

## **Declaración política sobre uso responsable de inteligencia artificial y autonomía en el ámbito militar**

Cada vez más Estados están desarrollando capacidades de inteligencia artificial (IA) en el ámbito militar, entre ellas, el uso de IA en la habilitación de funciones y sistemas autónomos. El uso de IA en el ámbito militar puede y debería ser ético, responsable e incrementar la seguridad internacional. El uso de IA en el ámbito militar debe cumplir con el derecho internacional correspondiente. En particular, el uso de IA en los conflictos armados debe efectuarse conforme a las obligaciones que los Estados han contraído en virtud del derecho internacional humanitario, incluidos sus principios fundamentales. El uso de las capacidades de IA en el ámbito militar debe estar sujeto a la rendición de cuentas, lo que supone un uso de IA en operaciones militares dentro de un marco de una cadena de comando y control responsable a cargo de seres humanos. Un enfoque del uso de IA en el ámbito militar basado en principios debería incluir un análisis preciso de los riesgos y beneficios, y debería también minimizar sesgos y accidentes involuntarios. Los Estados deberían tomar medidas apropiadas que garanticen un desarrollo, implementación y uso responsable de las capacidades de IA en el ámbito militar, particularmente en lo que respecta al uso de IA en la habilitación de funciones y sistemas autónomos. Tales medidas deberían ponerse en práctica en las etapas pertinentes del ciclo de vida de las capacidades de IA en el ámbito militar.

Los Estados que respaldan esta Declaración consideran que las medidas a continuación se deberían adoptar al desarrollar, implementar o utilizar capacidades de IA en el ámbito militar, particularmente con respecto al uso de IA en la habilitación de funciones y sistemas autónomos:

- A. Los Estados deberían asegurarse de que sus organizaciones militares adopten y pongan en práctica estos principios para desarrollar, implementar y utilizar las capacidades de IA de manera responsable.
- B. Los Estados deberían emprender acciones pertinentes, entre ellas, revisiones legales, que garanticen que las capacidades de IA en el ámbito militar se utilizarán de conformidad con sus respectivas obligaciones en virtud del derecho internacional, especialmente el derecho internacional humanitario. Los Estados deberían también analizar cómo se podrían utilizar las capacidades de IA para mejorar el cumplimiento del derecho internacional humanitario y para aumentar la protección de civiles y los bienes de carácter civil en los conflictos armados.
- C. Los Estados deberían garantizar que los altos mandos militares supervisen, de forma eficaz y adecuada, el desarrollo y la implementación de capacidades de IA en el ámbito militar que tengan aplicaciones de repercusiones graves, lo que incluye a título nominativo y no exclusivo esos sistemas de armas.
- D. Los Estados deberían tomar medidas proactivas para minimizar los sesgos no deseados en las capacidades de IA en el ámbito militar.
- E. Los Estados deberían garantizar que el personal pertinente tome las precauciones necesarias al desarrollar, implementar y usar las capacidades de IA en el ámbito militar, particularmente en lo que respecta a sistemas de armas que incorporan tales capacidades.

- F. Los Estados deberían garantizar que las capacidades de IA en el ámbito militar se desarrollen con metodologías, fuentes de datos, procesos de diseño y documentación que sean transparentes y verificables por el personal de defensa correspondiente..
- G. Los Estados deberían garantizar que el personal que utilice o apruebe el uso de capacidades de IA en el ámbito militar cuente con una capacitación que les permita comprender debidamente las capacidades y limitaciones de esos sistemas para poder tomar decisiones adecuadas basadas en el contexto sobre el uso de esos sistemas y mitigar el riesgo del sesgo de automatización.
- H. Los Estados deberían garantizar que todas las capacidades de IA en el ámbito militar tengan usos explícitos y claramente definidos, y que se desarrollen y piensen para cumplir con esas funciones previstas.
- I. Los Estados deberían asegurarse de que la seguridad, la vigilancia y la eficacia de las capacidades de IA en el ámbito militar estén sujetas a pruebas rigurosas y pertinentes, así como a garantías dentro de sus usos claramente definidos y a lo largo de todo su ciclo de vida. Durante los procesos de autoaprendizaje o actualización continua de las capacidades de IA en el ámbito militar, los Estados deberían garantizar que las medidas de seguridad fundamentales no se han degradado a través de procesos como, por ejemplo, la vigilancia.
- J. Los Estados deberían poner en práctica salvaguardas adecuadas que reduzcan los riesgos de fallas en las capacidades de IA en el ámbito militar, por ejemplo, la capacidad para detectar y evitar consecuencias no deseadas y la capacidad para retirar o desactivar sistemas emplazados que demuestren un comportamiento no deseado.

Para fomentar los objetivos de esta Declaración, los Estados que la respaldan:

- implementarán estas medidas al desarrollar, aplicar o utilizar capacidades de IA en el ámbito militar, incluidas las que habilitan funciones y sistemas autónomos;
- harán público su compromiso con esta Declaración y publicarán la información correspondiente relativa a su aplicación de estas medidas;
- apoyarán otras iniciativas adecuadas que garanticen que las capacidades de IA en el ámbito militar se usen de manera responsable y legítima;
- seguirán examinando cómo desarrollar, implementar y utilizar las capacidades de IA en el ámbito militar de manera responsable y legal entre los Estados que respaldan esta Declaración;
- promoverán la aplicación efectiva de estas medidas, y las refinarán, o establecerán nuevas medidas que los Estados que respaldan la Declaración consideren apropiadas; e
- involucrarán al resto de la comunidad internacional para promover estas medidas, por ejemplo, en otros foros sobre temas relacionados, sin perjuicio de las conversaciones en curso sobre temas conexos en otros foros.

Los Estados que respaldan esta Declaración reconocen que los conceptos de inteligencia artificial y autonomía están sujetos a diferentes interpretaciones. Para los fines de la presente Declaración, por inteligencia artificial podría entenderse la capacidad de las máquinas para realizar tareas que de otro modo requieren inteligencia humana. Esto podría abarcar, entre otras,

reconocer patrones, aprender a partir de la experiencia, extraer conclusiones, formular predicciones o generar recomendaciones. Una aplicación de IA podría guiar o modificar el comportamiento de un sistema físico autónomo o realizar tareas que permanecen exclusivamente en el terreno digital. Autonomía puede entenderse como un sistema con diferentes grados e implicar un sistema que funciona sin intervención humana tras su activación.