MARCO DE DERECHO INTERNACIONAL PÚBLICO Y USOS MILITARES DE LA INTELIGENCIA ARTIFICIAL EN LA UE

FRAMEWORK OF PUBLIC INTERNATIONAL LAW AND MILITARY USES OF ARTIFICIAL INTELLIGENCE IN THE EU

Francisco Lamas López[*]
Alfonso Peralta Gutiérrez[**]

Resumen: Este texto discute los problemas éticos y legales de usar Inteligencia Artificial (IA) en contextos militares. Es necesario tener una IA confiable que respete los derechos fundamentales, regulaciones y valores centrales, evitando causar daño no intencional. Los principios de beneficencia, no maleficencia, autonomía, justicia y explicabilidad proporcionan un marco para lograr una IA confiable. Es importante que el control humano sobre la IA garantice responsabilidad, supervisión y rendición de cuentas, y se aboga por un marco global para el uso de la IA que refleje el derecho humanitario internacional. El desarrollo y uso de Sistemas de Armas Autónomas Letales (*LAWS*) sin un control humano significativo es una gran preocupación y debe ser regulado para garantizar transparencia y regulación adecuada. La UE debería liderar la promoción de una estrategia integral sobre IA y defensa que enfatice la necesidad de una supervisión y control humano significativos. La regulación de la robótica y la IA debería reflejar los valores humanistas europeos y universales.

Abstract: *This paper discusses the ethical and legal issues of using Artificial Intelligence (AI) in military contexts. It emphasizes the need for trustworthy AI that respects fundamental rights, regulations, and core values while avoiding unintentional harm. The principles of beneficence, non-maleficence, autonomy, justice, and explicability provide a framework for achieving trustworthy AI. The paper highlights the importance of human control over AI to ensure responsibility, oversight, and accountability and advocates for a global framework for the use of AI that reflects international humanitarian law. The development and use of Lethal Autonomous Weapon Systems (LAWS) without significant human control is a major concern and must be regulated to ensure transparency and proper regulation. The EU should take the lead in promoting a comprehensive strategy on AI and defence that emphasizes the need for meaningful human oversight and control. The regulation of robotics and AI should reflect European and universal humanist values.*

Palabras clave: IA confiable, Usos militares de la IA, Derecho internacional público, Control humano, Regulaciones, Liderazgo de la UE, Sistemas de Armas Autónomas Letales (LAWS)

Fecha de recepción del trabajo: 25 de septiembre de 2023. Fecha de aceptación de la versión final: 9 de noviembre de 2023.

[*] Alférez de navío del Cuerpo de Ingenieros de la Armada. Doctor ENPC ParisTech. Responsable Técnico del Centro de Supervisión y Análisis de Datos de la Armada (CESADAR), Arsenal de Cartagena. Profesor *Computing and Artificial Intelligence Laboratory* (CAILab), Facultad de Ciencia y Tecnología. Universidad Camilo José Cela. Email: francisco.lamas@ucjc.edu

[**] Juez Especialista en IA y derecho digital. Juez del Juzgado de Primera Instancia e Instrucción nº 2 de Guadix Poder Judicial español. Investigador predoctoral en Universidad de Granada. Email: a.peralta@poderjudicial.es

DOI: 10.36151/reei.46.17

Keywords: *Trustworthy AI, AI Military uses, International public law, Human control, Regulations, EU leadership, Lethal Autonomous Weapon Systems (LAWS)*

I. INTRODUCCIÓN

En el presente artículo pretendemos analizar el estado del arte de las implicaciones éticas y legales del uso de la inteligencia artificial (IA) en contextos militares, haciendo hincapié en la necesidad de una IA confiable que respete los derechos y regulaciones fundamentales. Como metodología se ha elaborado a partir de una revisión exhaustiva de los instrumentos normativos internacionales existentes sobre el uso de la IA en contextos militares y el marco jurídico internacional aplicable. Se parte de una definición de inteligencia artificial y una identificación de los principios rectores éticos de la misma, para una mayor profundización en el estado de la cuestión del marco jurídico internacional aplicable al ámbito militar a través de un análisis documental. Con ello, se pretende identificar los principios rectores y las mejores prácticas para garantizar que la IA se utilice de manera responsable y ética en el ámbito militar y acabar formulando unas '*Conclusiones y Perspectivas futuras*'.

La razón de la elección de esta temática se debe a la presentación de este trabajo conjunto en el taller de la Agencia Europea de Defensa (EDA, por sus siglas en inglés) de la Comisión Europea titulado "*Trustworthy AI for Defence*" (IA confiable para Defensa) celebrado entre los días 5 al 7 de septiembre en Bruselas y defendido por el coautor Francisco Lamas López. Los resultados obtenidos en esta investigación corresponden a un nivel exploratorio y descriptivo. Esto es, una primera aproximación sobre el diseño normativo relativo a las leyes de la guerra concernientes al uso de la IA, las pruebas de concepto necesarias sobre soluciones técnicas éticas aplicadas a defensa y en última instancia el sistema de asunción de responsabilidades en caso de ausencia de supervisión humana.

El desarrollo de la Inteligencia Artificial (IA), la robótica y las tecnologías relacionadas está avanzando rápidamente y puede afectar directamente todos los aspectos de nuestras sociedades, incluyendo los valores y principios fundamentales. Por lo tanto, la Unión Europea (UE) y sus Estados miembros tienen la responsabilidad de asegurarse de que estas tecnologías se centren en lo humano, concebidas principalmente para ser utilizadas en servicio de la humanidad y el bien común, y que contribuyan al bienestar e interés general de sus ciudadanos. Es esencial proporcionar un marco legal adecuado y completo que aborde los aspectos éticos de estas tecnologías, así como la responsabilidad, transparencia y rendición de cuentas, especialmente para la IA, la robótica y las tecnologías relacionadas consideradas de alto riesgo. Este marco debe reflejar los valores humanistas europeos y universales inherentes que son aplicables a toda la cadena de valor en el desarrollo, implementación y uso de la IA.

La orientación ética es un buen punto de partida, pero no es suficiente para garantizar que las empresas actúen de manera justa y aseguren una protección efectiva de las personas. Por lo tanto, se necesita un análisis para determinar hasta qué punto las normas del derecho internacional se adaptan a estas tecnologías y destacar los desafíos y riesgos que representan para la autoridad del Estado, con el objetivo de poder ser gestionados adecuada y proporcionalmente. Un enfoque europeo armonizado sobre estos temas requiere la adopción de una definición

 DOI: 10.36151/reei.46.17

común de "IA" y la garantía de que se respeten los valores fundamentales de la UE, los principios de la Carta de los Derechos Fundamentales y la legislación internacional de derechos humanos. En este sentido, la Comisión Europea debe considerar tanto los aspectos civiles como militares del uso de la IA en su Libro Blanco.

La creación de un marco legal europeo para la IA y la robótica puede garantizar la certeza jurídica y la protección de los derechos humanos y, al mismo tiempo, fomentar la innovación y la competitividad de las empresas europeas en este campo. Asimismo, la UE debe desempeñar un papel líder en la gobernanza global de la IA y la robótica, ya que es un problema que afecta a todo el mundo y la UE tiene una responsabilidad particular de asegurar que se respeten los derechos humanos y la dignidad humana en el desarrollo y uso de estas tecnologías. Y es que como describe COTINO HUESO (2019)[1] es precisamente en la UE donde se apuesta por una ética confiable de la IA en el diseño y '*made in Europe*', para posicionarse frente a Estados Unidos y especialmente China.

También debe trabajar en estrecha colaboración con las organizaciones internacionales pertinentes, como la ONU y el Consejo de Europa, para desarrollar normas y principios internacionales para la IA y la robótica y promover el diálogo internacional para abordar cuestiones éticas, legales y de seguridad en este ámbito.

II. DEFINICIÓN DE INTELIGENCIA ARTIFICIAL

La Comisión Europea proporcionó una definición de Inteligencia Artificial en el libro blanco de la IA (COMISIÓN EUROPEA, 2020[2]) y se refiere a ella como sistemas que muestran un comportamiento inteligente al analizar su entorno y tomar acciones para lograr objetivos específicos, con cierto grado de autonomía. Estos sistemas pueden ser puramente basados en software, actuando en el mundo virtual, o estar integrados en dispositivos de hardware como robots o automóviles autónomos. Es decir, se puede afirmar tras esta definición que la IA es una disciplina que se enfoca en crear sistemas y programas capaces de simular funciones cognitivas y conductuales humanas, como la percepción, el razonamiento y el aprendizaje. Hay que subrayar que esta tecnología ha experimentado un rápido avance en los últimos años y está presente en varios campos, desde la robótica y el procesamiento de datos hasta la atención médica y la conducción autónoma.

En el contexto legal, la IA plantea diversos desafíos en términos de su regulación y uso responsable. La Comisión Europea ha identificado la necesidad de una estrategia integral para la IA que promueva su desarrollo al tiempo que garantiza la protección de los derechos fundamentales y la seguridad de los ciudadanos. En este sentido, la Directiva 2016/679 del Parlamento

1 COTINO HUESO, L., "Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y big data confiables y su utilidad desde el derecho," Revista Catalana de Dret Públic, núm. 58, 2019, pp. 29-48. DOI: doi.org/10.2436/rcdp.i58.2019.3303

2 COMISIÓN EUROPEA, "Libro blanco sobre inteligencia artificial: Un enfoque europeo para la excelencia y la confianza", 2020, https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

DOI: 10.36151/reei.46.17

Europeo y del Consejo de 27 de abril de 2016 (GDPR) *(PARLAMENTO EUROPEO, 2016[3])* establece los principios de protección de datos personales que deben ser respetados por cualquier tecnología, incluyendo la IA.

Además, la Unión Europea ha publicado una serie de recomendaciones no vinculantes para el diseño ético y responsable de los sistemas de IA, incluyendo la necesidad de garantizar la transparencia y la responsabilidad en el proceso de toma de decisiones, la inclusión de salvaguardas para prevenir la discriminación y la promoción de la confianza y la seguridad. Estas recomendaciones se reflejan en el Libro Blanco sobre IA de la Comisión Europea, publicado en 2020 *(COMISIÓN EUROPEA, 2020[4])*.

A nivel internacional, la IA también ha recibido atención de las Naciones Unidas (ONU), que ha publicado un conjunto de directrices éticas para la IA en el informe "*Ética en la Inteligencia Artificial*" *(NACIONES UNIDAS, 2019[5])*. Estas directrices establecen principios generales como la no maleficencia, la transparencia y la responsabilidad, así como recomendaciones específicas para su aplicación en diversos campos, como la salud, la educación y la justicia. Por parte de la UNESCO, también se ha publicado la "Recomendación sobre la ética de la inteligencia artificial" (UNESCO 2021)[6], el que se considera que fue el primer marco normativo universal sobre ética de la IA en la medida en que competen al mandato de la UNESCO y que aborda la ética de la IA como una reflexión normativa sistemática, basada en un marco integral, global, multicultural y evolutivo de valores, principios y acciones interdependientes. Fue adoptado por los 193 Estados miembros de la UNESCO en noviembre de 2021. Ese mismo año, como parte de su labor en materia de tecnología y derechos humanos, la Oficina del Alto Comisionado para los Derechos Humanos de la ONU publicó un informe (ACNUDH 2021)[7] en el que se analiza cómo la IA —incluidas la elaboración automática de perfiles, la toma de decisiones y otras tecnologías de aprendizaje para las máquinas— afecta al derecho a la intimidad y a otros derechos, incluidos los relativos a la salud, la educación, la libertad de movimiento, la libertad de reunión y asociación pacífica, y la libertad de expresión.

Por lo tanto, es importante reconocer que la IA es una tecnología con un enorme potencial, pero también presenta importantes desafíos legales y éticos. La regulación debe equilibrar la

---

3 PARLAMENTO EUROPEO, Directive 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), Consejo de la Unión Europea, Diario Oficial de la Unión Europea, L 119, 2016, pp. 1-88.

4 COMISIÓN EUROPEA, "Libro blanco sobre inteligencia artificial: Un enfoque europeo para la excelencia y la confianza", 2020, https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

5 NACIONES UNIDAS, "Ethics in artificial intelligence: Report of the Secretary-General", 2019. Disponible en: https://undocs.org/en/A/74/260

6 UNESCO, "Recomendación sobre la ética de la inteligencia artificial", 2021. Disponible en: https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa.locale=es

7 ACNUDH, Oficina del Alto Comisionado para los Derechos Humanos. "*Annual report of the United Nations High Commissioner for Human Rights and reports of the Office of the High Commissioner and the Secretary-General. The right to privacy in the digital age"*, 2021. Disponible en: https://www.ohchr.org/sites/default/files/HRBodies/HRC/RegularSessions/Session48/Documents/A_HRC_48_31_AdvanceEditedVersion.docx

DOI: 10.36151/reei.46.17

protección de los derechos fundamentales con la promoción de la innovación y el desarrollo económico. Las regulaciones europeas y los estándares internacionales proporcionan un marco de principios y recomendaciones para el diseño responsable de la IA que deben ser considerados por todas las partes involucradas en su desarrollo y uso.

III. PRINCIPIOS RECTORES

Según el Grupo de Expertos de Alto Nivel en Inteligencia Artificial (AI-HLEG) *(COMISIÓN EUROPEA, 2018*[8]*)*, la IA confiable tiene dos componentes: debe respetar los derechos fundamentales, la regulación aplicable y los principios y valores fundamentales, asegurando un "propósito ético", y debe ser técnicamente sólida y confiable, ya que, incluso con buenas intenciones, la falta de dominio tecnológico puede causar daños no intencionales. Si leemos el AI-HLEG *(COMISIÓN EUROPEA, 2018)*, 2018, los principios y valores en el contexto de la IA podrían resumirse en lo siguiente:

- **El principio de Beneficencia**: "*Hacer el bien*": Los sistemas de IA deben diseñarse y desarrollarse para mejorar el bienestar individual y colectivo. Los sistemas de IA pueden hacerlo generando prosperidad, creación de valor, maximización de la riqueza y sostenibilidad.
- **El principio de No maleficencia**: "*No hacer daño*": Los sistemas de IA no deben dañar a los seres humanos. Por diseño, los sistemas de IA deben proteger la dignidad, integridad, libertad, privacidad, seguridad y protección de los seres humanos en la sociedad y en el trabajo. Los sistemas de IA no deben amenazar el proceso democrático, la libertad de expresión, las libertades de identidad ni la posibilidad de rechazar servicios de IA.
- **El principio de Autonomía**: "*Preservar la agencia humana*": La autonomía de los seres humanos en el contexto del desarrollo de la IA significa libertad de subordinación o coerción por parte de los sistemas de IA. Los seres humanos que interactúan con sistemas de IA deben mantener una autodeterminación plena y efectiva sobre sí mismos. Además, para garantizar la agencia humana, se deben establecer sistemas que aseguren la responsabilidad y la rendición de cuentas.
- **El principio de Justicia**: "*Ser justo*": Los desarrolladores e implementadores deben asegurarse de que los individuos y los grupos minoritarios mantengan la libertad de sesgo, estigmatización y discriminación.
- **El principio de Explicabilidad**: "*Operar de manera transparente*": La transparencia es clave para construir y mantener la confianza de los ciudadanos en los desarrolladores de sistemas de IA y en los propios sistemas de IA. Tanto la transparencia tecnológica como la del modelo de negocio son importantes desde un punto de vista

8 COMISIÓN EUROPEA, "*High-Level Expert Group on Artificial Intelligence (AI-HLEG). Draft Ethics Guidelines for Trustworthy AI"*, Dirección General de Comunicación, 2018. Disponible en: https://digital-strategy.ec.europa.eu/en/library/draft-ethics-guidelines-trustworthy-ai

DOI: 10.36151/reei.46.17

ético. La transparencia tecnológica implica que los sistemas de IA sean auditables, comprensibles e inteligibles por seres humanos en diferentes niveles de comprensión y experiencia.

Lograr una IA confiable significa que los principios generales y abstractos deben ser mapeados en requisitos concretos para los sistemas y aplicaciones de IA. Los diez requisitos enumerados a continuación por el AI-HLEG *(COMISIÓN EUROPEA, 2018)* se derivan de los principios y valores fundamentales. Si bien todos son igualmente importantes, en diferentes dominios de aplicación e industrias, este contexto específico necesita estos requisitos:

1. Responsabilidad
2. Gobierno de Datos
3. Diseño para todos
4. Gobierno de la autonomía de la IA (Supervisión humana)
5. No discriminación
6. Respeto de la autonomía humana
7. Respeto por la privacidad
8. Robustez
9. Seguridad
10. Transparencia

IV. MARCO JURÍDICO INTERNACIONAL APLICABLE AL ÁMBITO MILITAR

La inteligencia artificial (IA) está revolucionando el mundo y puede ser aplicada en varios campos, incluyendo el militar. Los avances en IA, robótica y tecnologías "autónomas" han planteado una serie de cuestiones éticas cada vez más complejas y apremiantes. El AI-HLEG, Grupo Europeo sobre Ética en la Ciencia y las Nuevas Tecnologías *(COMISIÓN EUROPEA, 2020),* ha identificado preocupaciones sobre seguridad, prevención del daño y mitigación del riesgo. Además, han surgido preguntas sobre la responsabilidad moral humana. Estas tecnologías también plantean interrogantes sobre gobernanza, regulación, diseño, desarrollo, inspección, monitoreo, pruebas y certificación. La toma de decisiones democrática también es un tema crucial, que abarca decisiones sobre instituciones, políticas y valores subyacentes. Por último, es importante abordar cuestiones sobre la explicabilidad y transparencia de la IA y los sistemas "autónomos".

En cuanto al uso militar, los *Sistemas de Armas Autónomas Letales* (*LAWS, por sus siglas en inglés*) son definidos por el AI-HLEG *(COMISIÓN EUROPEA, 2018)* como armas que pueden operar sin un control humano significativo sobre las funciones críticas de seleccionar y atacar objetivos individuales. Si nos centramos en el ámbito militar y en los *LAWS*, hay un

 DOI: 10.36151/reei.46.17

amplio consenso en que el "*Control Humano Significativo*" es esencial para la responsabilidad moral. Esto significa que los humanos —y no las computadoras y sus algoritmos— deben mantener en última instancia el control y, por lo tanto, ser moralmente responsables.

1. Cuestiones de interpretación y aplicación del derecho internacional por la UE en las áreas de uso civil y militar y de autoridad estatal.

En particular, tratando sobre el uso de armas que involucren el uso de IA, una persona debe tener en todo momento los medios para corregir su curso, detenerla o desactivarla en caso de comportamiento imprevisto, intervención accidental, ciberataque o interferencia de terceros con tecnología basada en IA, o cuando terceros adquieren tal tecnología. Esto se destaca en la resolución del Parlamento Europeo del 20 de enero (2020/2013(INI)) (2021/C456/04) (PARLAMENTO EUROPEO, 2021[9]).

Esto plantea preocupaciones éticas fundamentales, como el hecho de que puede llevar a una carrera armamentista incontrolable a un nivel histórico sin precedentes, y puede crear contextos militares en los que se renuncia casi por completo al control humano y no se abordan los riesgos de mal funcionamiento. También se enfatiza la importancia de respetar el derecho internacional público, especialmente el derecho humanitario, que se aplica inequívocamente a todos los sistemas de armas y sus operadores. En este sentido debemos recordar que fruto de los distintos tratados y acuerdos internacionales y en especial los Convenios de Ginebra del Comité Internacional de la Cruz Roja en 1949 (CICR, 2019[10]) existen, al menos, 6 principios que se posicionan como la base del DIH que son: humanidad, distinción, limitación, precaución, necesidad militar y, proporcionalidad. Los Estados miembros deben cumplir con este requisito fundamental de proteger a la población civil o tomar medidas de precaución en caso de un ataque, ya sea militar o cibernético. La IA y las tecnologías relacionadas también pueden desempeñar un papel en la guerra irregular o no convencional.

Además, la resolución destaca que el uso de la inteligencia artificial (IA) brinda una oportunidad para fortalecer la seguridad de la Unión Europea (UE) y sus ciudadanos. Es esencial que la UE adopte un enfoque integrado en futuras discusiones internacionales sobre este tema. El Parlamento Europeo ha solicitado el desarrollo urgente de una posición legalmente vinculante común que aborde cuestiones éticas y legales de control humano, supervisión, responsabilidad e implementación de la ley internacional de derechos humanos, la ley humanitaria internacional y las estrategias militares.

Al reconocer que la dinámica de la moderna carrera armamentista entre los grandes estados militares para el desarrollo de sistemas de armas autónomos está superando el progreso en

---

9 PARLAMENTO EUROPEO, Resolución (2020/2013(INI)) (2021/C 456/04), P9_TA(2021)0009 "*Inteligencia artificial: cuestiones de interpretación y de aplicación del Derecho internacional"*, Resolución del Parlamento Europeo del 20 de enero de 2021 sobre inteligencia artificial: cuestiones de interpretación y aplicación del derecho internacional en la medida en que la UE se ve afectada en las áreas de usos civiles y militares y de autoridad estatal fuera del ámbito de la justicia penal, 2021.

10 CICR, Comité internacional de la Cruz Roja. (1949). Convenios de Ginebra de 12 de agosto de 1949.

DOI: 10.36151/reei.46.17

términos de normas comunes y marcos legales y su implementación y aplicación efectiva y universal, la comunidad de investigación de IA también debe integrar este principio en todos los sistemas basados en IA destinados para uso militar. En este sentido, ninguna autoridad puede derogar de tales principios o certificar dichos sistemas. Se reitera que la toma de decisiones autónomas no debe eximir a los humanos de su responsabilidad. Siempre, los individuos deben ser los responsables finales de los procesos de toma de decisiones para que la persona responsable de una decisión pueda ser identificada. Este enfoque es crucial para garantizar que las decisiones relacionadas con la seguridad y la defensa estén sujetas a control y responsabilidad humanos. La Resolución del Parlamento Europeo del 12 de febrero de 2021 sobre una estrategia europea para la inteligencia artificial (2020/2012(INI)) (PARLAMENTO EUROPEO, 2020)[11] establece principios importantes para el uso de la IA en los ámbitos militar y civil. Es esencial que la IA esté sujeta a un control humano adecuado, que se respete la ley internacional pública, que se apliquen las mismas condiciones para el uso de la IA en conflictos convencionales y no convencionales. Dentro de las no convencionales, como por ejemplo la "Zona Gris"[12] ("*Grey Zone*"), se refiere a un espacio o área de conflicto en la que las líneas entre la guerra y la paz no están claramente definidas. Dentro de las acciones de guerra no convencional también están encuadradas aquellas denominadas amenazas asimétricas, como el terrorismo, los ciberataques o las operaciones de desinformación o guerra psicológica (que no se adaptan estrictamente a una declaración formal de guerra entre estados). Para ello, la UE requiere adoptar un enfoque integrado en futuras discusiones internacionales y que la toma de decisiones autónomas no exima a los humanos de su responsabilidad. Y es que la inteligencia artificial y el uso innovador tecnologías emergentes, como argumenta ARGUMOSA (2019)[13] están íntimamente relacionadas con el concepto de guerra híbrida, al igual que los ataques informáticos o el empleo de la desinformación. De igual manera GALÁN (2018)[14] incluye la explotación de las vulnerabilidades tecnológicas dentro del concepto de conflicto híbrido. Este enfoque es crucial para garantizar que la IA se utilice de manera responsable y ética en los ámbitos militar y civil. A pesar de que somos conscientes de que como afirma MOLINER GONZÁLEZ (2019)[15], la dificultad de atribuir por ejemplo a los ataques cibernéticos la condición de acto de guerra.

---

11 PARLAMENTO EUROPEO, Resolución del Parlamento Europeo (2020/2012(INI)), de 12 de febrero de 2021, sobre "Una estrategia europea para la inteligencia artificial", 2020.

12 En la "*Grey Zone*", los actores estatales y no estatales pueden llevar a cabo actividades que están por debajo del umbral de una guerra abierta, pero teniendo un impacto significativo en la seguridad y la estabilidad. La guerra no convencional se encuentra frecuentemente en la zona gris, ya que implica el uso de métodos que no son plenamente militares pero que tienen la intención de lograr objetivos políticos y estratégicos.

13 ARGUMOSA PILA, J., "El discurso de la «Guerra Híbrida", Informe Cátedra de Estudios Estratégicos del Instituto Europeo de Estudios Internacionales, Madrid, 2019. Disponible en https://www.ieeiweb.eu/wp-content/uploads/2019/01/INFORME-CATEDRA-ESTUDIOS-ESTRATEGICOS-ENERO-2019.pdf

14 GALÁN, C., Amenazas híbridas: nuevas herramientas para viejas aspiraciones, Documento de trabajo 20/2018, Real Instituto Elcano, 2018

15 MOLINER GONZÁLEZ, J. A., La inteligencia artificial, aplicada a la defensa, Desafíos éticos en el uso militar de la inteligencia artificial. Ministerio de Defensa, Instituto Español de Estudios Estratégicos. Colecciones: Seguridad y Defensa, nº 79 2019, , p. 129.

DOI: 10.36151/reei.46.17

El Parlamento Europeo evaluó la necesidad de un control humano adecuado en el uso de la inteligencia artificial en contextos militares y civiles. El objetivo es garantizar que siempre haya un humano disponible para corregir el curso de la IA, detenerla o desactivarla en caso de comportamiento imprevisto, ciberataque o interferencia de terceros. Además, se enfatiza la importancia de respetar el derecho internacional público, especialmente en lo que respecta al derecho humanitario, que debe aplicarse inequívocamente a todos los sistemas de armas y sus operadores. Se debe destacar que la IA y las tecnologías relacionadas pueden desempeñar un papel capital en la guerra irregular o no convencional (como las descritas en el anterior párrafo). En el documento de la UE, Acción Externa, del 13 de junio de 2018, denominado "Una Europa que protege: Contraatacar las amenazas híbridas"[16] se expresa que *"las amenazas híbridas combinan actividades convencionales y no convencionales, militares y no militares que pueden ser utilizadas de manera coordinada por actores estatales o no estatales para lograr objetivos políticos específicos. Las campañas híbridas son multidimensionales, combinan medidas coercitivas y subversivas, utilizando herramientas y tácticas tanto convencionales como no convencionales. Están diseñados para ser difíciles de detectar o atribuir. Estas amenazas apuntan a vulnerabilidades críticas y buscan crear confusión para dificultar la toma de decisiones rápida y efectiva"*.

En igual sentido la COMISIÓN EUROPEA (2017)[17] en el "Informe conjunto al Parlamento Europeo y al Consejo relativo a la aplicación de la Comunicación conjunta sobre la lucha contra las amenazas híbridas" (2017) el cual en su Acción 12 da una especial importancia a aspectos cibernéticos de las amenazas híbridas. Por lo tanto, se propone que la investigación, el desarrollo y el uso de la IA en tales casos estén sujetos a las mismas condiciones que las establecidas para su uso en conflictos convencionales.

Es esencial que la UE adopte un enfoque integrado en futuras discusiones internacionales sobre la inteligencia artificial, destacando que la toma de decisiones autónomas no debe eximir a los seres humanos de su responsabilidad. Las personas siempre deben ser las responsables últimas de los procesos de toma de decisiones, de manera que se pueda identificar al ser humano responsable de una decisión.

En cuanto al uso de la inteligencia artificial en el entrenamiento y ejercicios militares, el informe reconoce su potencial, especialmente en vista de los ejercicios duales civil-militar de la UE. Sin embargo, destaca la necesidad de tener en cuenta los posibles riesgos durante todas las fases de diseño, desarrollo, pruebas, implementación y uso de sistemas basados en IA, en particular con respecto a las víctimas civiles accidentales, lesiones, pérdida de vidas y daños a la infraestructura civil, así como los riesgos relacionados con intervenciones no intencionales, manipulación, proliferación, ciberataques, interferencia de terceros con tecnología autónoma basada en IA o la adquisición de dicha tecnología por parte de terceros. La Unión Europea debería contribuir a la creación de un marco legal internacional para el uso de la inteligencia

16 ACCIÓN EXTERIOR DE LA UNIÓN EUROPEA, "A Europe that Protects: Countering Hybrid Threats", 2018.

17 COMISIÓN EUROPEA, informe Ref: JOIN/2017/030 final: Informe Conjunto al Parlamento Europeo y al Consejo relativo a "*La aplicación de la Comunicación conjunta sobre la lucha contra las amenazas híbridas – Una respuesta de la Unión Europea*", 2017.

DOI: 10.36151/reei.46.17

artificial, especialmente en los ámbitos militares, garantizando que se mantengan dentro de los límites establecidos por el derecho internacional y el derecho humanitario internacional, en particular las Convenciones de Ginebra de 1949 (CICR[18]) y sus Protocolos Adiciones de 1977[19/20]. Además, este marco debe estar en línea con los estándares de seguridad y los requisitos de protección al consumidor y nunca debe infringir o permitir violaciones de los dictados de la conciencia pública y la humanidad, tal como se establece en la cláusula Martens. Así lo establece el epígrafe 15 de la supracitada resolución del Parlamento Europeo del 20 de enero (2020/2013(INI)) (2021/C456/04) PARLAMENTO EUROPEO, 2021) que debemos poner en relación con el Informe de la ONU del Comité de Derecho Internacional de 1994 A/49/10[21], según el cual, a la vista de tratados o acuerdos internacionales que regulen específicamente el uso militar de la inteligencia artificial, *"las personas civiles y los beligerantes permanecen bajo la garantía y el régimen de los principios del derecho internacional preconizados por los usos establecidos, los principios de humanidad y los dictados de la conciencia pública"*. Como sabemos, esta idea sería recogida en 1899 por Frederic de Martens, quien indicó que las personas civiles y los combatientes quedan bajo la protección y el imperio de los principios del derecho de gentes derivados de los usos establecidos de los principios de humanidad y de los dictados por la conciencia pública. La denominada, como Cláusula de MARTENS que fue consagrada en el artículo 1.2 del Protocolo Adicional I de los Convenios de Ginebra de 1977.

La UE, junto con la ONU y la comunidad internacional, deberán liderar la promoción de este marco global para el uso de la IA, definiendo sistemas sólidos de monitoreo y evaluación para el desarrollo de tecnologías de IA, especialmente aquellas utilizadas con fines militares en estados autoritarios. Es importante destacar que cualquier robot de acción remota o de ayuda a la decisión basada en registro y tratamiento de datos utilizando modelos de IA (en su sentido amplio) no solo permitirá potencialmente mantener al personal militar a distancia del entorno donde se desenvuelve una acción, sino que también proporcionará una mejor protección personal en un escenario de guerra, así como por ejemplo en operaciones en entornos contaminados, extinción de incendios, desminado en tierra o en el mar, o defensa contra enjambres de drones. Se establece que, además de apoyar las operaciones, la IA también beneficiará al personal de las fuerzas armadas al procesar masivamente sus datos de salud y ampliar el alcance del monitoreo de la salud, identificar factores de riesgo relacionados con su entorno y condiciones de trabajo y proponer salvaguardas adecuadas para limitar el daño a su salud.

En el desarrollo, despliegue, uso y gestión de la inteligencia artificial, deben respetarse los derechos fundamentales, valores y libertades consagrados en los Tratados de la UE (Tratado de Lisboa y la Carta de los Derechos Fundamentales de la UE). Por lo tanto, los Estados miembros no deben desplegar sistemas de inteligencia artificial de alto riesgo que representen una amenaza para los derechos fundamentales. Los riesgos para los derechos fundamentales

18 CICR, Comité Internacional de la Cruz Roja. Convenios de Ginebra de 12 de agosto de 1949

19 CICR, Comité internacional de la Cruz Roja. Protocolo I Adicional a los Convenios de Ginebra de 1949 relativo a la protección de las víctimas de los conflictos armados internacionales, 1977.

20 CICR, Comité Internacional de la Cruz Roja. Protocolo II Adicional a los Convenios de Ginebra de 1949 relativo a la protección de las víctimas de los conflictos armados internacionales, 1977.

21 ONU, Informe de la Comisión de Derecho Internacional sobre la labor realizada en su 46.° período de sesiones (2 de mayo-22 de julio de 1994) 1994, GAOR A/49/10, p. 317.

 DOI: 10.36151/reei.46.17

que puedan surgir del uso de la IA por parte de las autoridades estatales y las instituciones, organismos, oficinas y agencias de la Unión Europea deben ser investigados más a fondo.

Finalmente, la UE debe facilitar la investigación y la conversación sobre las oportunidades de uso de la IA en la ayuda en caso de desastres, prevención de crisis y mantenimiento de la paz. Esto permitirá una exploración más detallada del potencial de la IA en estas áreas y asegurará su uso beneficioso y ético en el futuro. El Comité de Asuntos Exteriores del Parlamento Europeo ha adoptado un informe sobre la inteligencia artificial (IA) y su impacto en la seguridad y la defensa. El informe destaca la importancia de invertir en capacitación humana, en particular en habilidades digitales, para adaptarse a los desarrollos científicos que llevan a soluciones impulsadas por la IA, para personas en profesiones reguladas como actividades relacionadas con el ejercicio de los poderes de autoridad estatal, como la administración de justicia. Se pide a los Estados miembros y a la Comisión que tengan debidamente en cuenta esto al implementar la Directiva 2005/36/EC (PARLAMENTO EUROPEO, 2005[22]). Los sistemas de IA siempre deben cumplir con los principios de responsabilidad, equidad, gobernanza, precaución, responsabilidad, imputabilidad, predictibilidad, trazabilidad, dependencia, confiabilidad, transparencia, explicabilidad y proporcionalidad. Además, se destaca la importancia de que los sistemas basados en IA en los dominios de seguridad y defensa sean implementados con una comprensión integral de la situación por parte del operador humano, la predictibilidad, confiabilidad y resiliencia del sistema basado en IA, así como la capacidad del operador humano para detectar posibles cambios en las circunstancias y el entorno operativo e intervenir o terminar un ataque. Esto asegura que los principios del derecho internacional humanitario, en particular los principios de distinción, proporcionalidad y precaución en el ataque, se apliquen plenamente en toda la cadena de mando y control. De esta manera se trata de que los sistemas de armas autónomas y despliegue militar de la IA distingan —a través por ejemplo de reconocimiento facial o análisis automatizado satelital— entre las personas que participan en las hostilidades —esto es, los combatientes— y las personas civiles —no combatientes— y, a la vez, entre los bienes u objetivos civiles y objetivos militares, con la precisa finalidad que sólo los combatientes y objetivos militares sean objeto de ataque por los sistemas letales más modernos. En este sentido, parece difícilmente concebible el atribuir a máquinas y robots «la capacidad de seleccionar los objetivos y atacar a estos por su cuenta» por su falta de empatía[23]. Según MOLINER GONZÁLEZ (2019) se justifica la imposibilidad en que los sistemas autónomos y la IA que los dirige son incapaces de discernir las complejas situaciones que se pueden producir en el campo de batalla, como la posibilidad de que determinados objetivos hayan perdido su valor militar, o evaluar si un objetivo pretende atacar o rendirse. O según LÓPEZ-SÁNCHEZ, (2017)[24], por ejemplo, *«evaluar si un tanque es un objetivo militar o si el sistema de armas letal autónomo aceptaría su rendición no solo es cuestión de tener algoritmos inteligentes con altas capacidades de discernimiento. En su lugar, tenemos*

22 PARLAMENTO EUROPEO, Directiva 2005/36/CE del Parlamento Europeo y del Consejo, de 7 de septiembre de 2005, relativa al "Reconocimiento de cualificaciones profesionales", Consejo de la Unión Europea, Diario Oficial de la Unión Europea, L 255, 2005, pp. 22-142.

23 TRAVIESO, J., "Las consecuencias de mandar a la guerra a 'robots asesinos'", eldiario.es, 2015, p. 2. Disponible en: https://www.eldiario.es/sociedad/debate-torno-robots-asesinos_1_2718731.html

24 LÓPEZ-SÁNCHEZ, M., Some insights in artificial intelligence autonomy in military technology, 2017, p. 12. Disponible en: https://ttac21.net/2017/11/10/autonomy-in-future-military-and-security-technologies.

DOI: 10.36151/reei.46.17

que considerar los valores subyacentes que nosotros, como humanos desarrollando tales algoritmos, deberíamos ser capaces de instalar en ellos».

De igual manera, arguye Human Rights Watch[25] las *«armas completamente autónomas carecen de cualquier emoción que les pueda producir remordimiento si algún otro [humano] es castigado por sus acciones. Por lo tanto, el castigo de otros intervinientes no haría nada para cambiar la conducta del robot»*

Siguiendo la descripción de estos principios que realiza LOPEZ DÍAZ (2009)[26] y su adaptación a nuestra situación, el principio de proporcionalidad supondría que la acción militar autónoma sea razonable, no excesiva, incluyendo la prohibición de causar daños incidentales contra la población o bienes civiles, excluyendo toda forma de violencia excesiva o que no resulte indispensable para debilitar al adversario. Por último, aplicando el principio de precaución en el ataque, en todo conflicto armado deberá garantizarse el respeto y protección al medio ambiente, prohibiendo expresamente utilizarlo como un medio de combate. Las tecnologías emergentes como la IA suponen un alto consumo energético que puede acarrear su construcción y la preocupación por temas como el potencial consumo eléctrico mundial y dicho principio aplicado a este ámbito supondría una asunción de algoritmos e inteligencia artificial verde, eficiencia energética, y promoción de la sostenibilidad y energías renovables.

Los sistemas basados en inteligencia artificial no deben reemplazar la toma de decisiones humanas en los sistemas de armas letales autónomos (*LAWS*), definidos como sistemas de armas sin control humano significativo sobre las funciones críticas de selección y ataque de objetivos individuales. Se señala la necesidad de una posición común sobre *LAWS* para evitar el desarrollo, producción y uso de *LAWS* capaces de llevar a cabo ataques sin control humano significativo, así como la iniciación de negociaciones efectivas para su prohibición.

El «control humano significativo» pretende que sea siempre el ser humano el último responsable de la actuación de un sistema de armas en la amplia variedad de tareas militares que se le pueden encomendar: adquisición, seguimiento, identificación y preparación de objetivos; orientación de armas; selección y priorización de objetivos; determinación del momento de disparo; y detonación.

La clasificación de *LAWS* como una categoría específica de IA en el ámbito militar debe ser discutida y acordada a nivel internacional, en particular en el marco de la Convención de las Naciones Unidas sobre Ciertas Armas Convencionales. La Convención sobre prohibiciones o restricciones del empleo de ciertas armas convencionales que puedan considerarse excesivamente nocivas o de efectos indiscriminados (la Convención) aplica dos normas consuetudinarias generales del derecho internacional humanitario a armas específicas, a saber: (1) la prohibición de emplear armas que tienen efectos indiscriminados, y (2) la prohibición de emplear armas que causan daños superfluos. Estos límites, coincidentes con los principios de proporcionalidad, limitación y distinción, deberán aplicarse al desarrollo, producción y uso de *LAWS.*

---

25 HRW, Human Right Watch, "*Losing humanity: The case against killer robots*", 2012, p. 1. Disponible en: https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots

26 LOPEZ DIAZ, P., "Principios Fundamentales del Derecho Internacional Humanitario", REVISMAR Revista de Marina de Chile, núm. 3, 2009, pp. 230-238. Disponible en: https://revistamarina.cl/revistas/2009/3/lopez.pdf

DOI: 10.36151/reei.46.17

Se enfatiza la necesidad de una estrategia integral de la UE sobre AI y defensa que establezca un marco para la toma de decisiones sobre el uso de AI en defensa, y se hace un llamado a la Comisión para que promueva el diálogo y la cooperación entre los Estados miembros, investigadores, académicos, actores de la sociedad civil y el sector privado para garantizar la inclusión de los procesos de formulación de políticas sobre la regulación de la AI relacionada con la defensa. El desarrollo de sistemas de armas letales autónomos (*LAWS*) ha sido motivo de preocupación para la comunidad internacional, ya que su uso puede tener consecuencias desastrosas para la seguridad y la vida humana. Por esta razón, se han establecido regulaciones internacionales para regular su desarrollo y uso.

Se considera que las *LAWS* sólo deben ser utilizadas como último recurso y sólo son legales si están sujetas a un estricto control humano, es decir, si una persona puede tomar el control en cualquier momento, ya que la intervención y supervisión humana adecuada son esenciales en el proceso de toma de decisiones letales y los seres humanos siempre deben ser responsables de decidir entre la vida y la muerte. Los sistemas sin control humano en ningún momento deben ser prohibidos sin excepción y en cualquier circunstancia.

Una de las regulaciones que las *LAWS* deben cumplir es la Convención sobre Ciertas Armas Convencionales de la Asamblea General de las Naciones Unidas del 10 de octubre de 1980 (ONU, 1980[27]), que prohíbe el uso de armas consideradas "excesivamente perjudiciales". Las *LAWS*, al ser sistemas autónomos y potencialmente no estar sujetas al control humano, pueden considerarse armas que representan un peligro excesivo.

Se enfatiza que la inteligencia artificial utilizada en un contexto militar debe cumplir con un conjunto mínimo de requisitos, a saber, ser capaces de reconocer y aplicar los principios fundamentales del derecho internacional humanitario: ser capaz de distinguir entre combatientes y no combatientes en el campo de batalla, reconocer cuando un combatiente se rinde o está fuera de combate, no tener efectos indiscriminados, no causar sufrimiento innecesario a las personas, no estar sesgada ni formada a partir de datos intencionalmente incompletos y cumplir con los principios del derecho humanitario internacional, la proporcionalidad en el uso de la fuerza y la precaución antes de la intervención.

Además, se propone que las *LAWS* sean incluidas en la lista de armas sujetas a las disposiciones del Tratado sobre el Comercio de Armas del 2 de abril de 2013 (ONU, 2013[28]), que regula el comercio de armas convencionales y busca prevenir su proliferación incontrolada. Esto permitiría una mejor regulación de su comercio y una mayor transparencia en su uso. Esto implica que las *LAWS* no deben tener características que las hagan parecer humanos, ya que esto podría crear situaciones confusas y peligrosas. Se aplaude el acuerdo del Consejo y del Parlamento de la Unión Europea de excluir de las acciones financiadas por el Fondo Europeo de Defensa (FED) a las armas letales autónomas que no permiten un control humano significativo sobre las decisiones de selección de objetivos e intervención al llevar a cabo ataques. Esto

27 ONU, Asamblea General de las Naciones Unidas. Convención sobre ciertas armas convencionales de 10 de octubre de 1980, 1980.

28 ONU. Asamblea General de las Naciones Unidas. Tratado sobre el Comercio de Armas, de 2 de abril de 2013, 2013.

DOI: 10.36151/reei.46.17

significa que el uso, desarrollo o producción de *LAWS* sin control humano significativo no son elegibles para financiamiento del FED. En conclusión, existen varias regulaciones internacionales que buscan regular el desarrollo y uso de *LAWS*, debido al riesgo que representan para la vida y la seguridad humana. Es importante que estas regulaciones se apliquen y refuercen para garantizar una regulación adecuada y una mayor transparencia en su uso.

En este sentido, la Directiva 2006/42/CE (PARLAMENTO EUROPEO, 2006[29]) establece la definición de "máquina", que incluye robots y sistemas autónomos. En consecuencia, cualquier robot o sistema autónomo que cumpla con esta definición debe cumplir con los estándares y medidas de seguridad establecidos en esta Directiva. Las empresas que producen, importan o distribuyen robots y sistemas autónomos deben asegurarse de que cumplen con los requisitos de la Directiva, ya que son responsables de garantizar su seguridad. La Directiva establece que los robots deben ser diseñados y ensamblados de acuerdo con los principios de seguridad y protección de la salud de los trabajadores, consumidores y el medio ambiente.

Las medidas de seguridad que se deben implementar incluyen el uso de materiales no tóxicos y la inclusión de dispositivos de seguridad como sensores y sistemas de parada de emergencia. Además, se deben proporcionar instrucciones claras sobre el uso y reparación seguros de robots y sistemas autónomos. La Directiva 2006/42/CE (PARLAMENTO EUROPEO, 2006) establece una serie de requisitos que deben cumplir los robots y sistemas autónomos que cumplen con la definición de "máquina". Las empresas que producen, importan o distribuyen estos dispositivos son responsables de garantizar su seguridad y deben cumplir con las normas y medidas de seguridad establecidas en la Directiva. Esto garantiza que los robots y sistemas autónomos utilizados en la Unión Europea sean seguros para los trabajadores, los consumidores y el medio ambiente.

2. Marco para los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías relacionadas

La resolución del Parlamento Europeo de 20 de octubre de 2020 con recomendaciones a la Comisión sobre un marco para los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías relacionadas (2020/2012(INL)) (2021/C404/04) (PARLAMENTO EUROPEO, 2020)[30] en la sección de Seguridad y Defensa, evaluó que las políticas de seguridad y defensa de la Unión Europea y sus Estados miembros deben estar guiadas por los principios de igualdad, buena fe, paz, y convivencia de la Carta de las Naciones Unidas de 1945 y por una comprensión común de los valores universales de respeto a los derechos inviolables e inalienables de la persona, la dignidad humana, la libertad, la democracia, la igualdad y el Estado de derecho.

---

29 PARLAMENTO EUROPEO, Directiva 2006/42/CE del Parlamento Europeo y del Consejo, de 17 de mayo de 2006, relativa a las máquinas y por la que se modifica la Directiva 95/16/CE (refundición), Consejo de la Unión Europea, Diario Oficial de la Unión Europea, L 157, 2006, pp. 24-86.

30 PARLAMENTO EUROPEO. Resolución del Parlamento Europeo, de 20 de octubre de 2020 (2020/2012(INL)) (2021/C 404/04) "Marco para los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías relacionadas con recomendaciones destinadas a la Comisión sobre un marco para los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías relacionadas", 2020.

 DOI: 10.36151/reei.46.17

Esto fue uno de los resultados tras la aprobación por parte de la Reunión de Altas Partes Contratantes de la Convención de las Naciones Unidas sobre Ciertas Armas Convencionales de 2019 (CCW, por sus siglas en inglés) (ONU, 2019[31]) de once principios rectores establecidos por el Grupo de Expertos Gubernamentales sobre las Tecnologías Emergentes en el Ámbito de los Sistemas de Armas Autónomos Letales del Anexo III para el desarrollo y uso de sistemas de armas autónomas:

1. El derecho internacional humanitario se aplica plenamente a todos los sistemas de armas, incluyendo los sistemas autónomos letales.
2. La responsabilidad por las decisiones sobre el uso de estos sistemas debe mantenerse en manos de seres humanos, y no puede ser transferida a las máquinas.
3. La interacción entre humanos y máquinas debe asegurar que el uso de estos sistemas cumpla con el derecho internacional, considerando factores como el contexto operacional y las capacidades del sistema.
4. La rendición de cuentas por el desarrollo y uso de nuevos sistemas de armas debe cumplir con el derecho internacional, con un enfoque en el control humano.
5. Los Estados deben evaluar si el uso de nuevas armas podría estar prohibido por el derecho internacional en ciertas condiciones.
6. Se deben considerar aspectos como la seguridad física, las salvaguardias no físicas, la ciberseguridad y los riesgos de adquisición por grupos terroristas al desarrollar estos sistemas.
7. Las evaluaciones de riesgos y medidas de mitigación deben ser parte integral del ciclo de desarrollo de tecnologías emergentes en sistemas de armas.
8. Se sugiere explorar el uso de estas tecnologías para garantizar el cumplimiento del derecho internacional humanitario y otras obligaciones internacionales.
9. No se deben atribuir cualidades humanas a las tecnologías emergentes en sistemas de armas.
10. Los debates y medidas de política no deben obstaculizar el acceso a usos pacíficos de tecnologías autónomas inteligentes.
11. La Convención proporciona un marco adecuado para abordar la cuestión de las tecnologías emergentes en sistemas de armas autónomos letales, buscando un equilibrio entre las necesidades militares y humanitarias.

Lamenta, sin embargo, el Parlamento Europeo la falta de acuerdo sobre un instrumento jurídicamente vinculante que regule las armas autónomas letales con un mecanismo efectivo de cumplimiento.

---

31 ONU. Asamblea General de las Naciones Unidas. Convención sobre ciertas armas convencionales (CCAC): Reunión de las altas partes contratantes, 2019.

DOI: 10.36151/reei.46.17

Las tecnologías emergentes en los sectores de defensa y seguridad no reguladas por el derecho internacional deberían estar sujetas al principio del respeto a la humanidad y a los dictados de la conciencia pública y recomienda que cualquier marco europeo que regule el uso de sistemas basados en inteligencia artificial en el ámbito de la defensa, tanto en situaciones de combate como no combativas, respete todos los regímenes legales aplicables, en particular el derecho internacional humanitario y el derecho internacional de los derechos humanos. Reconoce que, en el actual contexto de guerra híbrida y avanzada, el volumen y la velocidad de la información durante las primeras etapas de una crisis pueden ser abrumadores para los analistas humanos, y que un sistema de inteligencia artificial podría procesar la información para asegurar que los tomadores de decisiones humanos rastreen todo el espectro de información dentro de un marco temporal apropiado para proporcionar una respuesta rápida.

Se aboga por un aumento de la inversión en inteligencia artificial europea para la defensa y la infraestructura crítica que la sustenta, en vista de los significativos esfuerzos de las potencias militares mundiales en investigación y desarrollo militar e innovación. Igualmente se subraya que la responsabilidad y la rendición de cuentas por la decisión de diseñar, desarrollar, desplegar y utilizar sistemas de inteligencia artificial deben recaer completamente en los operadores humanos, dado que debe haber una supervisión y control humano significativos sobre cualquier sistema de armas, y la intencionalidad humana en la decisión de utilizar la fuerza, en la ejecución de cualquier decisión de sistema de armas basado en IA que pueda tener consecuencias letales. Se destaca que el control humano debe ejercerse de manera efectiva sobre el mando y control de los sistemas basados en IA, de acuerdo con los principios de la participación, supervisión y control humano, en la conducción de operaciones militares. Igualmente, sistemas basados en IA deben permitir a los comandantes militares al frente de los ejércitos asumir plena responsabilidad y rendición de cuentas por el uso de la fuerza letal y ejercer el juicio necesario, que las máquinas no pueden tener y que debe basarse en la distinción, la proporcionalidad y la precaución, al tomar medidas letales o destructivas a gran escala mediante dichos sistemas. También se requiere establecer marcos de autorización y rendición de cuentas claros y trazables para el despliegue de armas inteligentes y otros sistemas basados en IA, utilizando características únicas del usuario, como especificaciones biométricas, para permitir el despliegue solo por personal autorizado.

3. Puntos de vista y recomendaciones sobre la Sexta Conferencia de Revisión de la Convención sobre Ciertas Armas Convencionales (*CCW*)

El documento de trabajo presentado por el Comité Internacional de la Cruz Roja el 8 de noviembre de 2021 (CICR, 2021)[32] sobre la Convención sobre Ciertas Armas Convencionales (CCW) es un pilar fundamental del derecho humanitario internacional. La CCW y sus Protocolos encarnan el principio básico de que el uso de medios y métodos de guerra no es ilimitado. Lo hacen prohibiendo o restringiendo el uso de armas convencionales que plantean

32 CICR, Comité Internacional de la Cruz Roja, "Puntos de vista y recomendaciones para la 6ª Conferencia de Examen de la Convención sobre ciertas armas convencionales", Documento de trabajo presentado por el Comité Internacional de la Cruz Roja, 6ª Conferencia de Examen de la Convención sobre Ciertas Armas Convencionales, 2021.

 DOI: 10.36151/reei.46.17

preocupaciones humanitarias, legales y éticas específicas, en particular las armas que pueden causar sufrimiento innecesario o que pueden tener efectos indiscriminados. Este documento de trabajo presenta las opiniones y recomendaciones del CICR sobre cuestiones de preocupación humanitaria pertinentes para la CCW (CICR, 2021), específicamente: la adhesión a la CCW y su aplicación nacional; minas que no son minas antipersonal; armas incendiarias y armas con efectos incendiarios; armas láser cegadoras y otros sistemas láser; restos explosivos de guerra; armas explosivas en zonas pobladas; sistemas de armas autónomas; y revisión de los avances en ciencia y tecnología, y revisión jurídica de nuevas armas, medios y métodos de guerra.

La 6ª Conferencia de Revisión de la CCW, que se celebró del 13 al 17 de diciembre de 2021 en Ginebra, es un momento clave para que las Altas Partes Contratantes hagan balance y construyan sobre el importante papel que la CCW ha desempeñado en la minimización del sufrimiento en los conflictos armados, con el fin de garantizar que la CCW siga siendo adecuada para el propósito a medida que la guerra evoluciona *(ICRC, 2021)*.

En su sección 8, el CICR aborda los sistemas de armas autónomas, un tema que ha sido planteado por muchas Altas Partes Contratantes durante las discusiones de la CCW desde 2014 y dentro del actual Grupo de Expertos Gubernamentales (GGE) desde la 5ª Conferencia de Revisión.

El debate mundial sobre el uso militar de la inteligencia artificial ha sido iniciado por la ONU y las reuniones para la Convención sobre Ciertas Armas Convencionales (CCW, Ginebra), donde varias de las Altas Partes Contratantes respaldaron el llamado principio de 'control humano significativo para los sistemas de armas autónomas afirmando que 'Los sistemas de armas autónomas que no requieren control humano significativo deben ser prohibidos' (Asamblea General de la ONU, 2016). La ONU también ha establecido un instituto de investigación especial en La Haya para estudiar la gobernanza de la robótica y la inteligencia artificial (UNICRI). Varias iniciativas y ONG que tienen como objetivo la IA y los sistemas 'autónomos' para el bien respectivamente hacen campaña por una prohibición de las armas 'autónomas', por ejemplo, la Fundación para la Robótica Responsable.

Gran parte del debate sobre la aceptabilidad moral de las armas "autónomas" y la responsabilidad legal y moral por el despliegue de estos sistemas tiene lugar en la Conferencia sobre Ciertas Armas Convencionales en Ginebra. Ahora es necesario prestar atención a las preguntas sobre cuál es la naturaleza y el significado de un "control humano significativo" sobre estos sistemas y cómo instituir formas de control moralmente deseables.

En la comprensión del CICR, los sistemas de armas autónomos, después de su activación inicial, seleccionan y aplican la fuerza a objetivos sin intervención humana. El usuario no elige el/los objetivo(s) específico(s) ni el momento y/o lugar preciso de la aplicación de la fuerza resultante. El desafío central con estas armas reside en la dificultad de anticipar y limitar sus efectos.

Desde una perspectiva humanitaria, el uso de sistemas de armas autónomos corre el riesgo de perjudicar a quienes se ven afectados por el conflicto armado, tanto civiles como combatientes '*hors de combat'*, y los sistemas aumentan el riesgo de escalada del conflicto. Desde una perspectiva legal, desafían la capacidad de las personas que deben aplicar las normas del DIH du-

 DOI: 10.36151/reei.46.17

rante la planificación, decisión y ejecución de los ataques para cumplir con sus obligaciones. Desde una perspectiva ética, este modo de funcionamiento corre el riesgo de sustituir efectivamente las decisiones humanas sobre la vida y la muerte por procesos de sensores, software y máquinas. Estas preocupaciones éticas son especialmente importantes cuando los sistemas de armas autónomos están diseñados o se utilizan para atacar directamente a personas.

Ante este contexto, el CICR ofreció recomendaciones a todos los Estados el 12 de mayo de 2021, incluyendo las Altas Partes Contratantes del CCW a la luz del mandato del GGE para la clarificación, consideración y desarrollo del marco normativo y operativo para los sistemas de armas autónomos. El CICR recomienda que los Estados adopten nuevas normas legalmente vinculantes sobre los sistemas de armas autónomos para garantizar que se mantenga un control y juicio humano suficiente en el uso de la fuerza. En opinión del CICR, esto requerirá prohibir ciertos tipos de sistemas de armas autónomos y regular estrictamente todos los demás. Los sistemas de armas autónomos impredecibles deben ser explícitamente descartados, en particular debido a sus efectos indiscriminados. Esto se lograría mejor con una prohibición de los sistemas de armas autónomos que estén diseñados o utilizados de tal manera que sus efectos no puedan entenderse, predecirse y explicarse adecuadamente.

El uso de sistemas de armas autónomos para atacar a seres humanos debería ser descartado. Esto se lograría mejor mediante la prohibición de los sistemas de armas autónomos que estén diseñados o utilizados para aplicar la fuerza directamente contra personas, en lugar de contra objetos. El diseño y uso de sistemas de armas autónomos que no fueran prohibidos deberían ser regulados, incluyendo una combinación de: imponer límites en los tipos de objetivos, como restringirlos a objetos que son objetivos militares por naturaleza; imponer límites en la duración, alcance geográfico y escala de uso, incluyendo para permitir el juicio y control humano en relación con un ataque específico; imponer límites en situaciones de uso, como restringirlos a situaciones donde no haya civiles u objetos civiles presentes; e imponer requisitos para la interacción humano-máquina, en particular para garantizar una supervisión y una intervención y desactivación oportunas y eficaces por parte de los seres humanos.

El CICR insta a las Altas Partes Contratantes en la Conferencia de Revisión a establecer un camino hacia la adopción de nuevas normas legalmente vinculantes sobre sistemas de armas autónomos, como mediante una decisión de negociar un nuevo Protocolo del CCW y fortalecer las normas existentes de DIH. Esto es, incluyendo el principio de distinción, las prohibiciones de ataques indiscriminados y desproporcionados, la obligación de tomar todas las precauciones factibles en el ataque y las normas de protección de combatientes fuera de combate, así como prohibiciones y regulaciones específicas que se encuentran en el Protocolo II (enmendado) del CCW, sobre la prohibición del uso, almacenamiento, producción y transferencia de minas antipersonal y sobre su destrucción, y la Convención sobre Municiones en Racimo.

4. Reglas de derecho civil sobre robótica

La resolución del Parlamento Europeo de 16 de febrero de 2017 con recomendaciones a la Comisión sobre las reglas de derecho civil sobre robótica (2015/2103(INL)). (2018/C 252/25)

DOI: 10.36151/reei.46.17

(PARLAMENTO EUROPEO, 2018)[33] reconoce que una nueva era requiere un impulso legislativo e incluso una definición generalmente aceptada de robots e inteligencia artificial que sea flexible y no obstaculice la innovación. Una regulación sobre responsabilidad, transparencia y responsabilidad que refleje intrínsecamente los valores humanistas europeos y universales y asuma el liderazgo europeo en una posible regulación que refleje valores éticos. Ve el potencial de la robótica y la inteligencia artificial para transformar los estilos de vida y las formas de trabajo, aumentar los niveles de eficiencia, ahorro y seguridad, y mejorar la calidad de los servicios, e incluso acepta la posibilidad de que a largo plazo la inteligencia artificial pueda superar la capacidad intelectual humana, y también reconoce varias preocupaciones sobre sus efectos directos e indirectos en la sociedad en su conjunto.

También se determina que la automatización requiere que aquellos involucrados en el desarrollo y comercialización de aplicaciones de inteligencia artificial incorporen características de seguridad y éticas desde el principio, reconociendo así que deben estar preparados para aceptar la responsabilidad legal por la calidad de la tecnología que producen. Europa no quiere quedarse atrás ya que varios países extranjeros, como Estados Unidos, Japón, China y Corea del Sur, están considerando medidas regulatorias en el campo de la robótica y la inteligencia artificial.

En el campo militar, hace solo dos menciones, en su considerando 64 destaca que las restricciones y condiciones establecidas en el Reglamento (CE) Nº 428/2009 del Parlamento Europeo y del Consejo (PARLAMENTO EUROPEO, 2009) sobre el comercio de artículos de doble uso (bienes, software y tecnología que se pueden utilizar tanto para aplicaciones civiles como militares o que podrían contribuir a la proliferación de armas de destrucción masiva) deberían extenderse a las aplicaciones de robótica.

Por último, pero no menos importante, prohíbe la modificación de robots para su uso como armas.

V. CONCLUSIONES Y PERSPECTIVAS

La evolución de la inteligencia artificial y su impacto en la cuarta Revolución Industrial plantea desafíos legales[34] que requieren respuestas adecuadas en el entorno militar dentro de la UE. La IA confiable tiene dos componentes principales: robustez técnica y propósito ético. Los principios de beneficencia, no maleficencia, autonomía, justicia y explicabilidad proporcionan un marco para garantizar que los sistemas de IA respeten los derechos fundamentales, las regulaciones aplicables y los principios y valores centrales, evitando el daño no intencional. Los requisitos concretos, como la responsabilidad, la gobernanza de datos y la transparencia, deben ser adaptados a contextos específicos para lograr una IA confiable. Al promover la pros-

33 PARLAMENTO EUROPEO. Norma 2015/2103(INL), 2018/C 252/25, "Normas de Derecho civil sobre robótica, Resolución del Parlamento Europeo", 2018.

34 HERRERA TRIGUERO, F., PERALTA GUTIÉRREZ, A., y TORRES LÓPEZ, L.S., "El derecho y la inteligencia artificial". Editorial UGR. 978-84-338-7049-0, 2022. https://editorial.ugr.es/libro/el-derecho-y-la-inteligencia-artificial_139323/

DOI: 10.36151/reei.46.17

peridad, la creación de valor y la sostenibilidad al tiempo que se protege la dignidad, la integridad, la privacidad y la autonomía humanas, la IA confiable puede beneficiar a individuos y a la sociedad en su conjunto.

El Comité Internacional de la Cruz Roja citado por MOLINER GONZÁLEZ (2019) plantea que el «control humano significativo» debería ser definido, planteado y resuelto en una norma legal[35], empezando por lograr una mayor precisión y consenso en los conceptos de «autonomía», «autonomía de las armas» y «armas autónomas», de cuya complejidad e importancia se ha tratado con anterioridad.

El uso de la IA en contextos militares y civiles ha planteado complejas preguntas morales sobre seguridad, responsabilidad moral, gobernanza, regulación y toma de decisiones democráticas. El Parlamento Europeo ha enfatizado la necesidad de control humano sobre la IA para garantizar la responsabilidad, supervisión y rendición de cuentas, así como la importancia de respetar el derecho internacional público y asegurarse de que la toma de decisiones autónomas no exima a los humanos de su responsabilidad.

Si bien la IA y las tecnologías relacionadas pueden desempeñar un papel en la guerra irregular o no convencional, la investigación, el desarrollo y el uso de la IA deben estar sujetos a las mismas condiciones que las establecidas para su uso en conflictos convencionales. La UE debería liderar en la promoción de un marco global para el uso de la IA, definiendo sistemas sólidos de monitoreo y evaluación e invirtiendo en la capacitación humana. Es esencial que se apliquen plenamente los principios del derecho internacional humanitario, responsabilidad, equidad, gobernanza, responsabilidad, previsibilidad, confiabilidad, transparencia, explicabilidad y proporcionalidad en toda la cadena de mando y control. El desarrollo y uso de sistemas de armas letales autónomas (*LAWS*) sin un control humano significativo es una preocupación importante para la comunidad internacional debido a las posibles consecuencias desastrosas para la vida y seguridad humana. Se han establecido recomendaciones internacionales para regular el desarrollo y uso de estos sistemas, y es esencial que se aprueben regulaciones, se cumplan y fortalezcan para garantizar un uso adecuado y transparencia. Además, se necesita una estrategia integral de la UE sobre IA y defensa, que enfatice la necesidad de supervisión y control humano significativos sobre cualquier sistema de armas, de acuerdo con los principios de involucramiento, supervisión y control humano. La regulación de la robótica y la IA también debe reflejar valores humanistas europeos y universales, y asumir un liderazgo europeo en una posible regulación que refleje valores éticos.

Se destaca la importancia de la IA confiable y se describen los principios que aseguran que los sistemas de IA respeten los derechos fundamentales, las regulaciones que deberían desarrollarse y los valores y principios fundamentales para evitar el daño no intencional. Las cuestiones éticas y legales con respecto al uso de la IA en contextos militares y civiles, enfatizando la

35 CICR, Comité Internacional de la Cruz Roja, "*Report of the expert meeting on autonomous weapon systems: Technical, military, legal and humanitarian aspects*", 2014. Disponible en: https://www.icrc.org/en/document/report-icrc-meeting-autonomous-weapon-systems-26-28-march-2014
Véase también "Autonomous weapons systems: Implications of increasing autonomy in the critical functions of weapons", (2016). https://www.icrc.org/en/publication/4283-autonomous-weapons-systems.

DOI: 10.36151/reei.46.17

necesidad de control humano sobre la IA y la importancia de respetar el derecho internacional humanitario, deberán ser convenientemente desarrolladas.

El uso de la IA en contextos militares es complejo y deben considerarse cuidadosamente para garantizar que los sistemas de IA sean confiables y respeten los derechos fundamentales, las regulaciones y los valores fundamentales. Es importante que la UE y la comunidad internacional desempeñen un papel activo en la promoción de un marco global para el uso de la IA y en la regulación del desarrollo y uso de sistemas letales de armas autónomas, al mismo tiempo que invierten en el desarrollo de la capacidad humana sobre estas cuestiones técnicas y éticas para asegurar que los humanos sigan estando en control del proceso de toma de decisiones. Así, se deberá regular no sólo las responsabilidades por crímenes de guerra en aquellos casos en que haya supervisión humana con un control humano significativo y se vulneren las reglas de guerra, sino también en aquellos casos autómatas, que debía haber esa supervisión y no la hubo.

Así, los resultados obtenidos en esta investigación corresponden a un nivel exploratorio y descriptivo. Una primera aproximación que requeriría una mayor profundización en un diseño normativo de leyes de guerra para IA, unas pruebas de concepto de soluciones técnicas éticas de defensa e incluso un sistema de asunción de responsabilidades en caso de ausencia de supervisión humana[36].

36 VIGEVANO, M. R., "Inteligencia artificial aplicable a los conflictos armados: límites jurídicos y éticos", Arbor, 197(800), 2021. Disponible en: https://doi.org/10.3989/arbor.2021.800002

DOI: 10.36151/reei.46.17