La Aeronave de Pruebas de Simulación de Vuelo Variable fabricada por Lockheed Martin, X-62A (VISTA), una aeronave de entrenamiento única en su clase es piloteada por un agente de IA el 13 de febrero de 2023 en la Base Aérea Edwards, California (aunque continuamente había pilotos de seguridad a bordo). La aeronave voló durante más de diecisiete horas y fue la primera vez que se utilizó IA en una aeronave táctica. (Foto: Kyle Brasier, Fuerza Aérea de EUA)

La inteligencia artificial en la guerra moderna

Innovación estratégica y riesgos emergentes

Ryan Atkinson, PhD

Traducción de Alexandro Bonilla, Army University Press

En los últimos años, la inteligencia artificial (IA) ha logrado notables victorias en contra de oponentes humanos, como AlphaZero en ajedrez, AlphaGo en Go y AlphaStar en StarCraft II. La Fuerza Aérea de EUA y la Agencia de Proyectos de Investigación Avanzada de Defensa (*Defense Advance Research Projects Agency, DARPA*) han creado AlphaDogfight para poner a prueba la IA contra un piloto humano. La IA atacó al piloto desde el frente con gran velocidad y precisión jugando a quién es más valiente «chicken game», «ganando 5-0 mediante maniobras agresivas y precisas que el piloto humano no pudo superar»[1]. Estos avances resaltan la creciente capacidad de la IA para desafiar y superar las habilidades humanas en escenarios complejos, enfatizando su potencial para reconfigurar los entornos competitivos y estratégicos.

Cada vez más, la toma de decisiones es automatizada y la participación humana es disminuida a medida que los sistemas autónomos tienen más control sobre las aeronaves. La Fuerza Aérea de EUA sometió a prueba un sistema de IA que piloteaba la aeronave táctica X-62A, también llamada VISTA[2]. Este importante logro en el desarrollo de sistemas de IA indica el potencial de futuras operaciones militares autónomas o semiautónomas.

Las tecnologías de doble uso son cada vez más importantes a medida que evolucionan las herramientas de IA, lo que plantea nuevos riesgos y oportunidades. Estas tecnologías pueden aplicarse a usos civiles que informen a las operaciones militares y viceversa. Por ejemplo, los precedentes y prácticas de la IA utilizados para dirigir anuncios en las redes sociales para campañas de marketing o políticas pueden servir de apoyo a la comunicación estratégica militar y a las operaciones psicológicas. Se desarrollarán nuevas medicinas, pero también nuevas armas químicas, lo que aumentará la necesidad de seguir investigando los riesgos y oportunidades relacionados[3]. Las tecnologías de doble uso siguen siendo un arma de doble filo para las aplicaciones de la IA.

La innovación en defensa y las asociaciones sólidas entre las fuerzas armadas y la industria son importantes. Las empresas emergentes de IA dentro de la industria de defensa ofrecen nuevas iniciativas de innovación entre aliados. Los casos críticos se encuentran a través de un énfasis en la colaboración dentro de la extensa red de titanes de la industria de defensa y nuevos innovadores emergentes. La IA está cambiando rápidamente la tecnología y las tácticas militares, además la naturaleza de doble uso de la tecnología supone un reto para el desarrollo de la IA aplicada en entornos militares.

Los sistemas de armas autónomos, que operan sin intervención humana directa, representan un avance significativo en la tecnología militar. Las aplicaciones militares benéficas incluyen casos específicos del Ejército, como los sistemas inteligentes de apoyo a la toma de decisiones y el reconocimiento asistido de objetivos, que pueden reducir la carga mental de los operadores, permitiendo una toma de decisiones más rápida[4]. Este enfoque proporciona ventajas, como tiempos de respuesta rápidos, la capacidad de operar en entornos de alto riesgo y un menor peligro para el personal humano.

Inteligencia generativa y enjambres coordinados

Las tecnologías emergentes relacionadas con los agentes generativos ofrecen aplicaciones de doble uso. Investigadores de Stanford y Google mostraron «agentes de software computacionales que simulan un comportamiento humano creíble», semejante a una pequeña ciudad de veinticinco agentes[5]. Se observó cooperación entre el grupo, lo que llevó a comportamientos sociales emergentes para «intercambiar información, formar nuevas relaciones y coordinar actividades conjuntas»[6].

La arquitectura permite a los agentes generativos «recordar, recuperar, reflexionar, interactuar con otros agentes y planificar a través de circunstancias que evolucionan dinámicamente»[7]. Los grandes modelos de lenguaje se utilizan para «complementar esas capacidades con el propósito de apoyar la coherencia del agente a más largo plazo, la capacidad para controlar la memoria que evoluciona dinámicamente y producir reflexiones de nivel superior de forma recurrente»[8].

Las democracias resilientes necesitan, intrínsecamente, mecanismos internos adaptables para ajustarse a los cambios y hacer frente a situaciones inesperadas de forma rápida. La aplicación de modelos de lenguaje a situaciones reales suele tener consecuencias

Ryan Atkinson, PhD, es becario postdoctoral en la Universidad de Carleton, Ottawa, Canadá. Su trabajo está financiado por la Red Canadiense de Defensa y Seguridad.

imprevistas y emergentes. Las democracias deben crear medidas de corrección proactivas para hacer frente a los riesgos emergentes asociados al uso generalizado de la IA generativa y los grandes modelos de lenguaje, que añaden un nivel adicional de retos para la seguridad. El abuso malintencionado de los modelos de lenguaje demuestra un inmenso desafío para las futuras elecciones y procesos democráticos[9].

Los riesgos asociados a las operaciones de influencia extranjera que utilizan deepfake (ultrafalso) vídeo y audio son cada vez más personalizados y específicos para cada caso. La investigación futura debe abordar la proliferación de operaciones de información patrocinadas por el Estado que utilizan la desinformación generada para fomentar «la incomprensión generalizada, las divisiones sociales e impactar negativamente en los sistemas económicos y políticos»[10]. La automatización también se ha aplicado al comportamiento grupal que involucra a drones que envían información a otros en el enjambre, proporcionando un inmenso valor para las operaciones militares. Las investigaciones sobre inteligencia de enjambre han incluido agentes autónomos para aplicaciones militares, y actualmente se están realizando pruebas en Estados Unidos y China[11].

El uso de drones autónomos con inteligencia artificial empleados en enjambres tiene un potencial significativo para causar una destrucción amplia y a gran escala en objetivos designados. Las fuerzas seleccionadas tendrían inmensas dificultades técnicas para defenderse de un primer ataque masivo y ampliamente coordinado contra múltiples objetivos. El empleo simultáneo de un gran número de drones podría sobrepasar las capacidades materiales de una fuerza defensora, así como el mando y control y la gobernanza civil de un defensor, en cuestión de horas, si no de minutos. Cabe destacar que, en junio de 2024, el Ejército Popular de Liberación de China llevó a cabo ejercicios con drones, incluyendo técnicas de enjambre, centrados en la incautación de islas que reflejaban de forma transparente las acciones que probablemente llevaría a cabo en una invasión de Taiwán. (Foto cortesía del Ejército de EUA/Shutterstock)

Los drones han planteado un importante desafío para el armamento convencional. En el Mar Rojo, un dron de 2 000 dólares derribó un misil de 2 millones de dólares[12]. En Ucrania, se están empleando drones de 400 dólares para destruir tanques de 2 millones de dólares[13]. Este marcado contraste subraya la creciente brecha entre el costo de los medios militares tradicionales y la asequibilidad y eficacia de la tecnología moderna de drones.

El crecimiento de la inteligencia artificial en China

Hasta el 2021, la industria china de la IA tenía un valor de 150 000 millones de yuanes (23 200 millones de dólares) y se espera que alcance más de 400 000 millones de yuanes (55 000 millones de dólares) en 2025[14]. El Plan de Desarrollo de la IA de Nueva Generación de China estableció el objetivo de que la IA contribuyera con 150 000 millones de dólares al PIB chino para el 2030[15]. En agosto de 2023, Pekín aprobó el lanzamiento público de tecnologías de IA generativa de las empresas chinas Tencent, Baidu, Huawei Technologies, Alibaba Group, JD.com, ByteDance, iFlytek y Kuaishou Technology[16].

Microsoft publicó un informe en septiembre de 2023 que mostraba cómo se utilizan las estrategias de IA generativa en las operaciones de influencia llevadas a cabo por la República Popular China (RPC)[17]. El Departamento de Justicia de Estados Unidos informó de la existencia de un grupo llamado Grupo de Trabajo del Proyecto Especial 912 dentro del Ministerio de Seguridad Pública de China que operaba una granja de trols en las redes sociales, que «creaba miles de personas falsas en línea y difundía propaganda del Partido Comunista de China (PCCh) dirigida a activistas prodemocráticos»[18].

El informe de Microsoft señalaba que, en marzo de 2023, las presuntas operaciones de influencia de la RPC «en las redes sociales occidentales han comenzado a aprovechar la IA generativa para crear contenido

(Foto de Adobe Stock)

visual», que «ya ha atraído a un mayor número de usuarios reales de las redes sociales»[19]. Las operaciones de información de China se harán más sofisticadas, a medida que las aplicaciones de la IA generativa se adapten cada vez más a objetivos específicos.

El informe describía los «estudios multilingües de celebridades en Internet» del PCCh, con 230 empleados de medios de comunicación estatales y afiliados que se hacen pasar por influencers independientes en las redes sociales, dirigidos a los medios de comunicación sociales occidentales[20]. Microsoft señaló que en 2022 y 2023, «nuevos influencers siguen debutando cada siete semanas en promedio»[21]. La Radio Internacional de China es una de las numerosas entidades que «reclutó, entrenó, promovió y financió» tales capacidades entre otras entidades mediáticas patrocinadas por el Estado para llegar a 103 millones de personas en cuarenta idiomas[22].

Entre las diversas plataformas en las que China ha centrado su actividad figuran empresas como Vimeo, Wattpad, Indeed, Rotten Tomatoes, Instagram, Quora, Medium, Facebook, Reddit, Tumblr, YouTube, Twitter/X, Pinterest, Blogger, TikTok, Flickr y LinkedIn[23]. Una red de influencia patrocinada pone de manifiesto un importante desafío, ya que las poblaciones occidentales pueden verse influidas por personalidades patrocinadas por gobiernos extranjeros, lo que ofrece la posibilidad de realizar operaciones de subversión a través de aplicaciones para compartir vídeos.

Microsoft proporcionó ejemplos de enero de 2022 acerca de una campaña alineada con el PCCh que tenía como objetivo «la organización no gubernamental española Safeguard Defenders después de que revelara la existencia de más de 50 comisarías de policía chinas en el extranjero»[24]. La campaña implantó 1 800 cuentas en plataformas de medios sociales y docenas de sitios web para difundir memes, vídeos y mensajes alineados con el PCCh en los que se criticaba a Estados Unidos y a otras democracias. Los mensajes se compartieron en holandés, griego, indonesio, sueco, turco y uigur, entre otros idiomas, en plataformas como Fandango, Rotten Tomatoes, Medium, Chess.com y VK.

Redes aliadas de innovación en defensa

Los países persiguen la superioridad tecnológica en IA para obtener ventajas competitivas en diversos ámbitos, como las capacidades militares, la productividad económica y la innovación tecnológica. En los últimos años, los aliados de la OTAN se han centrado de forma significativa en la innovación en defensa y los retos relacionados con ella. La OTAN publicó su

primera estrategia de IA en octubre de 2021[25]. En la Cumbre de Washington de julio de 2024 se publicó una estrategia de IA actualizada[26].

El Acelerador de Innovación en Defensa para el Atlántico Norte (*Defense Innovation Accelerator for the North Atlantic*, DIANA) de la OTAN trabaja con los gobiernos, la industria y el mundo académico para apoyar el desarrollo de tecnologías emergentes en América y Europa. El programa proporciona a los innovadores acceso a una red profesional para ayudarles a desarrollar un programa acelerador personalizado[27]. Además de la IA, la OTAN se ha centrado en numerosas tecnologías disruptivas emergentes, como los sistemas autónomos, tecnologías cuánticas, biotecnología y tecnologías de mejora humana, sistemas hipersónicos, el espacio, materiales y fabricación novedosos, energía y propulsión, y redes de comunicaciones de nueva generación[28].

DIANA entró en funcionamiento en el verano de 2023, donde lanzó su primera ronda de retos para fomentar la innovación en necesidades primordiales y específicas de seguridad con el fin de orientar el avance tecnológico[29]. En 2023, la OTAN lanzó la primera ronda de retos para apoyar el desarrollo de tecnologías de doble uso para resolver problemas de resiliencia energética, detección y vigilancia, e intercambio seguro de información[30].

DIANA lanzó cinco nuevos retos en 2024, que incluyen energía y potencia, seguridad de los datos e información, detección y vigilancia, salud humana y rendimiento, e infraestructuras primordiales y logística[31]. DIANA está comprometida a fomentar soluciones de vanguardia y reforzar las capacidades estratégicas de la OTAN en un panorama global cada vez más complejo. Estas iniciativas se ajustan a la necesidad primordial de innovaciones sólidas para la defensa y colaboraciones estratégicas esenciales para contrarrestar la rápida evolución de las aplicaciones militares de la IA. ■

Notas

1. «AlphaDogfight Trials Foreshadow Future of Human-Machine Symbiosis», Defense Advanced Research Projects Agency (DARPA), 26 de agosto de 2020, https://www.darpa.mil/news-events/2020-08-26.

2. «ACE Program's AI Agents Transition from Simulation to Live Flight», DARPA, 13 de febrero de 2023, https://www.darpa.mil/news-events/2023-02-13.

3. Fabio Urbina et al., «AI in Drug Discovery: A Wake-up Call», *Drug Discovery Today* 28, n.º 1 (enero de 2023): Artículo 103410, https://doi.org/10.1016/j.drudis.2022.103410.

4. David Oniani et al., «Adopting and Expanding Ethical Principles for Generative Artificial Intelligence from Military to Healthcare», npj *Digital Medicine* 6, n.º 1 (2 de diciembre de 2023): 1–10, https://doi.org/10.1038/s41746-023-00965-x.

5. Joon Sung Park et al., «Generative Agents: Interactive Simulacra of Human Behavior», arXiv, 5 de agosto de 2023, http://arxiv.org/abs/2304.03442.

6. *Ibid*., sec. 3.4.

7. *Ibid*., sec. 1.

8. *Ibid*.

9. Tom Di Fonzo, «What You Need to Know About Generative AI's Emerging Role in Political Campaigns», Tech Policy Press, 12 de octubre de 2023, https://www.techpolicy.press/what-you-need-to-know-about-generative-ais-emerging-role-in-political-campaigns/.

10. U.S. Department of Homeland Security, *Unveiling the Dark Art: Investigating the Nexus between Generative Artificial Intelligence and Foreign Malign Influence* (Washington, DC: U.S. Department of Homeland Security, 29 de septiembre de 2023), https://www.dhs.gov/sites/default/files/2023-09/23_0906_oia_GAI_ForeignMalignInfluence_508.pdf.

11. Matt Berg, «Killer Robot Swarms, an Update», Politico, 4 de enero de 2024, https://www.politico.com/newsletters/digital-future-daily/2023/02/07/killer-robot-swarms-an-update-00081623.

12. Laura Seligman y Matt Berg, «A $2M Missile vs. a $2,000 Drone: Pentagon Worried over Cost of Houthi Attacks», Politico, última actualización: 20 de diciembre de 2023, https://www.politico.com/news/2023/12/19/missile-drone-pentagon-houthi-attacks-iran-00132480.

13. Veronika Melkozerova, «The Future of Warfare: A $400 Drone Killing a $2M Tank», Politico, 26 de octubre de 2023, https://www.politico.eu/article/future-warfare-400-army-strike-drone-unit-2m-tank/.

14. Iris Deng, «Shenzhen Is First Chinese City to Draft Regulations Specifically for AI», *South China Morning Post* (sitio web), 30 de junio de 2021, https://www.scmp.com/tech/policy/article/3139319/shenzhen-chinas-silicon-valley-plans-turbocharge-local-ai-development.

15. Eamon Barrett, «AI in China: TikTok Is Just the Beginning», *Fortune* (sitio web), 20 de enero de 2020, https://fortune.com/longform/tiktok-app-artificial-intelligence-addictive-bytedance-china/.

16. Zhou Xin, «Too Late Now for US to Hold Back China in Global AI Race», Nikkei Asia, 24 de octubre de 2023, https://asia.nikkei.com/Opinion/Too-late-now-for-U.S.-to-hold-back-China-in-global-AI-race.

17. Microsoft Threat Intelligence, «Sophistication, Scope, and Scale: Digital Threats from East Asia Increase in Breadth and

Effectiveness» (Redmond, VA: Microsoft, septiembre de 2023), 6, https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RW1aFyW.

18. *Ibid*., 6.

19. *Ibid*.

20. *Ibid*., 7.

21. *Ibid*.

22. *Ibid*.

23. *Ibid*., 10.

24. *Ibid*.

25. «Summary of the NATO Artificial Intelligence Strategy», NATO, 22 de octubre de 2021, https://www.nato.int/cps/en/natohq/official_texts_187617.htm.

26. «NATO Releases Revised AI Strategy», NATO, 10 de julio de 2024, https://www.nato.int/cps/en/natohq/news_227234.htm.

27. «NATO DIANA Announces First Cohort of Innovators, Launches Call for Mentors», NATO, 4 de diciembre de 2023, https://www.nato.int/cps/en/natohq/news_220930.htm.

28. «Emerging and Disruptive Technologies», NATO, 22 de junio de 2023, https://www.nato.int/cps/en/natohq/topics_184303.htm.

29. «Defense Innovation Accelerator for the North Atlantic», NATO, 5 de julio de 2024, https://www.nato.int/cps/en/natohq/topics_216199.htm.

30. «NATO's Innovation Accelerator Becomes Operational and Launches First Challenges», NATO, 19 de junio de 2023, https://www.nato.int/cps/en/natohq/news_215792.htm.

31. «2024 DIANA Challenge Programme Call for Proposals», NATO, accedido el 22 de julio de 2024, https://www.diana.nato.int/resources/site1/general/2024_challenge_programme_web.pdf.