NOTA DE FUTURO 2/2018

8 de febrero de 2018

Pablo Márquez[*]

Inteligencia artificial y Seguridad Nacional

Nota de futuro: Inteligencia artificial y Seguridad Nacional

El presente documento es una síntesis y traducción de: Artificial Intelligence and National Security. A study on behalf of Dr. Jason Matheny, Director of the U.S. Intelligence Advanced Research Projects Activity (IARPA), escrito por Greg Allen & Taniel Chan. La versión original se puede encontrar en: https://www.belfercenter.org/sites/default/files/files/publication/AI%20NatSec%20-%20final.pdf

Resumen ejecutivo

- Investigadores en el ámbito de la Inteligencia Artificial (IA) han demostrado un progreso tecnológico importante durante los últimos cinco años. Dicho progreso continuará e incluso acelerará.
- La mayoría de los avances de investigación en IA se están haciendo en el sector privado y en el mundo académico.
- Capacidades existentes en la IA tienen un potencial importante para la seguridad nacional.
- El futuro progreso en la IA tiene el potencial de ser una tecnología transformadora para la seguridad nacional, como lo fueron las armas nucleares, la aviación, los ordenadores y la biotecnología.
- Los avances en la IA afectarán a la seguridad nacional a través del cambio en tres áreas: la superioridad militar, la superioridad informativa y la superioridad económica.
- Hemos analizado los cuatro casos anteriores de tecnologías militares transformadoras –nuclear, aeroespacial, cibernética y biotécnica– y hemos generado la "experiencia adquirida" o "lecciones aprendidas" (cinco lecciones) para la IA.
- Tomando un enfoque "gubernamental integral," hemos propuesto tres objetivos para la política de seguridad nacional de EE.UU. hacia la tecnología de IA y hemos propuestos 11 recomendaciones.

Introducción y enfoque del proyecto

Hay cuatro impulsores principales en el rápido progreso de la tecnología de IA:

1. Décadas de crecimiento exponencial en el rendimiento de la computación/informática.
2. El aumento de la disponibilidad de amplios conjuntos de datos con los que se pueden entrenar a las máquinas con sistemas de aprendizaje.
3. Avances en la implementación de las técnicas de aprendizaje en las máquinas.
4. Inversiones comerciales que han ido aumentando de manera rápida y significante.

* Alumno en prácticas de la Universidad Pontificia de Comillas

- La Inteligencia Artificial General: es decir, Inteligencia Artificial a escala y fluidez de un cerebro humano. La mayoría de los investigadores aún asumen que dicha capacidad en la IA todavía está a varias décadas de distancia.
- El progreso rápido en la IA probablemente tendrá un impacto en la seguridad nacional.
- Las cuatro tecnologías militares transformativas: Nuclear, cibernética, aeroespacial, biotecnológica.
- Nuestro Planteamiento – Parte 1: Analizar los posibles escenarios de desarrollo tecnológico relacionados con la IA y explorar como dichos escenarios pueden transformar la seguridad nacional.
- Nuestro Planteamiento – Parte 2: Evaluar las tecnologías militares transformativas previas para generar unas "lecciones aprendidas" para diseñar respuestas al surgimiento de un dominio tecnológico tan importante como la IA.
- Nuestro Planteamiento – Parte 3: Aportar recomendaciones de políticas relacionadas con la IA para preservar el liderazgo tecnológico de EE.UU., apoyar el uso pacífico de la IA, y mitigar los riesgos catastróficos.

Parte 1: El potencial de transformación de la inteligencia artificial

Implicaciones para la superioridad militar

Robótica y autonomía

En la guerra se han utilizado sistemas autónomos al menos desde la Segunda Guerra Mundial. La delegación del control humano a tales sistemas ha aumentado junto con el mejoramiento de las tecnologías instrumentales.

La dimensión del mercado, tanto para la robótica comercial como militar, se está incrementando de manera exponencial y los precios por unidad están disminuyendo de forma considerable.

El uso ampliado del aprendizaje automático, combinado con el crecimiento de mercado y la caída de precios, ampliará de manera formidable el impacto de los sistemas robóticos en la seguridad nacional.

Al igual que el impacto de la cibernética, el aumento de la utilización de la robótica y de los sistemas autónomos incrementará el poder de los actores no estatales y de los estados-nación.

A corto plazo, los avances en la IA probablemente permitirán que se pueda proporcionar un apoyo robótico más autónomo a los aviones de guerra y también acelerar el cambio de misiones de combate tripuladas a no tripuladas.

Llegará un punto en el que se superarán las limitaciones de tamaño, peso y poder, que ahora mismo limitan la autonomía avanzada, de la misma manera que los smartphones a día de hoy tienen tanta capacidad como lo que en su día fue el rendimiento máximo de un super-ordenador.

De medio a largo plazo, los sistemas robóticos y autónomos probablemente se equipararán al aumento de las capacidades tecnológicas que se han demostrado posibles por naturaleza.

A largo plazo, estas capacidades transformarán el poder militar y la guerra.

Ciberseguridad y guerra cibernética

Altos cargos de la seguridad nacional de Estados Unidos creen que la IA y el aprendizaje automático tendrán usos transformables para la ciberseguridad y la guerra cibernética.

Al igual que con toda automatización, la IA y el aprendizaje automático harán que el número de humanos necesitados para realizar tareas específicas en el ámbito cibernético disminuya.

La IA será útil a la hora de reforzar la defensa cibernética, puesto que vigilar los puntos flojos y monitorizar los sistemas se puede mejorar con automatización inteligente.

Sin embargo, la misma lógica sugiere que los avances en la IA permitirán que surjan mejoras en la ofensiva cibernética.

En el futuro inmediato, incorporar las aplicaciones de la tecnología de IA al ámbito cibernético beneficiará a los actores de carácter nación-estado poderosos. A largo plazo, los resultados del balance de poder no están claros, al igual que el balance a largo plazo entre la ofensiva y defensa cibernética.

Escenarios transformables potenciales

1. Armas autónomas letales forman la mayor parte de las fuerzas militares.
2. La inmensa cantidad de tecnologías perjudiciales hará que algunas plataformas militares se queden obsoletas.
3. Los asesinatos robóticos serán comunes y difíciles de atribuir.
4. Los artefactos explosivos improvisados (IED) móviles-robóticos darán unas capacidades baratas, parecidas a las de las municiones guiadas a precisión (PGM), a los terroristas.
5. El poder militar poco a poco se desvinculará de factores como el tamaño de la población y la fortaleza económica.
6. Las armas cibernéticas a menudo se utilizarán para matar.
7. La mayoría de los actores en el espacio cibernético no tendrán otra opción que habilitar niveles relativamente altos de autonomía o, si no, correrán el riesgo de ser derrotados en la competencia por adversarios que operen “a la velocidad de las máquinas”.
8. Interacciones inesperadas de los sistemas autónomos causaran “crisis precipitadas” ocasionales.
9. La participación del aprendizaje automático en los sistemas militares creará nuevos tipos de vulnerabilidades y nuevos tipos de ciberataques que tomarán como objetivo los datos sobre capacitación de los sistemas de aprendizaje automático.
10. El robo y la réplica de los sistemas militares y de inteligencia de la IA hará que el ciber-armamento caiga en manos equivocadas.

Implicaciones para la superioridad informativa

Recopilación y análisis de datos

Las agencias de inteligencia de Estados Unidos están repletas de mucha más información con potencial uso de datos crudos de inteligencia de lo que tienen capacidad de analizar.

El análisis de inteligencia asistido por ordenador, aprovechando el aprendizaje automático, pronto proporcionará capacidades notables, como poder fotografiar y analizar toda la superficie de la Tierra cada día.

Creación de datos y medios de comunicación

La utilización de la IA no solo se puede emplear para analizar datos, sino también para producirlos, incluyendo fotografías, vídeos y texto generados de manera automática.

En el futuro próximo, será posible que un aficionado o principiante genere falsificaciones fotorrealistas de vídeo de calidad HD, audio y documentos. Todo esto a gran escala.

La existencia de amplias capacidades de falsificación en la IA acabará con la confianza social, puesto que las pruebas o informaciones que antes eran fiables se convertirán en un factor de incertidumbre.

Escenarios transformables potenciales

1. Las medidas de vigilancia sobrecargada acaban con la guerra de guerrillas.
2. El país que tenga una ventaja significativa en el análisis de inteligencia basado en la IA conseguirá una ventaja a la hora de tomar y diseñar decisiones estratégicas decisivas.
3. La propaganda para los regímenes autoritarios e iliberales cada vez será mas indistinguible de la verdad.
4. Los problemas a los que se enfrenta la prensa libre y democrática con las noticias falsas empeorarán de manera dramática.
5. Las organizaciones de mando y control se enfrentarán a amenazas persistentes en su ingeniería social.
6. Combinado con los ciberataques y unidades de bots en las redes sociales, la falsificación de los medios de comunicación facilitada por la IA amenazará a la estabilidad de la economía y al régimen gubernamental.

Implicaciones para la superioridad económica

El sobre-alimentador de innovación

Puede que la Inteligencia Artificial sea una tecnología económica únicamente transformable, puesto que tiene el potencial para acelerar el ritmo de innovación y el crecimiento de la productividad de manera dramática.

1. La automatización de los experimentos científicos.
2. Sintetizar los hallazgos en miles de documentos científicos.
3. Generar y optimizar diseños técnicos de manera automática.

Automatización y desempleo

El *Informe 2016 de la Casa Blanca sobre la Inteligencia Artificial, la Automatización y la Economía* determinó que la automatización amenazará la existencia de millones de trabajos y que

las protestas laborales estarán más presentes en el futuro y podrán ser más permanentes que en casos previos.

Si la IA lleva como consecuencia al desplazamiento permanente de trabajadores, es posible que los países con ventajas tecnológicas se enfrenten al problema de "la maldición de los recursos", por el cual los propietarios del capital productivo se concentran en una élite muy pequeña y concentrada, mientras que la economía y la política se vuelven mucho más inestables.

- La composición de las industrias extractivas promueve la desigualdad y gobernanza deficiente.
- La redistribución de los ingresos provenientes de los recursos naturales toma el riesgo de caer bajo la corrupción gubernamental.
- La desigualdad promueve el conflicto político y civil.
- El éxito de la exportación de recursos naturales amenaza a otras industrias.

Escenarios transformables potenciales

1. La "maldición de los recursos" inducida por la automatización se extiende por los países desarrollados en el ámbito tecnológico.
2. El país que tiene una ventaja significativa en las tecnologías innovadoras, gracias a la IA, desarrollará una ventaja tecnológica y económica que se auto-reforzará a sí misma.
3. El sabotaje económico facilitado por la IA surgirá como un nuevo tipo de arma.

Parte 2: Aprendiendo de casos anteriores de tecnología transformable

Aspectos clave para la gestión tecnológica

1. Capacidad destructiva.
2. Perfil de coste.
3. Perfil de complejidad.
4. Capacidad de doble uso militar/civil.
5. Dificultad de espionaje y supervisión.

Tabla 1: Aspectos clave de la tecnología

NIVEL	Capacidad destructiva	Perfil de coste	Complejidad	Capacidad de doble uso militar/civil	Dificultad de espionaje y supervisión
Nuclear	Alto	Alto	Alto	Alto	Intermedio
Aeroespacial	Intermedio	Intermedio	Intermedio	Alto	Intermedio
Cibernética	Intermedio	Bajo	Intermedio	Alto	Alto
Biotécnica	Alto	Bajo	Alto	Alto	Alto

Enfoque gubernamental para la gestión tecnológica

Tabla 2: Enfoque de la gestión gubernamental de la tecnología

Nuclear	Esfuerzo supremo, desarrollo y utilización dirigido por el gobierno
Aeroespacial	Colaboración público-privada dirigida por el gobierno
Cibernética	El gobierno lo "siembra y lo cosecha"
Biotécnica	Restricción voluntaria

Enfoque de gestión gubernamental

1. Preservar el liderazgo tecnológico de Estados Unidos.
2. Apoyar el uso pacífico de la tecnología.
3. Gestionar los riesgos catastróficos.

Tabla 3: Resultados del enfoque de la gestión gubernamental de la tecnología

	1. Preservar el liderazgo tecnológico de EE.UU.	**2. Apoyar el uso pacífico de la tecnología**	**3. Gestionar los riesgos catastróficos**
Nuclear	Éxito parcial	Éxito parcial	Fracaso parcial
Aeroespacial	Éxito	Éxito	Éxito
Cibernética	Éxito	Éxito parcial	Fracaso parcial
Biotécnica	No aplicable	Éxito	Éxito parcial

Perfil tecnológico de la IA: ¿El peor escenario posible?

1. Capacidad destructiva: Alta
2. Perfil de coste: Diverso, pero potencialmente bajo
3. Perfil de complejidad: Diverso, pero potencialmente bajo
4. Potencial de doble uso militar/civil: Alto
5. Dificultad de espionaje y supervisión: Alta

Experiencia adquirida

Lección 1: Cambios tecnológicos radicales generan ideas radicales para las políticas del gobierno

Al igual que con previas tecnologías militares transformables, las implicaciones de seguridad nacional de la IA serán revolucionarias, no simplemente diferentes.

Los gobiernos alrededor del mundo considerarán, y algunos promulgarán, medidas normativas extraordinarias como respuesta, puede incluso que tan radicales como aquellas políticas que se consideraron durante las primeras décadas del armamento nuclear.

Lección 2: Las carreras armamentísticas a veces son inevitables, pero se pueden gestionar

En 1899, el temor colectivo a los bombardeos aéreos dio lugar a un tratado internacional que prohibía el uso de aeronaves militarizadas, pero dicha contención voluntaria se abandonó rápidamente y no consiguió evitar los combates aéreos durante la Primera Guerra Mundial.

El uso de IA en la guerra y en el espionaje probablemente será tan irresistible como lo fue en el caso de la aviación. Prevenir el uso militar masificado de la IA resultará prácticamente imposible.

Aunque las prohibiciones totales del uso de la IA en el ámbito de la seguridad nacional no sea una opción realista, se debería perseguir el objetivo más modesto de garantizar una gestión tecnológica segura y efectiva.

Lección 3: El gobierno debe, al mismo tiempo, promover y restringir la actividad comercial

El fracaso de reconocer la naturaleza de empleo de doble uso inherente de este tipo de tecnología puede cobrar vidas, como nos muestra el ejemplo del motor Nene de Rolls-Royce.

Tener la industria de tecnología digital más amplia y avanzada constituye una ventaja enorme para Estados Unidos. Sin embargo, la relación entre el gobierno y algunas instituciones de investigación destacadas de IA se encuentra repleta de tensiones.

Los encargados de formular políticas relacionadas con la IA deberán apoyar de manera efectiva los intereses de ambos grupos.

Lección 4: El gobierno debe formalizar sus objetivos para la seguridad tecnológica y prestar recursos adecuados

En cada uno de los cuatro casos, los responsables políticos se enfrentaron a cambios recíprocos entre seguridad y rendimiento, pero era más probable que el gobierno respondiera de manera más apropiada frente algunos riesgos en comparación con otros.

En todos los casos, los resultados de seguridad mejoraron cuando el gobierno creó organizaciones oficiales encargadas de mejorar la seguridad de sus respectivos ámbitos de tecnología y adecuar los recursos necesarios.

Dichos recursos no solo incluyen el aspecto de la financiación y los materiales necesarios, sino también el capital humano y la autoridad y acceso para ganar conflictos burocráticos.

Estados Unidos debería considerar la posibilidad de defender a las organizaciones oficiales de investigación y desarrollo encargadas de investigar y promover la seguridad de la IA en todos los departamentos gubernamentales y comerciales de la IA.

Lección 5: A medida que la tecnología va cambiando, también cambia el Interés Nacional de Estados Unidos

La reducción del coste y complejidad de las armas biológicas hizo que Estados unidos cambiara su estrategia de armas biológicas desde una posición de desarrollo agresivo a una posición de autocontención voluntaria.

En general, Estados Unidos tiene un interés estratégico en configurar el coste, la complejidad y los perfiles de ofensiva/defensiva de las tecnologías relacionadas con la seguridad nacional.

Como muestra el ejemplo de los aviones furtivos (sigilosos, indetectables por radar), las inversiones selectivas algunas veces permiten que Estados Unidos pueda influir el balance entre ofensiva/defensiva en ciertos ámbitos y construir una vanguardia tecnológica de larga duración.

Estados Unidos debería considerar como pueden configurar el perfil tecnológico de las aplicaciones militares y de inteligencia de la IA.

Parte 3: Recomendaciones para la inteligencia artificial y la seguridad nacional

Preservar el liderazgo tecnológico de EE.UU.

Recomendación 1: El Departamento de Defensa (DDD) debería realizar simulacros de combate enfocados en la IA, para identificar posibles innovaciones militares perjudiciales.

Recomendación 2: El DDD debería financiar análisis diversos y enfocados a largo plazo sobre la tecnología IA y sus implicaciones.

Recomendación 3: El DDD debería priorizar los sectores de gasto en investigación y desarrollo de la IA, que puedan aportar ventajas sostenibles y que mitiguen los riesgos clave.

Recomendación 4: La comunidad de inteligencia y defensa de Estados Unidos debería hacer grandes inversiones en capacidades ofensivas y defensivas “contra-IA”.

Apoyar el uso pacífico de la tecnología de IA

Recomendación 5: Se debería incrementar la financiación en investigación básica de IA a la Agencia de Proyectos de Investigación Avanzados de Defensa (DARPA), a la Actividad de Proyectos de Investigación Avanzados de Inteligencia (IARPA), a la Oficina de Investigación Naval, y a la Fundación Nacional para la Ciencia.

Recomendación 6: El Departamento de Defensa debería publicar una solicitud de información (RFI –*Request for Information* en inglés) sobre las capacidades de IA de doble uso.

Recomendación 7: Se debería dar recursos adicionales a In-Q-Tel para que pueda promover la colaboración entre la comunidad de seguridad nacional y la industria comercial de IA.

Mitigar riesgos catastróficos

Recomendación 8: El Consejo de Seguridad Nacional, el Departamento de Defensa,y el Departamento de Estado deberían investigar qué tipo de aplicaciones de IA debería restringir Estados Unidos con tratados.

Recomendación 9: El Departamento de Defensa y la Comunidad de Inteligencia deberían establecer organizaciones dedicadas a preservar la seguridad de la IA.

Recomendación 10: La Agencia de Proyectos de Investigación Avanzados de Defensa (DARPA) debería financiar la investigación a prueba de fallos y rendimiento de seguridad para los sistemas de IA.

Recomendación 11: El Instituto Nacional de Estándares y Tecnología (NIST) y la Agencia de Seguridad Nacional (NSA) deberían explorar las opciones tecnológicas para mitigar el riesgo catastrófico.