



UNIVERSIDAD  
**COMPLUTENSE**  
MADRID

**Inteligencia Artificial, Ética y Sistemas de  
Armas Automáticas en la Defensa Militar,  
Drones y Derecho Internacional**

© 2017

Esteban Yeray García Mederos

[Estebang@ucm.es](mailto:Estebang@ucm.es)

Mente Natural – Mente Artificial

Master de Epistemología de las Ciencias Naturales y Sociales

Facultad de Filosofía

UCM

## *Inteligencia Artificial en la Defensa Militar. (LAWS).*

Desde que se introdujo el término *Inteligencia Artificial (I.A.)*<sup>1</sup> en 1956 por John McCarthy en la gran conferencia realizada en la universidad de Dartmouth College junto a Claude Shannon y Marvin Minsky entre otros, se formuló rápidamente la pregunta que ya cuestionaban estos visionarios y que tanto nos inquietó posteriormente ¿Podrían las máquinas *pensar* como los seres humanos con todo lo que esto conlleva? Desde el punto de vista de la I.A. en las armas automáticas en defensa militar, y a partir de que La ONU propusiera su prohibición en la discusión que se llevó a cabo en 2015 en Reino Unido, se ha generado todo tipo de debates de carácter político, moral, jurídico, económico etc.

La definición propuesta por Marvin Minsky en su obra *Perceptrons*, afirma que la IA es el arte de construir máquinas capaces de hacer cosas que requieran inteligencia en caso de que fuesen hechas por los seres humanos, como sugiere Turing ¿Podrían las máquinas suplir a los seres humanos? Estas son las preguntas que lanzamos al futuro, y en vista de la problemática que se vislumbra a corto y medio plazo con respecto a su uso militar en este caso, se hace necesario adelantarse y prever sus posibles consecuencias, pues la tecnología avanza tan rápido que siempre llega tarde su reglamentación jurídica y el deber de cómo aplicarse. Por otra parte, la noción de *armas automáticas* que estamos analizando son aquellas “capaces de llevar a cabo una misión con intervención humana limitada o aún sin ella. Dependiendo del grado de implicación de la actividad humana, pueden ser semiautónomas o plenamente autónomas. Además, varía su letalidad. Algunas AW (*autonomous weapons*) operan de manera plenamente autónoma en actividades no letales, como vigilancia y reconocimiento... El diseño de las futuras AW les permitirá operar y alcanzar letalmente sus objetivos de manera autónoma sin intervención humana alguna...”<sup>2</sup>. Desde la dimensión más filosófica del debate, nos referimos con “*pensar*” en términos de la IA a sistemas expertos racionales-lógicos, que intentan emular el intelecto humano a través de cálculos con acciones como percibir, razonar y actuar<sup>3</sup>. Según la

---

<sup>1</sup> Minsky, M. y Papert, S.A. (1969): *Perceptrons*, MIT Press, Massachussets (USA).

<sup>2</sup> QUELHAS, Daniela: “La prolifération de robots-tués. Quelques problèmes juridiques et éthiques”, *Sentinelle bulletin* núm. 352, 16 junio 2013, pp. 1-14, p. 4 [[www.sentinelle-droit-international.fr](http://www.sentinelle-droit-international.fr)]).

<sup>3</sup> Patrick Henry Winston. *Artificial Intelligence*. Editor TBS (1992).

*teoría computacional de la mente*<sup>4</sup> pensar es manipular símbolos sintácticamente<sup>5</sup>, es decir, sostienen la tesis de que la mente funciona como un computador como afirma Jerry Fodor en su obra el *Lenguaje del pensamiento*<sup>6</sup>. De la misma manera, lo que llamamos pensar según la filosofía de las ciencias cognitivas, hace referencia también a manipular información<sup>7</sup>, a saber, “esta corriente de la psicología presupone para su investigación que los seres humanos, así como los animales superiores y también los computadores, son sistemas de procesamiento de información [...] también se le denomina *Psicología Computacional*.<sup>8</sup>

Estas son clases en las que se agrupan las definiciones sobre los robots de uso armamentístico militar según la LAWS<sup>9</sup>:

- 1.- Human in-the-Loop Weapons: Robots que pueden seleccionar objetivos y emplear la fuerza sólo bajo una orden humana.
- 2.- Human-on-the-Loop Weapons: Robots que pueden seleccionar objetivos y emplear la fuerza bajo la supervisión de un humano que puede anular las acciones del robot.
- 3.- Human-out -of -the-Loop Weapons: Robots que pueden seleccionar objetivos y emplear la fuerza sin ninguna interacción o decisión humana.

---

<sup>4</sup> Corriente que deriva de analogía tomada del *funcionalismo analítico* materialista, monista y fisicalista de Armstrong y Lewys, formado a consecuencia de la Teoría de la Identidad Psico-neuronal de los años 50 y 60 del siglo XX. Varios de los representantes más importantes de la teoría computacional de la mente son J. Fodor y H. Putnam.

<sup>5</sup> En la teoría computacional de la mente; también incluye la tesis semántica, entiendo el lenguaje desde símbolos más simples a relaciones más complejas. Sin embargo, nosotros comprenderemos el artículo desde un programa simbolista más débil como es el de John Searle, que afirma que de la sintaxis nunca podrá emergir la semántica.

<sup>6</sup> Fodor (1985) El lenguaje del pensamiento. Alianza Editorial. 1a edición inglesa 1975

<sup>7</sup> Según la *teoría matemática de la información* o de la comunicación de Claude E. Shannon (1948) en su artículo, *A Mathematical Theory of Communication*.

<sup>8</sup> Filosofía. Vol. Antropología, psicología y sociología, p. 223.

<sup>9</sup> Armas automáticas letales (leyes), en el que se estudiando dentro del marco del Derecho Internacional de los Derechos Humanos, Derecho Internacional Humanitario y dentro del marco del Derecho de los Conflictos Armados.

Según los expertos en robótica y juristas han reunido lo que consideran las ventajas de los sistemas de estas clases de armas automáticas según el Relator Especial del Consejo de Derechos Humanos, Christof Heyns<sup>10</sup>, entre las que acentúan:

- . Ofrecen una mayor protección de las fuerzas armadas propias (salvar vidas de los soldados y prevenir lesiones).
- . Multiplican la fuerza empleada.
- . Amplían el campo de batalla (facilitan la penetración tras las líneas enemigas y pueden mantenerse en el teatro de operaciones más tiempo; mucho más que las personas).
- . Poseen un tiempo de reacción menor que el de los seres humanos.
- . Nunca actuarán por pánico o venganza, ni por odio racial, etc.
- . Serán capaces, en el futuro, de emplear una fuerza menos letal, evitando muertes innecesarias; así, el desarrollo tecnológico puede ofrecer, como alternativas, la inmovilización o el desarme del objetivo<sup>11</sup>.

En otro orden de ideas, tenemos sus desventajas o problemas, como, por ejemplo, estos robots son incapaces de cumplir el Derecho Internacional Humanitario (DIH) que exige que la distinción entre combatiente y no combatiente, proporcionalidad y juicio en la respuesta dependiendo del contexto, y la necesidad militar de planificación, que exige subjetividad y capacidad de decisión que son herramientas importantes para proteger a los civiles de los efectos de la guerra.

Tomando una postura en el debate entre los tres grandes grupos de robótica armamentística militar dados, defenderemos los dos primeros grupos que son los más débiles, y rechazaremos la tercera clase de robots. Nuestras razones en la primera clase de robots tenemos, por ejemplo, los drones: Estos son muy efectivos y nuestro

---

<sup>10</sup> Cesáreo Gutiérrez Espada y María José Cervell Hortal (2013) SISTEMAS DE ARMAS AUTÓNOMAS, DRONES Y DERECHO INTERNACIONAL. ‘Revista del Instituto Español de Estudios Estratégicos’ no 2, pp. 23

<sup>11</sup> Estas armas (KAHN, Paul W.: “Imagining warfare”, European Journal of International Law”, 24 (2013), num. 1, pp. 199-226)

compromiso y responsabilidad sigue siendo humana. En el segundo grupo como afirma Cesáreo Guitierres Espada y José María Cervell Hortal, esta tecnología ya se utiliza en sincronización con humanos en gobiernos como EEUU, Israel o en zonas desmilitarizadas como son entre las dos Coreas. Estas dos clases de grupos ya generan muchísimos problemas éticos como muertes de civiles inocentes, pero son efectivos en sus objetivos. La tercera clase de tecnología robótica es la más problemática y a la que más se teme, hay algunos que proponen soluciones que consisten en implementar un gobernador ético como afirma Ronald Arkin. Un módulo ético que recoja y resuelva estas críticas de las que estamos hablando ¿Y en qué consiste este módulo? Afirma Arkin que no hay que pensar en abstracto, sino que hay que hablar de cosas concretas, un ejemplo claro es *el robot o la tabla* Ned Block (1981), una tabla lógica que tiene unas posibilidades de reacción pre-determinadas finitas, que se introducen en el sistema según la circunstancia, y así, según éstas puede determinar el robot (según sus Inputs) unas posibles respuestas diferentes (Outputs), este sistema ha sido propuesto por la LOW y las ROE<sup>12</sup>.

Por el contrario, las evidencias en los sistemas muestran que esto es imposible todavía, ya que “en relación con el *principio de distinción* se ha dicho que, a causa de la falta de idoneidad tecnológica de los sensores existentes en la actualidad, la incapacidad para comprender el contexto y la dificultad de aplicar en la práctica (mediante un programa informático) el estatuto del no combatiente, o la incapacidad para interpretar intenciones y emociones, sería muy difícil para un robot autónomo cumplir las exigencias del DIH sobre el mismo; y no digamos para identificar, en supuestos de conflictos asimétricos, quién es -y quién no- un combatiente”<sup>13</sup>. El principal problema para implementar este sistema ético sería ¿Cuáles son las proposiciones de la ética? Este lenguaje no existe, porque no es enunciativo, es decir es imposible saber a priori los enunciados del sentido

---

<sup>12</sup> LOW - Las leyes de la guerra. Regulación por las cuales se rigen las reglas del enfrentamiento entre gobiernos, esto prácticamente no se respeta, pero en cualquier caso hay unas reglas. Y ROE - Roe - Reglas de Enfrentamiento. Las formas de las roes, se pueden cambiar, son sistemas de producción. En caso del módulo ético tendría un conjunto de reglas de producción con una lógica (Sí ... entonces ...) y en el espacio después del entonces, disparo. Ejemplo, Si ...me fijan como objetivo, Entonces... disparo.

<sup>13</sup> Cesáreo Gutiérrez Espada y María José Cervell Hortal (2013) SISTEMAS DE ARMAS AUTÓNOMAS, DRONES Y DERECHO INTERNACIONAL. ‘Revista del Instituto Español de Estudios Estratégicos’ no 2, pp. 34

común ¿Cuál es la naturaleza de estas proposiciones? ¿Y en tal caso qué valores axiológicos introducimos? Si propones algoritmos utilitaristas, o lógicos de tipo “*Si X, entonces matas*”, esto no corresponde ni se puede hacer con todas las proposiciones de la ética, ya que el lenguaje se interpreta diferente por diferentes culturas, entre otros problemas. Entonces, cómo generamos un algoritmo para maximizar la utilidad, o ¿Cómo codificar y evaluar elementos prudenciales en cada situación nueva? Esto a día de hoy es impensable para un sistema experto armamentístico, que da más problemas que soluciones. No desde el punto de vista de la defensa de las LAWS para el gobierno o bando que quieran cumplir el objetivo, que pueden no considerar bajas de inocentes o problemas de cualquier tipo, sino según la DIH y la LOW, que son ante las que responden de algún modo las guerras internacionales. En resumen, los inconvenientes que se argumentan, es decir las razones en contra de proponer el tercer grupo de robots son:

- 1.- El agente humano capta la situación de un modo global mientras que el agente mecánico evalúa situaciones secuencialmente (sensación vs percepción)
- 2.- El agente humano involucra en sus decisiones componentes axiológicos, es decir, valorativos, los cuales parece difícil implementar en un robot.
- 3.- El agente humano afronta la toma de decisiones frente a situaciones nuevas, es decir, decide ante casos no previstos anteriormente. El agente mecánico necesita tener previstas con anterioridad todas las opciones posibles para tomar la decisión.
- 4.- Los sistemas de aprendizaje son radicalmente diferentes en el ente biológico (vicario) y en el mecánico.
- 5.- El agente humano emplea recursos semánticos en sus decisiones recurriendo a mecanismos como la metáfora y la analogía (razonamiento analógico), estrategias ambas difícilmente implementables en un robot.
- 6.- El agente humano toma en consideración componentes emocionales, prudenciales y de proporcionalidad difícilmente traspasables a un sistema experto regulador de este tipo de tecnologías.
- 7.- La ética es una parte esencial en el comportamiento humano y las proposiciones de la ética parecen imposibles o muy difíciles de formalizar en un sistema de reglas de producción.

La pregunta que se nos abre es ¿deberíamos de parar la investigación? Según el *principio de precaución*<sup>14</sup>, un concepto que propuso el consejo europeo en Niza (2000), que afirma y “respalda la adopción de medidas protectoras ante las sospechas fundadas de que ciertos productos o tecnologías crean un riesgo grave para la salud pública o el medio ambiente, pero sin que se cuente todavía con una prueba científica definitiva de tal riesgo”. Según este principio, deberíamos de llegar a acuerdos internacionales con arreglo de interés común, pero el problema es que no todos se someten a esta normativa. Como podemos observar, por ejemplo, en EEUU se rige por otro criterio Causal, que entiende que hasta que no haya una relación causal que afirme que la investigación esté dañando alguna de las dos esferas, a saber, la salud pública o el medio ambiente, no se puede parar la investigación. Continuando el argumento, y dando nuestra consideración se concluye de estos dos criterios que son factores contradictorios que hacen olvidar el principio de precaución, para así, empezar con una carrera armamentística internacional antes de su inicio, se intuye una posible competencia global por los beneficios de hacerse el primero con esta tecnología. Esto trae consigo otras consecuencias, que parecen periféricas al debate sobre el uso de la IA militar, pero que son elementos a tener en cuenta porque sin esta mirada holística no se puede comprender la magnitud del problema a la hora de abordar el problema que justo se desquebraja en sus complejidades. Otras de las consecuencias que prevén son:

1. Si el desarrollo de armas con IA continúa, la carrera armamentística Internacional (global) será imparable, además se puede crear un suministro estable para el mercado negro.
2. El armamento inteligente sería más barato y desecharable bajando los costos de ir a la guerra, lo cual incrementaría las posibilidades de que los países lucharan entre sí
3. La IA armada sería más fuerte y eficiente que nosotros cambiado drásticamente el tamaño de los conflictos militares<sup>15</sup>.

---

<sup>14</sup> Luis G. Jimenez Arias (2008). *Bioethics and the Environment*. p. 73.

<sup>15</sup> Artículo, video Newspaper Publicado el 2/10/2017 El riesgo de las armas con inteligencia artificial. (*20 Minutos*).

<http://www.20minutos.es/videos/tecnologia/isILNPW8-el-riesgo-de-las-armas-con-inteligencia-artificial/#xto=AD-15&xts=467263> Consultado el Lunes 11 de Diciembre de 2017.

Nuestra consideración más importante, es que se intuye que no se van a poder parar las investigaciones siguiendo el principio de precaución, pero sí se puede controlar éstas con muchos observadores para que no quebranten; el principio de Causalización o de Precaución según el riesgo y el caso particular. Hay que ser muy delicados, porque la robótica y la tecnología militar avanza a pasos agigantados y no puede cogernos de sorpresa. Estamos de acuerdo con respecto a los dos primeros grupos más débiles, aunque estos ya estén generando muchos problemas que tienen que ser revisados, esto no quiere decir que se introduzcan de manera incontrolable en todos los aspectos de nuestra vida, sino que “operen de manera plenamente autónoma en actividades no letales, como vigilancia y reconocimiento”<sup>16</sup>. Por ejemplo, en otro caso de usos probables, en el que sí estamos de acuerdo en que las máquinas sean totalmente autómatas en la actualidad, son la Defensa como; en las baterías de misiles programados para responder a los misiles lanzados, y los programas informáticos que responden a los ataques ciberneticos, siempre y cuando, la consecuencia pueda ser reversible. Esto es, que no se dispare por error cualquier alarma y se destruya, por ejemplo, medio mundo, sino que puedan actuar automáticamente en cuestiones de escasa consecuencia o relevancia en caso de error.

Se concluye, que “no se trata de limitar la inteligencia artificial, sino de introducir límites éticos en los robots, lograr que sean capaces de vivir en sociedad y, eso sí, rechazar de forma clara las armas autónomas sin control humano”, explica Francesca Rossi<sup>17</sup>. Son una clase de armas de alta tecnología militar aún inexistente, pero todo atisbo de imaginación sobre esta temática hace conjeturar sobre escenarios distópicos de ciencia ficción, en el cuál no queremos vivir, ni a corto, ni largo plazo.

---

<sup>16</sup> QUELHAS, Daniela: “La prolifération de robots-tués. Quelques problèmes juridiques et éthiques”, Sentinelle bulletin núm. 352, 16 junio 2013, pp. 1-14, p. 4 [[www.sentinelle-droit-international.fr](http://www.sentinelle-droit-international.fr)]).

<sup>17</sup> Artículo Carlos E. Cué / Alejandro Rebossio Buenos Aires Publicado 29/07/2015 Científicos contra los robots armados. (*El País*).

[http://www.coitt.es/index.php?page=noticias\\_sector\\_reg&ireg=0&icod=2979](http://www.coitt.es/index.php?page=noticias_sector_reg&ireg=0&icod=2979) Consultado el Miércoles 17 de Enero de 2018.

## **Bibliografía:**

- Armstrong D. M. *A Materialist Theory of the Mind*. Routledge & Kegan Paul. 1968.
- Cesáreo Gutiérrez Espada y María José Cervell Hortal *Sistemas de armas autónomas, drones y derecho internacional*. ‘Revista del Instituto Español de Estudios Estratégicos’ Nº 2. 2013.
- Fodor (1985) El lenguaje del pensamiento. Alianza Editorial. 1a edición inglesa 1975.
- KAHN, Paul W.: “Imagining warfare”, European Journal of International Law”, 24, Nº 1. 2013.
- Luis G. Jimenez Arias. *Bioethics and the Environment*. Editor Libros en red. 2008
- Mariano Moreno Villa. *Filosofía. Vol. Antropología, psicología y sociología*. Editorial MAD 2012.
- Minsky, M. y Papert, S.A. *Perceptrons*, MIT Press, Massachussets (USA). 1969.
- Patrick Henry Winston. *Artificial Intelligence*. Editor TBS 1992.
- QUELHAS, Daniela. “La prolifération de robots-tués. Quelques problèmes juridiques et éthiques”, Sentinelle bulletin núm. 352. 2013 [[www.sentinelle-droit-international.fr](http://www.sentinelle-droit-international.fr)]). Consultado el 18 de Enero de 2018.
- Shannon. Claude E. *A Mathematical Theory of Communication*. 1948.
- John Searle. *Mentes, Cerebros y ciencia*. Madrid: Cátedra. 1990.
- Rabossi, E. *La mente y sus problemas: temas actuales de la filosofía de la psicología*. Ed. Catálogos. Buenos Aires, 2004.

Artículo, video Newspaper Publicado el 2/10/2017 El riesgo de las armas con inteligencia artificial. (*20 Minutos*).

<http://www.20minutos.es/videos/tecnologia/isILNPW8-el-riesgo-de-las-armas-con-inteligencia-artificial/#xtor=AD-15&xts=467263> Consultado el Lunes 11 de Diciembre de 2017.

Artículo Carlos E. Cué / Alejandro Rebossio Buenos Aires Publicado 29/07/2015 Científicos contra los robots armados. (*El País*).

[http://www.coitt.es/index.php?page=noticias\\_sector\\_reg&ireg=0&icod=2979](http://www.coitt.es/index.php?page=noticias_sector_reg&ireg=0&icod=2979) Consultado el Miércoles 17 de Enero de 2018.