

## **El impacto social y humanitario de los sistemas de armas autónomas y de la inteligencia artificial**

Serie Minutas N° 109-23, 10/10/2023

*por Marek Hoehn*

### **Resumen**

*La presente Minuta fue elaborada para apoyar la participación de la delegación parlamentaria chilena en las sesiones de la Comisión Permanente de Paz y Seguridad Internacional durante la 147a Asamblea de la Unión Interparlamentaria (UIP) que se realizará en Luanda (Angola) entre los días 23 y 27 de Octubre de 2023.*

*En particular, esta Minuta entrega argumentos para el debate en el punto de la tabla "Audiencia de expertos sobre el tema El impacto social y humanitario de los sistemas de armas autónomas y de la inteligencia artificial".*

Disclaimer: Este trabajo ha sido elaborado a solicitud de parlamentarios del Congreso Nacional, bajo sus orientaciones y particulares requerimientos. Por consiguiente, sus contenidos están delimitados por los plazos de entrega que se establezcan y por los parámetros de análisis acordados. No es un documento académico y se enmarca en criterios de neutralidad e imparcialidad política.

## Tabla de contenido

Introducción.....	3
1. Deshumanización digital.....	3
2. Prejuicios algorítmicos.....	4
3. Pérdida de control humano significativo/ efectivo.....	6
4. Falta de juicio y comprensión humanos.....	7
5. Bajas civiles y daños colaterales.....	8
Bibliografía.....	8

## Introducción

Las armas autónomas son una realidad y no mera ciencia ficción. Es una certeza que su uso aumentará masivamente en los conflictos armados. Estas armas, capaces de tomar decisiones y actuar de forma independiente, plantean un serio problema ético que debe ser abordado de manera urgente. En este texto, exploraremos las implicaciones éticas de las armas autónomas, analizando el impacto que tienen en la vida humana y en la sociedad en su conjunto.

Al examinar este tema, nos adentraremos en un debate complejo y controversial, buscando comprender las implicaciones éticas detrás de la creación y el uso de estas armas. A medida que avancemos, nos enfrentaremos a preguntas fundamentales sobre la responsabilidad humana en el desarrollo y el despliegue de estas armas, así como a la necesidad de establecer regulaciones y límites claros en su uso.

En definitiva, este texto tiene como objetivo entregar argumentos para una reflexión profunda sobre el problema ético de las armas autónomas y promover un debate informado y sensible sobre este tema crucial en nuestra sociedad actual.

### 1. Deshumanización digital

Existe una preocupación por que el uso de armas autónomas pueda deshumanizar la guerra al reducir el papel del juicio y la toma de decisiones humanos.

La deshumanización digital en el contexto de las armas autónomas se refiere al proceso por el cual el uso de la tecnología y la inteligencia artificial (IA) en la guerra puede conducir al desprecio de la vida y la dignidad humanas. Implica la erosión de la agencia y la toma de decisiones humanas en favor de sistemas automatizados que pueden no reconocer o respetar adecuadamente los derechos y la humanidad de las personas.

Las armas autónomas, también llamadas "robots asesinos", son sistemas de armas que pueden funcionar sin control humano directo. Estos sistemas están diseñados para seleccionar y atacar objetivos de forma independiente, basándose en algoritmos, el entrenamiento previo y en las entradas de los sensores. Aunque sus defensores sostienen que las armas autónomas pueden reducir el número de víctimas humanas y aumentar la eficacia de la guerra, existen dudas sobre las implicaciones éticas y las posibles consecuencias de delegar decisiones de vida o muerte en máquinas.

La deshumanización digital se produce cuando el uso de armas autónomas conduce a un distanciamiento de las consecuencias humanas de la guerra. La dependencia de algoritmos y sistemas de inteligencia artificial puede dar lugar a una falta de *accountability*, es decir de responsabilidad por las acciones que llevan a cabo estas armas. El proceso de toma de decisiones se abstrae de la

experiencia humana, lo que puede conducir a una devaluación de la vida humana y a una erosión de los principios de humanidad y dignidad que sustentan el derecho internacional humanitario.

En el contexto de las armas autónomas, la deshumanización digital puede manifestarse de varias maneras:

Las armas autónomas se basan en algoritmos y sistemas de IA para tomar decisiones sobre a quién apuntar y atacar. Estos sistemas pueden no poseer el mismo nivel de juicio, empatía o comprensión de situaciones complejas que los operadores humanos. Como resultado, existe el riesgo de que el uso de armas autónomas pueda conducir a ataques indiscriminados o desproporcionados, causando daños a civiles o violando los principios de distinción y proporcionalidad en la guerra.

El uso de armas autónomas reduce el control humano y la supervisión del uso de la fuerza. Esto puede conducir a una falta de responsabilidad por las acciones realizadas por estas armas, así como a una disminución de la capacidad para garantizar el cumplimiento del derecho internacional humanitario y de las normas de derechos humanos. Los operadores humanos pueden llegar a desentenderse de las consecuencias de sus acciones, lo que puede llevar a ignorar las implicaciones éticas y legales del uso de la fuerza letal.

La creciente dependencia de la tecnología y la automatización en la guerra puede contribuir a una devaluación de la vida humana. Cuando las decisiones sobre objetivos y enfrentamientos se basan únicamente en parámetros técnicos y algoritmos, se corre el riesgo de pasar por alto o marginar el valor único y la dignidad de la vida humana. Esto puede tener profundas implicaciones éticas y humanitarias, especialmente en situaciones en las que los civiles corren el riesgo de sufrir daños.

En general, la deshumanización digital en el contexto de las armas autónomas pone de relieve la necesidad de considerar cuidadosamente las implicaciones éticas, jurídicas y humanitarias del uso de la IA y la automatización en la guerra. Subraya la importancia de garantizar que los valores humanos, el juicio y la responsabilidad sigan siendo fundamentales en los procesos de toma de decisiones relacionados con el uso de la fuerza.

## **2. Prejuicios algorítmicos**

Las armas autónomas se basan en algoritmos y *machine learning* para tomar decisiones, y existe la posibilidad de que el entrenamiento de la IA sea tendencioso o discriminatorio, dando lugar a resultados injustos.

Los sesgos algorítmicos en el contexto de las armas autónomas se refieren a la posibilidad de resultados discriminatorios o injustos que pueden surgir del uso de algoritmos sesgados en los procesos de toma de decisiones. Cuando los sistemas de armas autónomas se basan en algoritmos para seleccionar y atacar objetivos, los sesgos presentes en esos algoritmos pueden tener consecuencias

importantes.

Los sesgos algorítmicos pueden deberse a diversos factores, como datos de entrenamiento sesgados, un diseño defectuoso o sesgos involuntarios introducidos durante el proceso de desarrollo. Estos sesgos pueden manifestarse de diferentes maneras:

Los sistemas de armas autónomos pueden mostrar sesgos de género si los algoritmos utilizados para identificar objetivos o tomar decisiones se entrenan con datos que reflejan sesgos o estereotipos sociales. Por ejemplo, si los datos históricos utilizados para entrenar el algoritmo muestran una selección desproporcionada de individuos en función del género, el algoritmo puede perpetuar o amplificar estos sesgos.

Del mismo modo, los sistemas de armas autónomas pueden mostrar prejuicios raciales si los algoritmos utilizados para identificar objetivos o tomar decisiones se entrenan con datos que reflejan prejuicios raciales o discriminación. Si los datos históricos utilizados para entrenar el algoritmo muestran un ataque desproporcionado contra individuos por motivos de raza, el algoritmo puede perpetuar o amplificar estos sesgos.

Los sistemas de armas autónomos también pueden mostrar sesgos étnicos si los algoritmos utilizados para identificar objetivos o tomar decisiones se entrenan con datos que reflejan sesgos o prejuicios étnicos. Si los datos históricos utilizados para entrenar el algoritmo muestran un ataque desproporcionado contra individuos por motivos étnicos, el algoritmo puede perpetuar o amplificar estos sesgos.

Estos sesgos algorítmicos plantean importantes problemas éticos y humanitarios. Pueden dar lugar a una selección injusta o discriminatoria de los objetivos, lo que podría dar lugar a violaciones de los derechos humanos y del Derecho internacional humanitario. Además, los sesgos algorítmicos pueden exacerbar las desigualdades sociales existentes y perpetuar los sesgos y la discriminación sistémicos.

Abordar los sesgos algorítmicos en los sistemas de armas autónomas requiere una cuidadosa atención al diseño, desarrollo y despliegue de estos sistemas. Esto implica garantizar que los datos de entrenamiento sean diversos, representativos y libres de sesgos. Además, la supervisión y evaluación continuas de los algoritmos son necesarias para detectar y mitigar cualquier sesgo que pueda surgir durante el funcionamiento del sistema.

Es crucial reconocer y abordar los sesgos algorítmicos para garantizar que los sistemas de armas autónomas funcionen de manera que respeten los derechos humanos, promuevan la equidad y eviten exacerbar los sesgos y la discriminación existentes en la sociedad.

### **3. Pérdida de control humano significativo/ efectivo**

Las armas autónomas operan sin supervisión o control humano directo, lo que suscita preocupación sobre la responsabilidad y la capacidad de garantizar que las acciones realizadas por estas armas se ajustan a los valores humanos y a la ética.

Las armas autónomas implican la pérdida de un control humano significativo porque están diseñadas para funcionar sin intervención o supervisión humana directa. El concepto de control humano significativo se refiere a la capacidad de los seres humanos para comprender, influir y ser responsables de las acciones y decisiones de los sistemas autónomos.

Hay varias razones por las que las armas autónomas pueden conducir a una pérdida de control humano significativo:

Los sistemas de armas autónomos se basan en algoritmos y en la IA para tomar decisiones sobre la selección de objetivos y el enfrentamiento. Estos algoritmos se programan basándose en reglas y parámetros predefinidos, y funcionan de forma independiente sin intervención humana en tiempo real. Como resultado, los operadores humanos pueden tener una capacidad limitada o nula para intervenir o influir en las decisiones tomadas por el sistema autónomo.

Las armas autónomas pueden funcionar a gran velocidad y procesar grandes cantidades de información en tiempo real. Esta velocidad y complejidad pueden dificultar que los operadores humanos comprendan plenamente o sigan el ritmo del proceso de toma de decisiones del sistema autónomo. Como resultado, la capacidad de ejercer un control significativo sobre las acciones del sistema puede verse comprometida.

Las armas autónomas pueden carecer de la capacidad de comprender plenamente situaciones complejas y dinámicas del campo de batalla, incluida la presencia de civiles o circunstancias cambiantes. Los operadores humanos, en cambio, pueden ejercer su juicio y tener en cuenta factores contextuales a la hora de tomar decisiones. Sin un control humano significativo, existe el riesgo de que las armas autónomas tomen decisiones erróneas o inadecuadas que podrían causar daños a civiles o violaciones del derecho internacional humanitario.

El uso de armas autónomas plantea cuestiones sobre la responsabilidad de las acciones realizadas por estos sistemas. Si un sistema funciona de forma autónoma y causa daños o viola las normas éticas y legales, puede resultar difícil atribuir responsabilidades o hacer que las personas rindan cuentas por esas acciones. Esta falta de responsabilidad puede socavar los principios de transparencia, responsabilidad y justicia.

La pérdida de un control humano significativo sobre las armas autónomas plantea importantes problemas éticos, jurídicos y humanitarios. Puede conducir a una falta de responsabilidad, a un posible uso indiscriminado o desproporcionado de la fuerza y a la erosión de los valores humanos y del juicio en los procesos de toma de decisiones. Para hacer frente a estas preocupaciones, se están llevando

a cabo debates y esfuerzos para establecer marcos y normativas que garanticen el mantenimiento de un control humano significativo en el desarrollo, despliegue y uso de los sistemas de armas autónomas.

#### **4. Falta de juicio y comprensión humanos**

Las armas autónomas pueden carecer de la capacidad de comprender plenamente situaciones complejas, ejercer su juicio y evaluar la proporcionalidad y legalidad de sus acciones, lo que puede dar lugar a violaciones del derecho internacional humanitario. La falta de juicio y comprensión humanos en las armas autónomas es un problema importante con implicaciones éticas, jurídicas y humanitarias.

El juicio y la comprensión humanos son esenciales para tomar decisiones éticas en situaciones complejas. Las armas autónomas carecen de la capacidad de comprender plenamente los matices del comportamiento, las intenciones y el contexto humanos. Esto puede dar lugar a un posible uso indiscriminado o desproporcionado de la fuerza, con los consiguientes daños a civiles o violaciones de principios éticos como la proporcionalidad y la distinción.

Sin el juicio y la comprensión humanos, resulta difícil atribuir responsabilidades y hacer que los individuos rindan cuentas por las acciones de las armas autónomas. La falta de rendición de cuentas plantea problemas de transparencia, justicia y capacidad para abordar cualquier posible violación del derecho internacional humanitario o de las normas de derechos humanos.

Los operadores humanos tienen la capacidad de considerar los factores contextuales, evaluar la situación y ejercer su juicio a la hora de tomar decisiones. Las armas autónomas, en cambio, pueden carecer de la capacidad de adaptarse a circunstancias cambiantes o de comprender las complejidades de un campo de batalla dinámico. Esto puede dar lugar a decisiones inadecuadas o erróneas que pueden tener graves consecuencias.

Las armas autónomas que carecen de juicio y comprensión humanos pueden no priorizar adecuadamente la protección de los civiles ni cumplir los principios de distinción y proporcionalidad. Esto suscita preocupación por la posibilidad de que aumenten las víctimas civiles y los daños durante los conflictos armados.

La falta de juicio y comprensión humanos en las armas autónomas puede hacer que su comportamiento sea impredecible. Esta imprevisibilidad puede socavar la confianza en el uso de tales armas, tanto entre el personal militar como entre el público en general.

Abordar el problema de la falta de juicio y comprensión humanos en las armas autónomas requiere una cuidadosa consideración de las implicaciones éticas, legales y humanitarias. Se trata de garantizar que se mantengan el control y la supervisión humanos en los procesos de toma de decisiones, y que los sistemas autónomos se diseñen para funcionar dentro de los límites del derecho humanitario internacional y las normas de derechos humanos.

## 5. Bajas civiles y daños colaterales

Permitir que la IA tome decisiones sobre los objetivos suscita preocupación por la posibilidad de que aumenten las víctimas civiles y los daños colaterales inaceptables. El problema de las víctimas civiles y los daños colaterales surge cuando se permite que la IA tome decisiones en la guerra sin la suficiente supervisión humana. Los sistemas de IA pueden no poseer el mismo nivel de juicio, comprensión contextual o capacidad para distinguir entre combatientes y no combatientes que los operadores humanos. Esto puede provocar daños involuntarios a civiles y violaciones del principio de distinción, que exige distinguir entre objetivos militares y civiles.

Percepción pública y confianza: El uso de la IA en la guerra, sobre todo si provoca víctimas civiles y daños colaterales, puede erosionar la confianza y el apoyo de la opinión pública. La preocupación por las implicaciones éticas y las consecuencias imprevistas de los sistemas de IA puede provocar el escepticismo y la oposición del público a su uso en operaciones militares.

Abordar el problema de las víctimas civiles y los daños colaterales permitiendo la IA en la guerra requiere una cuidadosa consideración de las implicaciones éticas, legales y humanitarias. Implica mantener un control humano significativo, garantizar la transparencia y la rendición de cuentas, y desarrollar sistemas de IA que den prioridad a la protección de los civiles y cumplan el derecho internacional humanitario.

## Bibliografía

Chandler, Katherine: AI is Often Biased. Will UN Member States Acknowledge This in Discussions of Autonomous Weapon Systems? September 20, 2021, en: <https://theglobalobservatory.org/2021/09/ai-is-often-biased-will-un-member-states-acknowledge-this-in-discussions-of-autonomous-weapon-systems/>

United States Department of State: Artificial Intelligence (AI), en: <https://www.state.gov/artificial-intelligence/>

Maxwell, Paul: Artificial Intelligence is the Future of Warfare (Just Not in the Way you think", en: <https://mwi.westpoint.edu/artificial-intelligence-future-warfare-just-not-way-think/>

Sharkey, A. Autonomous weapons systems, killer robots and human dignity. *Ethics Inf Technol* 21, 75–87 (2019). <https://doi.org/10.1007/s10676-018-9494-0> <https://link.springer.com/article/10.1007/s10676-018-9494-0>

Lt. Col. John Cherry, Sqn. Ldr. Kieran Tinkler and Michael Schmitt: Avoiding Collateral Damage on the Battlefield, February 11, 2021, en: <https://www.justsecurity.org/74619/avoiding-collateral-damage-on-the-battlefield/>

Ethics and autonomous weapon systems: An ethical basis for ... •  
[https://www.icrc.org/en/download/file/69961/icrc\\_ethics\\_and\\_autonomous\\_weapon\\_systems\\_report\\_3\\_april\\_2018.pdf](https://www.icrc.org/en/download/file/69961/icrc_ethics_and_autonomous_weapon_systems_report_3_april_2018.pdf)

Fully Autonomous Weapons, en:  
<https://www.reachingcriticalwill.org/resources/fact-sheets/critical-issues/7972-fully-autonomous-weapons>

Heed the Call: A Moral and Legal Imperative to Ban Killer Robots ... •  
<https://www.hrw.org/report/2018/08/21/heed-call/moral-and-legal-imperative-ban-killer-robots>

Human Rights Watch: Killer Robots, en: <https://www.hrw.org/topic/arms/killer-robots>

Filippo Santoni de Sio/ Jeroen van den Hoven: Meaningful Human Control over Autonomous Systems: A Philosophical Account, Front Robot AI. 2018; 5: 15, Published online 2018 Feb 28. doi: 10.3389/frobt.2018.00015, en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7806098/>

Human Rights Watch: Mind the Gap, The Lack of Accountability for Killer Robots, en: <https://www.hrw.org/report/2015/04/09/mind-gap/lack-accountability-killer-robots>

Problems with autonomous weapons - Stop Killer Robots, en:  
<https://www.stopkillerrobots.org/stop-killer-robots/facts-about-autonomous-weapons/>

Pros and Cons of Autonomous Weapons Systems  
<https://www.armyupress.army.mil/Journals/Military-Review/English-Edition-Archives/May-June-2017/Pros-and-Cons-of-Autonomous-Weapons-Systems/>

The Emergence of Armed Drones and Today's Collateral Damage ...  
<https://www.fieldartillery.org/news/the-emergence-of-armed-drones-and-todays-collateral-damage-problem>

Sharkey, Noel: The impact of gender and race bias in AI, August 28, 2018, en:  
<https://blogs.icrc.org/law-and-policy/2018/08/28/impact-gender-race-bias-ai/>