

109/2018

26 de octubre de 2018

*Marc Valls Estefanell \**

La inteligencia artificial y su encaje  
en las Estrategias de Seguridad  
Nacional

[Visitar la WEB](#)

[Recibir BOLETÍN ELECTRÓNICO](#)

## La inteligencia artificial y su encaje en las Estrategias de Seguridad Nacional

### Resumen:

En el presente trabajo se plantea el encaje de la inteligencia artificial (IA) en las Estrategias de Seguridad Nacional (ESN). En primer lugar, el documento analiza brevemente sus conceptos clave: la IA y la ESN. A continuación, se analiza cuál es la situación actual de la IA en el sector de la seguridad y la defensa y qué papel juega España en él. En segundo lugar, se propone el encaje de la IA como amenaza y línea de acción estratégica en las ESN. En tercer lugar, se cuestiona si esta tecnología debería estar recogida como amenaza en este tipo de documentos. Por último, se formulan unas conclusiones en las que el autor deja claro su punto de vista al respecto de todos y cada uno de los temas previamente tratados.

### Palabras clave:

Inteligencia artificial, Estrategias de Seguridad Nacional, super inteligencia artificial, IA, ESN.

**\*NOTA:** Las ideas contenidas en los *Documentos de Opinión* son de responsabilidad de sus autores, sin que reflejen, necesariamente, el pensamiento del IEEE o del Ministerio de Defensa.

## *Artificial Intelligence and its fit in National Security Strategies*

### *Abstract:*

*This work presents the fit of Artificial Intelligence (AI) in National Security Strategies (NSS). Firstly, the document takes a brief look at the key concepts: AI and NSS. Then, it analyses which is the current situation of AI in the security and defense sector and which role Spain plays in it. Secondly, it proposes the fit of AI as a threat and as a strategic line of action in NSS. Thirdly, the document questions if this technology should be treated as a threat in this kind of documents. Finally, there are some conclusions where the author makes clear his point of view about each topic that has been discussed.*

### *Keywords:*

*Artificial Intelligence, National Security Strategies, Artificial Super Intelligence, AI, NSS.*

### **Cómo citar este documento:**

VALLS ESTEFANELL, Marc. *La inteligencia artificial y su encaje en las Estrategias de Seguridad Nacional*. Documento de Opinión IEEE 109/2018. [enlace web IEEE](#) y/o [enlace bie<sup>3</sup>](#) (consultado día/mes/año)

## Introducción

Son pocas las tecnologías que han supuesto un antes y un después en la historia de la humanidad y que han revolucionado la sociedad y su modo de hacer. Hoy nos encontramos ante una de ellas. Se llama inteligencia artificial (IA) y sus primeros pasos nos están aportando soluciones asombrosas en sectores tan diferentes como la automoción y la atención médica. No obstante, su crecimiento y desarrollo merece especial atención porque, al ser una tecnología con vida propia -piensa, aprende y toma decisiones por sí sola- puede plantear escenarios inesperados e incluso peligrosos para nuestra propia existencia.

En paralelo, cada vez son más los países que están desarrollando esta tecnología. Estados Unidos, China y Rusia están convencidos de que su liderazgo a nivel global será tan determinante como potente sea su propia IA. En definitiva, el futuro pasa por aquí, pero ¿a qué riesgos nos enfrentamos?

La IA plantea y planteará distintas amenazas para la seguridad nacional, pero el presente trabajo solo se centra en dos de ellas. La primera, el surgimiento de una super IA, es decir, que la IA llegue a ser más inteligente que los humanos en todos sus aspectos intelectuales. En este caso, el verdadero peligro es que esta decida terminar con la especie humana al considerarla un obstáculo en caso de que nuestros valores y objetivos no estén alineados con los suyos. La segunda amenaza contemplada es el desarrollo y uso de la IA con fines perversos y/o letales por parte de actores estatales y no estatales. Al mismo tiempo, se puede fijar una línea de acción estratégica en relación con esta tecnología como un instrumento para prevenir y combatir las amenazas que esta supone.

Este documento de opinión es el resultado de un resumen de mi tesis del Máster Universitario en Seguridad, Defensa y Geoestrategia de la Universidad a Distancia de Madrid (UDIMA). Este fue un trabajo exploratorio debido a la escasa literatura al respecto y, en consecuencia, la metodología empleada fue mayoritariamente cualitativa a base de un conjunto de cuatro entrevistas semiestructuradas realizadas a profesionales de renombre relacionados con los conceptos clave del presente trabajo: Alexander Velez-Green, investigador asociado del Programa de Defensa del *Center for a New American Security* (CNAS); Ramón López de Mántaras, director del Instituto de Investigación en Inteligencia Artificial (IIIA); Ángel Gómez de Ágreda, coronel del Ejército del Aire y

analista en el Ministerio de Defensa español; Guillem Colom Piella, doctor en Paz y Seguridad Internacional y profesor de Ciencias Políticas en la Universidad Pablo de Olavide de Sevilla.

Este es un documento muy abierto debido a la novedad y complejidad del tema y, al no haber una opinión mayoritaria al respecto, no es hasta las conclusiones donde dejo claro mi punto de vista al respecto de todos y cada uno de los temas previamente tratados.

## **Conceptos clave: inteligencia artificial y Estrategia de Seguridad Nacional**

### ***Inteligencia artificial: concepto, origen e historia***

La Real Academia Española (RAE) define la IA como la «disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico».

Existen muchas otras más definiciones sobre la IA. En este sentido, Peter Stone, experto en IA y profesor de la *University of Texas at Austin* de Estados Unidos, señala que no hay una definición exacta sobre IA porque es un fenómeno en constante evolución y que toma muchas direcciones<sup>1</sup>. En la misma línea, Ramón López de Mántaras, director del Instituto de Investigación en Inteligencia Artificial (IIIA), afirma que hay diversas aproximaciones y diferentes puntos de vista sobre qué es la propia IA y esto da lugar a posibles definiciones.

De todos modos, se podría definir la IA como la simulación de los procesos de la inteligencia humana por parte de sistemas informáticos. En otras palabras, hacer que las máquinas piensen, aprendan y tomen decisiones<sup>2</sup>. Hasta hace bien poco, solo introducíamos instrucciones en las máquinas y esperábamos un resultado. Lo que se está haciendo ahora es inducir las al aprendizaje para que aprendan solas. Este proceso se conoce como *Machine Learning* (ML).

---

<sup>1</sup> ICDDF, "Technical Bulletin National Police Chiefs' Council Data Communications Group Futures Group", NPCC DCG Futures Group, 2018.

<sup>2</sup> PASCUAL, Federico, "La inteligencia artificial te dejará sin trabajo? TEDxDurazno", TEDx Talks, 2014, disponible en: <https://www.youtube.com/watch?v=xnNvTwfZRM>. Fecha de la consulta 4.9.2018.

Por lo que respecta a los orígenes e historia de esta tecnología, se podría decir que Alan Turing y John McCarthy, junto con algunos de sus compañeros como Marvin Minsky y Claude Shannon, fueron los padres de la IA. Turing fue de los primeros que empezó a escribir sobre la posibilidad de diseñar y construir máquinas inteligentes a finales de la década de los 40 y principios de los 50 del siglo pasado. Por este motivo, Alan Turing fue un visionario, un precursor de la materia. No obstante, no fue hasta 1956, en el *Dartmouth Summer Research Project*, donde McCarthy propuso el nombre de inteligencia artificial definiéndola como la «ciencia e ingeniería de hacer máquinas inteligentes».

### ***Estrategia de Seguridad Nacional: concepto, origen e historia***

La función principal de la estrategia de seguridad nacional es cumplir con los objetivos de la política de seguridad, los cuales velan básicamente por la protección de los intereses nacionales. Este tipo de estrategia se aplica tanto en tiempos de paz como en tiempos de guerra y tiene como límites la política general (límite superior) y las estrategias departamentales de los gobiernos (límite inferior)<sup>3</sup>.

De este modo, se entiende por Estrategia de Seguridad Nacional «el arte mediante el que la política general dirige, orienta y coordina todas las energías y fuerzas de la nación para proteger los intereses nacionales contra riesgos, peligros, amenazas exteriores e interiores, a la vez que permite al país cumplir con sus compromisos internacionales»<sup>4</sup>.

En cuanto a los orígenes e historia de las ESN, son varios los factores que les dieron a luz. En primer lugar, las estrategias basadas en la defensa nacional (EDN) y en la disuasión convencional y nuclear -producto de la Guerra Fría- quedaron desfasadas ante el surgimiento e implantación de la globalización. Como consecuencia, se empezó a dar prioridad a las estrategias de seguridad nacional y colectiva, apostando por gestionar crisis, estabilizar zonas y velar por la paz. De este modo, las EDN pasaron a estar a un segundo nivel estratégico, porque entraron dentro del paraguas de las ESN<sup>5</sup>.

---

<sup>3</sup> BALLESTEROS, Miguel Ángel, “En busca de una Estrategia de Seguridad Nacional”, Ministerio de Defensa, 2016.

<sup>4</sup> *Ibid.*, p. 65.

<sup>5</sup> *Ibid.*

Paulatinamente, se fueron reconociendo nuevos riesgos y amenazas, que mayoritariamente no eran ni estatales ni militares, cosa que provocó que se empezara a revisar de qué modo los Estados podrían hacerles frente para así dar seguridad a sus ciudadanos. Sin embargo, no fue hasta los atentados del 11-S que se evidenció la presencia de nuevos enemigos y formas de ataque<sup>6</sup>.

En definitiva, el 11-S fue el detonante de un proceso de reflexión que llevaron a cabo varios Estados y que derivó con la elaboración de sus propias ESN, documentos que permitieron y permiten dar una respuesta integrada a los riesgos y amenazas mediante la coordinación de todos los instrumentos del Estado y de la nación para hacer frente a los nuevos peligros<sup>7</sup>.

### **España y la carrera armamentística de la inteligencia artificial**

La carrera armamentística de la IA ya ha empezado. China, Estados Unidos y Rusia se disputan la primera posición. Vladimir Putin está convencido de que quien domine la IA, dominará el mundo<sup>8</sup> y, por otro lado, China tiene previsto invertir 150 mil millones de dólares con el objetivo de ser el líder mundial en esta materia en 2030<sup>9</sup>. Pero ¿tiene intención España de participar en esta carrera?

Ante todo, el coronel Gómez de Ágreda, señala que, en general, tanto la IA como las demás tecnologías digitales son de uso dual<sup>10</sup>. Esto implica que no solamente tienen una aplicación militar, sino también civil, por lo que esta carrera es a nivel de todas las fuerzas del Estado y no únicamente armamentística.

---

<sup>6</sup> DE LA CORTE, Luís et al, "Seguridad nacional, amenazas y respuestas", Madrid: LID Editorial Empresarial, S.L., 2014.

<sup>7</sup> BALLESTEROS, Miguel Ángel, *op. cit.*

<sup>8</sup> MICÓ, Josep Lluís, "La guerra fría de la inteligencia artificial", La Vanguardia, 23 de marzo de 2018, disponible en: <http://www.lavanguardia.com/tecnologia/20180323/441857328514/hackers-rusia-eeuu-guerrafria.html>. Fecha de la consulta 4.9.2018.

<sup>9</sup> JUNCA, Gemma, "China quiere ser el líder en inteligencia artificial en 2030", La Vanguardia, 24 de julio de 2017, disponible en: <http://www.lavanguardia.com/tecnologia/20170724/4357022220/china-inteligenciaartificial-lider-2030.html>. Fecha de la consulta 4.9.2018.

<sup>10</sup> BRUNDAGE, Miles et al, "The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation", p. 16, febrero de 2018, disponible en: [https://img1.wsimg.com/blobby/go/3d82daa4-97fe-4096-9c6b-376b92c619de/downloads/1c6q2kc4v\\_50335.pdf](https://img1.wsimg.com/blobby/go/3d82daa4-97fe-4096-9c6b-376b92c619de/downloads/1c6q2kc4v_50335.pdf). Fecha de la consulta 9.10.2018.

Respecto a España, el coronel destaca que a nivel europeo ya se están dando algunos pasos para crear un proyecto que integre a varios países de la Unión con el fin de estar igualada en relación con los demás competidores. En este sentido, añade que lo preocupante de la evolución exponencial del ámbito digital es que estar a un paso por detrás significa estar a un 50 % de la capacidad de los Estados que están por delante. «Entonces, no ser el líder implica un diferencial enorme de capacidades», concluye.

El doctor Colom señala que muchos Estados están interesados en esta tecnología, pero, al final, la capacidad económica de estos termina siendo determinante para establecer una buena y contundente base en materia de I+D. En cuanto a España, Colom destaca diversos actores que están tratando la IA: universidades, grandes empresas como Telefónica que recopilan enormes cantidades de datos y los principales bancos españoles que operan a nivel mundial.

Velez-Green, recalcando que no conoce lo suficiente la base industrial y el sector tecnológico español, afirma que España -como miembro de la OTAN y como país innovador que es, al igual que otros Estados de tamaño similar- puede tener un importantísimo papel en la participación de esta carrera. Respecto a esta competición, el investigador estadounidense manifiesta que una de sus principales características es su progresión natural, es decir, a medida que los países vayan viendo como la IA les supone una ventaja militar sustancial, apostarán por ella y tratarán de exprimirla tanto como puedan. En este sentido, Velez-Green afirma que China reconoce que hay un tremendo potencial de crecimiento en cuanto a la IA y que la mejor manera de maximizar ese potencial y minimizar el riesgo de quedarse atrás es ser el primero en implementarlo.

## **Propuesta de encaje de la inteligencia artificial en las Estrategias de Seguridad Nacional**

### ***La inteligencia artificial como amenaza y desafío para la seguridad nacional***

Básicamente, la IA puede suponer dos tipos de amenazas, una de origen tecnológico y otra de origen humano. La amenaza tecnológica se puede dar si la IA se convierte en superinteligente, es decir, que esta llegue a ser superior a la inteligencia humana. En

cambio, la amenaza humana se materializa si la IA es desarrollada y usada con fines perversos y/o letales por parte de actores estatales y no estatales.

### El surgimiento de una super inteligencia artificial

Se entiende por super IA aquella inteligencia que va mucho más allá de la IA general porque no solo iguala a los seres humanos, sino que los supera en todos sus aspectos intelectuales. Este escenario es conocido como la «gran singularidad» y uno de sus principales peligros es que la super IA decida terminar con la especie humana<sup>11</sup> al considerarla un obstáculo en caso de que nuestros valores y objetivos no estén alineados con los suyos.

En su momento, Stephen Hawking ya advirtió del posible surgimiento de una super IA y Elon Musk también lo ha hecho en varias ocasiones. Hawking dijo que, aunque la IA que se ha desarrollado hasta el momento ha resultado ser beneficiosa para la humanidad, en un futuro esta podría suponer un peligro e incluso podría significar el fin de la raza humana<sup>12</sup>. En la misma línea, Musk ha llegado a decir que «con la IA estamos invocando al demonio» y que esta puede llegar a ser más peligrosa que las armas nucleares.

Nick Bostrom, filósofo y profesor de la Universidad de Oxford, está convencido de que en un futuro surgirá una super IA. Para explicarlo, equipara la historia de la humanidad con una urna llena de pelotas y afirma que cada pelota que los humanos sacamos de ella equivale a nuestro progreso social. Sin embargo, recalca que alguna pelota (en concreto, la super IA) puede llegar a ser muy perjudicial hasta el punto de suponer el fin de su descubridor: nosotros, los humanos. Bostrom señala que el problema es que aún no sabemos el modo de volver a poner la pelota que no nos guste en la urna, ya que los inventos difícilmente se pueden «desinventar»<sup>13</sup>. Por lo tanto, afirma que el reto está en

---

<sup>11</sup> GRACE, Katja et al, "When Will AI Exceed Human Performance? Evidence from AI Experts", p. 4, 3 de mayo de 2018, disponible en: <https://arxiv.org/pdf/1705.08807.pdf>. Fecha de la consulta 9.10.2018.

<sup>12</sup> EFE. "Stephen Hawking: «La inteligencia artificial podría significar el fin de la raza humana»", ABC, 3 de diciembre de 2014, disponible en: <http://www.abc.es/tecnologia/informatica-software/20141202/abci-stephen-hawking-peligros-inteligencia-201412021837.html>. Fecha de la consulta 4.9.2018.

<sup>13</sup> CORTÉS, Javier, "Entrevista Nick Bostrom: 'No tendremos una segunda oportunidad con la inteligencia artificial'" Retina, El País. 9 de diciembre de 2017, disponible en: [https://retina.elpais.com/retina/2017/12/02/tendencias/1512231406\\_905237.html](https://retina.elpais.com/retina/2017/12/02/tendencias/1512231406_905237.html). Fecha de la consulta 4.9.2018.



la transición que se lleve a cabo hasta este estadio, es decir, la era de la IA. Para ello, deja muy claro que probablemente no tendremos una segunda oportunidad con esta tecnología y, consecuentemente, es imprescindible que la desarrollemos de tal forma que esta comparta nuestros objetivos, ya que una vez saquemos la pelota de la urna será imposible volverla a meter dentro y empezar de cero<sup>14</sup>.

Una perspectiva totalmente contraria a la de Bostrom es la de López de Mántaras. El director del IIIA es muy crítico con la super IA y afirma que no existen argumentos científicos sólidos que nos permitan predecir que esto será factible y menos en unas escalas temporales -50, 100 o 200 años- en las cuales no tiene sentido especular. Al hablar del futuro, personalmente, no le gusta ir más allá de 10 o 15 años y deja claro que en estos intervalos de tiempo es absolutamente imposible el logro de una super IA. El director del IIIA destaca que, actualmente, las inteligencias específicas ya plantean problemas, por lo que no hace falta esperar a las super IA para que esta tecnología pueda suponer un problema.

Guillem Colom afirma que, aunque este sea un escenario de ciencia ficción, en un futuro todo es posible y, por lo tanto, nada es descartable. Gómez de Ágreda, en cambio, no ve posible el surgimiento de una super IA, «por lo menos a corto o medio plazo». Sin embargo, lo que sí ve, y le parece igual de preocupante, es el cómo puede llegar a cambiar a la especie humana una IA muy desarrollada. El coronel concluye que «esta tecnología cambiará la esencia del ser humano, no eliminará la especie, pero sí la hará mutar de alguna manera».

Velez-Green no descarta este escenario y cree que hay muchas razones para preocuparse de ello. En este sentido, ve fundamental que desarrolladores y políticos sean conscientes de este riesgo y que lo tengan en mente cada vez que se hable de avances en IA.

En mi opinión, lo más complejo que plantea esta potencial amenaza es que no sabemos qué tipo de forma y apariencia darle y esto dificulta mucho el hecho de tomar medidas al respecto porque no sabemos a qué nos enfrentamos. Si digital, material, humana, robótica... Las opciones en cuanto a su forma son infinitas y difícilmente tendremos la respuesta hasta presenciarla en persona. Como consecuencia de ello, mucha gente, y

---

<sup>14</sup> MÜLLER, Vincent C. y BOSTROM, Nick, "Future Progress in Artificial Intelligence: A Survey of Expert Opinion" <https://nickbostrom.com/papers/survey.pdf>. Fecha de la consulta 10.10.2018.

lo demuestran parte de los entrevistados, equiparan la super IA con las armas letales autónomas, pero esta será mucho más profunda y transversal que una simple arma, por muy autónoma y letal que pueda ser.

Sam Harris, filósofo y neurocientífico estadounidense, en su charla *TED Talk: Can we build AI without losing control over it?*<sup>15</sup>, señala que si la inteligencia es solo una cuestión de procesamiento de la información y seguimos mejorando nuestras máquinas, produciremos algún tipo de superteligencia. Personalmente, estoy convencido de ello y por este motivo mi postura es mucho más cercana a la de Bostrom, Musk y Hawking. Coincidiendo con Harris, una de las cosas más preocupantes en este momento es que la mayoría de los investigadores de IA afirman que el surgimiento de una super IA será probablemente dentro de 50 o 100 años, por lo que el motivo más usado para que no nos preocupemos es el tiempo. Ante esta postura, el filósofo deja claro que nadie parece darse cuenta de que tomar el tiempo como referencia es una incongruencia total y lo vuelve a remarcar: si la inteligencia es solo una cuestión de procesamiento de la información y seguimos mejorando nuestras máquinas, tarde o temprano produciremos algún tipo de superinteligencia y no tenemos idea de cuánto tiempo nos va a llevar crear las condiciones para hacerlo de forma segura. Volviendo a coincidir con Harris, «al final, 50 años no es tanto tiempo para lograr uno de los mayores desafíos al que nuestra especie se ha enfrentado» y, «una vez más, parece que no tenemos una respuesta emocional adecuada para lo que, con toda probabilidad, va a pasar».

El desarrollo y uso de la inteligencia artificial con fines perversos y/o letales por parte de actores estatales y no estatales

En la madrugada del 5 de enero de 2018, la base área rusa de Khmeimim, Siria, fue atacada por un enjambre de drones comerciales de no más de 1 000 € la unidad y que iban equipados con GPS y cargas explosivas. El ataque lo realizó un actor no estatal<sup>16</sup>.

---

<sup>15</sup> HARRIS, Sam, "Can we build AI without losing control over it?", TED Talks, 2016, disponible en: <https://www.youtube.com/watch?v=8nt3edWLgIq&t=615s>. Fecha de la consulta 4.9.2018.

<sup>16</sup> FLORES, Félix, "¿Quién ataca las bases rusas con drones?", La Vanguardia, 10 de enero de 2018, disponible en: <http://www.lavanguardia.com/internacional/20180110/434196478646/siria-quien-ataca-bases-rusasdrones.html>. Fecha de la consulta 4.9.2018.

Teniendo en cuenta este ejemplo, ¿es posible que ocurra lo mismo con la IA, es decir, que esta sea desarrollada y usada con fines perversos y/o letales por parte de actores estatales y no estatales?<sup>17</sup>

Guillem Colom afirma que «sin ninguna duda». Coincidiendo con él, Gómez de Ágreda está convencido de que vamos a ver el desarrollo y uso de la IA con fines perversos y/o letales por parte de actores estatales y, además, seguramente, a muy corto plazo. Sin embargo, respecto a los actores no estatales, cree que es bastante más difícil que estos lleven a cabo desarrollos importantes de IA. Aun así, advierte que, aunque no lleguen a desarrollar IA propia, sí que emplearán IA de uso dual<sup>18</sup> de una manera que no esté prevista y les permita realizar ataques, quizás no letales, pero sí por lo menos de mucho calado.

López de Mántaras, señala que el desarrollo y uso de la IA con fines perversos y/o letales ya existe y pone como ejemplos la puesta en marcha del polémico «crédito social»<sup>19</sup> establecido en China, los enjambres de drones y la creación de armas letales autónomas.

Velez-Green da por hecho el desarrollo y uso de la IA con fines perversos y/o letales y manifiesta que la premisa aquí es averiguar qué tipo de forma podrá adoptar la IA.

### ***La inteligencia artificial como línea de acción estratégica***

«Imaginad que recibimos un mensaje de una civilización alienígena: gente de la Tierra, llegaremos a vuestro planeta en 50 años, estad preparados». Este es el escenario que plantea Stuart Russell, científico experto en IA, para hacer reflexionar a la gente de que los humanos debemos de actuar y prepararnos para lo que pueda venir.

---

<sup>17</sup> HOROWITZ, Michael C., "Who'll want artificially intelligent weapons? ISIS, democracies, or autocracies?", Bulletin of the Atomic Scientists, 2016, disponible en: <https://thebulletin.org/2016/07/wholl-want-artificially-intelligent-weapons-isis-democracies-or-autocracies/>. Fecha de la consulta 9.10.2018.

<sup>18</sup> ALLEN, Greg y CHAN, Taniel, "Artificial Intelligence and National Security", Harvard Kennedy School Belfer Center for Science and International Affairs, p. 2, 2017, disponible en: <https://www.belfercenter.org/sites/default/files/files/publication/AI%20NatSec%20-%20final.pdf>. Fecha de la consulta 9.10.2018.

<sup>19</sup> El crédito social es el sistema a través del cual el Gobierno de Xi Jinping establece la confianza que deposita en cada uno sus conciudadanos. El crédito de una persona sube o baja en función de si los datos recopilados por el software encargado de rastrear y analizar su comportamiento en internet y redes sociales son favorables o no al Gobierno chino. Este rastreo no es únicamente digital, sino que también se deben de tener en cuenta los miles de cámaras instaladas en la vía pública que hacen un seguimiento de cada persona gracias a sus algoritmos de reconocimiento facial.

Siguiendo el consejo de Russell, si las dos amenazas citadas anteriormente fuesen contempladas en las ESN, sería necesario fijar los objetivos a cumplir para combatirlas. Para tal fin, se deberían establecer unas líneas de acción estratégicas al respecto:

1. Crear las condiciones necesarias para desarrollar una super IA de forma segura, es decir, que su sistema de valores y sistema moral sea el mismo que el nuestro. En este sentido, Bostrom apunta que es fundamental fijar cómo queremos que las máquinas interpreten los valores humanos a través del aprendizaje automático.
2. Establecer varios escenarios que podría plantear el surgimiento de una super IA: formas que podría adoptar, modos en los que podría reaccionar, etc. El objetivo es evitar el factor sorpresa.
3. Promover la creación de un organismo internacional que, semejante a los acuerdos de no proliferación existentes en el ámbito nuclear, fije unas líneas rojas de desarrollo de IA para Estados, empresas e investigadores especialistas en la materia y lleve a cabo un seguimiento de qué se está desarrollando y para qué finalidad. Este organismo también debe de elaborar una lista internacional de actores no estatales que puedan tener intención de desarrollar y/o usar IA con fines perversos y/o letales y haga un seguimiento de cada uno de ellos.
4. Fomentar e impulsar proyectos *AI for Good*, es decir, proyectos que desarrollen IA con fines benévolos<sup>20</sup>. Para hacer efectivo este punto, antes es fundamental que se cumplan las anteriores líneas de acción estratégicas con el objetivo de tener controladas las amenazas que supone la IA y así evitar que estas retrasen las aplicaciones beneficiosas de esta tecnología<sup>21</sup>.

---

<sup>20</sup> Executive Office of the President National Science and Technology Council Committee on Technology, "Preparing for the Future of Artificial Intelligence", p. 13, 2016, disponible en: [https://obamawhitehouse.archives.gov/sites/default/files/whitehouse\\_files/microsites/ostp/NSTC/preparing\\_for\\_the\\_future\\_of\\_ai.pdf](https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf). Fecha de la consulta 9.10.2018.

<sup>21</sup> BRUNDAGE, Miles, *op. cit.*, p. 10.

## ¿Debería la inteligencia artificial estar recogida como amenaza en las Estrategias de Seguridad Nacional?

Velez-Green considera que aún es demasiado pronto para contemplar la IA en las ESN como amenaza y desafío para la seguridad nacional, aunque señala que es algo en lo que ya se debería estar pensando.

López de Mántaras manifiesta que el desarrollo y uso de la IA con fines perversos y/o letales tendría que estar contemplado como amenaza en las ESN y piensa que se terminará haciendo. En cambio, respecto al surgimiento de una super IA, no opina lo mismo porque cree que nos debemos centrar en los problemas que la IA ya está planteando: las armas letales autónomas, el crédito social de China, los enjambres de drones, etc.

Colom considera que el desarrollo y uso de la IA con fines letales y/o perversos debería estar contemplado, pero no como amenaza, sino más bien como riesgo. En cuanto al surgimiento de una super IA, considera que aún es temprano para recogerlo sea en la categoría que sea.

Gómez de Ágreda manifiesta que es fundamental que la IA esté recogida en las ESN como reto tecnológico y como oportunidad, pero también como posible amenaza respecto a los cambios sociales y políticos que esta va a introducir.

En lo que sí coinciden la mayoría de los entrevistados es que la IA, de algún modo u otro, debe estar recogida en las ESN. También coinciden que si la IA aun no consta como amenaza en estos documentos es porque los Estados la ven más como una oportunidad que como una amenaza.

En la actualidad, algunas ESN ya recogen la IA, pero no como desafío, riesgo, amenaza u otra categoría propuesta por los expertos entrevistados, sino como mera tecnología emergente para tener en cuenta. Es el caso de la actual ESN española, la Estrategia de Seguridad Nacional 2017. De 128 páginas que tiene el documento, solo se cita la IA en dos de ellas (páginas 37 y 87). Otro ejemplo es la última ESN de Estados Unidos, de 2017. El documento de 68 páginas solo menciona la IA en tres de ellas (páginas 30, 44 y 45). Aunque existan otros documentos, como por ejemplo el *Worldwide Threat Assessment of the US Intelligence Community*, que sí reconocen la IA como amenaza, el verdadero reto aquí es que sea recogida como tal en las ESN.

## Conclusiones

La carrera armamentística de la IA ya se está librando. China, Estados Unidos y Rusia se disputan la primera posición. Esta competición se caracteriza por tener una progresión natural, ya que a medida que los países vayan viendo cómo la IA les supone una ventaja militar sustancial, apostarán cada vez más por ella y tratarán de exprimirla tanto como puedan.

Respecto a España y su papel en esta carrera, puede que llegue a tener un rol importante en ella teniendo en cuenta que es un país miembro de la OTAN y más aún si se termina formando un bloque europeo para competir con los chinos, estadounidenses y rusos.

El surgimiento de una super IA y el desarrollo y uso de la IA con fines perversos y/o letales por parte de actores estatales y no estatales son dos de las mayores amenazas que supone la IA. La primera de ellas es inevitable y no tenemos idea de cuánto tiempo nos va a llevar crear las condiciones para que no resulte ser letal. Además, no sabemos a qué nos enfrentamos teniendo en cuenta qué tipo de forma y apariencia adoptará por lo que es muy complejo tomar medidas al respecto. La segunda amenaza, en parte, ya es una realidad. La puesta en marcha del polémico «crédito social» en China, el desarrollo de enjambres de drones y la creación de armas letales autónomas son un ejemplo de ello. Por los motivos expuestos, es crucial contemplar estas dos amenazas en las ESN.

Además, también es necesario fijar los objetivos a cumplir para combatirlas. Para tal fin, se deben establecer unas líneas de acción estratégicas al respecto, por ejemplo: crear las condiciones necesarias para desarrollar una super IA de forma segura, fijar unas líneas rojas de desarrollo de la IA tanto para Estados como para empresas e investigadores, fomentar e impulsar proyectos *AI for Good*, etc.

En definitiva, existen motivos suficientes como para que la IA esté recogida en las ESN como una amenaza principal y, en consecuencia, esta tenga su propia línea de acción estratégica. Sin embargo, actualmente, no existe ningún tipo de consenso respecto a cómo catalogar y gestionar esta tecnología. En consecuencia, los Estados que la contemplan en sus respectivas estrategias no la recogen como amenaza, sino como una oportunidad dado el diferencial enorme que les proporciona respecto a los Estados que no disponen de ella. Probablemente esto siga igual porque si un Estado reconoce la IA

como amenaza en su ESN, esto le supondrá, por un lado, limitar automáticamente el desarrollo de esta tecnología, ya que sería contradictorio seguir produciendo una cosa que previamente se ha catalogado como amenaza, y, por otro lado, una enorme desventaja competitiva respecto a los demás Estados que sí la desarrollan.

Puestos a especular, también se debería tener muy presente el poder de la naturaleza. Al fin y al cabo, cualquier cosa es producto de ella, incluso todo aquello artificial. En este sentido, si termina surgiendo una super IA y su finalidad es letal, puede ser porque la naturaleza se encuentra desbordada ante una sobrepoblación cada vez mayor. De este modo, una superinteligencia podría ser el instrumento definitivo creado por la naturaleza, a través de los humanos y de la tecnología (ambos son producto de ella), para reequilibrar la superpoblación que hoy existe en el planeta Tierra o, incluso peor, para empezar de cero.

*Marc Valls Estefanell\**

Máster en Seguridad, Defensa y Geoestrategia (UDIMA)