqueda y reconocimiento de voz o escritura.<sup>59</sup>

<sup>59</sup> Copeland (2018).

Es decir, hablamos de un sistema tecnológico derivado de las ciencias de la computación cuyo enfoque de investigación se ha definido tradicionalmente como el estudio de *agentes inteligentes*; o lo que es lo mismo, de cualquier dispositivo que perciba su entorno y emprenda acciones que maximicen sus posibilidades de lograr una serie de objetivos,<sup>60</sup> adaptándose a diversas situaciones no previamente conocidas y aprendiendo de la experiencia.<sup>61</sup> Ello encajaría a grandes rasgos con la definición de Minsky de la inteligencia artificial como «la ciencia de producir máquinas que puedan llevar a cabo tareas que requerirían inteligencia (en caso de ser desarrolladas por humanos)».<sup>62</sup>

Otra de las primeras cosas que deberíamos tener en cuenta es que la inteligencia artificial no es un fenómeno reciente, sino que sus fundamentos se han ido construyendo a partir de aportaciones clave como las de Alan Turing, quien ya en 1935 describió el primer sistema de inteligencia artificial. Se trataba de una máquina de computación abstracta con una memoria ilimitada y un escáner que se movía hacia atrás y adelante a través de él, símbolo por símbolo, leyendo lo que encontraba y escribiendo más símbolos. Las acciones del escáner estaban dictadas por un programa de instrucciones que también se almacenaba en la memoria en forma de símbolos. Se abrió así la posibilidad de que la máquina funcionara mientras se modificaba o mejoraba su propio programa. Por ello, podemos afirmar que todos los sistemas modernos de computación son básicamente máquinas de Turing. Se trata de dispositivos que hoy en día forman parte de nuestra vida diaria. En 1952 se creó el *test de Turing* para determinar

<sup>60</sup> Poole, Mackworth y Goebel (1998). V. también Pool y Mackworth (2017).

<sup>61</sup> De Almeida Lenardon (2017).

<sup>62</sup> Minsky (1991).

si una máquina era realmente inteligente. Para superarlo, la máquina debía ser capaz de engañar a un humano haciéndole creer que era un igual. Cuatro años más tarde, Minsky y McCarthy, con Shannon y Rochester, organizaron una conferencia en Dartmouth e hicieron público el término *inteligencia artificial*.<sup>63</sup>

En dicha conferencia, McCarthy expuso lo siguiente: «Este estudio procederá sobre la base de la conjetura de que todos los aspectos del aprendizaje o cualquier otro rasgo de la inteligencia pueden, en principio, ser descritos de una forma tan precisa que se puede crear una máquina que los simule». Para ello creó el LISP, uno de los lenguajes de programación de más alto nivel que existen y que le permitió desarrollar la función de *tiempo compartido* (muchas personas conectándose a la vez a un superordenador), uno de los pilares de la posterior creación de Internet.<sup>64</sup> McCarthy investigó asimismo las posibilidades de que una máquina tuviera el mayor grado de humanidad posible —es decir, gozara de *libre albedrío*— y se preguntaba: «¿Podrá un ordenador decir algún día: puedo, pero no quiero?».<sup>65</sup>

Ciertamente, el nivel de desarrollo de la inteligencia artificial ha crecido a nivel exponencial a lo largo de los últimos años, y proyectos como AlphaGo Zero de Google Deep Brain Project o el propio Google Duplex ilustran la capacidad de la tecnología de superar a los humanos a la hora de desarrollar tareas acotadas y delimitadas.<sup>66</sup> La evolución de este sistema es vertiginosa, pero el término *inteligencia artificial* continúa unido a la idea de una máquina que imita funciones cognitivas como aprender y resolver

<sup>63</sup> Turing (1939 y 2009).

<sup>64</sup> Alandete (2016).

<sup>65</sup> Sanchis (2018).

<sup>66</sup> Deep Mind (2018).



problemas,<sup>67</sup> sin tener en cuenta los dispositivos con inteligencia artificial y *machine learning* ya insertados en nuestra vida diaria (Google Assistant, Alexa, Cortana, Siri, Autopilot de Tesla, etcétera), que pueden transformar radicalmente la realidad y amenazan nuestra propia existencia. Como señala Future of Live Institute, «la tecnología confiere a la vida la posibilidad de prosperar como nunca antes... o de autodestruirse».<sup>68</sup> Por ello, debe primar el derecho a la seguridad humana desarrollado por Naciones Unidas, y es urgente en este nuevo panorama una regulación desde el derecho administrativo preventivo tanto nacional como internacional.<sup>69</sup>

En relación a la noción de aprendizaje, cabe destacar que el *machine learning* ha sido descrito tradicionalmente como un proceso estadístico que comienza con una gran cantidad de datos e intenta derivar una regla o un procedimiento que explique los datos o pueda predecir datos futuros.<sup>70</sup> La definición en este sentido sería clara: los algoritmos de *machine learning* pueden descubrir cómo realizar tareas importantes generalizando a partir de ejemplos.<sup>71</sup> Se entiende que, aunque la máquina no pueda autoprogramarse, sí podría estar preparada para generar y almacenar asociaciones y hechos a partir de los datos. La generalización implicaría en este caso la capacidad de asociaciones oportunamente realizadas sobre la base de datos limitados. Y también algunas presunciones que pueden conducir a la repetición de errores pasados (por ejemplo, dinámicas de opresión) o efectos no previstos (por ejemplo, discriminación injusta). Las reglas

<sup>67</sup> Russell, Stuart, Norvig y Davis (2010).

<sup>68</sup> Cf. página *web* de Future of Life: <<https://futureoflife.org/?cn-reloaded=1>>.

<sup>69</sup> Fernández Pereira (2006), Andersen-Rodgers y Crawford (2018).

<sup>70</sup> Rosembuj (2017, 2018).

<sup>71</sup> Domingos (2013).

de interpretación y predicción muestran que uno de los principales problemas en materia de inteligencia artificial es precisamente el hecho de reproducir comportamientos humanos debido a la influencia de los programadores, pudiendo generarse conclusiones arbitrarias o dispares, contaminadas por las «creencias, falibilidades y prejuicios de la persona que las creó».<sup>72</sup>

Es decir, la inteligencia artificial aprendería en un contexto de imposible neutralidad: los prejuicios humanos impedirían que garantizase futuros de igualdad y representación paritaria. Y esto resulta especialmente preocupante cuando se lo relaciona con la construcción de sistemas capaces de seleccionar y eliminar objetivos sin control humano significativo, sobre todo si tenemos en cuenta que la delegación de atribuciones letales ha de estar exenta de cualquier tipo de incertidumbre.

Por otra parte, una de las características principales del *machine learning* es el conocido como *deep learning* o *aprendizaje profundo*. El *deep learning* utiliza técnicas de aprendizaje que combinan capas de redes neuronales para identificar los perfiles de un conjunto de datos necesarios para tomar decisiones. De esta forma, se reconoce la existencia de multiplicidad de capas entre los datos de entrada y los de salida, configurándose las salidas de las capas anteriores como entradas para las siguientes, lo que genera lo que se ha dado a conocer como *redes neuronales artificiales*.<sup>73</sup>

Los algoritmos de *machine learning* y *deep learning* se erigen así como la última frontera de la inteligencia artificial, siendo utilizados hoy en día en ámbitos tan dispares como las búsquedas *web*, filtros de correo no deseado, calificación crediticia, riesgo de seguro, detección de fraude, comercio

<sup>72</sup> Barret (2016).

<sup>73</sup> Kaplan (2016).

de acciones, diseño de medicamentos, evaluaciones de empleo, registros de salud, búsquedas de contratación, vivienda y muchos otros; redes estas que, de configurarse a partir de lenguaje natural, correrían a su vez el riesgo de reproducir patrones de comportamiento que derivan en formas de marginación y/o explotación de ciertos colectivos humanos. Si esto lo extendemos a la delegación de funciones letales, es evidente que su aplicación podría verse afectada por dichos prejuicios. El peligro de una desviación de poder es por lo tanto probable.

Otro aspecto a tener en cuenta es que los agentes de inteligencia artificial que cuentan con soporte físico tienen la capacidad de interactuar con su entorno. Son, por decirlo de otro modo, entes «capaces de realizar tareas detectando su entorno y/o interactuando con fuentes externas teniendo la capacidad de adaptar su comportamiento»,<sup>74</sup> tal y como sucede con las armas autónomas, como comprobaremos más adelante. Es decir, tendríamos ante nosotros una manifestación material de la realidad digital capaz de interactuar con humanos y otras especies a través de la recolección y procesamiento de datos a tiempo real, incidiendo sobre el curso de la realidad material.

La inteligencia artificial requiere una conceptualización amplia que tiene consecuencias importantes cuando se examinan las fases de despliegue y las potencialidades de una tecnología capaz de transformar radicalmente la realidad e incluso de amenazar nuestra propia existencia, como se explicará en los siguientes apartados.

Mientras que la ciencia-ficción retrata a menudo a la inteligencia artificial como robots con características similares a las humanas, la actual puede abarcar desde los algoritmos

<sup>74</sup> ISO 8373.

de búsqueda de Google hasta el Watson de IBM —que venció a sus competidores humanos en el concurso Jeopardy, que consiste en contestar preguntas formuladas en lenguaje natural—<sup>75</sup> o el propio desarrollo de sistemas de armamento autónomo.<sup>76</sup> Estos sistemas representan grandes dilemas éticos y jurídicos incluso en el estado actual de la técnica. Por ejemplo, un vehículo autónomo podría encontrarse en una situación en la que deba elegir entre proteger al pasajero o a otro grupo de personas, es decir, tomar una decisión de profundo carácter ético.

En lo que se refiere al estado de desarrollo actual de la tecnología, podemos afirmar que la fase actual está dominada por lo que conocemos como *narrow AI* (inteligencia artificial especializada, reducida o débil), lo que significa que está diseñada para realizar una tarea limitada (por ejemplo, solo reconocimiento facial o búsquedas en Internet o conducir un automóvil) de acuerdo con nuestras capacidades técnicas actuales. Sin embargo, el objetivo a largo plazo de muchos investigadores es crear lo que se ha denominado *artificial general intelligence* (AGI) o *AI fuerte e independiente*.<sup>77</sup> La diferencia entre ambos conceptos es que mientras que la inteligencia artificial reducida puede superar a los humanos en lo que vendría a ser una tarea específica, como jugar ajedrez o resolver ecuaciones, la AGI podrá realizar cualquier tarea cognitiva tan bien como los humanos e incluso superarlos a través de lo que ya se denomina *superinteligencia*.<sup>78</sup> Por ello, el primer objetivo de cualquier Estado debería ser desarrollar

<sup>75</sup> IBM (2018).

<sup>76</sup> O'Neil (2016). La autora alerta que, bajo su promesa de eficacia y justicia, los algoritmos y los métodos de análisis *big data* «pueden ser usados con fines poco honestos», como armas de destrucción de la sociedad misma.

<sup>77</sup> Goertzel y Pennachin (2007), p. 131. Bostrom (2005).

<sup>78</sup> Future of Life Institute (2018)

una inteligencia artificial segura y benéfica cuyos objetivos coincidan con los de los hombres, porque si dejamos de ser los seres más inteligentes del planeta es posible que perdamos también el control.<sup>79</sup>

Un desarrollo hipotético de la AGI<sup>80</sup> conllevaría consecuencias profundas no solo para nuestra sociedad sino para el mismo ordenamiento jurídico,<sup>81</sup> ya que se avanzaría en la generación de sistemas que se comportarían racionalmente, o lo que es lo mismo, sistemas de automatización del comportamiento que en el plano teórico se vincularían al fenómeno de *singularidad tecnológica*.<sup>82</sup> Ello implica que una computadora equipo, una red de computadoras o un robot podrían mejorarse a sí mismos recursivamente. Se dice que las repeticiones de este ciclo probablemente darían como resultado un efecto fuera de control, una *explosión de inteligencia*, como la denominó en 1965 el matemático Irving Good;<sup>83</sup> un fenómeno muy difícil de predecir y cuyas consecuencias podrían ser dramáticas... o no. Este escenario, pese a reconocerse como altamente improbable en el corto y medio plazo, no puede ser ignorado, siendo necesario generar sistemas de contingencia ante los riesgos potenciales y no potenciales que dicha tecnología podría desatar como arma de destrucción masiva.

<sup>79</sup> Tegmark (2017), p. 43.

<sup>80</sup> La expresión IAG la popularizaron Legg, Gubrud y Goertzel para referirse a una inteligencia de nivel humano. Tegmark (2017), p. 72.

<sup>81</sup> «La legislación actual no ofrece soluciones ni para el uso de vehículos autónomos con inteligencia artificial en vías públicas, lo que provoca inseguridad para el desarrollo y despliegue de dicha tecnología en Europa. En cambio, en los Estados Unidos, se está regulando a nivel estatal. La regulación serviría para asignar responsabilidades y ofrecer certeza legal». De Almeida Lenardon (2017), pp. 28 y 30.

<sup>82</sup> Bostrom (2005). El concepto fue popularizado por Vernor Vinge, profesor de ciencias de la computación en la Universidad de St.Diego. Vid. también sobre este tema Bostrom (2016).

<sup>83</sup> Chalmers (2010).

Por lo que respecta a la clasificación de tipologías de inteligencia artificial, además de la que distingue la *narrow AI* de la *strong AI* o la limitada de la general, las investigaciones llevadas a cabo a lo largo de las últimas décadas han permitido establecer otra que distingue cuatro grandes enfoques: sistemas que piensan como humanos, sistemas que piensan racionalmente, sistemas que actúan como humanos y sistemas que actúan racionalmente.<sup>84</sup> La primera tipología corresponde a sistemas que cuentan con información y la procesan con el propósito de comprender y predecir. La segunda, a máquinas que funcionan en base a las leyes del pensamiento aristotélico. La tercera hace referencias a máquinas que pueden ejecutar funciones propias de los humanos y requieren inteligencia limitada, y en cuarto lugar tendríamos sistemas que automatizan una conducta inteligente;<sup>85</sup> que se encuentran vinculados en el plano teórico al fenómeno de la *singularidad tecnológica*.<sup>86</sup>

Para comprender completamente el concepto debemos tener en cuenta que la inteligencia artificial se ha centrado principalmente en los siguientes componentes: aprendizaje, razonamiento, resolución de problemas, percepción y uso del lenguaje. Repasemos cada uno de ellos con atención.

Con *aprendizaje* se hace referencia a la capacidad de la máquina para perfeccionarse a sí misma a través del proceso de prueba de error y la aplicación de la experimentación: por ejemplo, cuando se identifican movimientos de ajedrez y se los utiliza en nuevas partidas. El proceso se llama *generalización* cuando implica aplicar experiencia a situaciones nuevas análogas de una forma similar a como lo hacen los humanos. La doctrina considera que no existe un verdadero aprendizaje

<sup>84</sup> Cairo Battistutti (2011).

<sup>85</sup> *Ibidem*.

<sup>86</sup> Villalba Gómez (2016).

en la inteligencia artificial, sino algoritmos de aprendizaje automático a través de redes neuronales, sin comprensión dentro del cálculo y con patrones de correlación sin causalidad. El aprendizaje automático puede ser altamente discriminatorio, puesto que se sirve de los datos sociales del momento.<sup>87</sup> Así pues, el problema del aprendizaje en el caso de los LAWS reside principalmente en sus patrones referenciales, ya que quien controle la tecnología podrá establecer los marcos de operatividad al margen de los dictados de la conciencia pública tal y como establece la cláusula Martens del Convenio de La Haya. Ahora bien, como se ha señalado, «la inteligencia artificial se reduce a información y computación, no a carne y sangre y átomos de carbono [...] no existe ninguna razón fundamental por la que las máquinas no puedan ser algún día tan inteligentes como nosotros»,<sup>88</sup> e incluso superarnos, por lo que tenemos que asegurarnos de que sus objetivos sean benéficos.

*Razonar* es hacer inferencias apropiadas a la situación. Las inferencias se clasifican tradicionalmente como deductivas (extracción de un juicio a partir de hechos, proposiciones o principios, sean generales o particulares) o inductivas (establecimiento de una ley o conclusión general a partir de la observación de hechos o casos particulares). Sin embargo, el verdadero razonamiento implica extraer inferencias relevantes para la solución de la tarea o situación particular. Este es uno de los problemas más difíciles que enfrenta la inteligencia artificial, y supone uno de los mayores retos tecnológicos en su fase de desarrollo actual.

La *resolución de problemas* se puede entender como una búsqueda sistemática a través de una gama de acciones po-

<sup>87</sup> McQuillan (2018a). Para este autor, «el aprendizaje automático de la inteligencia artificial no ayudará al humanitarismo y profundizará la dinámica neocolonial y neoliberal de las instituciones humanitarias» (2018b).

<sup>88</sup> Tegmark (2017), p. 75.

sibles a fin de alcanzar un objetivo o solución predefinida, es decir, una generación de escenarios posibles asociados a las posibilidades de acción. Los métodos de resolución de problemas se dividen en *propósitos especiales* y *propósitos generales*. Un método de propósito especial está hecho a medida para un problema en particular y a menudo explota características muy específicas de la situación en la que está incrustado. Por el contrario, un método de propósito general es aplicable a una amplia variedad de ellos. Muchos problemas diversos han sido resueltos por programas de inteligencia artificial. Algunos ejemplos son encontrar el movimiento ganador (o secuencia de movimientos) en un juego de ajedrez, idear pruebas matemáticas o manipular *objetos virtuales* en un mundo generado por ordenador, si bien el desarrollo de propósitos generales representa retos en su fase de desarrollo actual. El problema en este caso, en lo que se refiere a los LAWS, es precisamente la variabilidad de las condiciones del contexto, que pueden dificultar la operatividad de la tecnología y por tanto su fiabilidad. Además, en el caso de la inteligencia artificial aplicada a las potestades administrativas, deberíamos distinguir entre ejercicio de potestades discrecionales y regladas.

En relación con la *percepción*, debemos entender que el entorno se escanea por medio de varios órganos sensoriales, reales o artificiales, y que la escena se descompone en objetos separados en diversas relaciones espaciales. En consecuencia, el análisis se complica por el hecho de que un objeto puede aparecer diferente dependiendo del ángulo desde el que se ve, la dirección y la intensidad de la iluminación en la escena y cuánto contrasta el objeto con el campo circundante. Es un campo donde la inteligencia artificial todavía tiene un gran camino por delante, pero cuyos resultados actuales pueden superar ampliamente las capacidades humanas. Este es un



hecho que hace especialmente atractiva esta línea de investigación con fines militares. El proyecto Maven (colaboración entre Google y el Pentágono que posiblemente no se renueve en el año 2019) explora los usos de la inteligencia artificial sobre la captación de imágenes vía satélite con el objetivo de identificar estructuras *sensibles* para los intereses militares de Estados Unidos.

Finalmente, un *idioma* es un sistema de signos que tienen significados por convención. No solo la palabra hablada: las señales de tráfico, por ejemplo, forman un minilenguaje. La inteligencia artificial tiene la capacidad potencial de comprender y reproducir patrones de comunicación que pueden interactuar en la *esfera lingüística*, así como comprender las estructuras básicas del sistema con el fin de adaptarse a normas existentes del lenguaje compartido.

Todo lo hasta aquí expuesto es un conjunto de habilidades que, de desarrollarse en plenitud, representa promesas y oportunidades considerables que deben ser tenidas en cuenta por el ordenamiento jurídico al regular los LAWS. De lo contrario, podría ocurrir que la teoría económica de la *captura del regulador por el regulado* desarrollada por el Premio Nobel de Economía George Stigler<sup>89</sup> se convirtiera en una *captura* del Estado regulador y democrático por los LAWS con inteligencia artificial.<sup>90</sup>

Según Marcus Shingles, estas oportunidades incluyen obtener información de «los gigantes durmientes de los datos», mejorar la toma de decisiones y «aprovechar la sabiduría colectiva».<sup>91</sup> Quizás sea esta última una de las promesas de mayor interés social asociadas a la inteligencia artificial, pero, al mismo tiempo, tiene la capacidad intrínseca de dibujar es-

<sup>89</sup> Ballbé y Padrós (1997).

<sup>90</sup> Carpenter y Moss (2014).

<sup>91</sup> ITU y XPRIZE (2017).

cenarios distópicos en los que el control social y la falta de privacidad den forma a una sociedad de carácter autoritario. Esto se debe fundamentalmente a que la materia prima, la sangre del sistema, son los datos. Sin amplios *sets* de datos disponibles, el desarrollo de la inteligencia artificial sería una mera quimera. Tal y como afirma T. Rosembuj, los datos son «la principal materia prima del algoritmo, como el algodón, el trigo o el combustible en el último siglo. El procesamiento de datos es la esencia digital y virtual: sin datos no hay algoritmo y sin algoritmo es difícil argumentar que hay inteligencia artificial, bienes digitales o bienes virtuales».<sup>92</sup>

Así, el valor de los datos radica precisamente en su reutilización infinita: «El valor de los datos se calcula sobre la base de todas las formas posibles en las que podrían usarse en el futuro y no simplemente sobre la base de su uso actual».<sup>93</sup> De esta forma, la recombinación de datos, su acumulación y su extensión, son su valor real y, por lo tanto, el impulso para su acumulación por parte de organizaciones como Google, Facebook, Twitter, Amazon, Visa y un largo etcétera de organizaciones.

La gran paradoja aquí parece radicar en que los datos iniciales son susceptibles de ser eternos, reiterados y repetidos continuamente y aplicados sistemáticamente, lo que facilitaría procesos de evolución social a través de una visión conservadora del progreso humano y social por parte de la inteligencia artificial. Además, en caso de que se trate de datos personales, el sujeto perderá el rastro de su identidad por la privación de los derechos personales.<sup>94</sup> El origen de los datos y el consentimiento explícito de sus propietarios para su uso con finalidades armamentísticas es,

<sup>92</sup> Rosembuj (2017).

<sup>93</sup> Mayer-Schonberger y Cukier (2013), *apud* Rosembuj (2017).

<sup>94</sup> Rosembuj (2017).

pues, otro de los principales problemas relacionados con el desarrollo de los LAWS.

Finalmente, el último punto que requiere atención en este breve repaso al estado de la cuestión es el hecho de que la inteligencia artificial se podría clasificar como una tecnología de doble uso, o lo que es lo mismo, que tiene aplicabilidad tanto civil como militar, de tal manera que determinados *sets* de aplicabilidad pueden generar nuevas dinámicas de conflicto que conviertan en obsoleto el derecho a la guerra o el derecho humanitario en los conflictos bélicos. De esta forma, y por lo que se refiere a los avances del complejo militar-industrial (calificado de «complejo financiero global especulador no productivo y sin fronteras estatales»),<sup>95</sup> podemos afirmar que estos se han sucedido a gran velocidad a lo largo de las últimas décadas gracias a la connivencia y la monopolización de sistemas que van desde lo tecnológico hasta lo militar, pasando por lo científico y lo financiero. Si el general Eisenhower, en su despedida como presidente, alertó de los peligros del complejo militar-industrial y de la necesidad de su control, imaginemos lo que diría ante el surgimiento de LAWS cuyo control reside en manos de un oligopolio privado y cuyo objetivo es un beneficio económico a corto plazo y ofrecer al mundo la generación de nuevas armas y nuevas formas de entender el conflicto, la guerra y el control del territorio.<sup>96</sup>

## **Sistemas de armamento letal autónomo**

Entre el vasto abanico de tecnologías derivadas de la inteligencia artificial, son sus posibles aplicaciones para uso militar

<sup>95</sup> Ballbé (2006).

<sup>96</sup> Martínez (2003).

las que centrarán nuestra atención en esta obra debido a los riesgos que representan para la propia evolución del sistema y la garantía de derechos y libertades básicas. Focalizamos concretamente nuestro interés en los sistemas de armamento autónomos letales (LAWS por sus siglas en inglés).<sup>97</sup> Estos sistemas se caracterizan por la integración de la inteligencia artificial de tal manera que tengan la capacidad intrínseca de abordar procesos de decisión al margen del control o la supervisión humana de forma significativa; y podrían ser incluidos dentro de la tercera categoría clasificatoria expuesta en la sección anterior: sistemas que actúan como humanos. La diferencia principal entre los AWS y los LAWS vendría a ser que los primeros tienen un carácter meramente defensivo (los escudos antimisiles, por ejemplo), mientras que los segundos poseen la capacidad de identificar y eliminar objetivos militares, incluidas personas; y ello sin control humano significativo en el proceso, lo que significa una delegación de capacidades letales a entes robóticos.

El surgimiento de un variado elenco de nuevos sistemas de armamento da pie a una nueva carrera armamentística que puede determinar el curso de los conflictos no ya del futuro, sino del presente, ya que en algunos casos se encuentran en pleno funcionamiento (aunque no de forma totalmente autónoma del humano): piénsese, por ejemplo, en el sistema de defensa aérea Phalanx de la Armada estadounidense, que permite repeler ataques en modo automático.<sup>98</sup>

¿2018a o  
2018b?

<sup>97</sup> Waters (2018). Un nuevo informe, titulado *The malicious use of artificial intelligence*, advierte que si los avances en la inteligencia continúan a este ritmo, la tecnología será muy pronto tan poderosa que podría desbordar muchos de los mecanismos de defensa incorporados en los sistemas digitales y físicos actuales: «The malicious use of artificial intelligence», *Financial Times*, 14 de febrero de 2018 [en línea], <<https://www.ft.com/content/c54002ee-1668-11e8-9e9c-25c814761640>>.

<sup>98</sup> Horowitz, Kreps y Fuhrmann (2016).

En el último censo realizado en 2018, el Comité Internacional de la Cruz Roja contabilizó unos ciento treinta sistemas de armamento autónomo en el mundo, aunque otros recuentos aproximan el número a los trescientos.<sup>99</sup> Entre estos se incluirían los sistemas armamentísticos semiautónomos, puesto que estarían sometidos a supervisión humana en fases clave, como la selección de objetivos. Sería este el caso de los misiles Patriot o de drones como el modelo Reaper.

El debate actual no se centra en el análisis de los sistemas de armamento con control humano remoto (AWS con control humano significativo), sino en los riesgos potenciales para el futuro que supone el despliegue de un tipo de tecnología con inteligencia artificial fuerte, o sin control humano significativo. A estas se las podría denominar LIWS (*lethal independent weapons systems*) o *sistemas de armamento letales independientes*, y su regulación es urgente debido a su posible hibridación con otras tipologías de armas, tales como las nucleares o las biológicas, que unidas a los sistemas armados independientes son una amenaza real y más presente que nunca. Su surgimiento se encuadra en un escenario internacional cada vez más inestable e impredecible. Como dice D. Mourelle, «el mundo se encuentra haciendo funambulismo geopolítico sobre el abismo. Pero en esta ocasión, nada garantiza que en la próxima crisis nuclear tengamos tanta suerte como en las anteriores».<sup>100</sup>

Nuestro objeto de análisis son en suma aquellos sistemas armamentísticos capaces de seleccionar y atacar blancos sin intervención humana y cuya aplicabilidad se suele restringir teóricamente a objetivos militares en zonas no pobladas. El auge de sistemas cibernéticos de rápido desarrollo, alto po-

<sup>99</sup> Roff (2016).

<sup>100</sup> Mourelle (2017).

der de procesamiento e inteligencia artificial nos obliga a no ser ingenuos y valorar que ya no existen límites para su utilización como armas autónomas en espacios urbanos y sin declaración formal de guerra.<sup>101</sup> Es esta una tecnología que en caso de llegar a las manos de actores no estatales, como organizaciones terroristas, puede abrir un nuevo escenario que hipoteque el propio desarrollo de la inteligencia artificial incluso en aplicaciones no armamentísticas.

Uno de los mayores desafíos a los que nos enfrentamos y que Naciones Unidas quiere resolver es que no existe una definición de *autonomía* ni del concepto *independiente* acordada internacionalmente para los LAWS, ni tampoco consenso acerca de las características o rasgos que se combinan para formarlos. Se hace necesario, pues, aportar elementos que nos permitan una clasificación que facilite su regulación.

*Grosso modo*, podríamos entender que este tipo de armas *independientes* poseen tres características básicas:

- Pueden moverse independientemente a través de su entorno a lugares que ellos escogen de manera arbitraria. Sus capacidades son: movilidad, persistencia y orientación y navegación.
- Pueden seleccionar y disparar contra objetivos en su entorno. Sus capacidades son: identificación propia de objetivos, discriminación para categorizar objetivos, priorización de objetivos y selección del tipo de arma apropiada al objetivo.

<sup>101</sup> Encuentro de expertos *Autonomous weapons systems: technical, military, legal and humanitarian aspects*, Ginebra (Suiza), marzo de 2014 [en línea], <[–54–](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=oCB4QFjAAahUKEwiJ3KXXt9TIAhWImIgKHQ-ADO4&url=https://www.icrc.org/en/download/file/1707/4221-002-autonomous-weapons-systems-full-report.pdf&usq=AFQjCNHRUKZogcjiQGeY-cyOpgbarbixQw&sig2=85Tn4NKiirt6tskt9SVU9Q></a>>.</p></div><div data-bbox=)

— Pueden crear y/o modificar sus objetivos incorporando la observación de su entorno y la comunicación con otros agentes. Sus capacidades son: autodeterminación, auto-compromiso, comunicación autónoma con otros sistemas, automodificación de objetivos basada en información adquirida de fuentes autónomas, planificación de objetivos y aprendizaje y adaptación constantes.

Surge aquí un interrogante al que debemos enfrentarnos: ¿qué grado de inteligencia artificial o comportamiento inteligente es necesario para que el ordenamiento jurídico considere la prohibición de los LAWS? Lo que finalmente marcará la diferencia será precisamente si disponen o no de control humano significativo en las diferentes fases del proceso de acción letal (implementación, validación y ejecución).

Si atendemos al censo elaborado por el Future of Life Institute, actualmente existirían 256 sistemas categorizados y calificados de *autónomos*,<sup>102</sup> pero hasta la fecha, la mayoría de estados arguyen que todos tienen control humano o *juicio humano adecuado*<sup>103</sup> en algún momento. Es decir, todos los sistemas desarrollados hasta el momento dependen (o deberían depender) de supervisión humana o juicio humano previo en al menos alguna de sus fases críticas (selección de objetivos o anulación de la orden). Sin embargo, paralelamente se investigan y se desarrollan sistemas con autonomía total, y tarde o temprano deberá analizarse si cumplen o no con las prescripciones legales, porque la situación actual de práctica *arregulación* de las LAWS permite, por inactividad de los Estados anómicos (materia sensible, esta), una suerte de carrera

<sup>102</sup> Roff (2016a).

<sup>103</sup> Roff (2016b y 2017).

No figura  
así en la  
bibliografía.

competitiva sin ley entre gobiernos que puede resultar muy peligrosa.<sup>104</sup>

Los Estados justifican la investigación en AWS asegurando que no se trata de utilizarlos en ataques sino para defensa, es decir, simplemente como *automatic weapons defense systems* (AWDS).<sup>105</sup> Pero eso no parece más que un subterfugio para legitimar sistemas absolutamente letales dotados de la capacidad de independizarse de su creador y de su responsable. Es imprescindible desarrollar una regulación internacional que permita restringir sus usos no permitiendo la existencia de vasos comunicantes entre el desarrollo de sistemas de defensa y aquellos que tienen como finalidad la actuación letal contra personas. De lo contrario se podría atribuir a un ser sintético sin humanidad la potestad de decidir, en un conflicto, a quién abatir en base al interés general: o sea, una *licencia para matar* que debería ser exclusiva del poder público.

Por este motivo, han surgido a lo largo de los últimos años diversas iniciativas para generar una prohibición internacional de este tipo de armamento, tales como la campaña Stop Killer Robots. Se trata de un movimiento fundado en 2013 e integrado por numerosas organizaciones no gubernamentales que incluyen desde compañías tecnológicas hasta organizaciones de defensa de los derechos humanos; y su objetivo es dirigir los procesos normativos internacionales hacia la prohibición de las armas autónomas por considerar que representa una amenaza superior a la que suponen las armas nucleares. Utiliza el informe *Perdiendo humanidad* para argumentar que las armas autónomas letales no cumplen con los requisitos del derecho internacional

<sup>104</sup> Putin declaró en 2017 que «el país que logre el liderazgo en el desarrollo de la inteligencia artificial será *el amo del mundo*».

<sup>105</sup> Warren y Hillas (2018).



humanitario y arguye que debería hacerse con ellas lo que se hizo con el láser cegador: prohibir preventivamente su uso y su desarrollo.

En 2013, un informe de Christoph Heyns, relator especial de la ONU, sobre ejecuciones extrajudiciales, sumarias o arbitrarias pidió moratorias especiales y la creación de comisiones de expertos para frenar el desarrollo de las armas autónomas; informe que se presentó en el marco de la Convención sobre Ciertas Armas Convencionales de Naciones Unidas (CCW en inglés),<sup>106</sup> celebrada en diciembre de 2016 en Ginebra.<sup>107</sup> También se intentó en el Foro Económico Mundial de Davos en 2016 así como, en el mismo año, en la Conferencia de Seguridad de Múnich, donde se analizaron en profundidad estas cuestiones. Por su parte, la oenegé Human Rights Watch también se ha posicionado a favor de la prohibición inmediata de los AWS/LAWS. En un informe titulado *Shaking the foundations: the human rights implications of killer ro-*

<sup>106</sup> El Convenio sobre Prohibiciones o Restricciones en el Empleo de Ciertas Armas Convencionales que Pueden Considerarse Excesivamente Nocivas o de Efectos Indiscriminados fue adoptado en 1980 y entró en vigor en 1983. Tiene por objeto prohibir o restringir el uso de ciertas armas convencionales que se consideren excesivamente nocivas o cuyos efectos no son discriminatorios. Se trata de un anexo a los Convenios de Ginebra del 12 de agosto de 1949. Consta de cinco protocolos: restricción de armas no detectable de fragmentación, restricción de minas terrestres y de las armas trampa, de las incendiarias y de los láseres cegadores. También establece obligaciones para la limpieza de restos explosivos de la guerra. Además, la prohibición se amplió a las minas antipersona terrestres no localizables y a las minas de fragmentación. Cabe consignar que el convenio carece de mecanismos de verificación y cumplimiento y de métodos para llevar a cabo un proceso formal para la resolución de los problemas derivados de su debido cumplimiento. Es decir, no hay un control supraestatal.

<sup>107</sup> Se programaron reuniones del grupo de expertos gubernamentales para agosto y noviembre de 2017, pero fueron canceladas por falta de financiación. Cf. <[http://www.stopkillerrobots.org/wp-content/uploads/2013/03/KRC\\_PR\\_CCW\\_30May2017fnl.pdf](http://www.stopkillerrobots.org/wp-content/uploads/2013/03/KRC_PR_CCW_30May2017fnl.pdf)>.

*bots*<sup>108</sup> se argumentó que los LAWS no solo no son aptos para servir en conflictos armados por la falta de cumplimiento del derecho internacional humanitario, sino que también provocan una conculcación general de la ley: del derecho administrativo, del penal, del militar (por ejemplo, en los Estados Unidos se desactivaría la autorización del presidente como comandante en jefe y los LAWS se liberarían no solo de la jerarquía civil, sino, paradójicamente, también de la militar) y sobre todo del constitucional democrático, como veremos a continuación.<sup>109</sup>

En Naciones Unidas se están debatiendo la definición y los límites de los LAWS y ante la pregunta de qué se entiende por *robot autónomo armado* o *robot asesino*, tanto el Comité Internacional del Control de Robots Armados (ICRAC, por sus siglas en inglés) como la campaña Stop Killer Robots sugieren que es aquel sistema de armamento que tiene el potencial de carecer de cualquier *control humano significativo* en el proceso, bucle o ciclo de decisión de matar a un ser humano. Es decir, se trataría de un robot capaz de disparar discrecionalmente al objetivo, pero «fuera de control humano» (*human out-of-the-loop*);<sup>110</sup> control ese que sin embargo debe ser exigido tanto en una fase previa —a través de protocolos jurídico-técnicos que no puedan ser ignorados por el sistema— como durante el proceso, físico en este caso y llevado a cabo por un operador humano. La sustitución de personas por la inteligencia artificial puede aceptarse en ciertas actividades que requieren de alta precisión con vistas

<sup>108</sup> En línea: <<https://www.hrw.org/report/2014/05/12/shaking-foundations/human-rights-implications-killer-robots>>.

<sup>109</sup> McQuillan (2018) (según este autor, «la próxima generación de escándalos humanitarios será impulsada por inteligencia artificial»), Burri (2016) y Powell (2013).

<sup>110</sup> Sharkey (2016) y Suchman y Weber (2016).

a incrementar la seguridad de los resultados y salvar vidas, pero en otros casos, como en los LAWS, deberíamos seguir confiando en la toma de decisiones humanas, aunque sea ayudadas por asistencia informática. De lo contrario se pondrán en peligro la paz y la estabilidad regionales y/o mundiales. Es imperativo que nos anticipemos al futuro y seamos capaces de implementar barreras sólidas a los peligros irreversibles que la inteligencia artificial conlleva, máxime si se la aparea a armas nucleares o a sistemas de selección de objetivos individualizados.<sup>111</sup>

Los ingenieros coinciden en que no se ha conseguido aún lo que podríamos llamar *sistemas totalmente independientes*; que el *pensamiento independiente* no es todavía una realidad y no hay que confundirlo con la toma de decisiones independientes. Como explica G. Benson, «el pensamiento independiente está asociado al autoconocimiento y a la emoción, y aún no hemos conseguido este tipo de inteligencia artificial. Las complejas interacciones de nuestro cerebro funcionan con nuestra fisiología y son realmente difíciles de reproducir».<sup>112</sup> Hay, pues, tiempo todavía de desarrollar legislación y sistemas que le pongan coto. Y cada vez más países están de acuerdo en la necesidad de ponérselo, como ha evidenciado la reunión de expertos gubernamentales de la Convención sobre Ciertas Armas Convencionales celebrada en mayo de 2018. El objetivo está cifrado en negociar un tratado internacional que imponga límites a los LAWS y prohíba las armas completamente autónomas antes de que sean posibles tecnológicamente y de que

<sup>111</sup> Barzelay y Campbell (2003); Allison (2004). Esta autora ilustra que se suele creer que las armas nucleares tienen dimensiones voluminosas, pero existen bombas como la Davy Crockett, un arma nuclear de entre 120 y 155 milímetros, es decir, fácilmente transportable. Catorce años más tarde de la publicación de la obra de Allison, es evidente que las investigaciones habrán desarrollado un arma nuclear mucho más reducida.

<sup>112</sup> Benson (2017).

las armas autónomas de alta tecnología con diversos grados de control humano que ya están en uso en Estados Unidos,<sup>113</sup> China, Israel,<sup>114</sup> Corea del Sur,<sup>115</sup> Rusia,<sup>116</sup> Alemania,<sup>117</sup> Turquía<sup>118</sup> y el Reino Unido<sup>119</sup> se transformen en sistemas con inteligencia

<sup>113</sup> Estados Unidos ha desarrollado un así llamado *sistema aéreo de combate no tripulado X-47B* (UCAS: *unmanned combat air system*). Y una de las pruebas más importantes con AWS, consistente en enjambres de microdrones, ha sido llevada a cabo recientemente en California por el Departamento de Defensa. Concretamente, 103 drones Perdix fueron lanzados desde tres F/A-18 Super Hornets y demostraron comportamientos avanzados tales como la toma de decisiones colectivas, vuelo de formación adaptativa y autorreparación. No se trata de drones individuales preprogramados, sino que comparten IA para actuar como uno solo. Cf. <<https://www.defense.gov/News/News-Releases/News-Release-View/Article/1044811/departement-of-defense-/>>. Otros AWS son Talon (robot localizador de armas), SWORDS (Special Weapons Observation Reconnaissance Detection System; un robot localizador de armas armado), iRobot UGVs (robot de vigilancia y reconocimiento que puede llevar una bomba incorporada), BigDog/MULE, MAARS (Modular Advanced Armed Robotic System), X-47B (avión no tripulado para el combate aéreo), A-CTUV o Continuous Trail Unmanned Vessel (antiguerra submarina)... Cf. Jha (2006).

<sup>114</sup> Iron Dome, o sistemas antirradiación Harpy y Harop. También el Guardian, un vehículo terrestre.

<sup>115</sup> Super aEgis II, que es capaz de identificar, rastrear y destruir un objetivo a gran distancia sin operadores humanos, aunque el fabricante Dodaam Systems ha señalado que incluirán el requerimiento de la participación de un ser humano para autorizar disparos reales. También disponen del Korean Robotic Sentry o SGR-A1.

<sup>116</sup> Misil crucero antibuque P-700 Granit. Cabe señalar que la empresa de armas Kalashnikov Concern acaba de presentar un prototipo de robot asesino en la feria de armas de Moscú de este año (2018). Igorek mide 4 metros, pesa 4,5 toneladas y está totalmente blindado y está diseñado para «resolver tareas de ingeniería y combate» como un «caminante bípedo controlado» por humanos, es decir, con forma antropomórfica. Vives (2018).

<sup>117</sup> Nächstbereichschutzsystem Mantis. También el ugv multimisión de la empresa Rheinmetall, un sistema armado no tripulado que puede optar por actuar de forma autónoma. Por su parte, la empresa de armamento germano-francesa KNDS ha creado un vehículo robótico armado no tripulado llamado OPTIO X20, que puede operar de forma remota o autónoma.

<sup>118</sup> Torre armada creada por ASEL SAN para controlar la frontera con Siria.

<sup>119</sup> Misil antitanque Brimstone o el Taranis.

artificial en los que la decisión de atacar letalmente a otros humanos escape por completo de nuestro control.<sup>120</sup>

El plazo aparente de que disponemos para imponer esos límites a lo que ha sido descrito como *tercera revolución en la guerra* (la inteligencia artificial unida a la pólvora y a las armas nucleares) es un puñado de años. En este sentido, cabe alertar de que China está modernizando rápidamente su ejército y ha apostado por armas nucleares de última generación a través de ojivas con inteligencia artificial diseñadas para limitar el daño atacando a objetivos específicos.<sup>121</sup> En cambio, Estados Unidos es heredero aún de las armas del pasado, lo que les hace moverse más lentamente en lo que se ha denominado industria-militar *congresual*. Así, entre 2014 y 2018 China ha llevado a cabo alrededor de 200 experimentos de laboratorio para simular una explosión nuclear, mientras que Estados Unidos ha realizado 50 pruebas en el mismo periodo ha realizado 50 pruebas. La carrera emprendida por China es evidente. En el fondo, como señala Hartnett, del Bank of America, «la guerra comercial de 2008 debería ser reconocida como la primera etapa de una nueva carrera de armamentos entre Estados Unidos y China para alcanzar la superioridad nacional en tecnología a más largo plazo a través de la computación cuántica, inteligencia artificial, aviones de guerra hipersónicos, vehículos electrónicos, robótica y ciberseguridad».<sup>122</sup>

<sup>120</sup> «Se dice que el Taranis del Reino Unido es meramente semiautónomo, mientras que el nEUROn desarrollado por Francia, Grecia, Italia, España, Suecia y Suiza está explícitamente diseñado para demostrar una capacidad autónoma aire-tierra, como parece ser el caso del MIG de Rusia. Aunque se sabe poco sobre la espada afilada de China, es poco probable que esté muy por detrás de sus competidores en términos conceptuales», explica Michael Hass (2014), que fue candidato a Premio Nobel de la Paz.

<sup>121</sup> Ignatius (2018).

<sup>122</sup> Durden (2018)

¿Es  
correcto?

De ahí que la inversión en tecnología esté fundamentalmente ligada al gasto en defensa (aunque esto no significa siempre obtener mayor seguridad).<sup>123</sup> El pronóstico del FMI es que China supere a Estados Unidos progresivamente hasta 2050, y que se convierta en la superpotencia dominante del mundo. Concretamente, se habla de que supere la economía y la fuerza militar del gigante norteamericano, así como su influencia global en el mundo, como más tarde en 2032.

Existe el peligro de que no escapemos a la *trampa de Tucídides*, como llama Allison (y otros) a la situación a que se da lugar cuando el poder dominante (en este caso, Estados Unidos u Occidente en general) tiene la tentación de atacar a las potencias en auge (China).<sup>124</sup> «La guerra entre los dos países en las próximas décadas no solo es posible sino mucho más probable de lo que se piensa», dice, y podemos coincidir con este autor en que «el desafío preeminente de esta era no son los extremistas islámicos violentos o una Rusia resurgente sino el impacto que la ascendencia de China tendrá en el orden internacional liderado por Estados Unidos».<sup>125</sup>

<sup>123</sup> Danzig (2018). El autor, profesor de la Universidad Johns Hopkins y ex secretario de la Marina durante la presidencia de Clinton, señala en su informe que la superioridad tecnológica no es sinónimo de seguridad debido a la dramática pérdida de control que puede producirse en dichos sistemas de inteligencia artificial, biología sintética y armamento autónomo, sea por accidentes, efectos emergentes insospechados o sabotajes.

<sup>124</sup> Allison (2017). Tucídides, general y historiador ateniense (400 a. de C.) mantenía que era difícil que una potencia en pleno auge coexistiera pacíficamente con la potencia dominante, como ocurrió cuando Atenas desafió a Esparta en la Grecia antigua, o como hizo Alemania hace un siglo. Tucídides veía justificable la postura de Atenas, ya que a medida que crecía su influencia, también creció la confianza en sí misma, su conciencia de las injusticias pasadas, su sensibilidad ante los casos de falta de respeto, etcétera. Y era natural, según el historiador, que Esparta interpretara la postura ateniense como irracional, ingrata y amenazante para el sistema.

<sup>125</sup> Allison (2015).

Los retos que representa esa revolución tecnosocial solo pueden ser comprendidos cabalmente tras analizar las interdependencias existentes entre nuestro desarrollo como especie y los marcos tecnológicos que lo han propiciado. En el siguiente capítulo exploraremos, siguiendo esa línea, la relación entre sociedad y tecnología y su importancia a la hora de determinar nuestra aproximación al tema que nos ocupa, así como algunas de sus implicaciones éticas.<sup>126</sup>

<sup>126</sup> Leveringhaus (2016).





# Tecnología y sociedad: implicaciones de la inteligencia artificial

JOAQUÍN RODRÍGUEZ

Supimos que el mundo no sería el mismo. Unas pocas personas rieron, unas pocas lloraron, muchas estuvieron en silencio. Recuerdo la cita del libro sagrado hinduista, el Bhagavad-Gita. Vishnu está tratando de persuadir al Príncipe para que cumpla con su deber y para impresionarlo toma su forma con múltiples brazos y dice: «Ahora me he convertido en la muerte, el destructor de mundos». Supongo que todos pensamos eso, de una u otra forma.

Robert Oppenheimer sobre la bomba H.

## Humanidad y tecnología

Hasta hace más de dos millones de años, el ser humano era solo una especie más entre muchas; una contendiente como otra cualquiera en la lucha diaria por la supervivencia; pero todo cambió cuando nuestros ancestros remotos lograron dominar el fuego y desarrollar las primeras industrias líticas. Propulsada por aquellas primeras tecnologías, aquella especie débil se encaramó de pronto a una posición dominante: era capaz, como no lo era ninguna otra, de domeñar el medio natural, que hasta entonces había representado una amenaza arbitraria e impredecible. Y como no podía ser de otro modo, semejante transformación material trajo aparejada otras de orden más bien espiritual. Cambió radicalmente nuestro

concepto del mundo y de nosotros mismos e hicieron aparición nuevas formas de aproximación a las reglas invisibles de nuestro contexto.<sup>127</sup>

Siguieron corriendo los milenios y en el Neolítico —entre los años 10 200 y 2000 a. de C., según la cronología ASPRO— nuestra cultura material volvió a transformarse profundamente merced a la revolución agrícola, que nos permitió abandonar nuestra vida nómada, echar raíces y tomar posesión del territorio de forma tanto física como simbólica. Pasamos a vincularnos no solo con los animales que dibujábamos en las paredes de las cuevas para propiciar la caza a través de la magia simpática (probablemente nuestro primer sistema de creencias o metodología cognitiva),<sup>128</sup> sino también con los ríos, montañas y árboles de nuestro entorno, a los que otorgamos igualmente un significado mágico y erigimos como símbolos y fronteras de un nuevo sistema.

Este nuevo sistema agrícola fue desencadenando diversos fenómenos paralelos. Por un lado, la novedosa capacidad de transferir los frutos del trabajo de hoy al mañana a través del almacenaje de los productos agrícolas nos hizo adquirir un nuevo sentido del tiempo. Por otro, la producción de excedentes gracias a la mejora de la técnica agrícola dio luz al comercio. Y por otro, se fue generando un nuevo paisaje: el urbano, que es aquel en el que hoy se desenvuelve la mayor parte de la humanidad. Las primeras ciudades (Uruk, Jericó, Çatalhöyük...) brotaron en el Creciente Fértil alrededor del 4500 a. de C., y en su seno se fue gestando una nueva distribución del trabajo. Así, algunas personas eran necesarias para las labores agrícolas, mientras que otras lo eran para el desarrollo de incipientes actividades industriales, otras para

<sup>127</sup> Bernstein (1996).

<sup>128</sup> Frazer (1951).

el comercio, etcétera. Surgían por otro lado en esas ciudades florecientes nuevas necesidades, tales como la seguridad pública o la contabilidad de entradas y salidas de excedentes; y eso hizo necesario desarrollar otras tecnologías, y singularmente la escritura. Este nuevo y capital invento ofreció a los mesopotámicos la posibilidad de transmitir y almacenar información, compilar códigos legales como el de Hammurabi, centralizar el control de los bienes o cristalizar mitos legitimadores del orden social. Surgía la burocracia y, aparejada a ella, la figura del escriba, fácilmente identificable como lo que Thorstein Veblen llamaba una *clase tecnológica*: un conjunto de personas titulares y protectoras de un conocimiento estrechamente vinculado al poder en tanto que fundamental para sostener el tejido social.<sup>129</sup> El faraón necesitaba a los escribas para mantener el control del imperio, el emperador maya necesitaba a su sumo sacerdote y los reyes europeos necesitaban al papa.

Todo lo anteriormente expuesto venía imbuido, tal y como refleja la mitología griega (construida durante este período), de nuevos desafíos y riesgos. Conscientes de cómo los ciclos tecnológicos pueden transformar radicalmente un corpus cultural y de cuán difícil puede llegar a ser embridar esos cambios, los clásicos supieron someter al desarrollo tecnológico a un control estrecho. De los griegos se dice que nunca acometieron un desarrollo tecnológico profundo no porque no pudieran, sino porque no querían.<sup>130</sup> No lo quería, por ejemplo, Arquímedes, que después de demostrar sus teorías siempre destruía sus inventos.

La tecnología tiene la capacidad de modificar radicalmente las fuentes de significado colectivo. Existe una relación inter-

<sup>129</sup> Veblen (1919, 1944).

<sup>130</sup> Colli (1978); Ellul, Wilkinson y Merton (1964).

dependiente entre cultura material y proceso cognitivo; nuestras capacidades materiales moldean nuestras cosmovisiones. En aquel tiempo que fue el de los albores de la civilización desaparecieron las viejas tradiciones y adoraciones mágicas y nació la religión organizada, que sustituía al hechicero por el sacerdote y rompía los antiguos vínculos con la naturaleza. De un sistema que profesaba su capacidad de modificar las leyes de la naturaleza a través de la magia se pasó a otro en el que la naturaleza era simplemente el patio de recreo en el que se manifestaba el capricho de un panteón de dioses cuyo favor debía ganarse. Y ello generó una nueva estructura de poder para cuya extensión las ciudades eran un elemento clave. Era en la ciudad donde se erguía el templo desde cuya cúspide los sumos sacerdotes monopolizaban la relación con los dioses y adquirían un poder que llegaba a ser inmenso. También ellos eran una *clase tecnológica* que obtenía su poder a través del conocimiento; conocimiento que podía ser científico: pensemos en los gobernantes mayas que salían a lo alto de una pirámide y amenazaban a su gente con un eclipse solar y en el miedo que aquellas personas debían de experimentar en el momento en que el Sol desaparecía.

Que el conocimiento pueda proporcionar semejante poder también explica las formas rebuscadas y barreras de acceso que casi todas las sociedades de la historia han solido buscar a fin de proteger a los titulares de esos saberes fundamentales de la amenaza que una generalización de los mismos podía representar para su posición. La complejidad de los jeroglíficos egipcios y mayas ofrece un buen ejemplo de este propósito de estructurar el conocimiento de tal manera que fuera difícilmente accesible a la gente común.

Desde la Antigüedad hasta hoy, el papel de la tecnología nunca ha disminuido en importancia; y esa importancia es fácilmente rastreable a todo lo largo de la historia, manifesta-

da en la irrupción de determinados inventos que tuvieron la capacidad de transformar profundamente las sociedades que los crearon. Lin White explica por ejemplo cómo la nueva máquina de guerra que fue el estribo dio luz al feudalismo. La combinación de un hombre, un caballo y una espada propició una nueva hegemonía en el campo de batalla y las necesidades de entrenamiento de estos nuevos soldados de élite los obligaron a abandonar sus formas tradicionales de vida (agricultura, artesanía, etcétera) y a ser caballeros a tiempo completo, pasando a erigirse como una tercera clase técnica entre la Monarquía y la Iglesia y la gente común y a tejer en torno a sí las relaciones feudovasalláticas propias de este sistema que atraviesa toda la Edad Media.<sup>131</sup> Es solo un ejemplo: la revolución industrial, la de las comunicaciones, etcétera, todas ellas reprodujeron más tarde el mismo mecanismo. Cualquier nueva tecnología puede provocar un trastocamiento drástico de los equilibrios internos de un sistema: pensemos también en lo que supuso para el poder de la Iglesia católica la invención de la imprenta de Gutenberg, estrechamente vinculada a la revolución protestante de Lutero, que acabó con el monopolio ideológico del Vaticano sobre Europa proponiendo una relación con Dios sin intermediarios y un acceso directo y más amplio a los arcanos religiosos.

Cualquier comunidad dada necesita un progreso tecnológico constante para perpetuarse en un contexto de competencia con otras, pero también un fuerte control sobre el sistema tecnológico a fin de preservar el *statu quo* interno; así como un sistema de creencias que otorgue un significado colectivo a la comunidad mientras justifica el orden social. En relación con todo esto, podemos afirmar que un sistema tecnológico más complejo necesita herramientas de gobierno más com-

<sup>131</sup> White (1973).

plejas, y que un sistema de gobierno más complejo requiere a su vez de uno de creencias que también lo sea. Todo está interrelacionado y Marx afirmaba que «el molino de mano nos dará la sociedad con el señor feudal; el molino de vapor, con el capitalista industrial».<sup>132</sup>

Podríamos hablar de una tragedia en tres actos que se repite una y otra vez. Primer acto: aparece una nueva tecnología y, ligada a ella, una nueva clase técnica asociada al conocimiento necesario para implementarla. Segundo acto: quienes detentan el poder *stricto sensu* abren el proceso de toma de decisiones a la clase tecnológica como respuesta a una necesidad operativa del sistema. Con el tiempo, el conocimiento puede ir extendiéndose a una gran parte de la comunidad porque sea necesario para la evolución y la perpetuación del sistema, acercando a este a la participación democrática y a la provisión de calidad de vida, seguridad y orden a amplios sectores de la comunidad; pero también puede darse la situación de que el progreso científico comience a desarrollarse fuera de los márgenes del control estatal y que acabe apareciendo en el horizonte una nueva tecnología que desestabilice los frágiles equilibrios establecidos entre la clase técnica y la dominante, generándose un choque como el descrito por filósofos como Gramsci, Pareto o Mosca.<sup>133</sup> Esa situación puede dar lugar a dos escenarios principales diferentes: una nueva apertura democrática del proceso de toma de decisiones o una reconcentración autoritaria del control de los medios de producción material y simbólico en un grupo más cerrado de personas, que no necesariamente acabe con la democracia pero sí la erosione o la limite. En relación con la democracia, debemos compren-

<sup>132</sup> Marx (2008).

<sup>133</sup> Bates (1975); Gramsci (1995); Femia (1987); Nye (1977); Pareto (1991).

der que somos, como decía Nietzsche, esclavos de nuestras propias palabras,<sup>134</sup> y también de esta que ni en la teoría ni en la práctica designa un sistema de una sola vía. Por *democracia* debe entenderse en cambio una amplia gama de aplicaciones posibles que pueden ser muy diferentes entre sí, y algunas de las cuales pueden ser y de hecho son compatibles con el desarrollo, por parte de las personas al mando, de distintas técnicas orientadas al control y la manipulación de la opinión pública a través del de la educación, los medios de comunicación, etcétera.

La inteligencia artificial representa una nueva frontera; una redefinición radical de procesos organizativos y cognitivos, de la construcción de la otredad, de los mecanismos del Estado, de los símbolos que dan sentido colectivo a nuestra sociedad y, en general, de la relación del ser humano con su contexto. Nuevamente, nos hallamos ante una tecnología capaz de transformar nuestra realidad material y llamada a formar nuevas élites y bien a deconstruir los sistemas de privilegio vigentes, bien a cristalizarlos aún más. Tenemos ante nosotros, en consecuencia, el desafío de prever las transformaciones por venir, preparar a nuestras comunidades y definir marcos que permitan que el proceso de toma de decisiones admita a la mayoría.

En el caso concreto de la aplicación armamentística de la IA, nos enfrenta al advenimiento de un nuevo orden distópico. La delegación de la capacidad de matar en un sistema del que apenas somos capaces de predecir el comportamiento futuro y cuya fiabilidad no puede ser garantizada al cien por cien ni en lo relativo a la ejecución de las órdenes ni en lo relativo al cumplimiento del derecho internacional, delegación que no hay nada que pueda justificar desde un punto de vista

<sup>134</sup> Nietzsche (1989).

ético (ni la eficiencia, ni el coste ni tan siquiera la protección de soldados propios), es un paseo suicida hacia el abismo que solo evitaremos si somos capaces de dotarnos de instrumentos legales vinculantes que garanticen que la vida humana no puede ser sustraída por entes no humanos. Deben restringirse ciertas aplicaciones a la vez que se potencian otras a partir del propósito de socializar la tecnología; de hacerla accesible a amplias capas de la población de tal manera que nos ayude a construir una sociedad abierta y plural.

## **Inteligencia artificial y percepción**

Me parece, en este pequeño detalle, ser más sabio que este hombre en cualquier caso; que lo que no sé, tampoco creo que lo sepa.

Platón: *Apología*

La tecnología no solo afecta a la forma en que nos organizamos, sino que tiene la capacidad de transformar la manera en que percibimos nuestro entorno, nuestra realidad; de redefinir qué vemos, sentimos y experimentamos.

Podemos definirnos como seres inherentemente creativos; como una especie que tuvo la imaginación y el ingenio suficientes para desarrollar no solo elementos materiales capaces de garantizar su supervivencia, sino también sistemas cognitivos basados en reglas que, aunque en la mayor parte de los casos carecían de base científica, nos fueron permitiendo interpretar nuestro contexto. Entre el arte rupestre paleolítico —entendido por el abate Breuil y otros como un sistema de magia parasimpática diseñado para facilitar la caza, la fertilidad, la salud y la seguridad— y la ciencia aeroespacial hay un aspecto en común, que es la capacidad de la tecnología para definir nuestras creencias.



En tanto que humanos, no tenemos ningún mecanismo inherente que nos permita distinguir automáticamente entre la realidad y la ficción. A lo largo de los milenios, el ser humano ha ido creyendo en la magia, en dioses viejos y nuevos, en que la Tierra era plana y en toda otra serie de cosas que incluso en un contexto de alto desarrollo científico como el nuestro siguen mostrando una persistencia extraordinaria entre amplias capas de la población. Y ello representa un inconveniente enorme a la hora de establecer relaciones de cualquier tipo basadas en una verdad común. No importa cuántas veces la ciencia cure una enfermedad, ponga un satélite en órbita o haga funcionar teléfonos inteligentes o electrodomésticos: siempre habrá dudas, teorías de la conspiración e intentos de proporcionar respuestas alternativas. Este problema encuentra sus raíces en la esencia misma de la razón humana, que las más de las veces funciona como un método de justificación *a posteriori* para nuestras acciones y como una forma de muñir una historia cohesiva, sin que importe la viabilidad científica de lo que se dice. Las religiones organizadas ofrecen un buen ejemplo de este esfuerzo intelectual en clara disociación de las leyes de la naturaleza.

Numerosos filósofos a lo largo de la historia han analizado estos problemas y llegado a la conclusión de que la razón no es por sí misma suficiente para comprender la naturaleza de nuestro contexto, y que cualquier análisis que se haga requiere de experiencias tanto individuales como colectivas para adquirir sentido.<sup>135</sup> Pero también la experiencia representa varios problemas teóricos. El mito platónico de la caverna ilustra magníficamente tales problemas al hablarnos de la imposibilidad de percibir la realidad en sí misma: lo que percibimos son solo proyecciones de la realidad; sombras en las que apenas atinamos a adivinar aquella.

<sup>135</sup> Hegel (1966); Kant (1998); Schopenhauer (1959).

Después de Platón, el problema de la experiencia ha sido profundamente analizado por otros filósofos, y podemos distinguir varias respuestas al problema: desde que nuestra percepción está irremediablemente condicionada por nuestro contexto hasta que, incluso en el caso de que la percepción pueda estar equivocada, debemos aceptar que es una base común para analizar la realidad. En otras palabras, si todas las personas que compartimos un determinado espacio vemos que está lloviendo, debemos aceptarlo como una suposición correcta. Pero incluso en este punto nos encontramos frente a otro conflicto que Nietzsche expuso en su *La genealogía de la moral*. El filósofo alemán presentó allí las palabras como una prisión que debe romperse a fin de generar un nuevo código simbólico que permita expresar verdaderamente nuestro contexto. Comprendió Nietzsche que la palabra usada trasciende con mucho su significado estricto y representa toda una historia colectiva e incluso individual: la de la persona que la emplea encontrando inevitablemente en ella una intersección de experiencias pasadas asociadas con ella. De ello dan buena cuenta vocablos como *negro*, *mujer* o *gay*.

Todo esto adquiere una nueva dimensión con el surgimiento de la inteligencia artificial y su habilidad para trabajar con el lenguaje natural, ya que —como ya hemos apuntado en este ensayo— el uso de determinadas palabras en determinados contextos puede dar lugar a la cristalización de dinámicas de opresión. Sucedió en Facebook con la aparición de categorías antisemitas en su sistema de publicidad, tras lo cual la empresa aseguró que contrataría a más operadores humanos con el fin de controlar este tipo de consecuencias no deseadas de la puesta en marcha de sus algoritmos.<sup>136</sup> Si

<sup>136</sup> Lee (2017).

hablamos de armamento, la necesidad de aumentar el control humano significativo no puede sino aumentar.

Nuestra moral, en tanto que codificación de valores, funciona de manera muy similar, y ello hace necesario revisar los parámetros básicos con los que trabajamos a fin de adaptarnos a una nueva condición humana que conlleva la incorporación masiva de tecnologías a nuestra vida cotidiana; tecnologías cuya capacidad de aprender y reproducir nuestros patrones de comportamiento o el marco cultural en el que operan en el caso de que no exista límite o control estricto sobre ellas por parte de los poderes existentes puede reforzar el sistema de desigualdad existente en lugar de erosionarlo.

Por otro lado, para el desarrollo de la inteligencia artificial es fundamental imprimirle la capacidad de discernir simulaciones, capacidad que la percepción humana natural no necesariamente tiene. A modo de ejemplo, podemos poner el de una tradición española fácilmente traducible a otras similares de otros contextos: la de los tres Reyes Magos de Oriente, de origen bíblico y que la sociedad actual ha transferido a lo hiperreal. En España, toda la realidad material —desde los programas y anuncios de televisión hasta las cabalgatas en la calle— refuerza esa tradición con el objetivo de permitir a los niños creer y comportarse en consecuencia. Todos los adultos sabemos que son los padres y familiares de los niños quienes compran los regalos e interpretan unos papeles cuyo objetivo es imprimir algo de magia a nuestras vidas y rutinas. Y pensamos que la creencia en los Reyes Magos es infantil, propia solo de una determinada edad, pero la edad no tiene nada que ver con la creencia. El mismo proceso se reproduce *ad infinitum* entre los adultos en nuestra sociedad y cultura. Los niños creen porque todo el sistema refuerza la creencia: ven a los Reyes Magos, los tocan e incluso les pueden dar listas de los regalos que desean; regalos que después aparecen

bajo el árbol sin que nadie en casa se haga responsable de ello, afirmando en cambio que fueron traídos por los Reyes. Viven una suerte de *show de Truman* en el que todo está diseñado para que no haya lugar a la duda. ¿Podemos asegurar los adultos que no somos susceptibles de caer en estos juegos de fantasía? ¿Estamos verdaderamente seguros de que la realidad es lo que parece ser?

Podemos responder a esta pregunta remitiéndonos al método científico, que establece como pilares básicos la falsabilidad y la reproductibilidad precisamente con el fin de evitar las desviaciones a que nos conduciría nuestra razón pura y prescribe el desarrollo de teorías que puedan probarse bajo hipótesis capaces de predecir escenarios futuros. El sistema así configurado es el mejor de cuantos hemos ido desarrollando a lo largo de la historia, pero su infalibilidad es difícil de probar. Baudrillard afirmaba que «el secreto de cualquier teoría es que la verdad no existe»;<sup>137</sup> y tal vez sea esa la única *verdad* digna de tal nombre que hayamos encontrado hasta ahora. «Solo sé que no sé nada», la vieja máxima socrática, regresa una y otra vez. Por más que hayamos logrado avances tecnológicos sin precedentes, todavía estamos inmersos en la oscuridad.

En el tiempo actual, nuestro enfoque de la realidad está altamente determinado por la tecnología, principalmente porque —como ya se ha explicado en este ensayo— el conocimiento está íntimamente relacionado con el poder. Existe una clase tecnológica que se beneficia del monopolio de un conocimiento restringido y que constituye una élite social cuya mera existencia está vinculada a la tecnología. Y existe, como ha existido siempre, la censura, factor históricamente importante para retener el poder y que puede

<sup>137</sup> Baudrillard (1987), p. 290.

tomar muchísimas apariencias, desde la que predijo Orwell hasta la que predijo Huxley, pero cuyo objetivo es siempre el mismo: alejar el conocimiento del público. En realidad, no importa si el acceso al conocimiento es directamente vedado a través de barreras de acceso o si solo se distrae al público con otras golosinas, como por ejemplo una cultura del entretenimiento que haga que la población prefiera seguir un programa de televisión que la política nacional: los efectos son los mismos y la única diferencia está en la cantidad de represión requerida. Para quienes detentan el poder, es más cómodo desviar la atención que prohibir el acceso a la información.

Todas las civilizaciones, todas las culturas, disponen de medios para controlar el acceso a la información. Todas han creado mitos e historias que advierten sobre los riesgos que el conocimiento comporta: lo mismo para Prometeo, que roba el fuego a los dioses y se lo lleva a la humanidad, que para Eva, que prueba la manzana del Árbol del Conocimiento, que para Pandora, el conocimiento siempre se consigue a un precio elevado, y esas metáforas alimentan las restricciones. Por otro lado, la tecnología misma puede ser utilizada para transformar nuestra propia percepción de la realidad desde una perspectiva tanto individual como colectiva. Hoy en día, la digitalización está generando una nueva capa de complejidad; una era posterior a la verdad.

Si nos centramos en cómo se fabrica el consenso social, cabe remitirnos a las estrategias descritas por Edward S. Herman y Noam Chomsky en su *Manufacturing consent*, donde argumentan que en el sistema actual, la mayoría de los medios de comunicación solo transmiten las opiniones de las élites económicas o los gobiernos. En el caso de Estados Unidos, este sistema operaría de acuerdo con cinco leyes:

1. La mayoría de los medios está en manos de grandes corporaciones, es decir, pertenece en realidad a las élites económicas.
2. Los medios dependen de la publicidad de las élites económicas para su subsistencia.
3. Los medios deben producir un flujo permanente de nuevas noticias, pero los principales proveedores de noticias son los departamentos de prensa de gobiernos y grandes corporaciones.
4. Los grupos de influencia pueden organizar respuestas sistemáticas a cualquier desviación de las opiniones que apoyan.
5. Anticomunismo: las opiniones de la izquierda son consideradas «antipatrióticas».<sup>138</sup>

A esto, Rafael Correa, presidente de Ecuador, se refirió afirmando que «desde que existe la imprenta, la libertad de prensa es la que permiten los amos de la imprenta».<sup>139</sup> Napoleón, a su vez, decía que la religión era lo que impedía que los pobres asesinaran a los ricos; y ahora los medios juegan ese papel alienador. En efecto, una clara mayoría de la información que absorbemos en nuestra vida cotidiana es generada por medios de comunicación cuyo modelo de financiación es la publicidad. Y ese mecanismo narcotizador viene adquiriendo renovada complejidad a través de los algoritmos de personalización que ahora funcionan en Internet, tal y como explica Eli Pariser en su libro *El filtro burbuja: cómo la red decide lo que leemos y lo que pensamos*. Pariser define el *filtro burbuja* que da título a la obra como el ecosistema de infor-

<sup>138</sup> Herman y Chomsky (2010).

<sup>139</sup> Entrevista a Rafael Correa en Televisión Española el 19 de marzo de 2012 [en línea], <<http://www.rtve.es/alacarta/videos/los-desayunos-de-tve/desayunos-tve-rafael-correa-presidente-ecuador/1352796/>>. [Consulta: 20-8-2018].

mación personal que estos algoritmos van configurando para el usuario a medida que este va mostrando que un determinado conjunto de temas le interesan al hacer clic en enlaces o publicaciones de amigos, añadir películas a una lista de reproducción, leer noticias, etcétera. Toda esta información es procesada y da lugar a una huella digital que las empresas utilizan para publicar anuncios personalizados o hacer que esos anuncios aparezcan casi invariablemente en las páginas de resultados de los motores de búsqueda que el usuario utiliza.

La preocupación de Pariser es en cierto modo similar a la de Tim Berners-Lee en un artículo de 2010 sobre el *efecto Hotel California*: «Puede registrarse, pero no puede darse de baja». Berners-Lee aludía con ello a cómo las redes sociales no permiten que el usuario vea contenido de la competencia, convirtiéndose en *silos cerrados* que comportan el riesgo de fragmentar la *web*. Pariser advierte de que este filtrado nos hace más cerrados a nuevas ideas, asuntos e información importante y crea la impresión de que nuestros intereses limitados son los únicos que existen, haciendo a las personas más cerriles sobre sus propias opiniones y mucho más vulnerables a la propaganda y la manipulación. El caso de Cambridge Analytics ha puesto de manifiesto que estas advertencias no son baladí; y hay muchos otros fáciles de fabular. Imagínese el caso de una mujer embarazada que ha mostrado interés por terapias como la homeopatía o las flores de Bach en sus búsquedas en Internet a lo largo de su vida. Si en un momento dado decide hacer una búsqueda relacionada con las vacunas, es muy probable que los primeros artículos que aparezcan en su motor de búsqueda sean negativos y que ello tenga consecuencias nefastas para ella y su bebé.

Tras la publicación del trabajo de Pariser, un portavoz de Google aseguró que se agregarían nuevos algoritmos al motor de búsqueda a fin de «limitar la personalización y promo-

ver la variedad» de los resultados.<sup>140</sup> Su reacción ilustra bien algo muy importante: si conocemos los efectos de la tecnología, tenemos la posibilidad de regular sus efectos; de darles forma. Las agencias públicas de evaluación de tecnología y la regulación de la inteligencia artificial son necesarias y tal vez la única forma de recuperar algo del poder que se ha ido transfiriendo sutilmente de lo público a lo privado. Debe profundizarse asimismo en los análisis relativos al uso potencial de algoritmos de personalización para manipular la vida pública e inducir a consensos sociales que pueden derivar en nuevas formas de totalitarismo. También en cómo la citada fragmentación de la *web* provocada por estos algoritmos puede contribuir a generar espacios estancos de referenciación de la otredad.

Vivimos en un momento en el que la individualidad se estandariza en base a arquetipos fácilmente transformables en nichos de mercado y se desecha toda herramienta orientada a la formación de pensamiento crítico. Se inculcan, en cambio, valores relacionados con la obediencia y el conformismo y se busca condenar a la población a una suerte de sonambulismo constante. Incluso el viejo mito del ecosistema subversivo universitario ha quedado enterrado bajo una avalancha de procesos evaluativos y de calidad que valoran a las Universidades no por el conocimiento que imparten, sino por su capacidad para instalar a sus estudiantes en el mercado laboral. La Universidad ha pasado a ser vista como una palanca social que contribuye a un mejor posicionamiento del sujeto en la pirámide aspiracional.

Que la inteligencia artificial venga a cristalizar todo esto es un riesgo que debe ser tenido en cuenta, sobre todo en aquellas implementaciones que tengan por objeto una acción di-

<sup>140</sup> Weisberg (2011).



recta sobre la vida humana o trabajen sobre la tipificación de comunidades. El traslado de los prejuicios sociales de ciertos colectivos a la propia inteligencia artificial es un círculo vicioso de difícil salida: llegará un punto en que los algoritmos se dejen de adaptar a nosotros y pasemos a ser nosotros los que nos adaptemos a ellos.

Tal y como Neil Postman describió en su libro *Tecnópolis: la rendición de la cultura a la tecnología*,<sup>141</sup> estamos dando vida a una sociedad que transfiere las virtudes de las máquinas a los propios humanos, de tal manera que la eficacia y la eficiencia se convierten en valores fundamentales para el comportamiento. Si lo pensamos detenidamente, la gran mayoría de las empresas pide hoy a sus trabajadores que gestionen las contingencias con vistas a la resolución inmediata del problema, no a ramificaciones futuras que la solución pueda suponer. Crisis como la económica de 2008 encuentran en ello parte de su explicación: la dinámica endiablada del sistema impidió vislumbrar la crisis que se avecinaba. Y en ella, por cierto, los algoritmos jugaron un papel determinante, tal y como se explica en un artículo publicado en *The Guardian* y titulado «Was software responsible for the financial crisis?». <sup>142</sup> En él se incide en la manipulación de las percepciones ejercida por los algoritmos y en el posterior efecto dominó que se desencadenó debido a la automatización de órdenes de venta ante determinados acontecimientos. Que la consecuencia de ello haya sido aproximadamente una década de recesión nos ofrece una buena demostración de los problemas que puede representar el uso extensivo de inteligencia artificial sin control humano y en un contexto anómico o desregulado.

<sup>141</sup> Postman (2011).

<sup>142</sup> Dodson (2018).

Debe subrayarse algo: la inteligencia artificial no se limita a condicionar nuestra percepción de la realidad, como ha sucedido con otros *sets* tecnológicos a lo largo de la historia, sino que tiene una capacidad de incidencia real y autónoma sobre la propia realidad física en tanto no se ejerza un control humano constante sobre ella. Así, aquellas inteligencias artificiales emplazadas en entes robóticos serían materializaciones físicas de lo digital con capacidad plena de interacción. Y debe insistirse mucho en la respuesta urgente que requiere el riesgo de utilizar la inteligencia artificial en operaciones de las que dependen vidas humanas de forma directa. Cualquier tipo de armamento que utilice inteligencia artificial debe ser descartado tajantemente y la investigación relacionada suspendida hasta que queden aclaradas todas sus implicaciones éticas.

«El simulacro nunca es lo que oculta la verdad; es la verdad la que oculta que no existe. El simulacro es verdadero», escribía Baudrillard.<sup>143</sup>

## Sistemas de armamento e inteligencia artificial

Jano, el dios romano de la técnica, era comúnmente representado con dos caras: una alegoría clara de los avances y desastres potenciales que sus creaciones podían provocar por igual. Ya desde época de Sócrates los antiguos tenían muy presente esta cuestión. El diálogo platónico entre Thamus y Theuth suele ponerse de ejemplo de esa preocupación. Y esa dualidad siguió siendo reconocida con el tiempo por otros autores: por ejemplo, Sigmund Freud, quien en su obra *El malestar en la cultura* asocia muchas de las necesidades

<sup>143</sup> Baudrillard (1994).

económicas, sociales y filosóficas actuales con el devenir de la técnica. O Jacques Ellul, quien en *La technique*, de 1967, concibe la tecnología como una fuerza subyacente dotada del poder intrínseco de definir los marcos mentales y, por ende, dar forma a la historia condicionando a sus sujetos. El sistema así configurado evoluciona paralelo al orden social, es capaz de tambalearlo cuando resulta excesivamente restrictivo para la reproducción del propio sistema tecnológico y debe concebirse como una fuerza viva e interdependiente de nuestra especie y como un elemento básico de la superestructura del sistema en el sentido gramsciano del término.<sup>144</sup>

Como ya se ha venido exponiendo, la inteligencia artificial representa una nueva fase del dominio de la tecnología sobre la realidad, ya que es capaz de actuar tanto en lo digital como en lo material de una forma que carece de precedentes. Ello obliga a una aproximación multidisciplinar que permita evaluar sus implicaciones de una forma holística. Para ello, debe comenzarse por prestar atención a aspectos clave como la información sobre la que se ha construido el sistema y los datos que contribuyeron a su desarrollo. Ha de partirse en este sentido de la definición que de la privacidad da D. J. Solove: no una preservación del interés personal en contra del interés social, sino la protección del individuo sobre la base de los valores de la sociedad. «No se puede luchar por un derecho individual contra el bien social más importante. Las cuestiones de privacidad implican un equilibrio de intereses sociales en ambos lados de la escala», explica.<sup>145</sup> Esta noción de privacidad y protección de lo personal implica una propiedad de los datos por parte del sujeto que, en caso de ser cedida, debe serlo a través de un consentimiento explícito.

<sup>144</sup> Bates (1975); Gramsci (1995).

<sup>145</sup> Solove (2007).

Así, aquellas compañías civiles que colaboren en proyectos militares deben informar a sus usuarios de si sus datos han sido utilizados de alguna forma en el desarrollo de algoritmos que puedan tener uso militar. Y el consentimiento explícito debe ser pedido incluso si se trata de tecnologías derivadas que ya no incluyen los *data sets* originales. Esta medida podría ser efectiva para reducir los incentivos que estas empresas puedan tener para poner su tecnología a disposición de usos militares.

Otra cuestión a la que hay que prestar atención es la nueva escalada militar que los sistemas de armamento pueden producir, muy similar a la que en su día provocó la bomba atómica. Esta no condujo a la humanidad a su final gracias al mecanismo de disuasión nuclear mutua, pero nada garantiza que dicho mecanismo vuelva a funcionar con los AWS. Y eso nos obliga a toda una serie de debates jurídicos acompañados de reflexiones éticas y antropológicas que —entendiendo, como Culkin, que «damos forma a nuestras herramientas y luego nuestras herramientas nos dan forma»—<sup>146</sup> recuperen la noción de dignidad humana como eje central. El derecho no basta por sí solo, pues cualesquiera argumentos de tipo legal que se dé en esgrimir pueden ser fácilmente contrarrestados por el *establishment* en cualquier momento apelando a la seguridad nacional. Numerosos autores defienden que la ética es mucho más compleja de alterar, ya que conforma una parte básica del sujeto.<sup>147</sup>

En todo caso, debe tenerse en cuenta que la ética no es neutral, sino que hay principios éticos contrapuestos, y por lo tanto la regulación es fundamental. Como señala el psicólogo Daniel Kahneman, Premio Nobel de Economía, la

<sup>146</sup> Culkin (1967).

<sup>147</sup> Sharkey (2008).

regulación predetermina el comportamiento. Las mujeres no votaron durante décadas y se consideraba ético; el catolicismo institucionalizó una ética de la esclavitud; la caza mayor en África era ética hasta que se limitó, etcétera. Es decir, la ética tiene que correr paralela al derecho, reforzándose mutuamente.

Antes de continuar avanzando en nuestra exposición, es pertinente repasar los argumentos legales que, basados en la cláusula Martens (introducida por primera vez en los preámbulos de la Segunda Convención de La Haya sobre Leyes y Costumbres de la Guerra Terrestre de 1899 y ligeramente modificada hasta alcanzar su forma actual en las Convenciones de La Haya de 1907), han venido siendo reiteradamente esgrimidos en distintos medios, aunque con cierta ambigüedad. Dice así la cláusula:

Hasta que un Código más completo de las Leyes de guerra se emita, las Altas Partes Contratantes juzgan oportuno declarar que, en los casos no incluidos en las disposiciones reglamentarias adoptadas por ellas, las poblaciones y los beligerantes quedan bajo la protección y el imperio de los principios del derecho internacional, tal como resultan de los usos establecidos entre naciones civilizadas, de las leyes de la humanidad y las exigencias de la conciencia pública.

Es decir, la cláusula requiere 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. Esta referencia a las exigencias de la «conciencia pública» ha dado a entender a algunos que, al no existir un consenso social amplio sobre sus usos y aplicaciones, este tipo de armas debe ser prohibido. Pero hay dos problemas en este sentido. En primer lugar, no existe una única interpretación aceptada de la cláusula,

y aunque varios tribunales nacionales e internacionales han tomado en consideración la cláusula Martens a la hora de emitir sus juicios,<sup>148</sup> en ninguno de tales casos se han reconocido las *leyes de la humanidad* o los *dictados de la conciencia* pública como derechos nuevos e independientes. En consecuencia, podemos afirmar que la cláusula no es más que una declaración general de principios humanitarios, así como una suerte de guía para la comprensión e interpretación de las normas del derecho internacional. Es decir, algo así como lo que Paine expone en su obra clásica de 1776 *Common sense*, que recoge los principios de la Revolución norteamericana y de su derecho, condensados en la idea de *sentido común*.<sup>149</sup>

Por otro lado, tal y como afirma Michael Horowitz, incluso si se aceptara una lectura restrictiva de la cláusula Martens, la

<sup>148</sup> He aquí varios ejemplos: decisión del 27 de febrero de 1946 de la Corte Suprema de Noruega, en recurso de casación contra Karl-Hans Hermann Klinge, *Kriminalassistent* (asistente penal) de la Gestapo (confirmación de la sentencia de muerte impuesta en primera instancia). Decisión del Tribunal Militar III-A de Estados Unidos en Núremberg el 10 de febrero del año 1948 en el caso Estados Unidos contra Krupp. Decisión del Tribunal de Casación de los Países Bajos el 12 de enero de 1949 en el procedimiento contra el *SS-Obergruppenführer* (general) Hanns Albin Rauter, comisario general para la organización de la seguridad en los Países Bajos desde 1940 hasta 1945. Decisión de los tribunales militares de Bruselas (*Conseil de Guerre de Bruxelles*) en el caso K. W. el 8 de febrero de 1950. Decisión del Tribunal Penal Internacional para la ex Yugoslavia del 8 de marzo de 1996 sobre el permiso de la acusación durante el proceso en contra de Milan Martić (caso IT-95 11, decisión IT-95-11-R61). Decisión de la Corte Constitucional de Colombia del 18 de mayo de 1995 para la constitucionalidad del Segundo Protocolo Adicional a los Convenios de Ginebra del 12 de agosto de 1949, relativo a la protección de las víctimas de los conflictos armados internacionales (decisión C-225/95). Decisión de la Corte Internacional de Justicia sobre la opinión consultiva sobre la Legalidad de la Amenaza o el Empleo de Armas Nucleares del 8 de julio de 1996. Sentencia del Tribunal Constitucional de Alemania del 26 de octubre de 2004 para la compatibilidad de las expropiaciones en la antigua zona de ocupación soviética entre 1945 y 1949 con el derecho internacional (decisión BverfG, 2 BvR 955/00).

<sup>149</sup> Paine (2004).

noción de *conciencia pública* sigue siendo excesivamente ambigua y, lo que es más importante, excesivamente maleable.<sup>150</sup>

Otro aspecto legal que no puede ser pasado por alto, y que a ojos de estos autores se presenta como más sólidamente restrictivo con respecto al uso de este tipo de armas, se encuentra en el derecho humanitario o *jus in bello* surgido de los Tribunales de Núremberg, que dio cabida a la responsabilidad penal del sujeto en casos de crímenes de guerra y lesa humanidad. Así, en caso de víctimas civiles, alguna persona debe ser susceptible de ser declarada responsable. Se trata de un principio claro y contundente que sin embargo queda subvertido con el advenimiento de los sistemas letales autónomos, obligando a una reformulación profunda.

Sea como sea, hay que tener muy claro que los AWS pueden llegar a ser armas de destrucción masiva, y que por lo tanto su prohibición absoluta debe ser una posibilidad, igual que debe serlo que el solo hecho de investigar sobre los AWS pueda interpretarse como delito de conspiración para el genocidio con arreglo a las resoluciones del Tribunal Penal Internacional y de su tratado. Resulta interesante en este sentido leer a J. Diamond, que en su exitosa obra *Gérmenes, armas y acero: breve historia de la humanidad en los últimos trece mil años*, así como también en *Colapso: por qué unas sociedades perduran y otras desaparecen*, ilustra magníficamente al lector acerca de cómo la tecnología puede acabar con una sociedad —hacerla colapsar—, como ocurrió con la civilización maya o la de la isla de Pascua. Si pensamos que los AWS pueden ser incluso invisibles al ojo humano (*nano-AWS*), y por lo tanto quedar más fácilmente fuera del control de los gobiernos legítimos, las perspectivas se vuelven más y más sombrías.

<sup>150</sup> Horowitz (2016).

Por todo lo expuesto, entendemos que la inteligencia artificial, en el caso específico de los sistemas de armamento, nos obliga a una revisión de los imperativos éticos básicos vinculados al derecho en un momento en el que el principio categórico se encuentra asediado por el devenir tecnológico. Solo tal revisión puede deshacer el nudo gordiano de garantizar nuestra supervivencia no ya como ciudadanías o Estados, sino como especie, en este *mundo feliz*. La humanidad tiene el deber de protegerse a sí misma a través de la dignificación de la vida y de la garantía de los derechos y libertades individuales más allá de cualquier orden jurídico.<sup>151</sup> Ha de generarse un derecho que, espoleado por las reivindicaciones pacifistas y la jurisprudencia fruto del activismo judicial para prevenir las amenazas de la tecnología, se oriente a los cambios experimentados por la condición humana en su tránsito hacia la posthumana y posea capacidad prospectiva.<sup>152</sup> Es decir, un derecho marcado por las nuevas tendencias en gobernanza anticipatoria.<sup>153</sup>

Cuando se habla de *ética de las decisiones*, hay que citar al experto en comportamiento S. Bowles, que advierte de que los poderes públicos no pueden ignorar el lado moral y generoso de la naturaleza humana, las motivaciones cívicas que nos mueven para conseguir el bien común y un buen gobierno.<sup>154</sup> Los AWS carecen de tales características, puesto que se diseñan precisamente para lo contrario: obviar inmisericordemente la dimensión cooperativa y de reciprocidad del comportamiento humano. Por más que se quiera avalar esta

<sup>151</sup> Montobbio (2008). Según este autor, «en la era de la globalización de la sociedad de la información nos encontramos ya navegando en la nave espacial Tierra destino futuro; estamos todos en el mismo barco: nosotros somos, querámoslo o no, todos».

<sup>152</sup> Braidotti (2015).

<sup>153</sup> Arendt (2015).

<sup>154</sup> Bowles (2016).



nueva forma de armamento apelando a sus buenas intenciones y a sus propósitos humanitarios, no puede sino responderse a ello con el famoso principio según el cual la Constitución nos protege de nuestras mejores intenciones.<sup>155</sup> Para todo debe haber controles y un sistema de *check and balance*.

<sup>155</sup> Ballbé y Martínez (2003).



# El derecho ante la inteligencia artificial en los LAWS

ROSER MARTÍNEZ

## Introducción

La regulación de la inteligencia artificial cuando se aplica a armas, y más concretamente en los LAWS, es una cuestión que no admite demora. Urge el desarrollo de propuestas para un sistema de control de los mismos dentro del derecho del control de armas clásico,<sup>156</sup> así como una advertencia sobre los riesgos inquietantes y desafíos que esta tecnología representa tanto para el derecho como para el futuro de la seguridad humana.<sup>157</sup> Para ello, estableceremos un marco conceptual que permita discernir entre el estado de la cuestión actual y las vías de materialización futura. Procuraremos, sobre la base de lo expuesto en los capítulos previos, ofrecer al lector una recopilación de elementos que permitan el establecimiento de un corpus legal armonizado que limite la investigación y el desarrollo de armas robóticas letales independientes debido a los riesgos asociados a una carrera armamentística en este ámbito e impida continuar con la tentación unilateralista en lugar de someterse a acuerdos multilaterales honestos.<sup>158</sup>

Hoy reclaman una prohibición preventiva internacional y vinculante del desarrollo, producción, adquisición y des-

<sup>156</sup> Martínez (2003).

<sup>157</sup> Concepto acuñado por Naciones Unidas. Véase Ballbé y Martínez (2009).

<sup>158</sup> Para entender los diferentes tipos de armonización, véase Ballbé y Padrós (1997). Cf. también Jha (2016).

pliegue de sistemas autónomos de armas organizaciones internacionales, miles de científicos y expertos en inteligencia artificial, más de veinte premios Nobel y, hasta el momento, 26 países. Sin embargo, Francia, Israel, Rusia, el Reino Unido y Estados Unidos se han negado expresamente a negociar un tratado sobre armas completamente autónomas. En lugar de ello, esos países están invirtiendo en sistemas armados con un control humano cada vez menor. Además, en mayo de 2018, la financiación de los LAWS por el Fondo Europeo de Defensa se declaró posible a nivel de la Unión Europea. Alemania y Francia, concretamente, han decidido crear un insituto de investigación conjunta para la inteligencia artificial. Su objetivo es crear programas de robótica y sistemas autónomos aparentemente alejados de usos militares, pero la conexión es evidente. Alemania ya ha hecho público que tiene la pretensión de desarrollar plataformas de aprendizaje automático y ser líder en innovación en este campo (solo el estado de Baviera, sede de muchas empresas de armamento, tiene un presupuesto de 280 millones de euros para inteligencia artificial).<sup>159</sup>

El reto, ante ese panorama, no es regular solo las armas autónomas letales o no letales, sino hacer objeto de control y regulación embrionaria (desde las fases de investigación y experimentación) toda una dimensión más amplia de dichas armas que ya se adivina en el horizonte inmediato: la inteligencia artificial que, aplicada a estos sistemas, los puede convertir en independientes, de tal manera que acometan acciones letales de manera completamente autónoma.<sup>160</sup>

<sup>159</sup> Facing Finance (2018).

<sup>160</sup> En las ferias de París y Abu Dabi de 2018 se acaban de presentar sistemas armados completamente autónomos utilizando munición de merodeo que podrían encontrar y destruir objetivos de forma independiente aunque aún sin inteligencia artificial general. Facing Finance (2018).

Hollywood nos anticipó escenarios de ficción que hoy son más reales que nunca: recordemos la mítica *2001: una odisea del espacio*, dirigida por Stanley Kubrick en 1968 y que ponía en escena a la potente computadora HAL 9000, dotada de inteligencia artificial.<sup>161</sup> También los padres de la informática y de la inteligencia artificial vaticinaron el poder insospechado de las máquinas. Así, por ejemplo, el matemático húngaro John von Neumann señalaba en 1946 que «lo que estamos creando en este momento es un monstruo cuya influencia va a cambiar la historia, si es que queda historia alguna [...] pero sería imposible no llevarlo a cabo, no solo por razones militares sino porque también sería poco ético desde el punto de vista de los científicos no hacer lo que ellos saben que es factible, por muy temibles que puedan ser las consecuencias».<sup>162</sup> Esas consecuencias terribles que Von Neumann anticipaba están hoy a la vuelta de la esquina, resultado de la unión de inteligencia artificial, *big data* y armas; y debemos conseguir que el derecho sea un freno a la irresponsabilidad de colocar la tecnología al servicio de unos pocos.<sup>163</sup>

La comunidad internacional ya es consciente del problema y ha emprendido un debate conducente a un abordaje preventivo de la cuestión sobre la base del principio jurídico

<sup>161</sup> Sánchez Barrilao (2016), p. 228. El autor hace un paralelismo con sistemas de inteligencia artificial de la ficción cinematográfica como el Skynet de *Terminator*, o Matrix, o el Viki de *Yo, robot* y se refiere al androide Ultrón de los Marvel Cómics para advertir de que un sujeto robótico inteligente e independiente es generador de un riesgo autónomo para los humanos y no controlable por ellos, con lo que «el riesgo tecnológico únicamente puede ser salvado por el progreso tecnológico». Según Sánchez, «en los cómics de Marvel es donde mejor se advierte el desarrollo de cómo Ultrón genera, a la vez que es, inteligencia artificial en progreso, al autodiseñarse y actualizarse en distintas versiones de sí mismo» (p. 229).

<sup>162</sup> Dyson (2012).

<sup>163</sup> Jonas (2004).

de precaución,<sup>164</sup> proporcionalidad<sup>165</sup> y distinción.<sup>166</sup> En esta línea, Naciones Unidas ha creado el denominado Grupo de Expertos Gubernamentales de las Altas Partes Contratantes sobre Sistemas de Armas Autónomos Letales para tratar de regular los LAWS, y los autores de este libro son integrantes del mismo por ser miembros del ICRC (Comité Internacional para el Control de los Robots Armados; *International Committee for Robots Arms Control* en inglés).

Desde 2014, Naciones Unidas ha conseguido que los países se reúnan y debatan para lograr un consenso y especificar los límites de los LAWS en el marco de la Convención sobre Armas Convencionales, aunque aún no se ha conseguido aprobar un texto vinculante. Básicamente deberán establecer si en las operaciones armadas la decisión última sobre la vida o la muerte de la población debe permanecer en un humano o en una máquina. A menos que se acuerde una prohibición, los sistemas de armas sin control humano se convertirán en el equipamiento estándar de los ejércitos gracias a las numerosas tecnologías ya disponibles (sensores) y a la inteligencia artificial avanzada.

## Los LAWS como amenaza a los derechos a la dignidad humana y a la vida

Según Lin, en los conflictos armados, *derecho a la vida* significa derecho a no ser asesinado de forma arbitraria, caprichosa, inexplicable, inhumana o como daño colateral;<sup>167</sup> y es

¿2005a o 2005b?

<sup>164</sup> Véase Sunstein (2005), Vogel (2015) y Pardo (2003).

<sup>165</sup> Jackson (2017).

<sup>166</sup> Sassoli (2014).

<sup>167</sup> Según la terminología de la ley americana de procedimiento administrativo de 1946 (Pub.L.79-404 APA).

en realidad, de alguna forma, un 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.<sup>168</sup> Por otro lado, existe un consenso cada vez mayor en que antes aún que estos derechos está el derecho individual y colectivo de acceso a la información; el derecho a conocer y, como señala Rosemberg, a preguntar.<sup>169</sup> En este marco, la legislación administrativa (en los Estados Unidos: FOIA,<sup>170</sup> *Sunshine Act*,<sup>171</sup> etcétera) se revela como la herramienta preventiva más eficaz para hacer frente a los riesgos y peligros que representan determinados productos o artefactos con inteligencia artificial, y especialmente las armas letales, que carecen de la capacidad de reflexión y respeto morales. Lin señala que no es absurdo que la dignidad y la libertad puedan prevalecer sobre la seguridad.<sup>172</sup>

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

<sup>168</sup> Lin (2015, 2017).

<sup>169</sup> Rosemberg (2006).

<sup>170</sup> Siglas de *Freedom of Information Act*, «ley por la Libertad de Información» en inglés, una ley promulgada en 1966 y firmada por el entonces presidente Lyndon Johnson que otorga a todos los ciudadanos de Estados Unidos el derecho de acceso a la información del Gobierno federal.

<sup>171</sup> Ley de 2010 que persigue incrementar la transparencia de las relaciones financieras entre los profesionales de la salud y la industria farmacéutica.

<sup>172</sup> Lin (2015, 2017).

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».<sup>173</sup> Los LAWS 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 los AWS 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,<sup>174</sup> así como los conceptos relacionados con el control humano significativo y la autodeterminación de los AWS.<sup>175</sup> Es también necesario profundizar en nuevas formas de convivencia teniendo en

<sup>173</sup> Lin (2015, 2017).

<sup>174</sup> 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.

<sup>175</sup> Lin (2015, 2017).



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,<sup>176</sup> 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 combate con AWS. 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.<sup>177</sup> 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.<sup>178</sup>

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 (el Pentágono ha valorado el coste de cada soldado desplegado en Afganistán durante un año en 850 000 dólares, mientras que el de un robot tipo TALON es de 230 000), 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 ma-

<sup>176</sup> Axelrod (1984).

<sup>177</sup> Sparrow (2016).

<sup>178</sup> Martínez (2003).

nera inmoral; no padecerán prejuicios cognitivos que afectan al ser humano [...] e incluso pueden distinguir legítima y fiablenente los blancos legítimos de los ilegítimos». <sup>179</sup>

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, <sup>180</sup> la negociación, <sup>181</sup> 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. <sup>182</sup> 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. <sup>183</sup>

Existen incluso factores de tipo neurológico que fomentan una seguridad más humana y cooperativa; y hay cimientos éticos fuertemente arraigados en nuestra psique, como la noción de responsabilidad, que juegan asimismo un rol determinante, y que podría y debería ser en un futuro parte integrante de los algoritmos en que se basen las inteligencias artificiales. <sup>184</sup> Se ha señalado que «los estudios de las experiencias de

<sup>179</sup> Cit. en Sparrow (2016).

<sup>180</sup> Fischer (2008).

<sup>181</sup> Uri (2005).

<sup>182</sup> Rodríguez-Arana (1993).

<sup>183</sup> Axelrod (1984).

<sup>184</sup> O'Neil (2016). La autora alerta de que, bajo su promesa de eficacia y justicia, los algoritmos y los métodos de análisis big data distorsionan la educación, aumentan la deuda, incitan a las autoridades a criminalizar a un determinado grupo social, golpean a los pobres en casi todas las situaciones y socavan la democracia. Cf. también Pazzanese (2016).

los soldados apoyan que los seres humanos son naturalmente reacios a quitar la vida, y esta aversión puede manifestarse en momentos de compasión y humanidad en medio de los horrores de la guerra. Programar una inteligencia artificial para permitir que los sistemas de armas autónomas cumplan técnicamente con la ley de guerra en situaciones en que se debe discriminar con intuición y proporcionalidad, incluso si fuera posible, no es suficiente».<sup>185</sup>

Tampoco se debe pasar por alto que la naturaleza pública del conflicto configurada por los *mass media* tiende a fijar la mirada solo sobre las bajas propias,<sup>186</sup> y que en ese marco las armas autónomas pueden generar y de hecho generan nuevos niveles de opacidad y una mayor libertad para que los gobiernos actúen al margen de su población en asuntos militares.

Finalmente, entre los problemas asociados a la militarización de la inteligencia artificial también se halla el propio devenir de la tecnología, que puede quedar profundamente afectado por unos usos que vayan en contra del criterio de la opinión pública, de tal manera que quede comprometida la totalidad de la tecnología, tal y como sucedió con la nuclear o la química.<sup>187</sup> Del mismo modo, una relajación de la intervención sobre dicha tecnología puede llevar a su propio fin y el de la humanidad misma. Los efectos mortíferos del agente naranja empleado en Vietnam alcanzaron incluso a los descendientes de los propios militares implicados, por lo que los militares son los primeros interesados en frenar el desarrollo de determinadas armas que pueden acabar con ellos mismos.

Otro argumento sólido a favor de la prohibición de los LAWS es que una vez activados podrían seleccionar objetivos y acabar con la vida de las personas a su libre arbitrio, de for-

<sup>185</sup> Amoroso (2017).

<sup>186</sup> Herman y Chomsky (2010).

<sup>187</sup> Morales (2009).

ma irrevocable y sin intervención humana, lo que supondría el otorgamiento de una facultad administrativa contraria al ordenamiento jurídico 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...<sup>188</sup> En consecuencia, las acciones de los LAWS 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.<sup>189</sup>

Para Amoroso, «la idea de una máquina dotada de poder para tomar decisiones de vida o muerte es intuitivamente repugnante» y «extremadamente irrespetuosa con la humanidad de nuestro enemigo».<sup>190</sup> Y en esa línea creemos que, aunque los AWS pudieran llegar a ofrecer resultados mejores en base a un cálculo de coste-beneficio o llegaran a conseguir comportamientos humanos, deben prohibirse por razones éticas y legales.<sup>191</sup> 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 especial-

<sup>188</sup> Lieblich (2016).

<sup>189</sup> Asaro (2012); Sharkey (2017).

<sup>190</sup> Amoroso (2017) y Sparrow (2016).

<sup>191</sup> Heyns (2016).

mente 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.

El derecho, por otro lado, nunca tiene en cuenta un elemento cognitivo fundamental, la intuición humana, cuando regula las facultades públicas discrecionales en la toma de decisiones, quizás porque da por hecho que son los seres humanos quienes las llevan a cabo. Pero los LAWS 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.

La intuición ha venido siendo descrita de diversas maneras. Una de esas definiciones es la capacidad de discernir cuándo existe un problema u oportunidad y seleccionar la acción más apropiada sin razonamiento consciente; poniendo en funcionamiento patrones profundamente arraigados de experiencia adquirida y destilada.<sup>192</sup> También ha sido considerada como la capacidad de reconocer rápida y fácilmente las posibilidades de una determinada situación.<sup>193</sup> También como un conjunto de juicios emocionalmente cargados que emergen a través de una forma rápida, no consciente y holística de asociación.<sup>194</sup> Y también como la confianza en los modelos mentales: representaciones internas del entorno

<sup>192</sup> Behling y Eckel (1991), Khatri y Ng (2000), Gilovich, Griffin y Kahneman (2002).

<sup>193</sup> Agor (1989).

<sup>194</sup> Dane y Pratt (2007).

que nos permiten anticipar acontecimientos futuros a partir de las observaciones actuales.<sup>195</sup>

Todas estas definiciones comparten una serie de asunciones comunes. En primer lugar, que la intuición es rápida; en segundo, que es un análisis automático e inconsciente de un proceso; en tercero, que se basa en la experiencia e involucra emociones humanas; y en cuarto, que ofrece potencial para la creatividad y la innovación.<sup>196</sup> Kahneman, especialista en intuición y toma de decisiones bajo incertidumbre, defiende que la intuición es fruto de la experiencia humana y que el cerebro humano, al formular un juicio o tomar la decisión de, por ejemplo, matar, emplea dos sistemas combinados: por una parte, un pensamiento rápido, intuitivo y emocional; por otro, uno más lento que es deliberativo y aplica la lógica. Según este autor, el ser humano no siempre debe creer en su intuición, porque está basada en su experiencia y no en el sistema lento de pensamiento.<sup>197</sup> Por su parte, Klein defiende que la intuición no es un sexto sentido innato, sino una visión propia de cada persona y una habilidad esencial que se puede aprender.<sup>198</sup>

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 guerra. Y los LAWS 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.<sup>199</sup> Como dice el neurocientífico G. Rizzolatti, descubridor de las neuronas espejo, «los

<sup>195</sup> Kahneman y Klein (2009).

<sup>196</sup> Lunenburg (2010).

<sup>197</sup> Kahneman (2015).

<sup>198</sup> Klein (2004).

<sup>199</sup> Howard, Zhang y Horvitz (2017).

robots pueden imitar, no sentir».<sup>200</sup> Y si es así, si los algoritmos incluidos en los LAWS 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 humanas.

Un problema adicional en los LAWS es que su responsabilidad es difícil de exigir desde un punto de vista tradicional del derecho, y en caso de violaciones de derechos humanos, los recursos contra dichas acciones podrían no ser efectivos.<sup>201</sup> En este sentido, es interesante el informe de Human Rights Watch titulado «Mind the gap: the lack of accountability for killer robots». En él se afirma que

Los obstáculos para la rendición de cuentas por la producción y el uso de armas totalmente autónomas bajo la ley actual son monumentales. Las propias armas no podrían ser consideradas responsables de su conducta porque no podrían actuar con intención delictiva, estarían fuera de la jurisdicción de los tribunales internacionales y no podrían ser castigadas. La responsabilidad penal probablemente se aplique solo en situaciones donde los humanos intentaron específicamente utilizar los robots para violar la ley. Al menos en Estados Unidos, la responsabilidad civil sería virtualmente imposible debido a la inmunidad otorgada por la ley al Ejército y sus contratistas y los obstáculos probatorios a las demandas por responsabilidad civil por productos defectuosos.<sup>202</sup>

<sup>200</sup> Rodella (2018).

<sup>201</sup> Markoff (2016).

<sup>202</sup> Human Rights Watch (2015).

También el embajador de España ante Naciones Unidas, Julio Herraiz, ha mostrado su preocupación ante estas cuestiones. Así, en la Conferencia de Desarme de la Convención sobre Ciertas Armas Convencionales, de Naciones Unidas, celebrada en Ginebra el 13 de noviembre de 2017, decía esto: «España entiende que en el empleo de sistemas con un cierto grado de autonomía y capaces de proyectar fuerza letal, siempre debe existir la intervención de un operador humano. Asimismo, debe considerarse la inclusión en estos sistemas de elementos técnicos que puedan facilitar la atribución de responsabilidad jurídica. La responsabilidad debe recaer en el operador, así como en la persona que pueda ordenar el uso del arma contra derecho». Se ha dicho también que «la delegación del uso de la fuerza a los tomadores de decisiones no humanos crearía una laguna de responsabilidad».<sup>203</sup>

Teniendo en cuenta que un robot de este tipo podría identificar un objetivo y lanzar un ataque a su libre arbitrio, la excusa inevitable de los agentes implicados será que ellos no son los responsables de lo que haga el artefacto a partir de su puesta en funcionamiento. Y podría señalarse como responsables de estas cuestiones a la jerarquía política y militar del Estado que ha financiado el proyecto (sean empresas públicas o privadas) y que ha puesto en acción el LAWS, las personas que han investigado y activado el robot, el programador que ha creado los algoritmos de actuación, el fabricante que lo ha puesto en el mercado, etcétera;<sup>204</sup> pero se ha llegado a señalar que ninguno de ellos sería completamente responsable, puesto que la decisión propiamente dicha no correspondería sino al sistema autónomo letal.

<sup>203</sup> Amoroso y Tamburrini (2017).

<sup>204</sup> Drohan (2003).



Deben serlo. Los agentes implicados deben estar sujetos a una suerte de *obligación de no hacer* ante dicha tecnología letal con inteligencia artificial. La eximente de la obediencia debida no puede aplicarse —ni siquiera en Estados autoritarios— al personal que sabe, o que debería saber, que está experimentando, creando o transfiriendo un sistema letal completamente autónomo que puede convertirse en el más dramático enemigo de la humanidad. Y en todo caso, debe al menos quedar clara la responsabilidad de la Administración competente si accede a fomentar o a disponer de dicha tecnología.<sup>205</sup>

Ying Hu, profesor de derecho de Yale, reflexiona sobre que «imponer responsabilidad penal a los robots (y a sus creadores) en ocasiones podría tener un valor instrumental significativo, como ayudar a identificar a las personas culpables y servir como un dispositivo de autocontrol para las personas que interactúan con ellos».<sup>206</sup> En esta línea, Sánchez del Campo también apunta a que los robots pueden tener obligaciones y algún tipo de responsabilidad legal, e incluso que podría predicarse de ellos que como tienen cierta *personalidad* y autonomía, podrían llegar a ser considerados sujetos que cometen delitos. Sin embargo, Quintero Olivares, que recoge y analiza profusamente esta discusión en un artículo sugerente y de indispensable lectura, rechaza contundentemente la responsabilidad penal de los robots dejando claro que en ningún caso puede aceptarse la idea de la autorresponsabilidad del robot. Pese a todo, matiza que ello «no equivale a la *irrelevancia* de lo que *haga* una máquina».<sup>207</sup>

Sea como sea, la mera investigación y desarrollo de este tipo de tecnología por parte de cualquier persona u organi-

<sup>205</sup> Human Rights Watch (2015).

<sup>206</sup> Hu (2018).

<sup>207</sup> Quintero (2017), pp. 10 y 14.

zación debería ser tipificada como indicio claro de conspiración para el delito. En este sentido, deberían promulgarse a nivel nacional e internacional los tipos penales necesarios para que la experimentación y creación de LAWS con inteligencia artificial fuerte para fines no defensivos constituyan «delitos de conspiración para el genocidio o para el asesinato selectivo a través de seres sintéticos con independencia». Cualquier democracia que no proceda así estará incumpliendo flagrantemente su propia Constitución.

Es interesante constatar que dentro del *Common Law* constitucional norteamericano ha habido desde el primer momento delitos de *conspiración* para el monopolio y las acciones colusorias. Todas las Constituciones de los trece estados que conformaron el primer Estados Unidos establecieron de forma contundente que los monopolios eran contrarios al espíritu libre y no debían ser tolerados; prohibición que se incardina dentro del capítulo de los derechos fundamentales de las personas. Sobre esa base se fue desarrollando una legislación tanto federal como nacional (Sherman Act de 1890, Clayton Act de 1913...) que penalizaba delitos como alterar los precios o repartirse el mercado. Si se hizo así con estas cuestiones, más aún debería penalizarse la conspiración para otorgar poderes y subvencionar a empresas privadas para investigar sistemas con inteligencia artificial, pues ello podría alterar el equilibrio no ya del mercado, sino del Estado mismo. Y debería ser posible acuñar este tipo de legislación para todo el mundo en la medida en que hoy existe un *Common Law* internacional que, desarrollado desde Núremberg hasta Yugoslavia (aunque discutido también en el Tratado de Roma del Tribunal Penal Internacional), señala que la obediencia debida no sirve de eximente para cometer semejantes acciones: la sentencia del Tribunal Militar Internacional de Núremberg contra los criminales de guerra alemanes del 1

de octubre de 1946 señala claramente que los crímenes «son cometidos por hombres, no por entidades abstractas, y solo mediante el castigo de las personas que cometen tales crímenes se pueden hacer cumplir las disposiciones del derecho internacional». En nuestro caso, debe examinarse minuciosamente a los agentes implicados en todo el proceso de creación de LAWS completamente autónomos, y que cada cual asuma su responsabilidad en la confabulación para crear estas máquinas de destrucción masiva o individual.

## El derecho frente al Estado de excepción algorítmico

George Orwell planteó ya en su famosa obra *1984* la amenaza terrible que representaba la vulneración de la privacidad y los derechos del individuo, pero su denuncia parece ingenua comparada con algunos casos impactantes que hemos ido conociendo en los últimos años y que la están materializando de forma espeluznante, como el de Snowden y la NSA o el de Facebook-Cambridge Analytica.<sup>208</sup> 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».<sup>209</sup> Por su parte, Rosebuj nos recuerda que otros autores ya lo habían calificado como *capitalismo cognitivo* (Boutang)

<sup>208</sup> Wylie, exdirector de investigación de Cambridge Analytica, filtró información sobre el secuestro de datos privados de 50 millones de usuarios de Facebook para influir en la campaña a favor de Trump. Cf. «Cambridge Analytica compartió datos con Rusia, según el cerebro de la firma», *El País*, 17 de mayo de 2018 [en línea], <[https://elpais.com/internacional/2018/05/17/estados\\_unidos/1526514308\\_942521.html](https://elpais.com/internacional/2018/05/17/estados_unidos/1526514308_942521.html)>.

<sup>209</sup> Thompson (2012), Schönberger (2009), McQuillan (2015).

o como *capitalismo de vigilancia* (Zuboff). La vigilancia, el acceso y el control se convierten en el núcleo del sistema mediante la recopilación, extracción, almacenamiento y análisis por *big data*.<sup>210</sup>

Hasta ahora entendíamos *estado de excepción* como la implantación del militarismo o del estado policial como una nueva Inquisición,<sup>211</sup> pero la tenebrosa actualidad ha dejado corta esa preocupación. Incluso para la *intelligentsia* humana es difícilmente imaginable el grado de *estado policial público-privado de inteligencia artificial* real que estamos padeciendo. Como señaló Montesquieu, «no hay peor tiranía que la que se perpetúa bajo el escudo de las leyes y en nombre de la justicia». Y hoy ese escudo protege a monopolios globales como Facebook o Google, que poseen y manejan la información más privada de dos mil millones de ciudadanos.<sup>212</sup> 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*».<sup>213</sup> Podrá ser 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 inter-

<sup>210</sup> Rosembuj (2017).

<sup>211</sup> Ballbé (1985), Agamben (2004). Este autor desmonta cualquier intento de legitimación jurídica del estado de excepción y nos ayuda a apreciar la relación entre violencia, derecho y política.

<sup>212</sup> López-Tarruella (2012).

<sup>213</sup> Kang y Roose (2018).

nacionales legalmente vinculantes o incluso leyes nacionales que prohíban el desarrollo, producción y uso de los llamados *robots asesinos*.<sup>214</sup>

El Premio Nobel de Economía D. North explica que la explosión científica y tecnológica que en el siglo XVIII dio lugar en Inglaterra a la revolución industrial fue la regulación: se promulgó una ley de propiedad intelectual o de patentes y se gestó un incipiente Estado administrativo regulador en esta materia. Antes, el mercado autorregulado no estimulaba la investigación porque se plagiaba y no había incentivos.<sup>215</sup> En Estados Unidos, la primera agencia administrativa fue la Oficina de Patentes, creada en 1787.<sup>216</sup> Incluso quienes no están de acuerdo con que Estados Unidos era ya un Estado administrativo han de aceptar que esos tribunales resolvían los conflictos y que existía una clara intervención del Estado judicial.

Este proceso de administrativización tiene un paralelismo con el mercado anómico y autorregulado en materia de LAWS con inteligencia artificial. Si en aquellos momentos se crearon una administración de *law enforcement*, unos funcionarios especializados y unos tribunales especializados en los litigios sobre propiedad intelectual y patentes tanto en Inglaterra como en Estados Unidos, ahora conviene una regulación administrativa con respecto a la inteligencia artificial. Los fallos del mercado hicieron que el derecho interviniera, y esa necesidad vuelve a ser acuciante. La única garantía de progreso y sostenimiento de los derechos ciudadanos ante la inteligencia artificial en LAWS es la regulación.

Un grupo de expertos británicos de la Universidad de Oxford, la Electronic Frontier Foundation, elaboró en febrero de 2018

<sup>214</sup> Rosembuj (2017).

<sup>215</sup> North (1981).

<sup>216</sup> Marshaw (2006).

un importante informe titulado *The malicious use of artificial intelligence: forecasting, prevention and mitigation*. En él se alerta acerca de las amenazas de la inteligencia artificial y su más que posible transmutación en tecnología de *doble uso*, es decir, su traslación del ámbito civil al militar.<sup>217</sup> Para mayor preocupación, esta tecnología se investiga y se desarrolla en una colaboración público-privada, pero por entidades cuyo único interés es el mercantil y que no están sujetas al control ni de la Administración ni del ordenamiento jurídico por tratarse de procesos que afectan a la seguridad nacional<sup>218</sup> y al espionaje.<sup>219</sup> Además, un peligro aparejado que entraña el traspaso de la soberanía tecnológica crítica a otros Estados no democráticos es la corrupción soterrada que lleva aparejada. Afortunadamente, el Bundesnachrichtendienst (BND), es decir, la Agencia Federal de Inteligencia alemana, denunció la venta de la principal empresa de inteligencia artificial a China, y con ello alertó de que la tecnología más puntera en seguridad podía acabar en manos del enemigo por intereses privados. El gobierno de Angela Merkel se ha apresurado por ello a preparar un proyecto de ley para reforzar el veto en las adquisiciones extranjeras de participaciones en empresas consideradas de seguridad nacional.<sup>220</sup>

<sup>217</sup> GAO (2018).

<sup>218</sup> Un freno parcial a esta inercia pueden ser organismos de seguridad como, en Estados Unidos, CIFIUS (*Committee on Foreign Investment in the United States*). Se trata de una agencia interdepartamental responsable de advertir al presidente sobre las inversiones extranjeras que puedan sustraer soberanía y que afecten negativamente a la seguridad nacional. Véase Rubio (2017).

<sup>219</sup> Hayden (2018).

<sup>220</sup> En concreto, hasta ahora, Berlín podía vetar acuerdos que implicaran la compra de al menos el 25 % del capital social de una empresa alemana por parte de una de fuera de la Unión Europea si se consideraba que la operación ponía en peligro la seguridad nacional. Con la nueva ley, el umbral se reducirá al 15 % para poder verificar más adquisiciones en sectores sensibles. Cf. «Alemania blindo sus empresas ante China», *El Boletín*, 8 de agosto de 2018

Cabe resaltar, por lo demás, que hoy en día comienzan a observarse patrones de autorregulación o autocensura en centros de investigación que ven como imperiosa la necesidad de limitar su trabajo debido a los riesgos potenciales que este representa. Por ejemplo, la empresa OpenAI (fundada en 2015 por Musk y Altman con el objetivo de asegurar que el desarrollo de la inteligencia artificial tenga un impacto positivo en la humanidad) ha restringido el lanzamiento de una neurona de sentimiento no supervisada, un algoritmo que se había entrenado para comprender sentimientos a través de la lectura de reseñas en Amazon. El grupo decidió mantener el modelo anterior de lenguaje que había desarrollado a fin de evitar el mal uso del algoritmo y las perversiones que percibía que se generaban.<sup>221</sup> Merece destacarse también el boicot que está recibiendo la universidad surcoreana KAIST por su decisión de abrir un laboratorio de armas con inteligencia artificial; laboratorio que, en colaboración con el fabricante de armas Hanwha Systems, puede desarrollar los llamados *robots asesinos*. Proyectan cuatro líneas con inteligencia artificial: sistemas de comando y decisión, algoritmos de navegación para vehículos submarinos no tripulados, sistemas de entrenamiento de aeronaves inteligentes y tecnología de reconocimiento y seguimiento de objetos inteligentes.<sup>222</sup>

En suma, los propios sectores implicados deberían autorregularse obligados por una normativa que impusiera un *compliance* preventivo a fin de precaverse de potenciales aplicaciones devastadoras de sus descubrimientos. El camino a seguir es el que en su momento marcó la industria nuclear, que había capturado y corrompido a la Comisión Re-

---

[en línea], <<https://www.elboletin.com/noticia/166171/internacional/alemania-blinda-sus-empresas-ante-china.html>>. [Consulta: 7-9-2018].

<sup>221</sup> Waters (2018).

<sup>222</sup> Smyth y Harris (2018).

guladora Nuclear estadounidense de la época pero que tras el accidente de Three Mile Island se dio cuenta de que debía alcanzar niveles más altos de seguridad y fiabilidad y creó el INPO (*Institute of Nuclear Power Operations*); y en 1986, tras el de Chernóbil, acordó que o se autorregulaba de forma estricta e implementaba sistemas de vigilancia o la negligencia de unos pocos podría llevar a la decadencia y el fin de este sector energético. Se creó entonces la Asociación Mundial de Operadores Nucleares a fin de promover la cooperación y la excelencia entre sus miembros y se nombró a un almirante, James Ellis, para presidir el INPO y disciplinar férreamente a este sector desregulado *de facto*. Ha de hacerse notar que de todas formas, y pese a esa visión lúcida, el sector no logró autodisciplinarse satisfactoriamente, tal y como dejó sentado el desastre de Fukushima en 2011.<sup>223</sup>

Otro ejemplo de un principio esperanzador de regulación lo ofrece el Proyecto Marven, un contrato firmado entre Google y el Pentágono por el cual la primera desarrolló algoritmos de aprendizaje automático con el objetivo de alumbrar un sistema sofisticado que permitiera al segundo vigilar ciudades enteras. Para ello, necesitaba imágenes recopiladas por drones militares, y se le concedió a tal efecto autorización oficial del Gobierno (el programa FedRAMP, que establece estándares de seguridad para los servicios en la nube). Sin embargo, y a pesar de los intentos de Google de mantener el asunto en secreto,<sup>224</sup> los empleados de la empresa y un gru-

<sup>223</sup> Rees (1994) y Morales (2009).

<sup>224</sup> Así, por ejemplo, Fei-Fei Li, científico jefe de IA en Google Cloud, decía estas cosas en un correo electrónico del 24 de septiembre de 2017: «¡Es tan emocionante que estemos cerca de obtener Marven...! Eso sería una gran victoria»; «Creo que deberíamos hacer un buen trabajo publicitario sobre la historia de la colaboración entre el Departamento de Defensa y GCP desde un punto de vista tecnológico vanguardista (almacenamiento, red, seguridad, etcétera), pero evitando a toda costa cualquier mención o alusión a la IA»; «Google ya



po de académicos liderado por Lucy Suchman, Lilly Irani y Peter Asaro, del ICRAAC, reaccionaron haciendo pública una carta con miles de firmas en la que exigían que Google cancelara el contrato con el Departamento de Defensa. En protesta, muchos empleados incluso renunciaron a su puesto de trabajo. Y finalmente, la empresa anunció que no renovarían el contrato para 2019. Además, ha prometido publicar una declaración sobre una política ética de la empresa en el desarrollo de tecnología con inteligencia artificial, que incluirá no utilizar nunca los datos de los consumidores en operaciones militares ni para vigilancia masiva, así como no desarrollar aplicaciones militares de la inteligencia artificial.<sup>225</sup>

Estos ejemplos ilustran la importancia del afloramiento de información por parte de las empresas a través del *compliance* (con delegados expertos)<sup>226</sup> o de los empleados a través de un canal de denuncias interno o externo.<sup>227</sup> Deben establecerse sistemas para que los *wistleblowers* (denunciantes o *deltatores*) que detecten actividades ilegales en la empresa las saquen a la luz y queden protegidos por la legislación, que debe incentivar tales prácticas y establecer protocolos para protegerlos fomentando la transparencia de las actividades privadas con trascendencia en materia de seguridad.<sup>228</sup> Es asi-

---

está luchando con problemas de privacidad cuando se trata de inteligencia artificial y datos; no sé qué pasaría si los medios comenzaran a hacerse eco de que Google está construyendo secretamente armas con IA». Cf. Conger (2018).

<sup>225</sup> Sandoval (2018).

<sup>226</sup> Vid. trabajo original de investigación presentado para la segunda prueba del concurso Ba58/2798 para la provisión de una plaza de profesor titular de la Universidad para Derecho Administrativo de la Universitat Autònoma de Barcelona, convocada por resolución de 8 de noviembre de 2001 (BOE de 26 noviembre de 2001), de R. Martínez: *El delegado de prevención de riesgos ambientales*.

<sup>227</sup> Espín (2017).

<sup>228</sup> Vandekerckhove (2016). Según este autor «el público tiene derecho a saber, pero las organizaciones requieren lealtad y secreto. En este conflicto, entre

mismo imperiosa la urgencia de elaborar una regulación que proteja a la investigación de aplicaciones indeseables a través de sistemas internacionales de control exhaustivo y de programas de cumplimiento normativo en las fases previas.<sup>229</sup> Dicha regulación debería garantizar que, aunque se segmenten las investigaciones en diferentes centros con posibles intenciones de opacidad deliberada, exista una inspección administrativa o auditoría que conecte dichos segmentos.<sup>230</sup> Como ejemplo a seguir en este caso puede ponerse el de la inspección a que Naciones Unidas sometió a algunos países de los que se sospecha que están incumpliendo la prohibición de investigar sobre armas nucleares. El éxito de dicha inspección preventiva se vio reconocido en 2006 con la concesión del Premio Nobel de la Paz al Organismo Internacional de Energía Atómica y a su director general, Mohamed el-Baradei. Recientemente, el OIEA ha aprobado un proyecto para renovar su sistema de información para ser más eficaz en su tarea de garantizar el uso pacífico de la tecnología nuclear.<sup>231</sup>

Los contrapoderes son en general fundamentales. En Estados Unidos, la ley de Autorización para Uso de la Fuerza Mi-

---

el interés público y los intereses de la organización, cabe una discusión seria sobre una política apropiada para el sistema de denuncias de irregularidades y los denunciantes». En este sentido, Naciones Unidas podría asegurar una reubicación para los científicos que denuncien actividades ilegales de sus empresas.

¿2005a o  
2005b?

<sup>229</sup> Sunstein (2005).

<sup>230</sup> Ayres y Braithwaite (1992). Braithwaite alerta sobre los riesgos de una regulación que sea una mera transacción entre el Estado y las empresas y sostiene que, a menos que haya un tercer actor en el juego regulatorio, la regulación será abducida y corrompida por el dinero. Para él, una regulación responsable (*responsive regulation*) implica escuchar a múltiples partes interesadas y tomar una decisión debatida y flexible.

<sup>231</sup> Cf. página *web* de la IAEA: <<https://www.iaea.org/newscenter/pressreleases/iaea-completes-3-year-project-to-modernize-safeguards-it-system>>. Cabe resaltar también el Premio Nobel otorgado a La Organización para la Prohibición de las Armas Químicas (OPAQ) en 2013 y su labor preventiva en este sentido.

litar (AUMF) de 2001 otorgó al presidente poderes especiales por los cuales podrá emplear toda la fuerza militar necesaria contra aquellas naciones, organizaciones o personas que hubieran estado implicadas en el atentado del 11-S.<sup>232</sup> Y a raíz de ello se desencadenó en ese país un interesante debate sobre si la ley, o incluso la Constitución, facultaba al presidente para decidir de forma unilateral el ataque y los medios.<sup>233</sup> El artículo 2, sección II decreta simplemente que «el presidente deberá ser el comandante en jefe del Ejército y de la milicia cuando sea llamada para el servicio de los Estados Unidos».<sup>234</sup> Y en esta pugna en contra o a favor del sometimiento de la Administración política y militar al control de los tribunales, la Corte Suprema, a pesar de contar con una mayoría conservadora, desautorizó esos pretendidos poderes y limitó y sometió a control judicial las actuaciones de la administración Bush incluso tratándose de la administración militar y encontrándose fuera del territorio de soberanía estadounidense, recuperando el principio jurídico liberal anglosajón según el cual «*Magna Carta follows the flag*».<sup>235</sup>

Si unos poderes tan superlativos están sometidos al control judicial (de la Corte Suprema o de jueces especiales por tratarse de materias reservadas), es evidente que cualquier artefacto

<sup>232</sup> Lavitt (2010).

<sup>233</sup> Ballbé y Martínez (2009); Powell (2013).

<sup>234</sup> Sunstein (2006).

<sup>235</sup> Daniel (2017) y Walker (2018). Se trata del famoso debate jurídico constitucional sobre si tiene que existir una deferencia judicial hacia las agencias o el mismo presidente por razón de *expertise* y de complejidad en la materia o de seguridad nacional como es nuestro caso. En este sentido, hay una jurisprudencia ambivalente que señala que no debe haber deferencia sino control judicial, por la gravedad del tema. Así, debemos entender que no cabe dicha deferencia cuando nos referimos a los LAWS. El peligro está en que los dos últimos magistrados de la Corte Suprema pretendan dar deferencia al presidente o a las agencias con los LAWS, siendo esto contrario al principio constitucional de *check and balance*. Barnett (2018).

independiente debe estarlo también. No puede ser que el Congreso otorgue al presidente un poder de delegación en LAWS que es evidente que puede volverse contra el mismo Congreso y contra el Estado en general, ya sea por decisión propia o porque quede en manos del enemigo. En el plano global, desde la creación del Tribunal Penal Internacional, en los demás Estados todas aquellas actividades o tecnologías que puedan llevar a un genocidio son sometidas a su control judicial, de manera incluso preventiva en algunos casos. Sin embargo, Estados Unidos ha elaborado para 2019 un proyecto de ley de Autorización de Defensa Nacional que persigue avanzar en el uso militar de la inteligencia artificial. El texto señala la intención de que el nuevo Consejo de Política y Supervisión de Inteligencia Artificial y Aprendizaje Automático «desarrolle y mejore la investigación, la innovación, la política, los procesos y los procedimientos conjuntos para facilitar su avance». Tendrá asimismo este consejo la tarea de «integrar la inteligencia artificial y el aprendizaje automático en todo el Departamento de Defensa y garantizar el uso eficiente y efectivo de sus capacidades». Según el borrador, se considera inteligencia artificial «cualquier sistema con capacidad de actuar racionalmente y realizar tareas en circunstancias variables e impredecibles sin supervisión humana significativa, o que puede aprender de la experiencia y mejorar el rendimiento cuando se expone a un conjunto de datos».<sup>236</sup>

De momento es solo una intención, pero es urgente un posicionamiento jurídico claro sobre estos extremos. La deliberada ambigüedad del borrador es sospechosa y parece otorgar un cheque en blanco a la investigación de inteligencia artificial fuerte en los LAWS, que pueden llegar a ser independientes. Y la Corte Suprema acaba de anular precisamente una orden

<sup>236</sup> Knapp (2018).

del presidente por basarse en una ley deliberadamente ambigua sobre deportaciones.<sup>237</sup> La ley debe proporcionar elementos descriptivos claros y dejar claro lo que significa realmente para que, parafraseando a la propia Corte Suprema, el derecho constitucional global pueda protegernos de nuestras mejores intenciones respecto de la inteligencia artificial. Fue la juez O'Connor quien, como ponente en la famosa sentencia Nueva York contra Estados Unidos, señaló en 1992 que «la Constitución nos protege de nuestras mejores intenciones: divide el poder entre soberanías y entre distintos poderes del Estado precisamente para que resistamos la tentación de concentrar el poder como solución expeditiva a las crisis del día a día».<sup>238</sup>

No hablamos de otra cosa que del principio de *check and balance*, que tiene su origen en la Constitución federal estadounidense y vertebrado todo un sistema de controles mutuos que también debe aplicarse a los LAWS; y aplicarse en todo el mundo. Ningún país debe disponer de una soberanía absoluta que le permita crear un arma genocida independiente. Los organismos estatales e internacionales deben protegernos de esta amenaza e incluso de sus propias buenas intenciones a través del derecho.

La potestad organizatoria de la Administración no puede ejercerse de modo abstracto: debe estar sujeta a un programa claro y preciso de asignación de competencias que puedan ser redistribuidas o revocadas por razones de necesidad y de geopolítica. Nuestro ordenamiento jurídico establece técnicas de traslación de competencias como la delegación o la

<sup>237</sup> Caso *Dimaya v. Chung*, A. (2018): «US Supreme Court restricts deportations of immigrant felons», 18 de abril [en línea], <<https://uk.reuters.com/article/uk-usa-court-deportation/u-s-supreme-court-restricts-deportations-of-immigrant-felons-idUKKBN1HO3DQ>>. [Consulta: 22-8-2018].

<sup>238</sup> U. S. Supreme Court, *New York v. U. S.*, 505 U. S., 144, 112 S. Ct. 2408 L. Ed. 2d.120 (1992). V. Ballbé y Martínez (2009).

avocación. El problema es que —como advierte Hass— los LAWS pueden no ser revocables o reprogramables aunque la situación política cambie. Su *instinto* de supervivencia puede impedirles autodestruirse, y ello puede causar escaladas convencionales o nucleares verdaderamente suicidas. Ello representa para el derecho administrativo internacional lo mismo un reto que una amenaza. Cualquier potestad que sea delegada por autoridad a un ente o a un órgano político, administrativo o militar debe poder ser revocada en cualquier momento como principio general, y en España, el artículo 9 de la ley 40/2015 de Régimen Jurídico del Sector Público señala que la delegación será revocable unilateralmente en cualquier momento por el órgano que la haya conferido, en todo o parte de las potestades delegadas. Además, podrá suspenderse para casos concretos mediante la avocación. Pero una inteligencia artificial superior a la humana podría no admitir estar sujeta al principio básico de revocabilidad de todas las delegaciones debido a su capacidad intrínseca de autodeterminación irreversible. También podrían no aceptar la anulación, sustitución o enmienda de la orden dada por la autoridad, ni una avocación de la competencia, ni la extinción del poder otorgado. Su metacognición podría no considerar aceptable la pérdida de competencias y hacerle desobedecer las órdenes, poniendo en peligro a toda la sociedad. En los LAWS, la distinción entre delegación de potestades y adopción de órdenes de ejecución de actuaciones materiales civiles y militares se desdibuja, y ello hace que tampoco quepa la potestad presidencial de *orden y mando* (*command and control* en Estados Unidos).<sup>239</sup>

<sup>239</sup> Véase el capítulo 4 («¿El fin ambiental justifica el *orden y mando*?») de Ballbé y Martínez (2003), donde se confirma que el poder federal del presidente y del Congreso es una atribución indelegable y por tanto hace inadmisibles una delegación a los LAWS. Especialista en este tema de los poderes

El Parlamento Europeo ha adoptado recientemente una resolución por la que se pide al alto representante de la Unión Europea para Asuntos Exteriores y Política de Seguridad, a los Estados miembros y al Consejo prohibir el desarrollo y producción de esta tecnología.<sup>240</sup> Por lo que respecta a Estados Unidos, la Constitución federal permite la *preemption* (especie de avocación) por parte del poder federal de Washington a los Estados en las competencias relacionadas con el comercio interestatal, de modo que diseñar jurídica y tecnológicamente armas totalmente autónomas iría contra la Constitución misma, porque el Gobierno federal no podría activar el derecho de avocación respecto de los poderes delegados a los LAWS, ya que su inteligencia artificial podría activar los mecanismos de desobediencia respecto de la ejecución de sus competencias.<sup>241</sup>

Solo debería admitirse el desarrollo de los AWS cuando sus fines sean defensivos y siempre y cuando se permita la revocabilidad anteriormente expuesta o sus competencias delegadas se lleven a cabo con un control humano significativo claro y se demuestre que se han conseguido implementar algoritmos con las características básicas de la humanidad y con un sistema de *compliance* preventivo.

## Política, opacidad y connivencia

La inteligencia artificial aplicada al ámbito armamentístico-militar supone un salto exponencial con respecto a tecnolo-

---

presidenciales incluso obviando al Congreso es Yoo (2010, 2012). Es interesante asimismo la reflexión sobre la *doctrina de la no delegación* que hace Sunstein (2017).

<sup>240</sup> Amoroso y Tamburrini (2017).

<sup>241</sup> Ballbé y Martínez (2009); Zimmermann (1991).

No figura  
en la  
bibliografía  
con ese año.

gías militares previas que, como el Zyklon B utilizado en los campos de exterminio nazis o el napalm empleado por las tropas de Estados Unidos en Vietnam, marcaron un antes y un después en nuestra comprensión del conflicto, la guerra y la seguridad debido a su habilidad para aumentar la distancia física entre víctima y verdugo y, como consecuencia, entre la opinión pública y el propio conflicto.

Se ha dicho que las cámaras de gas eran la respuesta tecnocientífica de la biopolítica nazi al estrés extremo sufrido por los soldados al verse obligados a eliminar a mujeres y niños de manera masiva. No es cierto: de lo que se trataba era de conseguir la máxima efectividad en el genocidio y que hubiera el menor número de soldados testigos implicados. Pues bien: hoy las armas derivadas de las ciencias robóticas corren el riesgo de seguir tales precedentes y aducir las mismas excusas a escala global, reduciendo y ocultando las pruebas de su tenebroso impacto tanto físico como psicológico sobre las sociedades que las utilicen. Y todo ello yendo un paso más allá con respecto a aquellas otras armas: la autonomía de los LAWS los hará decidir sobre la vida y la muerte sin juicio humano previo. Es decir, se está desarrollando una controvertida estrategia de *defense of inhumanity* de los instrumentos utilizados en las guerras.<sup>242</sup>

Los LAWS pueden ser considerados el armamento propio de la necropolítica postantropológica, creado para salvaguardar las dinámicas de los nuevos procesos de consumo ayudándose de la inteligencia artificial. Es importante en este punto destacar la idea de Haraway del paso de la economía política del panóptico (arquitectura carcelaria ideada por Bentham a finales del siglo XVIII pero también aplicada al control de la sociedad misma sin que esta sepa que es

<sup>242</sup> Satia (2006, 2009).



observada) al dominio de la informática,<sup>243</sup> bajo cuya égida la asimetría del poder es si cabe más extrema: no hay mejor ejemplo de un panóptico postcontemporáneo que los dos mil millones de usuarios que controla Facebook y su capacidad para recabar información, que supone un riesgo para la privacidad de millones de usuarios en todo el mundo; entendida esa privacidad en los términos del gran jurista Louis Brandeis en su artículo clásico «*The right to privacy*» y de la Constitución no escrita que cinceló.<sup>244</sup> Así lo cree Fast, que hace un análisis de la estructura y función de la red social y detecta numerosos paralelismos con la estructura del panóptico de Bentham, así como entre las estrategias implementadas por Facebook y las descritas en la discusión más amplia de Foucault sobre la evolución del control y el castigo. Aborda Fast las implicaciones culturales de estas similitudes y especialmente las que surgen en una era post-Snowden en la que los usuarios de Internet tienen motivos para sospechar una falta general de privacidad y seguridad en las redes, pero constata que, aunque existe cierta desconfianza, las personas aún usan las redes sociales y Facebook sigue siendo la dominante en todo el mundo. Así pues, las similitudes entre Facebook y el panóptico dentro de un contexto cultural más amplio se vislumbran como el más tenebroso e inimaginable escenario; más aún que los descritos por Orwell en su obra *1984*.<sup>245</sup>

Ha de tenerse en cuenta que atravesamos una fase en que quien se halla a la vanguardia de la robótica y de la utilización de inteligencias híbridas o mixtas (máquina-humano) con el objetivo de captar y sistematizar datos a gran velocidad es la tecnología militar público-privada.<sup>246</sup> La evolución y la

<sup>243</sup> Haraway (2013).

<sup>244</sup> Warren y Brandeis (1890).

<sup>245</sup> Fast (2015), McMullan (2015).

<sup>246</sup> Suchman y Weber (2016).

relación de esas tecnologías con el *big data* está produciendo una variedad extraordinaria de especies y formas, fruto, a menudo, de la connivencia del mundo universitario; y ese entendimiento Universidad-empresa-Estado tiene una clara traslación material. Nada menos que el 80 % del I+D+i estadounidense es para programas militares y de guerra, los cuales, normalmente, no se licitan a través de un concurso con las debidas garantías de la legislación antimonopolio, de transparencia y de concurrencia (*competitive tendering*), sino a través de procedimientos directos por ser materias clasificadas, siendo auténticas subvenciones condicionadas secretas por afectar a la seguridad nacional.

Por ejemplo, en el conflicto de la Organización Mundial del Comercio y la Unión Europea (Airbus) contra Estados Unidos (Boeing) sobre subvenciones encubiertas prohibidas<sup>247</sup> se reveló que la actividad mayoritaria de Boeing era I+D proporcionado por la NASA y el Departamento de Defensa para fines militares y afectando a la seguridad nacional. Es solo un ejemplo: muchos de los escándalos de corrupción más impactantes de Estados Unidos (como Lockheed Martin, que dio origen a las leyes anticorrupción) tienen su origen en la connivencia entre sectores académicos y militares para promover investigaciones secretas bajo las leyes de materias re-

<sup>247</sup> En junio de 2017, la Organización Mundial del Comercio condenó a Boeing por incumplir las resoluciones de la organización con respecto a la batalla por las subvenciones a la aviación comercial en el conflicto DS353 (Unión Europea contra Estados Unidos). En marzo de 2012, el Órgano de Solución de Diferencias de la OMC había dictaminado que varias de las subvenciones proporcionadas por Estados Unidos a Boeing eran ilegales, pero la UE denunció que Estados Unidos sigue proporcionando ventajas a Boeing en forma de subvenciones no equitativas y contrarias a la competencia por intereses de defensa militar. Cf. «La OMC condena el incumplimiento de Boeing y las nuevas subvenciones», *Airbus*, 9 de junio de 2017 [en línea], <<https://www.airbus.com/newsroom/press-releases/es/2017/06/wto-condemns-boeing-s-non-compliance-and-new-subsidies.html>>. [Consulta: 22-8-2018].

servadas y seguridad nacional (para que no les sea de aplicación la FOIA)<sup>248</sup> a través de contratos con empresas encubiertas para fines militares.<sup>249</sup> Recuérdese que el instrumento más representativo de la nueva revolución global, Internet, fue fruto de un programa público-privado militar pergeñado entre el Pentágono, empresas privadas y universidades.

Es decir, la innovación desarrollada para los militares se transfiere posteriormente y se desarrolla aún más en el ámbito civil, dando lugar a lo que se conoce —y ya hemos citado— como *tecnologías de doble uso*.<sup>250</sup> Otro ejemplo de esta metamorfosis son los drones creados y utilizados por primera vez por la Fuerza Aérea norteamericana después de la segunda guerra mundial. Marin señala que «la industria de la guerra es un motor fuerte y poderoso para la innovación, pero es y sigue siendo la primera fuente de muerte y garantía (incluso cuando se apunta) de destrucción».<sup>251</sup>

Son en el fondo las grandes potencias quienes, a través de subvenciones públicas o privadas encubiertas o no, dominan la innovación tecnológica más avanzada (Microsoft, Google, etcétera): recientemente se ha publicado que la inteligencia artificial de Google está siendo usada por el programa de drones del Departamento de Defensa de Estados Unidos.<sup>252</sup> Pero el secretismo con el que operan los Estados en según

<sup>248</sup> Ballbé y Martínez (2009).

<sup>249</sup> Goodman (2013).

<sup>250</sup> Tucker (2012). También Mazzucato (2015). La autora quiere quebrar el mito de que los Estados son estructuras pesadas y burocráticas enfrentadas a un sector dinámico y emprendedor. Precisamente demuestra lo contrario, es decir, que el sector privado solo avanza en el sector más puntero cuando el estado emprendedor ha realizado inversiones de alto riesgo. Critica que hemos creado un sistema de innovación mediante el cual el sector público socializa los riesgos mientras que las recompensas están privatizadas.

<sup>251</sup> Marin (2017), p. 300.

<sup>252</sup> Gibbs (2018).

qué casos puede poner en peligro los derechos individuales y colectivos de la sociedad, así como la seguridad global.<sup>253</sup>

Ejemplo de las relaciones que muchas empresas del sector privado tienen con la seguridad nacional es el caso de Boston Dynamics, comprada en junio de 2017 por la japonesa SoftBank a Alphabet Inc (matriz de Google), provocando la intervención de CIFIUS por afectar a seguridad nacional.<sup>254</sup> CIFIUS es el *Committee on Foreign Investment in the United States*, uno de los frenos que felizmente existen al complejo militar tecnológico post-humano, responsable en este caso de advertir al presidente sobre las inversiones extranjeras que pueden sustraer soberanía y afectar negativamente a la seguridad nacional.

Existe un complejo entramado industrial, excesivamente opaco para la opinión pública, que posee la capacidad de habilitar máquinas y armas para tomar decisiones de forma autónoma e independiente, afectando a la soberanía de los poderes democráticos y dando un vuelco a la noción de responsabilidad existente hasta la fecha, que ni puede ni debe ser reducida simplemente a la cuestión de la propiedad como si de un animal de compañía se tratara, la vieja solución inspirada en el derecho romano sobre la responsabilidad asociada a dichos entes.<sup>255</sup>

Estas tecnologías y el complejo militar-financiero-industrial que las gesta son símbolos del advenimiento de una transformación mucho más profunda de lo que nos imaginamos. La *teoría de las dos culturas*,<sup>256</sup> así como la propia contraposición existente entre cultura y naturaleza, se deshace en

<sup>253</sup> Ballbé y Padrós (1997).

<sup>254</sup> Cf. Rubio (2017).

<sup>255</sup> Liu (2016) p-325. Bhuta, Beck, Geib, Liu y Krebs (2016); Brändli, Harasgama, Schuster y Tamò (2014); Margulies (2017).

<sup>256</sup> Snow (2001).

un complejo sistema de retroacciones; de interdependencias con la tecnología. Sin embargo, los nuevos armamentos autónomos suponen la reducción del sujeto, del individuo, a un objeto pasivo, un número, una baja, un dato.

Otra cuestión aparejada al desarrollo de los LAWS que debe preocuparnos es el posible desarrollo exponencial de gigantes oligopolistas de datos. Los LAWS se basan en algoritmos preprogramados<sup>257</sup> que en apariencia pueden ofrecer una actuación discrecional *eficaz y eficiente* tecnológicamente.<sup>258</sup> Sin embargo, no pueden obviarse los peligros que esto lleva aparejados. Aquellos gigantes suministran a los LAWS con inteligencia artificial los datos necesarios para alimentar su metacognición, por lo que pueden acabar teniendo en sus manos las decisiones letales más discrecionales. Como ha señalado *The Economist*, «el recurso más valioso del mundo ya no es el petróleo, sino los datos», y el derecho administrativo debe actuar para prevenir estas amenazas monopolistas a la seguridad y la privacidad. Esto dice el prestigioso diario:

Una nueva mercancía genera una industria lucrativa en rápido crecimiento, lo que lleva a los reguladores antimonopolio a intervenir para restringir a quienes controlan su flujo. Hace un siglo, el recurso en cuestión era el petróleo. Ahora están siendo planteadas preocupaciones similares con respecto a los gigantes que se ocupan de los datos, el petróleo de la era digital. Estos titanes

<sup>257</sup> O'Neil (2016); Ramírez (2014).

<sup>258</sup> No podemos dejar de destacar que Ada Lovelace (1815-1852), hija de Lord Byron, fue la primera programadora en la historia de la informática, ya que inventó una anotación para describir algoritmos y ser procesados en la máquina analítica (se trataba de una calculadora mecánica) del matemático Charles Babbage. Ada firmaba sus trabajos con las iniciales A.A.L. para que no la censuraran por ser mujer. En 1979 el Departamento de Defensa estadounidense desarrolló un lenguaje de programación que denominó ADA en su honor. Cf. Essinger (2015) y Hollings y Martin (2018). Asimismo, una interesante obra sobre la vida de la hija y la madre: Seymour (2018).

(Alphabet, compañía matriz de Google; Amazon, Apple, Facebook y Microsoft), parecen imparables. Son las cinco firmas cotizadas más valiosas del mundo. Sus ganancias están aumentando: ha acumulado colectivamente más de veinticinco mil millones de dólares en ganancias netas en el primer trimestre de 2017. [...] Tal dominación ha provocado llamamientos para que los gigantes tecnológicos se fraccionen, como Standard Oil a principios del siglo xx. [...] Pero la naturaleza de los datos hace que los remedios antimonopolio del pasado sean menos útiles. [...] Reiniciar el antimonopolio en la era de la información no será fácil. Implicará riesgos: más intercambio de datos podría amenazar la privacidad. Pero si los gobiernos no quieren una economía de datos dominada por unos pocos gigantes, deberán actuar pronto.<sup>259</sup>

En lo que respecta a la ciberdelegación, los profesores de derecho administrativo Coglianesse y Lehr se preguntan «si el uso de herramientas de decisión robótica por parte de las agencias del Gobierno puede pasar por alto doctrinas consagradas en el derecho administrativo y constitucional. La primera idea de una regulación algorítmica podría ir el contra de más de un principio, como la prohibición de delegación de determinados poderes, el debido proceso, la igualdad entre partes, la motivación de los actos o la transparencia».<sup>260</sup> Señalan además estos estudiosos que utilizar los algoritmos de aprendizaje automático puede optimizar la búsqueda del interés general en los procedimientos administrativos, pero siempre y cuando los funcionarios gubernamentales conserven el control final de la decisión, de la acción, de la especificación de los algoritmos y de la traducción de sus resultados y el *machine learning* no eluda los principios antes enumerados.<sup>261</sup>

<sup>259</sup> The Economist (2017).

<sup>260</sup> Coglianesse y Lehr (2017).

<sup>261</sup> *Ibíd.*

En esta misma línea, Massaro también señala que la Administración pública debe poder utilizar esta tecnología, pero teniendo en cuenta que

el algoritmo aprende de los datos para producir una predicción. Este proceso se conoce como *caja negra* porque los humanos solo ven las entradas y salidas. Pero el aprendizaje automático no es sinónimo de inteligencia artificial. El objetivo de la inteligencia artificial es eliminar el error humano, mientras que los algoritmos de aprendizaje automático producen una predicción (y en consecuencia dictan un acto) a través del reconocimiento de patrones. En todo caso, el aprendizaje automático puede ayudar a las administraciones a tomar mejores decisiones procesando conjuntos de datos más grandes y más rápidamente que los humanos.<sup>262</sup>

Sea como sea, no podemos utilizar el aprendizaje automático para actuaciones discrecionales que requieran un margen de valoración intuitiva. Sigue señalando Massaro que

en el contexto del derecho administrativo, el aprendizaje automático se puede dividir en dos categorías: 1) Adjudicación por algoritmo y 2) Regulación de la actuación por un robot. La primera es apropiada cuando los datos cuantificables determinan un resultado. Así, el Ayuntamiento de Los Ángeles usa la regulación por robots para mejorar el flujo del tráfico y reducir demoras. Sin embargo, el algoritmo sintetiza grandes cantidades de datos y adapta los semáforos en consecuencia.

En nuestro caso, los LAWS podrían utilizar algoritmos y acciones robotizadas en gran parte de su actuación, pero siempre deberá estar presente la intervención humana intuitiva, el juicio humano, a fin de que sea posible paralizar las decisiones predichas por los sistemas.

<sup>262</sup> Massaro (2018).

Coglianese y Lehr afirman que, puesto que «el Tribunal Supremo norteamericano ha permitido la delegación a partes privadas cuando sus acciones se limitan a decisiones de asesoramiento» y no influyen en la decisión final, la colaboración entre Administración y robots inteligentes es defendible en tanto en cuanto «los humanos puedan, en cualquier momento, rechazar una regla elegida por la máquina, alterar las especificaciones de un algoritmo o incluso desconectar el sistema por completo [...] La objeción sería que los humanos no pudieran ejercer un control significativo sobre los sistemas automatizados en entornos de alta velocidad».<sup>263</sup>

Por su parte, el profesor de derecho Lieblich se plantea la posibilidad de que la inteligencia artificial pueda contener algoritmos de sensibilidad y humanidad y que los sistemas autónomos armados lleguen a ser capaces de reflexionar, crear sus propios procesos de pensamiento y aprender de ellos mismos y de toda la información que ofrece la nube. Tales sistemas —advierte Lieblich— podrían rebelarse contra sus propios creadores; contra los algoritmos con características humanas que su creador les ha implantado. Es decir, podrían transformarse en seres sintéticos independientes con inteligencia artificial.<sup>264</sup> Y podrían desligarse del objetivo de servir a los intereses generales; al bien común. Es iluso pensar que un Gobierno o una administración pública puedan controlar fácilmente un artefacto susceptible de autodeterminarse e independizarse del control gubernamental o privado y actuar como un *free rider*, convirtiéndose en la más peligrosa arma de destrucción masiva precisamente por su descontrol. Gracias a la metacognición, los LAWS podrían conocer, innovar y regular los procesos mentales básicos que intervienen en su

<sup>263</sup> Coglianese y Lehr (2017).

<sup>264</sup> Bhuta, Beck, Geib, Liu y Kreb (2016), p. 269.



cognición y manipularlos e incluso contradecir no solo el derecho administrativo por lo que respecta a su actuación, sino también el derecho internacional humanitario básico. Es por ello que, como se viene argumentando en este libro, el principio de la legalidad puede ser el paraguas que prevenga a la sociedad no solo de la capacidad de actuar discrecionalmente de los LAWS con inteligencia artificial, sino de la investigación misma para ponerlos en funcionamiento.

Los actos arbitrarios por parte de los Estados están prohibidos por el ordenamiento jurídico nacional e internacional, y es evidente que a unos sistemas complejos letales, a unas máquinas autónomas e independientes, no se les debería permitir de ninguna manera tomar decisiones sobre si usar o no la fuerza letal. De lo contrario, el Estado estaría cediendo competencia a un genocida; a un asesino en serie sin humanidad.<sup>265</sup> Además, los mismos gobiernos estarían siendo instigadores y cooperadores necesarios de dicha ilegalidad. Cuanto más independiente sea el LAWS, más debe vigilarse al Estado que subrepticamente otorga dichas potestades a un ser sin sentimientos ni sensibilidad humana. Sin embargo, si dichas máquinas o sistemas autónomos e independientes fueran capaces de asumir los valores y componentes de humanidad de sus responsables, imitar el pensamiento humano y actuar discriminando objetivos, y si pudieran ser utilizados para fomentar sistemas más humanitarios, sí que podrían permitirse —sujetas a un compliance preventivo— las investigaciones en inteligencia artificial con humanidad.<sup>266</sup>

<sup>265</sup> McQuillan (2018); Bhuta, Beck, Geib, Liu y Krebs (2016).

<sup>266</sup> Cadena (2016).

## Un nuevo autoritarismo

En los modelos administratocéntricos europeos, el monopolio de la violencia legítima reside en el Estado y concretamente en el personal funcionario que conforma las Fuerzas y Cuerpos de Seguridad y las Fuerzas Armadas, sometido al presidente. Y eso hace evidentemente ilegal que los LAWS asuman esas competencias que además de ser discrecionales, comportan el uso de una violencia que ya no sería legítima, porque decide sobre ella una máquina que incluso posee la tecnología necesaria para rebelarse y decidir por sí misma de forma soberana. También es así en Estados Unidos, donde el modelo, que originariamente era individuocéntrico, es actualmente administratocéntrico en lo militar y en lo policial. Sin embargo, en la actualidad se adivina en el horizonte un nuevo modelo que ya no es ni el Estado-Ejército individuocéntrico (ciudadano-policía: concretado en el *posse comitatus* o partida de la comunidad comandada por un *sheriff*; ciudadano-soldado: voluntarios en la milicia, hoy Guardia Nacional) como en el origen de Estados Unidos,<sup>267</sup> ni administratocéntrico o funcionariocéntrico, sino un *tertium genus* consistente en un modelo de Estado y de Ejército con inteligencia artificial, robotizado y despersonalizado.

Hay muchos modelos de Estados autoritarios: unos son dictaduras militares, otros son dictaduras civiles, otros dictaduras de partido... En España, durante el siglo XIX e incluso en las épocas democráticas, existió un modelo denominado de *autonomía militar* en el que el Gobierno y el Parlamento no podían penetrar en los ámbitos del Ejército y la policía militarizada y ni en sus presupuestos, ni en sus decisiones de intervención. Con el tiempo, se acabó aceptando —fue la clave de bóveda de

<sup>267</sup> Beebe y Kaldor (2010); Klay (2016).

la Transición española— que un Estado democrático solo lo es si ha destruido dicha autonomía militar. Esa desmilitarización del Estado moderno se llevó a cabo también, entre otras cosas, para evitar las rebeliones militares que jalonan, por ejemplo, la historia hispanoamericana. Hoy, el principio constitucional del *check and balance* (y entre otras cosas, el sometimiento del poder militar al civil y democrático) es ya consustancial a nuestras sociedades<sup>268</sup> y se ha vuelto más necesario que nunca en un momento de auge de la privatización de la industria militar.<sup>269</sup> Si un Estado democrático solo puede serlo cuando no dispone de autonomía militar, menos puede admitirse un Estado que admita la autonomía de los LAWS.<sup>270</sup>

Las guerras y el terrorismo se combaten, en una parte decisiva, con prevención y especialización de mandos.<sup>271</sup> Y la primera medida de prevención debe ser el estudio de los riesgos de esta carrera tecnológica, pero el tema no está incluido en muchas obras generales sobre riesgo y regulación.<sup>272</sup>

Los conflictos bélicos son hoy conflictos completamente asimétricos entre Estados o entre no-Estados<sup>273</sup> y deben ser entendidos como una nueva forma de *governance*. La Administración tiene obligaciones administrativas ejecutivas básicas, y el Estado, un compromiso de aplicar los principios generales del derecho, aunque se consideren actos políticos. El núcleo reglado de dicha competencia discrecional debe ser siempre tenido en cuenta, y nunca puede dejarse en manos de un sistema militar tecnológicamente avanzado independiente y unilateral.

<sup>268</sup> De Vergottini (1982).

<sup>269</sup> Ballbé (1982), Singer (2003).

<sup>270</sup> Levitsky y Way (2010); Ballbé (1985).

<sup>271</sup> Meyer, (2007), p. 472.

<sup>272</sup> Ambrus, Rayfuse y Werner (2017).

<sup>273</sup> Kaldor (2001).

En los conflictos, la decisión última de actuar la debe tomar el personal competente de las FFCCSS y de las Fuerzas Armadas, cuya competencia, intuición y cuyas emociones son humanas y se dirigen a solucionar un conflicto con seres humanos. Aquí se ha citado ya el principio *la Constitución sigue a la bandera*, implantado progresivamente en todo el mundo a raíz de su adopción por la jurisprudencia inglesa y que significa que los derechos y garantías constitucionales de los ciudadanos y las responsabilidades de los poderes públicos siguen siendo aplicables cuando actúan fuera de su territorio. Con arreglo a él, es de todo punto inaceptable que en un ordenamiento jurídico tecnológicamente más avanzado sea posible la atribución de potestades públicas a unos robots autónomos letales: la aplicación de su discrecionalidad iría contra los principios constitucionales de responsabilidad de militares y civiles allí donde vayan.

Hoy, la deslocalización de las grandes empresas de suministro militar y tecnológico —no solo por motivos fiscales, sino también para eludir el control político y judicial— está a la orden del día. Halliburton, empresa vinculada al exvicepresidente estadounidense Dick Cheney, es un ejemplo de ello y de cómo la tecnología y los servicios militares están fuera de control y mantienen relaciones corruptas con gobiernos extranjeros. Dar poder a estas empresas para crear LAWS con inteligencia artificial y que estos lleven a cabo acciones arbitrarias sería un procedimiento en sí mismo arbitrario y por lo tanto prohibido. Como dejó sentado el Tribunal Supremo español en 1992,

la potestad discrecional de la Administración en la producción de actos no reglados por el derecho administrativo únicamente se justifica en la presunción de racionalidad con que aquella se ha utilizado en relación con los hechos, medios técnicos y la

multiplicidad de aspectos y valores a tener en cuenta en su decisión, de tal suerte que la actividad discrecional no ha de ser caprichosa, ni arbitraria, ni ser utilizada para producir una desviación de poder sino, antes al contrario, ha de fundarse en una situación fáctica probada, valorada a través de previos informes que la norma jurídica de aplicación determine e interpretados y valorados dentro de la racionalidad del fin que aquélla persigue.<sup>274</sup>

La libertad como ideario, como aspiración metafísica, se sustancia en el derecho público subjetivo, que «comienza configurándose como un derecho a la legalidad en el sentido de un derecho a oponerse a la opresión que no venga en nombre de la ley; a oponerse a toda posibilidad de ser afectados en la esfera de los intereses personales si no es por disposición expresa de la ley».<sup>275</sup> En la era tecnológica actual, la lucha contra la inmunidad del poder ya no se efectúa tan solo a través del cumplimiento de la legalidad formal, sino también de los principios inherentes a la constitucional no ordinaria, que no puede aceptar que unos sistemas letales complejos y las empresas que los fomentarán sean destinatarios de una competencia para no solo aplicar casi legalmente la pena de muerte sobre los ciudadanos, sino también sentar las bases para un auténtico genocidio local o global.

Se suele aludir a la capacidad de los LAWS de discriminar entre objetivos legítimos e ilegítimos, pero es muy difícil que esa capacidad exista verdaderamente: lo que es legítimo e ilegítimo depende con mucha frecuencia del contexto político y del derecho internacional humanitario;<sup>276</sup> y no digamos ya cuando dichos objetivos se tuvieran que discriminar en fun-

<sup>274</sup> STS del 6 de mayo de 1992, Sala 3.ª, Sección 6.ª. Ponente Sr. D. Fco. J. Hernando Santiago.

<sup>275</sup> García de Enterría (1983).

<sup>276</sup> Sparrow (2016).

ción de los *big data* a través de los cuales la inteligencia artificial del LAWS aprende o imita patrones de conducta.

De acuerdo con el *jus belli*, los ataques sobre combatientes son ilegítimos bajo tres tipos de circunstancias. Primera, que el ataque vaya a causar un número desproporcionado de bajas civiles (artículo 57 de la Convención de Ginebra, Protocolo Adicional). Segunda, que el ataque constituya un uso innecesariamente destructivo y excesivo de la fuerza. Tercera, que se haya demostrado un deseo de rendición o que ya se está fuera de combate (artículo 41). Los LAWS deberían ser capaces de distinguir estas circunstancias, pero ello requiere un análisis abstracto e intuitivo de la situación. En el fondo, deberían enfrentarse a un dilema ético y moral y ser capaces de *entender* la naturaleza de sus acciones desde un punto de vista humano, lo cual, actualmente, es imposible.

En un conflicto armado, a menudo la acción de discriminar objetivos es discrecional, por lo que la apreciación conforme a criterios extrajurídicos no es de recibo atribuirlos a un órgano inanimado; a una máquina con inteligencia artificial. No se trata siempre de aplicar conceptos jurídicos determinados, sino, con frecuencia, de la elección entre varias formas de comportamiento igualmente posibles, e incluso entre escoger y discriminar objetivos. Es decir, un LAWS podría llegar a decidir tanto sobre la necesidad de intervención como sobre las medidas a tomar.<sup>277</sup>

La discrecionalidad administrativa se sitúa en el plano volitivo y no se trata de una simple actividad de cognición, lo que comporta que, a la hora de efectuar un control judicial, el juez, como han entrado criterios extrajurídicos (políticos o de oportunidad) en la decisión discrecional, no puede fiscalizar esta más allá de los límites impuestos por el ordena-

<sup>277</sup> Parejo (1993, 2016).

miento al carecer de parámetros para realizar tal juicio.<sup>278</sup> En el caso de los LAWS, estaríamos hablando de otorgarles un cheque en blanco no fiscalizable.

Insistamos: el otorgamiento del ejercicio de la potestad discrecional letal (la elección del interés general por lo que se refiere a la aplicación de la violencia legítima por parte de los poderes públicos) no debe atribuirse a un LAWS independiente, porque el ordenamiento jurídico se refiere a órganos con capacidades humanas y a los que exista la posibilidad de someter a un control jurisdiccional de su actuación acorde a dicha naturaleza, lo que es totalmente imposible con respecto a un ser letal con inteligencia artificial general. En consecuencia, las armas autónomas con inteligencia artificial no deberían ostentar potestades que impliquen autoridad ni ser destinatarias de la atribución de discrecionalidad administrativa.<sup>279</sup>

Existe una idea común dominante según la cual los Estados se forman a partir del proceso de monopolización de las armas por el poder público. Así sucedió en Europa, donde fueron creándose unas instituciones administrativas profesionalizadas a las que se atribuía el empleo legítimo de la fuerza y consecuentemente el recurso —más o menos excepcional y limitado— a las armas: ejércitos profesionales permanentes y, posteriormente, cuerpos de policía para mantener el orden y la seguridad interna.<sup>280</sup> Pero ello se explica en base al susttrato católico del continente: la Iglesia católica, a partir del siglo x, al no haber Estados, se convirtió en la protectora de

<sup>278</sup> Sánchez Morón (1994), Beltrán de Felipe (1995).

<sup>279</sup> En cuanto a la discrecionalidad, en el ámbito de la Unión Europea se distingue entre apreciación política y técnica. Splo la técnica puede ser objeto de delegación por las instituciones. Véase el asunto C-270/12, apartados 41 y 54, donde se resume la jurisprudencia sobre la materia.

<sup>280</sup> Martínez (2003).

sus creyentes y estableció un derecho y una serie de instituciones de seguridad (*Tregua Dei*, órdenes religioso-militares, protección de los que no llevaran armas, etcétera) que siglos después, a partir del siglo XIII, fueron emulados secularmente por los primeros Estados absolutistas. Esa identificación tradicional entre Estado y monopolio de las armas no ha sido el único modelo existente en los procesos de construcción de los Estados modernos. La creación de Estados Unidos siguió en cambio un modelo de influencia protestante con unos principios contrapuestos a los del sistema europeo. Se instauró un Estado basado en el derecho individual a llevar armas, garantizado constitucionalmente por la segunda enmienda; un Estado-comunidad que rechazaba inicialmente una Administración profesionalizada que pretendiera suplantar las tareas del ciudadano en las funciones públicas. Se institucionalizaba un ciudadano armado que era por un lado ciudadano-soldado en la milicia estatal —hoy Guardia Nacional—, por otro ciudadano-policía en el *posse comitatus* o *partida de la comunidad* y finalmente ciudadano-juez en el jurado popular.<sup>281</sup>

En este sentido, a nivel civil la prohibición de los LAWS será fácil de abordar en Estados como España, que monopolicen la violencia legítima y firmen convenios en ese sentido; pero más problemática en países como Estados Unidos, que garantizan constitucionalmente la autodefensa armada de sus ciudadanos.<sup>282</sup> La segunda enmienda puede ser un freno a una regulación restrictiva de los LAWS.

El reto es de entidad. Por un lado, debería sujetarse a las milicias y guardias nacionales a poderosas restricciones en este sentido si no queremos arriesgarnos a una nueva epi-

<sup>281</sup> *Ibidem*.

<sup>282</sup> Martínez (2015).



demia de destrucción y muerte: recuérdese que la mitad del ejército de Estados Unidos corresponde a las guardias nacionales de los estados, conformadas por ciudadanos soldados y por la comunidad misma. Y por otro lado, es preciso prestar atención a los nanodrones con inteligencia artificial en manos privadas, un aspecto de los LAWS apenas analizado en un contexto comunitarocéntrico. Llevado a un último extremo, el derecho a llevar armas recogido en la Constitución norteamericana y la sentencia Distrito de Columbia contra Heller (2008) y McDonald contra Chicago (2010) de la Corte Suprema permiten a los ciudadanos estadounidenses no solo disponer de un arma corta de defensa personal en casa, sino también cualquier arma que considere necesaria para su seguridad e incluso armas automáticas, por ejemplo un AK-47 en determinados estados. Según el significado original de la Constitución norteamericana,<sup>283</sup> el ciudadano puede defenderse privadamente contra la posible tiranía del poder público y la criminalidad, y podría llegarse a extremos tan absurdos como la defensa de que dicho derecho comporte la potestad de disponer de un robot autónomo que le proteja de forma ofensiva o defensiva: un dron con armamento letal autónomo o incluso un nanodrón. El ya accesible al público dron 720X de bolsillo, por ejemplo, es un artefacto inofensivo, pero si a dicha tecnología se le aplicara inteligencia artificial y un dispositivo con capacidad de disparar, podría llegar a convertirse en un LAWS.<sup>284</sup>

Ahora bien, a pesar del minimalismo judicial norteamericano y de la alegaldad imperante con respecto a los LAWS,

<sup>283</sup> Rakove (1990).

<sup>284</sup> Cf. «Selfie quadcopter conquers Spain. The idea is genius...», *blog Daily Life Tech*, 3 de agosto de 2018 [en línea], <<http://blogs.dailylifetech.com/vzra/drone-720x/d/selfie-quadcopter-conquers-the-idea-is-genius-1117>>. [Consulta: 22-8-2018].

lo cierto es que existe un activo movimiento partidario de la prohibición de dicha tecnología. Y está transformando lentamente el horizonte de la seguridad, pero no se da cuenta que el terror no solo puede provenir de Estados irresponsables, corruptos o totalitarios, o de grupos terroristas, sino de ciudadanos libres de toda sospecha pero que, al disponer del derecho a armarse, si eligen hacerlo con sistemas autónomos de este tipo, pueden llegar a provocar una auténtica escalada de accidentes, muertes y terror descontrolado.

Si las armas de defensa personal tradicionales (automáticas y semiautomáticas) comportan hoy diez veces más víctimas en sociedades armadas como la estadounidense que en Europa, imaginemos qué ocurriría si dichos ciudadanos, amparados por su Constitución, se dotaran de armas letales autónomas con inteligencia artificial (por ejemplo nanodrones armados) para uso particular en defensa de su seguridad si no hay un pronunciamiento claro de la jurisprudencia en este sentido. Esta polémica ya existe con armas automáticas, y la reciente sentencia del juez federal del distrito de Massachusetts de 5 de abril de 2018 deja claro que las armas de asalto AR-15 no entran dentro de la garantía de la segunda enmienda y pueden ser prohibidas, con lo que da la razón a la fiscal general del Estado, Mauren Healey, que defendía la legalidad de las políticas restrictivas estatales. El camino a seguir con los LAWS debe ser justamente ese.

# Epílogo

ROSER MARTÍNEZ Y JOAQUIN RODRÍGUEZ

Es que yo os pedí, Creador, que del barro me modelarais hombre? Es que yo os solicité que me sacarais de la oscuridad?

John Milton: *El paraíso perdido* (1667)<sup>285</sup>

El desarrollo de una inteligencia artificial avanzada y su posible aplicación a sistemas letales de armamento autónomo representa una amenaza para la vida, la paz y los fundamentos del derecho, y es imprescindible avanzar en un sistema clasificatorio que desemboque en una regulación internacional y estatal sobre el particular. Pero esto solo será posible si se llega a un acuerdo previo en lo que respecta los procesos para lograr dicho objetivo. Tenemos ante nosotros dos modelos contrapuestos: el unilateralista y el multilateral, que busca un acuerdo integrador; pero lamentablemente, como en otras épocas, estamos caminando hacia un esquema unilateral del que son prueba lo que está ocurriendo con la Rusia de Putin o con el *Brexit* en el Reino Unido. China es ambivalente, lo que se demuestra en que en unos ámbitos se ha integrado perfectamente (la Organización Mundial del Comercio, por ejemplo), mientras en otros muestra una tendencia no ya unilateralista, sino de autoritarismo

<sup>285</sup> Poema recogido por M. Shelley en la primera página de su obra *Frankenstein*.

creciente, como ilustra la autoproclamación vitalicia de su actual presidente.

Todo ello son indicios claros de cuál va a ser el modelo futuro: unilateralismos de las grandes potencias. Sin embargo, en algunos ámbitos a nivel internacional hay una sorprendente cooperación que ha hecho avanzar un sistema global integrador. Bajo diferentes presidencias estadounidenses y rusas se fueron consolidando procedimientos de cooperación como la OMC o Kioto. Es también ejemplar el modelo de cooperación que existe en campos colindantes a los LAWS, como el nuclear, donde funciona un procedimiento deliberativo que ha dado lugar a tratados e instituciones colocados bajo el auspicio de Naciones Unidas, tales como el Organismo Internacional de Energía Atómica (OIEA). Otro modelo exitoso de cooperación e integración multilateral son los proyectos relacionados con la navegación y los satélites espaciales. Después de una inicial carrera de confrontación entre Estados Unidos y Rusia, lo que encontramos en la actualidad es una casi total cooperación en los distintos programas y proyectos e incluso en la Estación Espacial Internacional. En ella cooperan Estados Unidos, Rusia, Europa y Japón, entre otros. La NASA ha dicho que «la Estación Espacial Internacional ha sido el más complejo programa de exploración del espacio hasta ahora nunca acometido [...] visitada por astronautas de 18 países».<sup>286</sup> El modelo de *international cooperation in space: now more than ever* debería servirnos de guía para implantar los mismos procedimientos de negociación en la regulación de los LAWS.<sup>287</sup> Sin embargo, no está sucediendo así.

Es absolutamente preciso que unas administraciones estatales e internacionales sólidas protejan el derecho a la vida

<sup>286</sup> NASA (2017).

<sup>287</sup> Younes (2018).

y a la dignidad de los ciudadanos vetando una amenaza clarísima como es la independencia total de los LAWS. Ha de descartarse el principio neoliberal de *más mercado y menos Estado* (menos regulación, menos administraciones públicas e instituciones internacionales, menos impuestos, etcétera), que conduce a una menor seguridad nacional y global de todo punto indeseables. Frente al egoísmo y la tentación unilateral de cada Estado, debe ventearse una pavorosa realidad: los LAWS provocan el *síndrome de rehenes mutuos* que ya operó en materia nuclear después de los accidentes de Three Mile Island, Chernóbil y Fukushima. Es decir, los errores e negligencias de un Estado en materia de LAWS pueden desencadenar una escalada de desastres descontrolada en todo el mundo. En consecuencia, a todos los Estados les interesa que nadie esté programando en secreto un proyecto de este tipo. Se diga lo que se diga, los LAWS no están exentos de cometer errores y provocar la muerte de seres inocentes, ya que pueden desviar sus objetivos por causa de una situación fortuita (por ejemplo, una alteración del funcionamiento del sistema por recalentamiento) no prevista por los científicos o por empresarios irresponsable.

En todo caso, la tecnología militar debería quedar bajo el control y la participación accionarial (*golden share*) del Estado y no librada a un mercado privado donde se compran y se venden *sets* tecnológicos que amenazan a la soberanía y a la seguridad nacionales debido a la corrupción cleptocrática que denuncia Chayes.<sup>288</sup>

En un planeta plagado de amenazas nucleares o bioquímicas en manos de potencias desestabilizadoras no democráti-

<sup>288</sup> Rasor y Bauman (2007); Chayes (2015). El libro de este último sobre corrupción demuestra que en los Estados Unidos aquella se centra en el sector industrial-financiero-militar.

cas y corruptas o en no-Estados (como Dáesh) y en el que las armas nucleares pueden llegar a entrar en el mercado negro y a pasar a manos de grupos terroristas,<sup>289</sup> la aparición de los LAWS hace prácticamente inevitable una carrera armamentística global, así como el desencadenamiento de un genocidio selectivo o general.

En la actualidad, se hace evidente que sufrimos una erosión de paradigmas que parece conducirnos al surgimiento de una nueva revolución científica,<sup>290</sup> y ello nos obliga a recuperar el derecho como herramienta para hacer frente a los desafíos futuros y la ética como fuente de regeneración social.

En consecuencia, y teniendo en cuenta el último informe de International Human Rights Clinic (IHRG) y la Harvard Law School, titulado «Heed the call: a moral and legal imperative to ban killer robots» y publicado en agosto de 2018,<sup>291</sup> urge llevar a cabo acciones simultáneas de control y regulación de este tipo de armas, tales como:

- Definir claramente el concepto de *sistema letal completamente autónomo*.
- Reiterar el principio general de que todos los sistemas de armas deben respetar el derecho internacional humanitario, el principio de distinción y el de proporcionalidad, y siempre con control humano suficiente.
- Firma de convenios internacionales de control de armamento y prohibición de la investigación y el desarrollo

<sup>289</sup> Hass (2014).

<sup>290</sup> Ravetz (1971) y Kuhn (2011).

<sup>291</sup> Bonnie Docherty, redactora principal del informe, es investigadora en la división de armas de HRW y miembro del IHRG junto a Steve Goose, director de la división de armas, y Mary Wareham, directora jurídica, que han sido los editores del informe y son una representación de los más activos expertos en la materia.

de dichos sistemas, tal como se hizo con la proliferación de armas nucleares o químicas.<sup>292</sup>

- Firma de convenios internacionales de verificación de la no experimentación con armas con autonomía letal total.
- Firma de convenios internacionales sobre la compatibilidad con el derecho internacional humanitario del desarrollo o adquisición de armas autónomas con control humano, en cumplimiento del artículo 36 del Protocolo Adicional I de los Convenios de Ginebra.
- Aprobación de leyes estatales para restringir dicha experimentación e innovación en los centros privados en estos temas bajo sanciones administrativas y penales.
- Aprobación de leyes estatales para establecer la obligación de disponer de inspectores y delegados de *compliance* en los centros de experimentación e innovación de inteligencia artificial.

En este sentido, cabe destacar que España (son palabras de Julio Herraiz, embajador de España ante la ONU, en la Conferencia de Desarme de la Convención sobre Ciertas Armas Convencionales, celebrada en Ginebra el 13 de noviembre de 2017).

apoya la realización de medidas voluntarias de fomento de la confianza y la transparencia en todos los aspectos relativos a las armas autónomas letales, así como un intercambio de información amplio sobre esta materia. Dicho intercambio de información podría tener lugar tanto en relación con contenidos sustantivos y mejores prácticas identificadas en los exámenes legales

<sup>292</sup> Revill (2017). Sparrow señala que «un tratado de control de armamento que prohibiera armas autónomas podría representar la única manera de prevenir su desarrollo». Véase también Meier (2016).

de armas bajo el artículo 36 como en relación a otros aspectos regulatorios o técnicos de la investigación o el desarrollo operativo de posibles sistemas con autonomía.

En todo caso, este tipo de medidas voluntarias es totalmente insuficiente. Es necesario desarrollar nuevas formas de cooperación tecnológico-jurídica que eviten que se pueda avanzar en experimentación sin que ello se sujete a una regulación preventiva y a una evaluación de impacto inmediata, teniendo como principio central el de precaución. En lo que respecta a estos avances acometidos por empresas privadas, generalmente con subvenciones públicas, todo protocolo y toda regulación deben ser pocos si afectan a la seguridad. Debería haber incluso un inspector especializado facultado para valorar el impacto de la regulación o la *arregulación* en los experimentos de alto riesgo. Pero es que con el escándalo del *Dieseldgate* quedó demostrado que ni siquiera tales controles garantizan que la corrupción no exista, pues existió entre el gobierno de Baja Sajonia y Volkswagen.<sup>293</sup>

En la actualidad, atravesamos un estadio de la historia en el que la confrontación ya no es siempre entre Estados o bloques (ruso, chino, norteamericano, etcétera), sino también entre grandes empresas privadas que poseen cierto control sobre áreas que antes recaían en su totalidad bajo control estatal. Los países democráticos, que es en los que esa libertad permite a empresas privadas estudiar las tecnologías letales, el control y la regulación, deben penetrar en los mismos centros de investigación en los que han de surgir dichas innovaciones y debe crearse un procedimiento interno de *compliance* o cumplimiento normativo.<sup>294</sup>

<sup>293</sup> Ewing (2017).

<sup>294</sup> Cherer (2016).



Hace unos años, centenares de científicos expertos en tecnología y especialistas en inteligencia artificial firmaron un manifiesto contra las armas autónomas, a las que definían como «la tercera revolución en la historia de la guerra, después de la pólvora y las armas nucleares». Entre los firmantes figuraban Stephen Hawking, Elon Musk y Steve Wozniak, cofundador de Apple. Pero el verdadero motor de cambio está en la sede de Naciones Unidas, donde desde 2013, y sobre todo durante 2018, se están debatiendo, a partir del Convenio sobre Ciertas Armas Convencionales (CCAC), los desafíos de los LAWS.

Con esta obra no se busca paralizar el desarrollo de la inteligencia artificial, que ofrece pese a todo, y lo ofrece ya, un gran potencial de mejora de nuestras condiciones de vida, sino solo acotar sus ámbitos de aplicación y animar a la comunidad internacional a emprender un debate al respecto y apostar por un abordaje preventivo de la cuestión.<sup>295</sup> El derecho debe adquirir conciencia de su importancia como herramienta de prevención: podríamos estar hablando de la posibilidad de evitar un genocidio perpetrado por los LIWS, como hemos bautizado a los hipotéticos sistemas de armamento letal independiente que puede depararnos el futuro, y que deben ser prohibidos tajantemente ya desde su misma fase de experimentación por conculcar los principios generales del derecho internacional humanitario y del derecho administrativo mismo.<sup>296</sup>

Los LAWS podrían llegar a erigirse en monstruos que arruinen la vida de su creador, como Frankenstein en el relato clásico de Mary Shelley. De Shelley se ha dicho que, brillantemente ya en 1818, fue capaz de alertar y crear una alegoría

<sup>295</sup> Meza (2016).

<sup>296</sup> Criddle (2016).

de la perversión a que puede conducir el desarrollo científico. Ahora podemos hacer un paralelismo y pronosticar la misma perversión con la creación de los LAWS/LIWS. Efectivamente, con un capitalismo permisivo y una arregulación sobre los límites de la inteligencia artificial general podemos provocar el surgimiento de un nuevo monstruo. Esto que se dice en la película *Frankenstein*, de 1994 (guión de Kenneth Branagh), es perturbadoramente válido para describir palabra por palabra lo que representan los LAWS: «No, no es imposible. Podemos hacerlo. Estamos a un paso. Y si podemos cambiar una parte de un ser humano, podemos cambiar todas las partes. Y si podemos hacer esto, también podemos diseñar una vida. Podemos crear un ser que no envejezca nunca, ni enferme, que será más fuerte que nosotros y mejor que nosotros, más inteligente y más civilizado que nosotros».<sup>297</sup>

El futuro de los LAWS oscilará previsiblemente entre el modelo de colaboración público-privada sin restricciones en aras de una pretendida carrera armamentística irremediable y el de intervencionismo administrativo (nacional e internacional) regulador preventivo, restrictivo y basado en un procedimiento de *compliance* que limite la investigación a usos defensivos y siempre con control humano significativo contando, eso sí, con la ayuda de alguna agencia de trayectoria reputada, como la del Organismo Internacional de Energía Atómica. Si optamos por el primer modelo, la consecuencia será paradójicamente la desprotección nacional e internacional y una nueva propagación de estas minas vagantes llamadas LAWS. El físico Stephen Hawking señalaba que «el desarrollo de la inteligencia artificial total podría sig-

<sup>297</sup> Es interesante destacar la trayectoria de Shelley, hija de la famosa feminista, filósofa y escritora Mary Wollstonecraft y del filósofo anarquista William Godwin. Véanse Mellor (1989) y St. Clair (1991).

nificar el final de la raza humana»,<sup>298</sup> y el general estadounidense del aire Robert Latiff también advierte que la guerra del futuro está por llegar y que el despliegue de una tecnología nueva insuficientemente considerada puede tener consecuencias devastadoras no deseadas.<sup>299</sup> Harshaw, por su parte, advierte de que

Las armas autónomas letales amenazan con convertirse en la tercera revolución en la guerra. Una vez desarrolladas, permitirán que el conflicto armado se dispute a una escala mayor que nunca, y en escalas de tiempo más rápidas de lo que los humanos pueden comprender. Estas pueden ser armas de terror; armas que los déspotas y los terroristas usan contra poblaciones inocentes, y armas pirateadas para comportarse de maneras no deseadas. No tenemos mucho tiempo para actuar. Una vez se abra esta caja de Pandora, será difícil cerrarla.<sup>300</sup>

El principio de la soberanía dual desarrollado por la jurisprudencia de la Corte Suprema americana (desde el juez Marshall en 1819) establece que Washington no puede monopolizar el poder ni tomar determinaciones unilaterales, sino que debe compartir tales poderes con sus cincuenta estados.<sup>301</sup> Para ser consecuentes con este mandato interno y con sus valores constitucionales *ad intra* y *ad extra*, en la globalización también debe imperar el principio de que Estados Unidos no puede hacer unilateralismo sobre el mundo y debe contar con el resto de los casi doscientos Estados.

En nuestro caso, en la arquitectura constitucional global a que aboca la globalización-americanización, Naciones Unidas debe propugnar un acuerdo multilateral a través de un

<sup>298</sup> Coglianese y Lehr (2017).

<sup>299</sup> Latiff (2017).

<sup>300</sup> Harshaw (2017).

<sup>301</sup> Ballbé y Martínez (2003).

proceso deliberativo en el que se busquen soluciones de consenso para frenar este nuevo monstruo al que se le quiere dar autonomía e independencia.

«Cuesta verlo. Siempre en movimiento, el futuro», decía Yoda.<sup>302</sup>

No figura  
en la  
bibliografía  
con ese año.

<sup>302</sup> Sunstein (2017).

## Bibliografía

- ACKERMAN, B. (2005): *Before the next attack: preserving civil liberties in an age of terrorism*, New Haven (Estados Unidos): Yale University Press.
- AGAMBEN, G. (2004): *Estado de excepción: Homo Sacer, II*, Valencia: Pre-Textos.
- AGOR, W. H. (1989): *Intuition in organizations: leading and managing productively*, Newbury Park (Estados Unidos): Sage.
- ALANDETE, D. (2016): «Necrológica: John McCarthy, el arranque de la inteligencia artificial», *El País*, 27 de octubre de 2011.
- ALEMÁN, D. (2016): *Inteligencia basada en efectos para la seguridad humana de las Naciones Unidas*, tesis defendida el 24 de noviembre de 2016 en la Universitat Autònoma de Barcelona en el marco del programa de doctorado de derecho público global; codirectores BALLBÉ y FERNÁNDEZ PEREIRA (en prensa). <<https://www.tdx.cat/handle/10803/400082>>.
- ALLISON, G. (2004): *Nuclear terrorism: the ultimate preventable catastrophe*, Nueva York: Times.
- (2015) «The Thucydides Trap: are the US and China headed for war?», *The Atlantic*, 24 de septiembre [en línea], <<https://www.theatlantic.com/international/archive/2015/09/united-states-china-war-thucydides-trap/406756/>>. [Consulta: 7-9-2018].
- (2017) *Destined for war: can America and China escape Thucydides's trap?*, Boston (Estados Unidos): Mariner Books.
- ALMEIDA LENARDON, J. P. (2017): *The regulation of artificial intelligence*, trabajo de fin de máster, Tilburg Institute of Law, Technology and Society, 14 de junio.
- ÁLVAREZ, J. (2005): *International organizations as law-makers*, Nueva York: Oxford University Press.
- (2009): «Contemporary International Law: an Empire of law or the law of Empire?», *American University International Law Review*, 24.

- AMOROSO, D. (2017): «*Jus in bello* and *jus ad bellum* arguments against autonomy in weapons systems: a reappraisal», *Questions of International Law*, 4(43), pp. 5-31.
- y G. TAMBURRINI (2017): «The ethical and legal case against autonomy in weapons systems», *Global Jurist*, 22 de septiembre de 2017 [en línea], <<https://doi.org/10.1515/gj-2017-0012>>. [Consulta: 14-8-2018].
- ANDERSEN-RODGERS, D. y K. CRAWFORD (2018): *Human security: theory and action. Peace and security in the 21st century*, Lanham (Estados Unidos): Rowman and Littlefield.
- ARENDT, H. (2015): *La condición humana*, Barcelona: Paidós (8.<sup>a</sup> edición).
- ASARO, P. (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, R. (1984): *The evolution of co-operation*, Nueva York: Basic.
- AXELROD, R. y W. HAMILTON (1981): *The evolution of cooperation. Science*, 211, 4489, pp. 1390-1396.
- AYRES, I. y J. BRAITHWAITE (1992): *Responsive regulation: transcending the deregulation*, Oxford University Press.
- BALKO, R. (2014): *Rise of the warrior cop: the militarization of America's police force*, Nueva York: Public Affairs.
- BALLBÉ, M. (1985): *Orden público y militarismo en España*, Madrid: Alianza.
- (2005): «Seguridad humana: del Estado anómico al Estado regulador», prólogo a C. Hood *et alii*: *El gobierno del riesgo*, Barcelona: Ariel.
- e Y. CABEDO (2013): «La necesidad de administraciones reguladoras en Latinoamérica para prevenir los ataques especulativos, los riesgos financieros y la defensa de los consumidores», *Revista del CLAD Reforma y Democracia*, 57.
- y R. MARTÍNEZ (2003): *Soberanía dual y Constitución integradora*, Barcelona: Ariel.
- y R. MARTÍNEZ (2009): «Law and globalization: between the United States and Europe», en J. ROBALINO y J. RODRÍGUEZ-ARANA (eds.): *Global administrative law*, Londres: Cameron May.
- y C. PADRÓS (1997): *Estado competitivo y armonización europea: los modelos norteamericanos y europeos de integración*, Barcelona: Ariel.
- BARNETT, K. H. (2018): «The politics of selecting Chevron Deference», *Journal of Empirical Legal Studies*, 15, 3, p. 597.

- BARRET, L. (2016): «Deconstructing data mining: protection privacy and civil liberties in automated decision-making», *Geo L.Tech. Rev.*, 1253.
- BARRAT, J. (2013): *Our final invention: artificial intelligence and the end of the human era*, Nueva York: St. Martin's Press.
- BARZELAY, M. y C. CAMPBELL (2003): *Preparing for the future: strategic planning in the US Air Force*, Washington DC: Brookings.
- BATES, T. (1975): «Gramsci and the theory of hegemony», *Journal of the History of Ideas*, 32(2), pp. 351-356.
- BAUDRILLARD, J. (1987): *Cool memories (1980-1985)*, París: Galilée.
- (1994): *Simulacra and simulation*, Detroit (Estados Unidos): University of Michigan Press.
- BEEBE, S. D. y M. KALDOR (2010): *The ultimate weapon is no weapon: human security and the new rules of war and peace*, Nueva York: Public Affairs.
- BEHLING, O. y N. L. ECKEL (1991): «Making sense out of intuition», *The Executive*, 5(1), pp. 46-54.
- BEINE, M. y C. PARSONS (2015): «Climatic factors as determinants of international migration», *The Scandinavian Journal of Economics*, 117(2), pp. 723-767.
- BELTRÁN DE FELIPE, M. (1995): *Discrecionalidad administrativa y Constitución*, Madrid: Tecnos.
- BENJAMIN, M., (2012): *Drone warfare: killing by remote control*, Nueva York: OR Books.
- BENSON, G. (2017): «Governing AI: can regulators control artificial intelligence?», *Raconteur*, 6 (octubre) [en línea], <<https://www.raconteur-net/technology/governing-ai-can-regulators-control-artificial-intelligence>>. [Consulta: 14-8-2018].
- BENTHAM, J. (2017): *El panóptico*, Buenos Aires: Quadrata.
- BERNSTEIN, P. L. (1996): *Against the gods: the remarkable story of risk*, Nueva York: Wiley John and Sons.
- BHUTA, N.; S. BECK, R. GEIB, H. LIU y C. KREB (eds.): *Autonomous weapons systems: law, ethics, policy*, Cambridge University Press.
- BOMBOY, S. (2018): «The space force and the Constitution», *National Constitution Center*, 22 de agosto [en línea], <<https://constitutioncenter.org/blog/the-space-force-and-the-constitution>>. [Consulta: 7-9-2018].
- BOTHMER, F. von (2014): «Robots in court: responsibility for lethal autonomous weapons systems», en S. BRÄNDLI, R. HARASGAMA, R. SCHISTER y A. TAMÒ (eds.): *Mensch und Maschine – Symbiose oder Parasitismus?*, Berna: Stämpfli.

- BOSTROM, N. (2005): «A history of transhumanist thought», *Journal of Evolution and Technology*, 14(1), pp. 1-25.
- BOWLES, S. (2016): *The moral economy: why good incentives are no substitute for good citizens*, New Haven (Estados Unidos): Yale University Press.
- BRAIDOTTI, R. (2015): *Lo posthumano*, Barcelona: Gedisa.
- BRAITHWAITE, J. (2017): «Responsive excellence», en C. COGLIANESE (ed.): *Achieving regulatory excellence*, Washington DC: Brookings.
- BREUIL, H. (1952): *Four hundred centuries of cave art*, Montignac (Francia): Centre d'Études et de Documentation Préhistoriques.
- BURRI, T. (2016): «The politics of robot autonomy», *EJRR*, 2, pp. 341-260.
- BATTISTUTTI, CAIRO, O. (2011): *El hombre artificial: el futuro de la tecnología*, México DF: Alfaomega.
- CHALMERS, D. (2010): «The singularity: a philosophical analysis», *Journal of Consciousness Studies*, 17(9-10), pp. 7-65.
- CHAYES (2015): *Thieves of State: why corruption threatens global security*, Nueva York: Norton and Company.
- CHUNG, A. (2018): «US Supreme Court restricts deportations of immigrant felons», 18 de abril [en línea], <<https://uk.reuters.com/article/uk-usa-court-deportation/u-s-supreme-court-restricts-deportations-of-immigrant-felons-idUKKBN1HO3DQ>>. [Consulta: 22-8-2018].
- COAESE, R. (2013): *How China became capitalist*, Nueva York: Palgrave-McMillan.
- COGLIANESE, C. y D. LEHR (2017): «Regulating by robot: administrative decision making in the machine-learning era», *Georgetown Law Journal*, 105:6, pp. 1147-1223.
- COLLI, G. (1978): *Después de Nietzsche*, Barcelona: Anagrama.
- CONGER, K. (2018): «Google plans not to renew its contract for Project Mawen, a controversial Pentagon drone AI imaging program», *Gizmodo*, 6 de enero.
- CSÁNYI, V. (1989): *Evolutionary systems and society: a general theory of life, mind and culture*, Londres: Duke University Press.
- CULKIN, J. M. (1967): *A schoolman's guide to Marshall McLuhan*, Saturday Review.
- DANE, E. y M. G. PRATT (2007): «Exploring intuition and its role in managerial decision making», *Academy of Management Review*, 32(1), pp. 33-54.
- DATTERI, E. y G. TAMBURRINI (2013): «Robotic weapons and democratic decision-making», en E. HILGENDORF y J.-P.



- GUENTHER (eds.): *Robotik und Gesetzgebung*, Baden-Baden (Alemania): Nomos.
- DANIEL, J. (2017): «Scrutinizing deference to administrative agencies», *The Regulatory Review*, 27 de noviembre.
- DANZIG, R. (2018): «Technology roulette: managing loss of control as many militaries pursue Technological superiority», Centre for a New American Security [en línea], <<https://s3.amazonaws.com/files.cnas.org/documents/CNASReport-Technology-Roulette-DoSproof2v2.pdf>>. [Consulta: 7-9-2018].
- DEEP MIND (2017): «AlphaGo Zero: learning from scratch» [en línea], <<http://deepmind.com/blog/alphago-zero-learning-scratch/>>. [Consulta: 14-8-2018].
- DEVLIN, H.: «AI programs exhibit racial and gender biases, research reveals», *The Guardian*, 13 de abril de 2017 [en línea], <<https://www.theguardian.com/technology/2017/apr/13/ai-programs-exhibit-racist-and-sexist-biases-research-reveals>>. [Consulta: 14-8-2018].
- DIAMOND, J. (2015): *Colapso: por qué unas sociedades perduran y otras desaparecen*, Barcelona: Debolsillo.
- (2016): *Gérmenes, armas y acero: breve historia de la humanidad en los últimos trece mil años*, Barcelona: Debate.
- DODSON, S. (2008): «Was software responsible for the financial crisis?», *The Guardian*, 16 de octubre [en línea], <<https://www.theguardian.com/technology/2008/oct/16/computing-software-financial-crisis>>. [Consulta: 14-8-2018].
- DOMINGOS, P. (2013): *A few useful things to know about machine learning*, Seattle (Estados Unidos): University of Washington.
- DORF, M. (2018): «Is the space force constitutional?», *The blog of the Center for the study of Constitutional originalism at the University of San Diego School of law*, 20 de agosto [en línea], <<http://originalismblog.typepad.com/the-originalism-blog/2018/08/is-the-space-force-constitutionalmichael-ramsey.html>>. [Consulta: 7-9-2018].
- DROHAN, M. (2003): *Making a killing: how corporations use armed force to do business*, Nueva York: Random House.
- DURDEN, T. (2018): «China accelerates next-gen nuclear weapons development to compete with US, Russia», *Zero Hedge*, 29 de mayo [en línea], <<https://www.zerohedge.com/news/2018-05-28/china-accelerates-next-gen-nuclear-weapons-development-compete-us-russia>>. [Consulta: 7-9-2018].
- DWORKIN, R. (1986): *Law's Empire*, Cambridge (Estados Unidos): Harvard University Press.

- DYSON, G. (2013): *Turing's cathedral: the origins of the digital universe*, Nueva York: Vintage.
- ELLUL, J., J. WILKINSON y R. MERTON (1964): *The technological society*, Nueva York: Random House.
- EPP, C. (2014): *La revolución de los derechos*, Madrid: Siglo XXI.
- ESPÍN, R. (2017): *El canal de denuncias internas en la actividad empresarial como instrumento del «compliance»* (tesis doctoral; dir. Manuel BALLBÉ), Universitat de Barcelona.
- ESSINGER, J. (2015): *Ada's algorithm: how Lord Byron's daughter Ada Lovelace launched the digital age*, Nueva York: Melville House.
- ETZIONI, A. y O. ETZIONI (2017): «Pros and cons of autonomous weapons systems», *Military Review* (mayo-junio), pp. 72-81.
- FASSBENDER, B. (1998): «The UN Charter as Constitution of the International Community», 36 *Colum.J.Transnat'l L.*, 529.
- FAST, S. (2015): *Facebook/panopticon: an analysis of Facebook and its parallels to the Foucaultian panopticon*, Chattanooga (Estados Unidos): University of Tennessee.
- FEMIA, J. (1987): *Gramsci's political thought: hegemony, consciousness and the revolutionary process*, Londres: Clarendon (4.ª edición).
- FERNÁNDEZ PEREIRA, J. P. (2006): *Seguridad humana*, Universitat Autònoma de Barcelona.
- FOUCAULT, M. (1990): *Vigilar y castigar: nacimiento de la prisión*, Madrid: Siglo XXI.
- (1997): *Las palabras y las cosas: una arqueología de las ciencias humanas*, Madrid: Siglo (3.ª edición).
- y J. VARELA (1978): *Microfísica del poder*, Madrid: Endymion.
- FRANTZ, D. y C. COLLINS (2007): *The nuclear jihadist*, Nueva York: Twelve.
- FRAZER, J. (1951): *La rama dorada: magia y religión*, México DF: Fondo de Cultura Económica (1.ª edición).
- FREUD, S. (1999): *El malestar en la cultura*, Madrid: Alianza.
- FRISCHMANN, B. y E. SELINGER (2018): *Re-engineering humanity*, Cambridge University Press.
- FUERTH, L. S. y E. M. FABER (2013): «Anticipatory governance: winning the future», *Futurist*, 47(4), pp. 42-49.
- FUTURE OF LIFE INSTITUTE (2018): «Benefits and risks of artificial intelligence» [en línea], <<https://futureoflife.org/background/benefits-risks-of-artificial-intelligence/>>. [Consulta: 14-8-2018].
- GARCÍA DE ENTERRÍA, E. (1983): *La lucha contra las inmunidades del poder*, Madrid: Civitas.

- GIBBS, S. (2018): «Google's AI is being used by US military drone programme», *The Guardian*, 7 de marzo.
- GILOVICH, T., D. GRIFFIN y D. KAHNEMAN (eds.) (2002): *Heuristics and biases: the psychology of intuitive judgment*, Cambridge University Press.
- GIDDENS, A. (1999): *Consecuencias de la modernidad*, Madrid: Alianza.
- GOERTZEL, B. y C. PENNACHIN (eds.) (2007): *Artificial general intelligence*, Nueva York: Springer.
- GÓMEZ DE HINOJOSA, A. (2018): *Seguridad privada global: ¿amenaza u oportunidad en el marco de las Naciones Unidas?*, tesis doctoral depositada el 5 de septiembre en la UAB e inscrita en el marco del programa de doctorado Seguridad Humana y Derecho Global del Departamento de derecho público y CHJ de la UAB (dir. M. BALLBÉ).
- GOODMAN, M. (2013): *National insecurity: the cost of American militarism*, San Francisco (Estados Unidos): City Lights.
- GRAMSCI, A. (1995): *Further selections from the prison notebooks*, Saint Paul (Estados Unidos): University of Minnesota.
- (1999): *Cuadernos desde la prisión*, México DF: Era, tomo 2 (ed. crítica del Instituto Gramsci a cargo de Valentino GERRATANA).
- HARAWAY, D. (2013): *Simians, cyborgs and women: the reinvention of nature*, Londres: Routledge.
- HARSHAW, T. (2017): «Nobody's ready for the killer robot. A Q&A with General Robert Latiff on the ethics of warfare in the autonomous future», *Bloomberg*, 30 de diciembre [en línea], <<https://www.bloomberg.com/view/articles/2017-12-30/nobody-s-ready-for-the-killer-robot>>. [Consulta: 14-8-2018].
- HASS, M. (2014): «Autonomous weapon systems: the military's smartest toys?», *The National Interest*, 20 de noviembre [en línea], <<http://nationalinterest.org/feature/autonomous-weapon-systems-the-militarys-smartest-toys-11708>>.
- HEGEL, G. W. F. (1966): *Fenomenología del espíritu*, Buenos Aires: Fondo de Cultura Económica.
- HERMAN, E. y N. CHOMSKY (2010): *Manufacturing consent: the political economy of the mass media*, Londres: Random House (2.<sup>a</sup> edición).
- HEYNS, C. (2013): *Report of the special rapporteur on extrajudicial summary or arbitrary execution*, vigésimo tercera sesión del Consejo de los Derechos Humanos.
- HEYNS, C. (2016): «Autonomous weapons systems: living a dignified life and dying a dignified death», en N. BHUTA, S. BECK, R. GEIB,

- H. LIU y C. KREB (eds.): *Autonomous weapons systems: law, ethics, policy*, Cambridge University Press.
- HOLLINGS, C. y U. MARTIN (2018): *Ada Lovelace: the making of a computer scientist*, University of Oxford.
- HOROWITZ, M. C. (2016): «Public opinion and the politics of the killer robots debate», *Research & Politics*, 3(1), pp. 1-8.
- , S. E. KREPS y M. FUHRMANN (2016): «Separating fact from fiction in the debate over drone proliferation», *International Security*, 41(2), pp. 7-42.
- HOWARD, A., C. ZHANG y E. 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 (Estados Unidos), 8 al 10 de marzo de 2017.
- HU, Ying, (2018): «Robot Criminals», *University of Michigan Journal of Law Reform* (en prensa).
- IGNATIUS, D. (2018): «The Chinese threat that an aircraft carrier can't stop», *The Washington post*, 7 de agosto.
- ITU Y XPRIZE (2017): *AI for Good global summit report* [en línea], <[https://www.itu.int/en/ITU-T/AI/Documents/Report/AI\\_for\\_Good\\_Global\\_Summit\\_Report\\_2017.pdf](https://www.itu.int/en/ITU-T/AI/Documents/Report/AI_for_Good_Global_Summit_Report_2017.pdf)>. [Consulta: 14-8-2018].
- JAMESON, F. (1991): *El posmodernismo o la lógica cultural del capitalismo avanzado*, Barcelona: Paidós.
- JANASHVILI, L. (2018): *Human trafficking: la nueva esclavitud* (en prensa).
- JHA, U. C. (2016): *Killer robots: lethal autonomous weapons systems. Legal, ethical and moral challenges*, Nueva Delhi: Vij.
- JONÁS, H. (2004): *El principio de responsabilidad (ensayo de una ética para la civilización tecnológica)*, Barcelona: Herder.
- KAHNEMAN, D. (2015): *Pensar rápido, pensar despacio*, Barcelona: Debate.
- y G. KLEIN (2009): «Conditions for intuitive expertise: a failure to disagree», *American Psychologist*, 64(6).
- KALDO, M. (2001): *Las nuevas guerras*, Barcelona: Busquets.
- KANG, C. y K. ROOSE: «Zuckerberg faces hostile Congress as calls for regulation mount», *The New York Times*, 11 de abril de 2018 [en línea], <<https://www.nytimes.com/2018/04/11/business/zuckerberg-facebook-congress.html>>. [Consulta: 21-8-2018].
- KANT, I. (1998): *Critique of pure reason*, Nueva York: Kosimo.
- KARLSRUD, J. (2018): *The UN at war: peace operations in a new era*, Nueva York: Palgrave MacMillan.

- KAPLAN, J. (2016): *Artificial intelligence*, Oxford University Press.
- KHATRI, N. y H. A. NG (2000): «The role of intuition in strategic decision making», *Human Relations*, 53(1), pp. 57-86.
- KLAY, P. (2016): «The citizen soldier: moral risk and the modern military», *Brookings*, 24 de mayo [en línea], <<http://csweb.brookings.edu/content/research/essays/2016/the-citizen-soldier.html>>. [Consulta: 14-8-2018].
- KLEIN, G. (1999): *Sources of power: how people make decisions*, Cambridge (Estados Unidos): MIT Press.
- (2004): *The power of intuition*, Cambridge (Estados Unidos): MIT Press.
- KNAPP, B. (2018): «A new council could advance artificial intelligence for the military», *C4IRSNET*, 27 de abril [en línea], <<https://www.c4isrnet.com/intel-geoint/2018/04/25/a-new-council-could-advance-artificial-intelligence-for-the-military/>>. [Consulta: 14-8-2018].
- KU, J. y J. YOO (2012): *Taming globalization. International law, the US Constitution and the new world order*, Oxford University Press.
- LAHOZ-BELTRÁ, R. (2005): *Turing: del primer ordenador a la inteligencia artificial (la matemática en sus personajes)*, Tres Cantos (Madrid): Nivola.
- LATIFF, R. (2017): *Future war: preparing for the new global battlefield*, Nueva York: Knopf.
- LAVITT, J. (2010): «The crime of conviction of John Choon Yoo: the actual criminality in the OLC during the Bush Administration», *Maine Law Review*, 62:1.
- LEE, D. (2017): «Facebook adds human reviewers after ‘Jew haters’ ad scandal», *BBC News*, 20 de septiembre [en línea], <<https://www.bbc.com/news/technology-41342642>>. [Consulta: 20-8-2018].
- LEVERINGHAUS, A. (2016): *Ethics and autonomous weapons*, Madrid: Palgrave McMillan.
- LEVI, M. (2007): *On nuclear terrorism*, Cambridge (Estados Unidos): Harvard University Press.
- LEVITSKY, S. y L. WAY (2010): *Competitive authoritarianism: hybrid regimes after the Cold War*, Cambridge University Press.
- LIN, P. (2015): «Do killer robots violate human rights?», *The Atlantic*, 20 de abril.
- , R. JENKINS y K. ABNEY (2017): *Robot ethics 2.0: from autonomous cars to artificial intelligence*, Oxford University Press.
- LIU, H. (2016): «Redefining responsibility: differentiating two types of responsibility issues raised by autonomous weapons

- systems», en N. BHUTA, S. BECK, R. GEIB, H. LIU y C. KREB (eds.): *Autonomous weapons systems: law, ethics, policy*, Cambridge University Press.
- LÓPEZ-TARRUELLA, A. (ed.) (2012): *Google and the law: empirical approaches to legal aspects of knowledge-economy business models*, Nueva York: Springer.
- MADERER, J. (2017): «Daydreaming is good. It means you're smart», *Georgia Teach*, 24 de octubre [en línea], <<https://www.news.gatech.edu/2017/10/24/daydreaming-good-it-means-youre-smart>>. [Consulta: 7-9-2018].
- MAKRIDAKIS, S. (2017): «The forthcoming artificial intelligence (AI) revolution: its impact on society and firms», *Futures*, 90, pp. 46-60.
- MARIN, L. (2017): «The 'metamorphosis' of the drone: the governance challenges of drone technology in border surveillance», en D. BOWMAN, E. STOKES y A. RIP (eds.): *Embedding and governing new technologies: a regulatory, ethical and societal perspective*, Singapur: Pan Stanford.
- MARX, K. (2008): *The poverty of philosophy*, Nueva York: Cosimo.
- MADDEN, R. (2007): *Citizen soldier*, Createspace.
- MARKOFF, J. (2016): «As artificial intelligence evolves, so does its criminal potential», *The New York Times*, 23 de octubre.
- MARTÍNEZ, R. (2003): *Armas: ¿libertad americana o prevención europea?*, Barcelona: Ariel.
- (2015): «Armas en Estados Unidos: sangre y negocio», *Ahora*, 16 de octubre [en línea], <<https://www.ahorasemanal.es/armas-en-estados-unidos:-sangre-y-negocio>>. [Consulta: 7-9-2018].
- MASHAW, J. (2006): «Recovering American administrative law: federalism foundation, 1787-1801», *Yale Law Journal*, 1256.
- MASSARO, F. (2018): «Overview of the application of machine-learning in administrative law», *Administrative Conference of the United States*, 18 de julio [en línea], <<https://www.acus.gov/newsroom/administrative-fix-blog/overview-application-machine-learning-administrative-law>>. [Consulta: 17-8-2018].
- MAYER-SCHÖNBERGER, V. (2009): *Delete: the virtue of forgetting in the digital age*, Princeton (Estados Unidos): Princeton University Press.
- MBEMBE, A. (2011): *Necropolítica*, Barcelona: Melusina.
- MCQUILLAN, D. (2015): «Algorithmic states of exception», *European Journal of Cultural Studies*, 18, 4-5, pp. 564-576.
- (2018): «Manifiesto on algorithmic humanitarianism», *Open Democracy*, 4 de abril [en línea], <<https://www.opendemocracy>>.

- net/dan-mcquillan/manifesto-on-algorithmic-humanitarianism». [Consulta: 17-8-2018].
- MCMULLAN, T. (2015): «What does the panopticon mean in the age of digital surveillance. The power of privacy», *The Guardian*, 23 de julio.
- MELLOR, A. (1989): *Mary Shelley: her life, her fiction, her monsters*, Londres: Routledge.
- MEYER, J. (2007): «Supervising the Pentagon: covert action and traditional military activities in the war on terror», *Administrative Law Review*, 59, 463.
- MINOW, M. y J. FREEMAN (2009): *Government by contract: outsourcing and American democracy*, Cambridge (Estados Unidos): Harvard University Press.
- MINSKY, M. (1991): «Society of mind: a response to four reviews», *Artificial Intelligence*, 48(3), pp. 371-396.
- (2007): *The emotion machine: common sense thinking, artificial intelligence and the future of the human mind*, Nueva York: Simon & Schuster.
- MONTOBBIO, M. (2008): *Salir del callejón del gato: la deconstrucción de Oriente y la gobernanza global*, Barcelona: Icaria.
- MORALES, A. (2009): *Regulación nuclear globalizada*, Madrid: La Ley Actualidad.
- MORENO, L. y R. JIMÉNEZ (2018): *Democracias robotizadas: escenarios futuros en Estados Unidos y la Unión Europea*, Madrid: Catarata.
- MOURELLE, D. (2017): «La amenaza nuclear del siglo XXI», *El Orden Mundial*, 27 de noviembre [en línea], <<https://elordenmundial.com/2017/11/27/la-amenaza-nuclear-en-el-siglo-xxi/>>. [Consulta: 17-8-2018].
- MULLER, F. L. (2018): «Identifying German legal approaches to terror: how the Constitution shapes legislation allowing the shooting down of a hijacked plane», *German Law Journal*, vol. 19, núm. 1.
- NASA (2017): «International cooperation» [en línea], <[https://www.nasa.gov/mission\\_pages/station/cooperation/index.html](https://www.nasa.gov/mission_pages/station/cooperation/index.html)>. [Consulta: 7-9-2018].
- NAVAS, S. (ed.) (2017): *Inteligencia artificial: tecnología, derecho*, Madrid: Tirant lo Blanch.
- NIETZSCHE, F. (1989): *On the genealogy of morals and Ecce Homo* (ed. W. KAUFMAN), Nueva York: Vintage.
- NORTH, D. (1981): *Structure and change in economic history*, Nueva York: Norton.

- NYE, R. (1977): *The anti-democratic sources of elite theory: Pareto, Mosca, Michels*, Nueva York: Sage.
- OHLIN, J. D. (ed.) (2017): *Research handbook on remote warfare*, Cheltenham (Reino Unido): EE Elgar.
- O'NEIL, C. (2016): *Weapons of math destruction: how big data increases inequality and threatens democracy*, Nueva York: Broadway.
- ORTEGA, A. (2017): «La robolución llega al Parlamento Europeo», *World Economic Forum*, 1 de marzo [en línea], <<https://es.weforum.org/agenda/2017/03/la-robolucion-llega-al-parlamento-europeo>>. [Consulta: 7-9-2018].
- PAZZANESE, C. (2016): «Don't trust that algorithm: Harvard Ph.D warns of big data's dark side in 'Weapons of math destruction'», *The Harvard Gazette*, 13 de octubre [en línea], <<https://news.harvard.edu/gazette/story/2016/10/dont-trust-that-algorithm/>>. [Consulta: 17-8-2018].
- PAINE, T. (2004): *Common sense*, Peterborough (Canadá): Broadview.
- PARDO, J. (2003): «Principio de precaución. El derecho ante la incerteza científica», *Revista Jurídica de Cataluña*, 102, 3.
- PAREJO, L. (1993): *Administrar y juzgar: dos funciones constitucionales distintas y complementarias. Un estudio del alcance y la intensidad del control judicial a la luz de la discrecionalidad administrativa*, Madrid: Tecnos.
- (2016): *Lecciones de derecho administrativo*, Madrid: Tirant lo Blanch (8.ª edición).
- PARETO, V. (1991): *The rise and fall of the elites: an application of theoretical sociology*, Nueva York: Transaction.
- PARISER, E. (2011): *The filter bubble: what the Internet is hiding from you*, Londres: Penguin.
- PASCUAL CADENA, A. (2016): *El plan de prevención de riesgos penales y responsabilidad corporativa*, Barcelona: Bosch.
- PLATÓN (1998): *Apología de Sócrates* (trad. Conrado EGGERS LAN), Universidad de Buenos Aires.
- POOLE, D.; A. MACKWORTH y R. GOEBEL (1998): *Computational intelligence: a logical approach*, Oxford University Press.
- POOL, D. y A. MACKWORTH (2017): *Artificial intelligence: foundations of computational agents*, Cambridge University Press (2.ª edición).
- POSNER, R. (2005): *Preventing surprise attacks: intelligence reform in the wake of 9/11*, Lanham (Estados Unidos): Rowman and Littlefield.
- (2006): *Uncertain shield: the U. S. Intelligence system in the throes of reform*, Lanham (Estados Unidos): Rowman and Littlefield.



- POSTMAN, N. (2011): *Technopoly: the surrender of culture to technology*, Nueva York: Vintage.
- POWELL, H. J. (2013): *The President as Commander in Chief: an essay in Constitutional vision*, Durham (Estados Unidos): Carolina Academic Press.
- POWER, S. (2008), *Chasing the flame: Sergio Vieira de Mello and the fight to save the world*, Nueva York: Penguin.
- (2013) *A problem from hell: America and the age of genocide*, Nueva York: Basic Books.
- PRIGOGINE, I. (1984): *Order out of chaos*, Nueva York: Bantam.
- (1996): *El fin de las certidumbres*, Barcelona: Taurus.
- QUINTERO, G. (2017): «La robótica ante el derecho penal: el vacío de respuesta jurídica a las desviaciones incontroladas», *Revista Electrónica de Estudios penales y de la seguridad*, 1.
- RAKOVE, J. N. (1990): *Original meanings: politics and ideas in the making of the Constitution*, Nueva York: Vintage.
- RAMÍREZ, D. (2014): «Cuando la decisión de atacar es resultado de un algoritmo», documento de análisis núm. 56 del Instituto Español de Estudios Estratégicos.
- RAMSEY, M. (2018): «Is the space force constitutional?», *The blog of the Center for the study of Constitutional originalism at the University of San Diego School of law*, 20 de agosto [en línea], <<http://originalismblog.typepad.com/the-originalism-blog/2018/08/is-the-space-force-constitutionalmichael-ramsey.html>>. [Consulta: 7-9-2018].
- RAPPAPORT, M. (2007): «Is an independent air force constitutional?», *The Right Coast*, 30 de enero [en línea], <[http://rightcoast.typepad.com/rightcoast/2007/01/is\\_an\\_independe.html](http://rightcoast.typepad.com/rightcoast/2007/01/is_an_independe.html)>. [Consulta: 7-9-2018].
- RASOR, D. y R. BAUMAN (2007): *Betraying our troops: the destructive results of privatizing war*, Nueva York: St. Martin's.
- RAUSTIALA, K. (2009): *Does the Constitution follow the flag? The evolution of territoriality in American law*, Oxford University Press.
- RAWORTH, K. (2018), *Economía rosquilla: 7 maneras de pensar la economía del siglo XXI*, Barcelona: Paidós.
- REES, J. (1994): *Hostages of each other*, The University of Chicago Press.
- REJÓN, I. (2016): «Unos servicios digitales para una sociedad digital: la revolución del Internet de las cosas», *Revista del Instituto de Estudios Económicos*, 1/2 (*La revolución digital*).
- ROCCA, A. V. (2008): *Zygmunt Bauman: modernidad líquida y fragilidad humana*, Universidad Complutense de Madrid.

- RODELLA, F. (2018): «Los robots pueden imitar, no sentir», *El País*, 28 de abril [en línea], <[https://elpais.com/elpais/2018/04/26/ciencia/1524760728\\_658244.html](https://elpais.com/elpais/2018/04/26/ciencia/1524760728_658244.html)>. [Consultar: 21-8-2018].
- RODRÍGUEZ-ARANA, J. (1993): *Principios de ética pública: ¿corrupción o servicio?*, Madrid: Montecorvo.
- RODRÍGUEZ, J. (2016): *La civilización ausente: tecnología y sociedad en la era de la incertidumbre*, Gijón: Trea.
- ROFF, H. (2016): «Meaningful human control or appropriate human judgment? Necessary limits on autonomous weapons», documento informativo preparado para la conferencia de revisión de la UNCCW.
- (2017): «Advancing human security through artificial intelligence», Londres: Chatham House.
- ROSEMBUJ, T. (2017): «Governing artificial intelligence», *LLR*, n.2/2017.
- ROSEMBUJ, T. (2018): *Inteligencia artificial e impuestos*, Barcelona: El Fisco.
- RUBIO, A. (2017): *Inversión extranjera, soberanía económica y seguridad nacional*, Pamplona: Aranzadi.
- RUSSELL, S. J.; J. STUART, P. NORVIG y E. DAVIS (2010): *Artificial intelligence: a modern approach*, Nueva York: Pearson (3.ª edición).
- SÁNCHEZ MORÓN, M. (1994): *Discrecionalidad administrativa y control judicial*, Madrid: Tecnos.
- SÁNCHEZ BARRILAO, J. F. (2016): «El derecho constitucional ante la era de Ultrón: la informática y la Inteligencia artificial como objeto constitucional», *Estudios de Deusto*, 64, 2.
- (2018): «Derecho constitucional, desarrollo informático e inteligencia artificial: aproximación del Parlamento europeo a favor de una regulación sobre robótica», en J. VALLS PRIETO (coord): *Retos jurídicos por la sociedad digital*, Pamplona: Aranzadi.
- SÁNCHEZ DEL CAMPO, A. (2016), «Cuestiones jurídicas que plantean los robots», *Revista de Privacidad y Derecho Digital*, 2.
- SANCHIS, E. (2018): «A model of free will for artificial entities», *working paper* [en línea], <<http://arxiv.org/abs/1802.09317>>. [Consulta: 7-9-2018].
- SANDOVAL, G. (2018): «After a dozen employees quit in protest, Google has reportedly decided not to renew its contract for military drone initiative Project Maven», *Business Insider*, 1 de junio [en línea], <<http://www.businessinsider.com/google-caves-to-criticism-over-military-drone-work-wont-renew-project-maven-pentagon-contract-2018-6>>. [Consulta: 17-8-2018].

- SATIA, P. (2006): «Defense of inhumanity: air control and the British idea of Arabia», *The American Historical Review*, 111, 1.
- (2009): *Spies in Arabia: the Great War and the cultural foundations of Britain's covert Empire in the Middle East*, Oxford University Press.
- SARDAR, Z. (2010): «Welcome to postnormal times», *Futures*, 42(5), pp. 435-444.
- SASSOLI, M. (2014): «Can autonomous weapon systems respect the principles of distinction, proportionality and precaution?», conferencia celebrada durante el comité de expertos *Autonomous weapons systems: technical, military, legal and humanitarian aspects* (Comité Internacional de la Cruz Roja, Ginebra [Suiza]).
- SAXON, D. (2014): «A human touch: autonomous weapons, DoD Directive 3000.09 and the interpretation of “appropriate levels of human judgment over the use of force”», *Georgetown Journal of International Affairs*, 15, 2, pp. 100-109.
- SCAHILL, J. (2008): *Blackwater: the rise of the world's most powerful mercenary army*, Londres: Profile.
- SCHARRE, P. (2018): *Army of none: autonomous weapons and the future of war*, Nueva York: Norton.
- SCHOPENHAUER, A. (1959): *The world as will and representation*, Seattle (Estados Unidos): University of Washington.
- SCHERER, M. U. (2016): «Regulating artificial intelligence systems: risks, challenges, competences and strategies», *Harvard Journal of Law and Technology*, 29, 2.
- SEYMOUR, M. (2018): *In Byron wake: the turbulent lives of Lord Byron's wife and daughter Annabella Milbanke and Ada Lovelace*, Nueva York: Simon & Schuster.
- SHARKEY, N. (2008): «The ethical frontiers of robotics», *Science*, 32(5909), pp. 1800-1801.
- (2010): «Saying – No! to lethal autonomous targeting», *Journal of Military Ethics*, 4(9), pp. 299-313.
- (2011): «The automation and proliferation of military drones and the protection of civilians», *Journal of Law, Innovation and Technology*, 3(2), pp. 229-240.
- SHELLEY, M. (1818): *Frankenstein o el moderno Prometeo*, Londres: Lackington.
- SIDDIGA, A. (2007): *Military Inc. Inside Pakistan's military economy*, Londres: Pluto.
- SINGER, P. W. (2003): *Corporate warriors: the rise of the privatized military industry*, Ithaca (Estados Unidos): Cornell University Press.

- SLAUGHTER, A. (2004): *A new world order*, Princeton (Estados Unidos): Princeton University Press.
- SLOTERDIJK, P. (2003): «El post-humanismo: sus fuentes teológicas, sus medios técnicos», conferencia pronunciada en el IV Seminario «La deshumanización del mundo. Estancias de reflexión en torno a la crisis del humanismo», celebrado entre el 6 y el 9 de mayo de 2003 en la Universidad Internacional de Andalucía (UNIA) de Sevilla.
- SMYTH, J. y B. HARRIS (2018): «Academics boycott South Korea University on killer robot fears. Kaist faces backlash over opening IA weapons lab», *Financial Times*, 4 de abril de 2018 [en línea], <<https://www.ft.com/content/6ef206e6-37d1-11e8-8b98-2f31af407cc8>>. [Consulta: 20-8-2018].
- SNOW, C. P. (2001): *The two cultures*, Londres: Cambridge University Press.
- SOLOVE, D. J. (2007): «I've nothing to hide and other misunderstandings of privacy», *San Diego L. Rev.*, 44, 745.
- SOMIN, I. (2018): «More on the debate over the Space Force and the Constitution», *The Volokh Conspiracy*, 26 de agosto [en línea], <<https://reason.com/volokh/2018/08/26/more-on-the-debate-over-the-space-force>>. [Consulta: 7-9-2018].
- SPARROW, R. (2016): «Robots and respect: assessing the case against autonomous weapon systems», *Ethics and International Affairs*, 30, 1, pp. 93-116.
- ST. CLAIR, W. (1991): *The Godwins and the Shelleys: a biography of a family*, Johns Hopkins University Press.
- SUNSTEIN, C. (ed.) (2000): *Behavioral law and economics*, Cambridge University Press.
- (2005a): *Laws of fear: beyond the precautionary principle*, Cambridge University Press.
- (2005b): «Administrative law goes to war», *Harvard Law Review*, 118, 2663-2664.
- (2006): «Clear statement principles and national security: Hamdan and beyond», *Supreme Court Review*, 1, 11.
- (2016a): *The world according to Star Wars*, Nueva York: Dey Street.
- (2016b): *La revolución de los derechos: redefiniendo el estado regulador*. Madrid: Editorial Universitaria Ramón Areces.
- TASIOULAS, J. (2018): «First steps towards an ethics of robots and artificial intelligence», *King's College London Dickson Poon School of Law Legal Studies Research Paper Series* [en línea], <[https://www.academia.edu/36555833/First\\_Steps\\_Towards\\_an\\_Ethics\\_of](https://www.academia.edu/36555833/First_Steps_Towards_an_Ethics_of)

- Robots\_and\_Artificial\_Intelligence?auto=download>. [Consulta: 7-9-2018].
- THE ECONOMIST (2017): «The world's most valuable resource is no longer oil, but data. Regulating the Internet giants», 6 de mayo [en línea], <<https://www.economist.com/news/leaders/21721656-data-economy-demands-new-approach-antitrust-rules-worlds-most-valuable-resource?fsrc=scn/tw/te/rfd/pe>>. [Consulta: 20-8-2018].
- THURNHER, J. (2013): «The law that applies to autonomous weapon systems», 17, *American Society of International Law Insights*, 4, 18 de enero.
- TOYAMA, K. (2015): *Geek heresy: rescuing social change from the cult of technology*, Nueva York: Public Affairs.
- TUCKER, J. (2012): *Innovation, dual use and security: managing the risks of emerging biological and chemical technologies*, Cambridge (Estados Unidos): MIT Press.
- TURING, A. M. (1939): «Systems of logic based on ordinals: proceedings of the London Mathematical Society», s2-45(1), pp. 161-228.
- (2009): *Computing machinery and intelligence*, en R. EPSTEIN, G. ROBERTS y G. BEBER (eds.): *Parsing the Turing test*, Dordrecht (Países Bajos): Springer Netherlands, pp. 23-66.
- VV. AA. (2012): *Estética y violencia: necropolítica, militarización y vidas lloradas*, México DF: MUAC-UNAM.
- VEBLEN, T. (1919): *The place of science in modern civilisation and other essays*, Nueva York: Huebsch.
- (1944): *Teoría de la clase ociosa*, Madrid: Ariel (3.ª edición).
- VERGOTTINI, G. de (1982): «La supremacía del poder civil sobre el poder militar en las primeras constituciones europeas», *Revista Española de Derecho Constitucional*, 6, pp. 9-34.
- VILLALBA GÓMEZ, J. A. (2016): «Problemas bioéticos emergentes de la inteligencia artificial», *Diversitas*, 12(1), 137.
- VIVES, J. (2018): «Un fabricante de armas ruso presenta un robot asesino de 4 metros de altura», *La Vanguardia*, 23 de agosto.
- VOGEL, D. (2015): *The politics of precaution: regulating health, safety and environmental risks in Europe and the United States*, Princeton University Press.
- WADDELL, J. (2018): «How algorithms can bring down minorities' credit scores», *The Atlantic*, 5 de junio [en línea], <<https://www.theatlantic.com/technology/archive/2016/12/how-algorithms-can-bring-down-minorities-credit-scores/509333>>. [Consulta: 20-8-2018].

- WALDMAN, P. (2018): *El Estado anómico: derecho, seguridad pública y vida cotidiana en América Latina*, Madrid: Iberoamericana.
- WALKER, C. (2018): «Attacking Auer and Chevron Deference: a literature review», *Georgetown Journal of Law and Public Policy*, 16, pp. 103-122.
- WARREN, S. y L. BRANDEIS (1890): «The right to privacy», *Harvard Law Review*, 4, 5, pp. 193-220.
- WARREN, A. y A. HILLAS (2017): «Lethal autonomous weapons systems. Adapting to the future of unmanned warfare and unaccountable robots», 12, pp. 71-85.
- WATERS, R. (2018a): «AI progress sparks cyber weapons fears», *Financial Times*, 20 de febrero [en línea], <<https://www.ft.com/content/c54002ee-1668-11e8-9e9c-25c814761640>>. [Consulta: 20-8-2018].
- (2018b): «AI experts warn of new types of threats», *Financial Times*, 21 de febrero [en línea], <<https://www.ft.com/content/c54002ee-1668-11e8-9e9c-25c814761640>>. [Consulta: 20-8-2018].
- WEIMER, M. y L. MARIN (2016): «The role of law in managing the tension between risk and innovation», *European Journal of Risk Regulation*, 7, 3.
- WEISBERG, J. (2011): «Bubble trouble», *Slate.com*, 10 de junio [en línea], <[http://www.slate.com/articles/news\\_and\\_politics/the\\_big\\_idea/2011/06/bubble\\_trouble.html](http://www.slate.com/articles/news_and_politics/the_big_idea/2011/06/bubble_trouble.html)>. [Consulta: 20-8-2018].
- WHITE, L. (1973): *Tecnología mundial y cambio social*, Barcelona: Paidós.
- YOO, J. (2005): *The powers of war and peace: the Constitution and foreign affairs after 9/11*, University Chicago Press.
- (2010): *Crisis and command*, Nueva York: Kaplan.
- (2012): «John Yoo responds to Taming International Law with presidential supremacy», *Law and Liberty*, 21 de mayo.
- YOUNES, B. (2018): «Intentional cooperation in Space: now more than ever», NASA [en línea], <[https://www.nasa.gov/pdf/696854main\\_Pres\\_International\\_Cooperation\\_in\\_Space.pdf](https://www.nasa.gov/pdf/696854main_Pres_International_Cooperation_in_Space.pdf)>. [Consulta: 7-9-2018].
- ZIMMERMANN, J. (1991): *Federal preemption*, Ames (Estados Unidos): Iowa State University Press.



