



ARMAS AUTÓNOMAS
LETALES Y SU
PROHIBICIÓN:
UN PLAN DE ESTUDIO

FERNANDA CARLES

 CAMPAIGN TO STOP
KILLER ROBOTS

TE
DIC.

**Armas
Autónomas
Letales
y su prohibición:
un plan de estudio**

Fernanda Carles



TECNOLOGÍA &
COMUNIDAD

TEDIC es una Organización No Gubernamental fundada en el año 2012, cuya misión es la defensa y promoción de los derechos humanos en el entorno digital. Entre sus principales temas de interés están la libertad de expresión, la privacidad, el acceso al conocimiento y género en Internet.

Desarrollo: Fernanda Carles

Edición de estilo: Luis Alonzo Fulchi

Diseño de tapa: Adriana Peralta

Diagramación: Horacio Oteiza

ASUNCIÓN - PARAGUAY - 2021



Esta obra está disponible bajo licencia
Creative Commons Attribution 4.0 Internacional (CC BY SA 4.0)
<https://creativecommons.org/licenses/by-sa/4.0/deed.es>

Tabla de contenidos

Introducción	7
Objetivos	7
Robots de aplicación militar y de vigilancia	8
<i>¿Qué problemas puede llegar a traer este tipo de tecnología?</i>	10
¿Qué son los robots?	13
Componentes de los sistemas móviles autónomos	13
<i>Sensores y actuadores</i>	14
Programas: ¿qué es la inteligencia artificial, y cómo se diferencia de la inteligencia humana?	15
Ejemplos del uso problemático de la Inteligencia Artificial	17
<i>Dinámica 1 - Analizando casos problemáticos de implementación de Inteligencia Artificial</i>	18
Sesgos: género, racismo e interseccionalidad	19
<i>¿Es la inteligencia artificial neutral?</i>	19
<i>Dinámica 2 - ¿Cómo actuamos cuando la Inteligencia Artificial toma el control?</i>	21
El razonamiento humano para el control de armas	22
<i>Dinámica 3 - Niveles de control de armas y cómo impacta en el proceso de decisión humano</i>	23
Argumentos legales contra las armas autónomas	24
<i>Argumentos legales: leyes de guerra y leyes civiles</i>	24
<i>La cláusula de Martens</i>	25
<i>Derecho Internacional de los Derechos Humanos</i>	25
<i>¿Quién es responsable por las acciones de estas máquinas?</i>	26
¿Cuál es el futuro de las armas autónomas letales, y cómo afectaría a la seguridad global?	27
¿Qué podemos hacer para apoyar la campaña “Stop Killer Robots”?	29
<i>Cartas científicas abiertas y el rol de la prensa</i>	31
Referencias	33

Introducción

La Campaña Contra Robots Asesinos (“Stop Killer Robots”) creada en 2012, es una alianza de organizaciones no gubernamentales que trabaja para prohibir las armas totalmente autónomas y mantener el control humano sobre el uso de la fuerza. Las armas autónomas son aquellas que podrían seleccionar y atacar objetivos sin intervención humana. Es decir, podrían elegir por sí mismas entre la vida y la muerte de una o muchas personas.

Actores de gobierno, sociedad civil y academia de más de 68 países están llamando a su prohibición, por considerar que desde una perspectiva ética y de derechos humanos es inaceptable que un robot tome decisiones de vida o muerte de manera autónoma, ya que esto puede poner en riesgo las vidas de personas inocentes y atenta contra la dignidad humana.

“Armas Autónomas Letales y su prohibición: un plan de estudio” es una serie de resúmenes y dinámicas que busca brindar herramientas para construir el debate alrededor de esta problemática, y concientizar sobre los diversos aspectos técnicos, sociales, éticos y políticos alrededor del uso de tecnologías de inteligencia artificial para la toma de decisiones sobre la vida humana.

Objetivos

- Contextualizar acerca del uso, las capacidades y limitaciones actuales de la robótica, la inteligencia artificial y los sistemas de biometría computarizada.
- Sensibilizar sobre los problemas existentes en la construcción de sistemas de inteligencia artificial en cuanto a sesgos con una mirada interseccional, además de sus riesgos en la toma de decisiones sobre la vida humana, y en específico los potenciales problemas de la utilización de armas autónomas letales.
- Resaltar la tendencia global de desarrollo y adquisición de tecnologías armamentistas y de vigilancia que utilizan tecnologías de inteligencia artificial por parte de actores estatales y civiles.
- Concientizar sobre los argumentos de la campaña “Stop Killer Robots” en contra de esta tendencia.
- Concientizar en el rol que debe cumplir la comunidad científica y los medios de comunicación en la lucha por la prohibición del desarrollo de armas autónomas letales.

Robots de aplicación militar y de vigilancia

La robótica es un elemento fundamental de la fabricación avanzada desde hace más de medio siglo. A medida que los robots y sus equipos periféricos se vuelven más sofisticados, fiables y miniaturizados, estos sistemas se utilizan cada vez más para fines militares y policiales.

Las aplicaciones militares siguen creciendo a un ritmo acelerado debido a la demanda alimentada por la inversión gubernamental. Las funcionalidades de estas aplicaciones muchas veces se centran en tareas de reconocimiento de terreno, vigilancia, desactivación de explosivos y detección de minas, mulas de terreno robóticas y exoesqueletos que funcionan como armaduras para soldados¹. Aunque estas aplicaciones pueden pensarse como inofensivas, han posibilitado el desarrollo de lo que se podrían denominar “nuevas armas”.

Los sistemas militares robóticos de ofensiva, y en particular los drones, han conseguido abaratar drásticamente las operaciones militares mientras han incrementado significativamente el volumen de negocio de la industria militar. El sector de las empresas que fabrican y exportan drones militares y el instrumental de alta tecnología que se utiliza en las guerras modernas se encuentra en pleno auge².

Las siguientes imágenes muestran los países donde estos desarrollos se están llevando adelante, y los que hoy en día tienen disponible este tipo de tecnología.

FIGURA 1: Países fabricantes de drones militares y países importadores. Fuente: Anexo 1 y Rusi.org, “Armed Drones in the Middle East”: <https://drones.rusi.org/countries/>; Gettinger, Dan (2019). The Drone Databook, Center for the Study of the Drone at Bard College. Disponible en: <https://dronecenter.bard.edu/files/2019/10/CSD-Drone-Databook-Web.pdf>

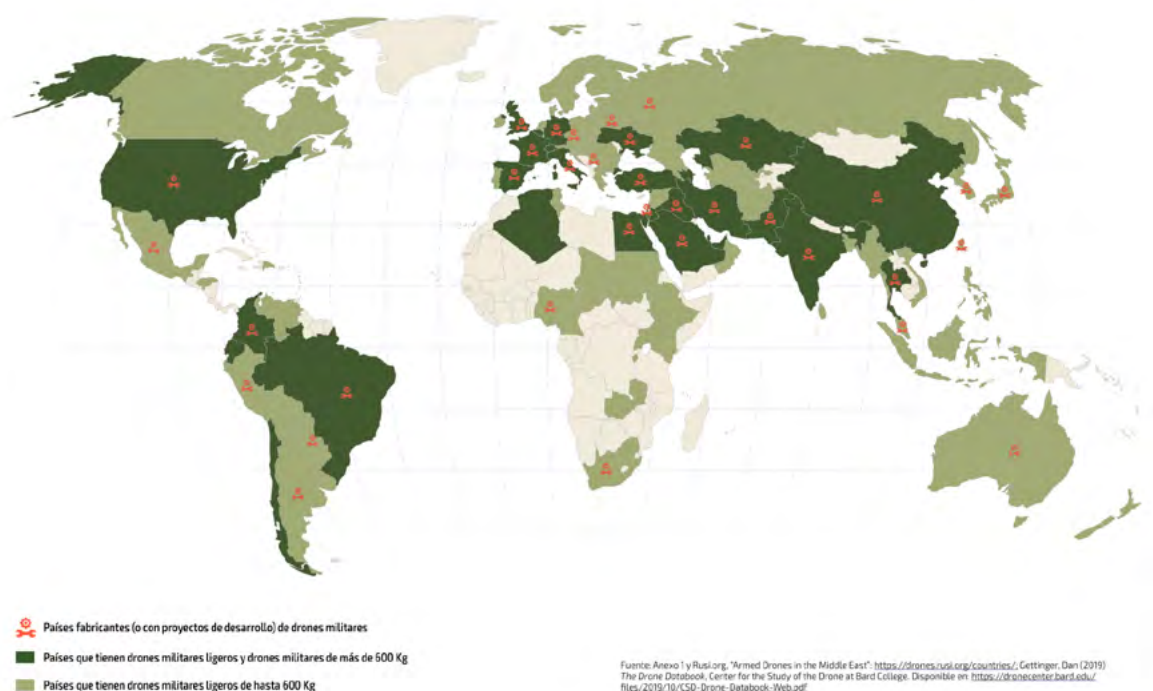
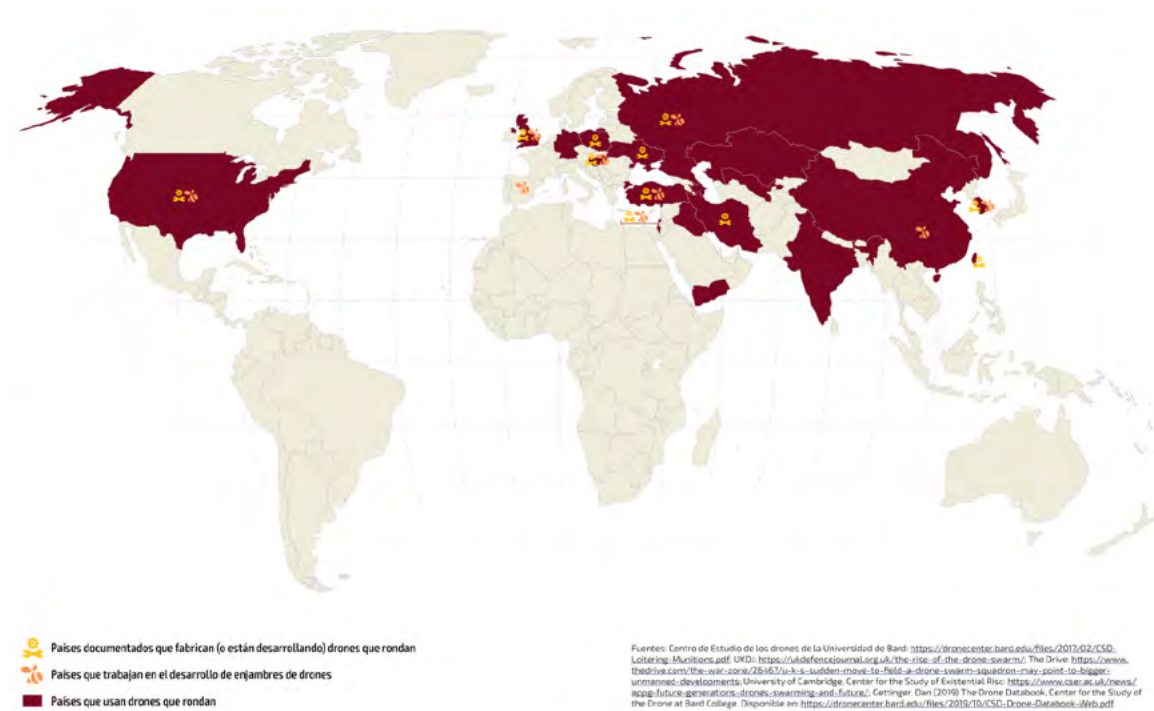


FIGURA 2: Países que disponen de drones que rondan (“Loitering drones”) y que desarrollan enjambres de drones. Fuente: Centro de Estudio de los drones de la Universidad de Bard: <https://dronecenter.bard.edu/files/2017/02/CSD-Loitering-Munitions.pdf>; UKDJ: <https://ukdefencejournal.org.uk/the-rise-of-the-drone-swarm/>; The Drive: <https://www.thedrive.com/the-war-zone/26467/u-k-s-sudden-move-to-field-a-drone-swarm-squadron-may-point-to-bigger-unmanned-developments>; University of Cambridge, Center for the Study of Existential Risk: <https://www.cser.ac.uk/news/appg-future-generations-drones-swarming-and-future/>; Gettinger, Dan (2019) The Drone Databook, Center for the Study of the Drone at Bard College. Disponible en: <https://dronecenter.bard.edu/files/2019/10/CSD-Drone-Databook-Web.pdf>



¿Qué problemas puede llegar a traer este tipo de tecnología?

El proceso de aceleración tecnológica en el que se ve inmersa nuestra sociedad ha transformado de manera irreversible al mundo. En este contexto, el desarrollo de la inteligencia artificial supone un nuevo horizonte de posibilidades, con beneficios en todos los campos del conocimiento y actuar humanos.

Sin embargo, estas nuevas herramientas también pueden presentar graves riesgos en términos de derechos civiles y humanos. Específicamente en el ámbito militar, surge el siguiente cuestionamiento: ¿Podrían ser los robots asesinos parte de nuestro futuro? ¿Qué implicaría la implementación de esta tecnología?

A menos que se establezcan restricciones, las armas autónomas letales se desplegarán en los próximos años y no en décadas. Varios precursores demuestran claramente la tendencia hacia sistemas de armas cada vez más autónomos.

Ejemplos de aplicaciones en terreno de armas autónomas letales

SRG-A1

Fabricado por: Samsung (Corea del Sur)

Vendido a: Corea del Sur

Este robot estacionario, armado con una metralleta y un lanzagranadas, opera sobre la frontera entre Corea del Norte y Corea del Sur. Puede detectar seres humanos utilizando sensores infrarrojos y software de reconocimiento de patrones. El robot tiene modalidades tanto supervisadas como no-supervisadas. Puede identificar y atacar intrusos, con la posibilidad de disparar.



FIGURA 3: SRG-A1. Fuente: <https://defensereview.com/samsung-sgr-a1-armedweaponized-robot-sentry-or-sentry-robot-remote-weapons-station-rws-finally-ready-for-prime-time/>

Sea Hunter

Fabricado por: DARPA (USA). Construido por Vigor Industrial
Vendido a: EEUU (En desarrollo)

Esta nave de guerra de 40m de longitud está diseñada para cazar submarinos enemigos y puede funcionar sin contacto con un operador humano por periodos de entre 2 y 3 meses. La versión actual se encuentra armada, y se planea construir flotillas sin tripulación en los próximos años.



FIGURA 4: Sea Hunter. Fuente: [https://en.wikipedia.org/wiki/Sea_Hunter#/media/File:Sea_Hunter_gets_underway_on_the_Willamette_River_following_a_christening_ceremony_in_Portland,_Ore._\(25702146834\).jpg](https://en.wikipedia.org/wiki/Sea_Hunter#/media/File:Sea_Hunter_gets_underway_on_the_Willamette_River_following_a_christening_ceremony_in_Portland,_Ore._(25702146834).jpg)

Harpy

Fabricado por: Israel Aerospace Industries (Israel)
Vendido a: China, India, Israel, Corea del Sur y Turquía

Este misil de merodeo de 2,1 m es lanzado desde un vehículo en tierra. Está armado con una cabeza explosiva de 15 kg. El Harpy puede merodear por un periodo de hasta 9 horas a la vez, buscando señales de radares enemigos. Detecta, ataca y destruye automáticamente estos emisores.

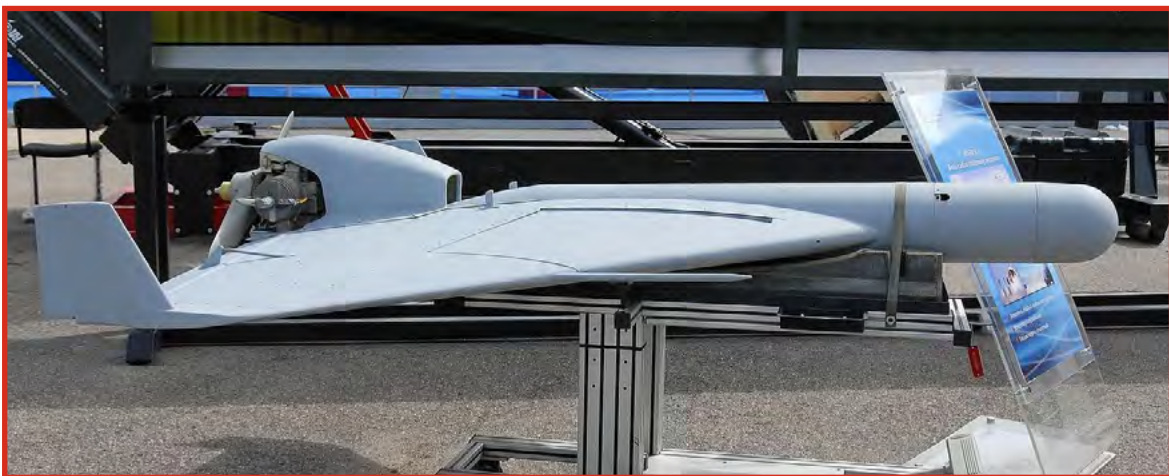


FIGURA 5: Harpy. Fuente: https://es.wikipedia.org/wiki/IAI_Harpy#/media/Archivo:Paris_Air_Show_2007-06-24_n25.jpg

nEUROn

Fabricado por: Dassault Aviation (France)

Vendido a: En desarrollo

Este avión de combate no tripulado de 10 metros de largo puede volar de forma autónoma durante más de 3 horas para la detección, localización y reconocimiento autónomo de objetivos terrestres. El Neuron tiene capacidades de ataque totalmente automatizadas, ajuste de objetivos y comunicación entre sistemas.



FIGURA 6: nEUROn. Fuente: <https://www.dassault-aviation.com/wp-content/blogs.dir/2/files/2017/05/nEUROn-resize.jpg>

Debido a serias implicaciones éticas, legales y de seguridad que trae consigo el desarrollo de este tipo de armas, en Octubre del 2012, diversas organizaciones, entre ellas la Human Rights Watch, el Comité Internacional por el Control de Armas Robóticas (ICRAC, por sus siglas en inglés) y otras cinco organizaciones no-gubernamentales fundaron la Campaña “Stop Killer Robots” (Paremos los robots asesinos), para dar una respuesta coordinada por parte de la sociedad civil a esta problemática.

El objetivo de la campaña “Stop Killer Robots” no ha cambiado desde su creación. La misma trabaja por la prohibición preventiva del desarrollo, la producción y el uso de armas totalmente autónomas que tengan como objetivos personas. Además, aboga por la regulación de aquellas que tengan otros objetivos conservando el control humano significativo.

En posteriores secciones, se explorarán los argumentos técnicos, legales y sociales que enmarcan esta campaña.

¿Qué son los robots?

La palabra “robot” fue utilizada por primera vez por el escritor de ficción Karel Čapek en el año 1920, en su obra de ficción “RUR” (Rossum’s Universal Robots)³. Esta palabra, derivada de *robotnik*, del checo “trabajador forzado”, es hoy en día utilizada universalmente para referirse a cualquier tipo de aparato mecánico construido para la realización de tareas físicas de manera autónoma. En su obra, Čapek describe el objetivo de la construcción de sus “robots” como la búsqueda de la liberación de la humanidad de tareas peligrosas y repetitivas. A través de esta herramienta, se nos concedería un nuevo impulso y una nueva realidad, con una vida mejor y menos mecánica.

Aunque el uso de la palabra “robot” no comenzó hasta el siglo XX, existen numerosos ejemplos a lo largo de la historia de máquinas construidas para emular movimientos y realizar tareas reservadas para seres vivos.⁴ Estas máquinas no solo buscaban emular movimientos y realizar tareas reservadas para los seres vivos (con un énfasis especial en los seres humanos), sino que además, sus creadores invertían grandes esfuerzos en moldearlas para que se vieran como tal.

Es así que los robots responden a dos necesidades o deseos humanos: la liberación de tareas triviales, y la comprensión, réplica y potenciación de sistemas que tengan las mismas capacidades (o capacidades similares) a los de los seres vivos. Pero, ¿qué tan cerca nos encontramos de este objetivo?

Componentes de los sistemas móviles autónomos

Todo sistema autónomo móvil tiene tres ingredientes esenciales. Cada uno de estos asiste en la interacción del sistema autónomo con el mundo físico: detectar un entorno (sensores), analizar los elementos de este ambiente (programas), y realizar una acción física a partir de este análisis (actuadores). Estos elementos no son únicos de los sistemas robóticos, sino que podemos ver que también los sistemas vivos que interactúan con el medio, incluido el ser humano, cuentan sensores (sentidos), programas (sistema nervioso) y actuadores (sistema esqueleto-muscular). Nos preguntamos, entonces, ¿cuentan los sistemas robóticos actuales, estos “trabajadores forzados”, con sensores, actuadores y programas con las mismas capacidades que los humanos? ¿De qué manera se diferencian estas capacidades de las máquinas a las nuestras?

Sensores y actuadores

Los **sensores** son dispositivos, módulos, máquinas o subsistemas que tienen el propósito de detectar eventos y cambios en el ambiente y enviarlos a un procesador. En este sentido, los humanos contamos con al menos cinco sentidos principales (olfato, gusto, tacto, vista y audición)⁵, además del sentido del espacio (propiocepción)⁶. Los sensores utilizados por los sistemas robóticos proveen análogos a los sentidos humanos, tales como micrófonos, cámaras con sistemas de visión artificial, entre otros, que recogen información del ambiente con una granularidad ampliamente superior, y pueden monitorear otros fenómenos para los cuales los humanos no tenemos sentidos específicos.



FIGURA 7: Diferentes tipos de sensores. Fuente: Open Source Robotics Foundation - CC BY 3.0: <http://www.ros.org/news/2011/03/sensors-supported-by-ros.html>

Por ejemplo, los sistemas robóticos pueden contar con tecnologías que permiten la detección de emisiones electromagnéticas infrarrojas y ultravioletas (fuera del espectro visible humano), orientación física respecto al eje magnético terrestre, medición exacta de distancia entre objetos, presencia, identidad y concentración de sustancias químicas, entre otros. Es también interesante reconocer que los sistemas robóticos pueden estar distribuidos en un espacio de gran amplitud, recogiendo y centralizando información de distintas áreas al mismo tiempo. Esto es reconocible, por ejemplo, en los edificios automatizados, que recogen información del ambiente de todo el edificio para regular temperatura, abrir puertas, accionar mecanismos, etc. Además, en teoría, la extensión de estos sistemas interconectados no tiene límites. Esto implica que los sistemas robóticos pueden ser construidos para recoger una amplia gama de información de su entorno.

Así como los sensores recogen información del ambiente, la manera que tienen los sistemas que interactúan con el ambiente de ejercer acciones sobre el mismo es a través de los actuadores. **Un actuador** es un componente de un sistema responsable por el movimiento y control de un mecanismo. Existe una diversa gama de actuadores que dan diferentes tipos de movimiento a los sistemas robóticos.

Actualmente, los mecanismos de movimiento desarrollados por empresas como Boston Dynamics⁷ han sido capaces de emular con extraordinaria precisión movimientos y acciones anteriormente reservadas para los seres vivos. Sin embargo, es interesante notar que la variedad de capacidades de movimiento, velocidad, precisión y fuerza de los sistemas robóticos es extremadamente extensa. Existen en la actualidad robots de transporte como drones aéreos, terrestres y acuáticos, robots con capacidad de levantar toneladas⁸, robots flexibles que utilizan mecanismos de locomoción inspirados en insectos que pueden ingresar en zonas imposibles para los humanos⁹ y otros con precisión suficiente para realizar cirugías complejas¹⁰.

Tanto los sensores como actuadores permiten a las máquinas autónomas sentir el ambiente y realizar movimientos que superan en gran medida las capacidades humanas. En ese sentido, las máquinas autónomas se han convertido en un gran aliado en la resolución de problemas, y cumplen con el cometido original de los robots de apoyar en la liberación humana de ciertos tipos de trabajo.

Sin embargo, es importante recordar que uno de los componentes principales que hacen a un robot es la capacidad de autonomía. Una vez que el robot toma información del ambiente a través de los sensores, su capacidad de análisis le permite tomar decisiones acerca de qué manera debe utilizar sus actuadores. Es aquí en el procesamiento y razonamiento automático, en la toma de decisiones, donde las máquinas quedan cortas.

Programas: ¿qué es la inteligencia artificial, y cómo se diferencia de la inteligencia humana?

Los sensores y actuadores son el enlace entre el programa, “la mente” de las máquinas autónomas, y el mundo físico. Aunque los actuadores y sensores tienen la capacidad de influir en el espacio de innumerables maneras, la ciencia de computación no ha sido capaz de replicar aún el razonamiento humano.

La inteligencia artificial (IA) es una rama de la ciencia de datos que busca construir máquinas inteligentes capaces de realizar tareas que usualmente requieren de inteligencia humana y cognición. Estas máquinas inteligentes están capacitadas para aprender de la experiencia y los datos históricos, analizar el entorno que las rodea y realizar las acciones correspondientes. Es una tecnología interdisciplinaria que utiliza conceptos y herramientas de múltiples campos como la ciencia computacional, ciencias cognitivas, lingüística, psicología, neurociencia y matemáticas.

En contraste, la inteligencia humana se refiere a la capacidad intelectual de los seres humanos que nos permite pensar, aprender de diferentes experiencias, comprender conceptos complejos, aplicar la lógica y la razón, resolver problemas matemáticos, reconocer patrones, hacer inferencias y tomar decisiones, retener información y comunicarnos con otros seres humanos.

Lo que hace que la inteligencia humana sea única es que está respaldada por emociones abstractas como la autoconciencia, la pasión y la motivación que permiten a los humanos realizar tareas cognitivas complejas.

Aspecto	Inteligencia Humana	Inteligencia Artificial
Naturaleza	Pretende adaptarse a nuevos entornos utilizando una combinación de diferentes procesos cognitivos. El cerebro humano es analógico.	Pretende construir máquinas que puedan imitar el comportamiento humano y realizar acciones similares a las humanas. Los procesadores de IA son digitales.
Funcionamiento	Utiliza el poder computacional del cerebro, memoria, y habilidad de pensamiento.	Utiliza datos provistos e instrucciones específicas provistas por un programador humano que alimenta el sistema.
Poder de aprendizaje	La inteligencia humana se construye a partir del aprendizaje sobre diversos incidentes y experiencias pasadas a través del método de ensayo y error. Estos ensayos se realizan a lo largo de la vida y experiencia humana.	Las máquinas no pueden pensar. El aprendizaje se da a partir de datos y entrenamiento continuo, pero el proceso de pensamiento humano aún no ha podido ser emulado. Aunque los sistemas impulsados por la IA pueden realizar tareas específicas con gran acierto, pueden tardar años en aprender un conjunto de funciones completamente diferente para un nuevo ámbito de aplicación.

TABLA 1: Tabla comparativa - Inteligencia Artificial VS Inteligencia Humana

El poder de decisión de los sistemas de IA se basa principalmente en eventos, los datos de entrenamiento, y cómo éstos se relacionan a un evento en particular. Las máquinas de IA no pueden entender el concepto de “causa y efecto”, ya que no cuentan con sentido común.

Ejemplos del uso problemático de la Inteligencia Artificial

A continuación, se presenta una breve recopilación de sistemas de Inteligencia Artificial problemáticos que han surgido en la última década. Los problemas que presentan estas tecnologías son tanto por errores en el diseño como por prácticas éticas cuestionables de las empresas creadoras.

- **Vehículo autónomo de Uber mata a peatón:** la primera muerte relacionada a un vehículo autónomo en la vía pública ocurrió en el 2018 en Tempe, Arizona. El vehículo de Uber se encontraba en modo autónomo, con un conductor humano de seguridad al volante. Según el reporte preliminar del accidente realizado por la US National Transportation Safety Board¹¹ Uber descubrió que su software de conducción autónoma decidió no tomar ninguna acción luego de que los sensores del vehículo detectaron a un peatón.
- **IBM Watson para atención sanitaria:** en 2013 IBM desarrolló la primera aplicación comercial para recomendaciones de tratamiento de cáncer. La misma terminó siendo un fracaso, ya que los médicos presentaron quejas de que el sistema realizaba recomendaciones para los pacientes que podrían tener consecuencias severas y hasta fatales¹².
- **Herramienta de reclutamiento de Amazon presenta sesgos de género:** el departamento de recursos humanos de Amazon utilizó un software de contratación basado en IA entre el 2014 y el 2017, para ayudar a revisar currículums y realizar recomendaciones. Se descubrió que el software favorecía a candidatos masculinos.
- **Predicción de IQ a partir de rostros:** la empresa israelí de aprendizaje automático Faception realizó una afirmación controversial de que su tecnología de IA podía analizar imágenes faciales y estructura ósea para predecir el coeficiente intelectual, la personalidad e incluso tendencias violentas de las personas¹³.

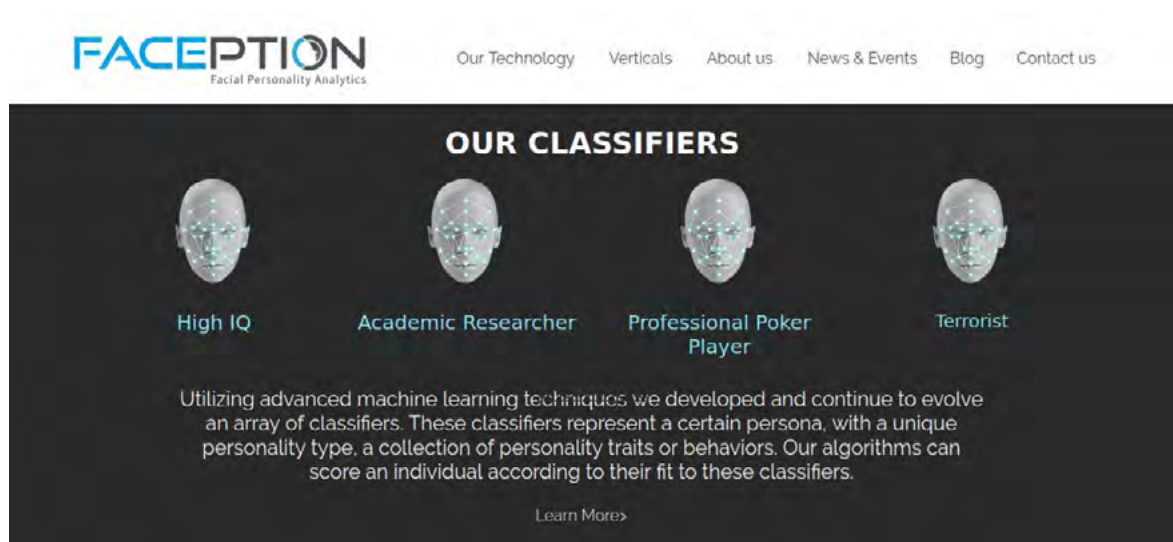


FIGURA 8: Página oficial de la empresa Faception <https://www.faception.com/>

DINÁMICA 1

Analizando casos problemáticos de implementación de Inteligencia Artificial

Objetivo:

- Generar un debate acerca de la construcción de sistemas de inteligencia artificial, sus posibles sesgos y consecuencias de implementación inadecuada.

Desarrollo:

En base a los dos últimos ejemplos de uso cuestionable de la inteligencia artificial (“Herramienta de reclutamiento de Amazon presenta sesgos de género” y “Predicción de IQ a partir de rostros”), el facilitador genera un debate con los participantes acerca de por qué estas tecnologías pueden considerarse como problemáticas.

Preguntas disparadoras:

- La herramienta de reclutamiento de personal automatizada de Amazon utilizó la base de datos de contrataciones de la empresa de los últimos diez años para entrenar su sistema de Inteligencia Artificial.
 - ¿Por qué la herramienta recomendaba mayoritariamente a hombres?
 - ¿Qué impacto pudo haber tenido esto en los resultados que arrojaba?
- En el caso de Faception, la empresa asegura que su sistema de IA tiene la capacidad de identificar rasgos de personalidad e inteligencia de las personas a partir de sus rasgos faciales.
 - ¿Qué tanta validez científica tiene esta afirmación, y qué resultados podría llegar a arrojar teniendo en cuenta factores como el género y la raza?
 - ¿Dejarías que Faception analice tu rostro?
- En su página web, la empresa Faception menciona que esta tecnología podría ser utilizada para control de fronteras, vigilancia, detección de potenciales terroristas, el escaneo de vídeos para detectar personas sospechosas en investigaciones criminales, y detección de amenazas en estadios y otras áreas concurridas.
 - ¿Qué tipo de impacto puede llegar a tener esta tecnología en la vida de las personas?
 - ¿Qué otras aplicaciones podrían darse a esta herramienta?

Sesgos: género, racismo e interseccionalidad

¿Es la inteligencia artificial neutral?

Cuando se habla de tecnología, la idea de que la misma es esencialmente neutral, y que su estructura y funcionalidades son independientes a los sentimientos y pasiones humanos puede parecer hasta intuitiva. Sin embargo, es importante reconocer que la tecnología está construida por personas, y que los creadores de tecnología plasman en sus diseños sus ideas y percepciones no sólo técnicas, sino también sociales y políticas, muchas veces de manera no intencional. La tecnología tiene la capacidad de reproducir las actitudes y desigualdades presentes en nuestras sociedades.

Esto es especialmente relevante cuando hablamos de tecnologías que tienen un fuerte componente de Inteligencia Artificial. Esta tecnología utiliza como insumo principal grandes bases de datos para entrenar al sistema, y es común encontrar que estas bases de datos no representan la diversidad de personas y voces que existen en el mundo, y muchas veces presentan sesgos de género y raza. **Una gran preocupación respecto a la construcción de sistemas de armas autónomas letales tiene que ver con el impacto diferenciado que podrían tener estas armas hacia ciertos grupos de la población, con consecuencias letales.**

Existen numerosos ejemplos de esta problemática en diversos tipos de aplicaciones. En un estudio del Media Lab del Massachusetts Institute of Technology (MIT)¹⁴, se encontró que diversas tecnologías de reconocimiento facial basados en técnicas de IA sólo funcionaban de manera adecuada para cierto tipo de personas. El estudio indica que en hombres de piel clara el error de reconocimiento facial es del 1%, en hombres de piel oscura es del 19%, y en mujeres de piel oscura es del 35%. Es decir, de ser utilizada este tipo de herramienta en la composición de un arma autónoma, las personas de piel oscura –y en particular las mujeres de piel oscura– enfrentarían un riesgo mayor de un error de reconocimiento que los hombres de piel clara, con posibles consecuencias letales. Actualmente, el reconocimiento facial tiene un sesgo androcéntrico y racista que puede reproducir y amplificar la discriminación: en este ejemplo, por el género y el color de piel.

Sería muy probable que las armas autónomas tuvieran también, un sesgo discriminatorio contra las personas con discapacidad, neuro diversas y personas de la comunidad LGBTI+, cuya apariencia, comportamiento y diversidad de condiciones podrían no corresponder a las normas consideradas por las personas que se dedican al desarrollo de tecnología y de armas autónomas. Los prejuicios humanos están presentes en los algoritmos y los datos que utilizamos para entrenar programas de aprendizaje automático, y a menudo reflejan los sesgos de nuestra propia sociedad: género, clase y raza.

Es importante mencionar que este problema está relacionado con las brechas de género, clase y raza en torno al acceso al conocimiento necesario para implementar este tipo de sistemas, así como las personas o grupos de personas con los recursos para implementarlos y tomar decisiones al respecto. En la actualidad, un porcentaje extremadamente reducido de la población mundial tiene las habilidades y educación necesarias para crear programas de IA, y la mayoría de estas personas han nacido en circunstancias bastante privilegiadas. Así mismo, una estimación de WIRED con Element AI encontró que sólo el 12% de los principales investigadores de aprendizaje automático son mujeres¹⁵.

DINÁMICA 2

¿Cómo actuamos cuando la Inteligencia Artificial toma el control?

Objetivo:

- Generar un debate acerca del nivel de confianza que se asigna a los sistemas de inteligencia artificial y las posibles consecuencias de errores en el sistema.

Desarrollo:

En base al ejemplo presentado “Vehículo autónomo de Uber mata a peatón”, el facilitador genera un debate con los participantes acerca de por qué se dio este accidente fatal:

- **Vehículo autónomo de Uber mata a peatón:** la primera muerte relacionada a un vehículo autónomo en la vía pública ocurrió en el 2018 en Tempe, Arizona. El vehículo de Uber se encontraba en modo autónomo, con un conductor humano de seguridad al volante. Según el reporte preliminar del accidente realizado por la US National Transportation Safety Board Uber descubrió que su software de conducción autónoma decidió no tomar ninguna acción luego de que los sensores del vehículo detectaron a un peatón.

Preguntas disparadoras:

- Según el reporte preliminar del accidente realizado por la US National Transportation Safety Board, el vehículo autónomo contaba con un conductor humano en el momento del accidente, cuyo rol era monitorear el trayecto del vehículo e intervenir de ser necesario. El conductor humano podía intervenir en el curso del vehículo moviendo el volante, freno, acelerador, o presionando un botón de desconexión o desactivación.
 - ¿Cuáles son posibles motivos por los cuales el conductor humano no intervino?
 - ¿Podemos pensar en errores tanto humanos como de máquina, que dieron lugar a este accidente?

El razonamiento humano para el control de armas

Esta sección está basada en el artículo “Human Control of Weapons Systems” de Noel Sharkey (International Committee for Robot Arms Control), publicado en el “Campaign to Stop Killer Robots - Campaigner’s Kit”

¿Qué tipo de control humano es necesario para garantizar medidas de precaución cuando se evalúa la importancia de los objetivos potenciales en conflictos bélicos?

De acuerdo al artículo “Human Control of Weapons Systems” de Noel Sharkey, una distinción establecida en la psicología humana divide el razonamiento humano en dos tipos. Los **procesos automáticos rápidos** son necesarios para llevar a cabo tareas rutinarias como conducir una bicicleta, evitar el tráfico o jugar un deporte. Este tipo de razonamiento es vital para reaccionar rápido y llevar adelante tareas que no necesitan pensamiento consciente.

Por otro lado, los **procesos deliberativos más lentos** son necesarios para el razonamiento reflexivo. Esto es importante para tomar decisiones de alto impacto como las decisiones diplomáticas, médicas o judiciales.

El razonamiento automático es esencial en nuestro funcionamiento normal diario, pero tiene una serie de inconvenientes cuando se trata de tomar decisiones importantes como las necesarias para determinar la legitimidad de un objetivo:

- Niega la ambigüedad y suprime la duda: el razonamiento automático humano salta a conclusiones. Una respuesta sin ambigüedades aparece instantáneamente sin cuestionamientos. No hay búsqueda de interpretaciones alternativas o incertidumbres. Si algo parece un objetivo legítimo, en circunstancias ambiguas, el razonamiento automático tendrá la certeza de que es legítimo.
- Infiere e inventa causas e intenciones: el razonamiento automático rápidamente inventa historias causales coherentes uniendo fragmentos de información disponible. Por ejemplo, en el contexto de un conflicto armado, un grupo de personas subiendo cajas a un camión puede iniciar una historia causal de que las cajas están llenas de rifles.
- Está predispuesto a creer y confirmar: el razonamiento automático favorece la aceptación acrítica de las sugerencias y mantiene un fuerte sesgo. Si un ordenador sugiere un objetivo a un operario, el razonamiento automático por sí solo haría muy probable su aceptación. Esto es el sesgo de automatización. El sesgo de confirmación selecciona la información que confirma una creencia previa.
- Se concentra en evidencia existente e ignora evidencia faltante: el razonamiento automático construye historias explicativas coherentes sin tener en cuenta las pruebas o la información contextual que pueda faltar. Facilita la sensación de coherencia que nos hace confiar en aceptar la información como verdadera. Por ejemplo, un hombre que dispara un rifle puede considerarse un objetivo hostil, cuando un rápido vistazo a su alrededor podría revelar que está disparando a un lobo que caza su ganado.

DINÁMICA 3

Niveles de control de armas y cómo impacta en el proceso de decisión humano

Objetivo:

- Generar un debate acerca de los sesgos de razonamiento y sus posibles consecuencias en distintos niveles de control humano sobre armas

Desarrollo:

El facilitador presenta a los participantes una tabla de clasificación de niveles de control humano sobre armas. La misma es completada por los participantes a través del debate.

Nivel	¿Qué sesgos del razonamiento humano pueden aparecer en este tipo de control?	¿Qué tipo de consecuencias negativas puede traer este nivel de control?	¿Qué tan aceptable es este nivel de control de armas letales?
Nivel 1 Un humano delibera sobre un objetivo antes de iniciar cualquier ataque.			
Nivel 2 Un programa provee una lista de posibles objetivos y un humano elige cuál atacar.			
Nivel 3 Un programa selecciona un objetivo y un humano debe aprobar la selección antes de atacar.			
Nivel 4 Un programa selecciona un objetivo y un humano tiene un periodo restringido de tiempo para parar el ataque.			
Nivel 5 Un programa elige un objetivo e inicia un ataque sin intervención humana.			

Argumentos legales contra las armas autónomas

Esta sección está basada en el artículo “Legal Arguments” de Bonnie Docherty y Matthew Griechen (Harvard Law School International Human Rights Clinic (IHRC)), publicado en el “Campaign to Stop Killer Robots - Campaigner’s Kit”.

Argumentos legales: leyes de guerra y leyes civiles

De ser desarrolladas las armas autónomas letales, el alcance de las mismas probablemente no se limitará a aplicaciones bélicas, sino que existe el riesgo de que los Estados utilicen estas tecnologías para aplicaciones policiales. Los robots asesinos amenazan con violar tanto la Ley Humanitaria Internacional (LHI) como la Ley Internacional de Derechos Humanos.

El Derecho Internacional Humanitario, también conocido como las leyes de la guerra, gobernaría el uso de las armas autónomas letales en el campo de batalla. Como estas armas operarían sin un control humano significativo, entrarían en conflicto con dos reglas fundamentales de la LHI:

- En primer lugar, el derecho internacional convencional y el artículo 48 del Protocolo Adicional I de los Convenios de Ginebra obligan a las partes beligerantes a distinguir entre civiles y soldados y entre objetos civiles (como casas o escuelas) y objetivos militares. Las armas que no pueden hacer esas distinciones se consideran “indiscriminadas” e ilegales.

Los robots asesinos podrían encontrar importantes obstáculos para cumplir la norma de distinción. Diferenciar entre civiles y soldados, especialmente en una época en la que los combatientes suelen mezclarse con la población local, depende de algo más que de reconocer un uniforme. También depende de entender las intenciones de una persona a través de pistas como el tono de voz, las expresiones faciales o el lenguaje corporal. Los seres humanos están mejor equipados para entender esos matices que las máquinas.

- En segundo lugar, el derecho internacional consuetudinario y el artículo 51(5)(b) del Protocolo Adicional I exigen que las partes beligerantes sopesen la proporcionalidad de un ataque. Esta norma prohíbe los ataques en los que el daño previsto a la población civil y a los bienes de carácter civil es excesivo en relación con la ventaja militar prevista. La proporcionalidad no es una cuestión matemática. Depende del contexto, y la prueba es si un “comandante militar razonable” habría considerado lícito lanzar el ataque.

Las armas totalmente autónomas no podrían replicar el juicio humano necesario para evaluar la proporcionalidad de un ataque específico. Dado que los programadores no pueden prever de antemano la infinidad de escenarios que pueden surgir en el campo de batalla, las armas totalmente autónomas se encontrarán con circunstancias imprevistas y cambiantes. Sin embargo, a diferencia de los humanos, estas máquinas no podrían aplicar el razonamiento y la experiencia humanos a la hora de equilibrar los factores relevantes de esta prueba subjetiva.

La cláusula de Martens

Esta cláusula es una provisión de la Ley Humanitaria Internacional que enlaza la ley y la ética. La misma declara que en ausencia de un tratado específico sobre un tema, las personas están protegidas por la “costumbre”, “los principios de humanidad”, y “los dictámenes de la conciencia pública”. Esta cláusula crea un estándar moral con el cual juzgar las armas autónomas letales.

Las armas autónomas letales plantean graves problemas en virtud de los principios de humanidad. Los humanos están motivados a tratar entre ellos de manera humana, ya que pueden sentir compasión y empatía por las experiencias de otras personas. Las armas autónomas letales, en contraste, no tendrían la capacidad emocional que subyace al trato humano. Estas armas no pueden entender el valor de una vida y la significancia de su pérdida. Determinarían a quién matar basadas en algoritmos, y no podrían considerar el valor inherente de una víctima individual.

Los dictámenes de la conciencia pública, que se refieren a guías morales compartidas, argumentan en contra de las armas autónomas letales de manera similar. En una encuesta de opinión pública de diciembre del 2018 llevada a cabo en 26 países, más del 60% de las personas respondieron que están en contra de los robots asesinos¹⁶.

Derecho Internacional de los Derechos Humanos

El Derecho Internacional de los Derechos Humanos es una rama del derecho desarrollada para promover y proteger los derechos humanos a nivel internacional, regional y nacional, que aplica tanto durante tiempos de paz como de conflicto armado. Las armas autónomas tienen el potencial de violar dos derechos humanos fundamentales:

- Primero, bajo el Artículo 6 del Pacto Internacional de Derechos Civiles y Políticos, todas las personas tienen el derecho fundamental a la vida, es decir, no pueden ser privados arbitrariamente de sus vidas. Matar sólo es lícito cuando es necesario para proteger la vida humana, constituye un último recurso y se aplica de forma proporcional a la amenaza.

Este juzgamiento es específico al contexto, y los robots asesinos no tendrían las cualidades humanas de empatía y juicio necesarias para realizar estas determinaciones en situaciones no previstas.

- En segundo lugar, el principio de la dignidad humana sustenta la legislación sobre derechos humanos. Toda vida humana tiene valor y merece respeto.

Como ya se ha dicho, delegar las decisiones sobre la vida y la muerte en máquinas que no pueden apreciar plenamente el valor de la vida humana socavaría la dignidad humana.

¿Quién es responsable por las acciones de estas máquinas?

Tanto el derecho internacional humanitario como el derecho internacional de los derechos humanos exigen responsabilidad individual por los actos ilícitos. Dicha responsabilidad personal ayuda a disuadir de futuras violaciones, a la vez que proporciona retribución a las víctimas de daños pasados. Sin embargo, responsabilizar a una persona por los actos ilícitos de un arma totalmente autónoma sería un reto y, en la mayoría de los casos, casi imposible.

Un robot en sí mismo no podría ser considerado responsable según la ley. Los delitos implican tanto un acto (causar la muerte) como un estado mental (la intención). Aunque un arma totalmente autónoma podría cometer el acto, como máquina carecería del estado mental. Además, las armas totalmente autónomas no podrían ser castigadas porque, a diferencia de los humanos, no pueden experimentar sufrimiento.

En la mayoría de los casos, los humanos también podrían escapar de la responsabilidad criminal por las acciones de los robots. Aunque un comandante que desplegara un arma totalmente autónoma con la clara intención de cometer un delito podría ser declarado culpable, sería jurídicamente difícil hacerle responsable de las acciones imprevisibles de una máquina autónoma.

Los programadores y fabricantes probablemente podrían eludir la responsabilidad en una demanda civil:

- En algunos países, como en Estados Unidos, los fabricantes de armas son inmunes a demandas siempre que sigan especificaciones del gobierno y no engañen deliberadamente a las Fuerzas Armadas.
- Adicionalmente, probar que un producto es defectuoso requiere de una investigación en profundidad realizada por expertos, y el acceso completo a los al código y registros de la máquina puede ser extremadamente difícil, ya que en la mayoría de los casos estas tecnologías podrían ser código cerrado.
- Finalmente, las demandas civiles requieren de amplios recursos económicos y de tiempo, especialmente para víctimas que viven lejos de los países donde estas máquinas son fabricadas.

En conclusión, las armas autónomas letales no solamente tienen el potencial de violar la ley internacional, sino que también podrían permitir que los comandantes, operadores, programadores y fabricantes escapen de la responsabilidad de las violaciones que puedan ocurrir.

¿Cuál es el futuro de las armas autónomas letales, y cómo afectaría a la seguridad global?

Las preocupaciones alrededor de la implementación de sistemas de armas autónomas son varias, y tienen que ver mayormente con dos cuestiones principales. En primera instancia, la velocidad: la rapidez con la cual este tipo de armas puede actuar y además ser distribuida alrededor del mundo, avizora graves consecuencias de generación de conflictos accidentales.

La segunda tiene que ver con que, en los nuevos conflictos, las armas autónomas son utilizadas para generar terror: opresión automatizada, uso por parte de actores no estatales, venta en mercados ilegales y ciber ataques a sistemas autónomos armados, son algunos riesgos que debemos tener en cuenta a la hora de pensar en qué cantidad de recursos estamos brindando como sociedad a la creación de armas automatizadas.

El Comité Internacional para el Control de Armas Robóticas (ICRAC) cita los siguientes aspectos como algunos de los más preocupantes cuando hablamos de los riesgos de la implementación de armas autónomas letales para la seguridad global:

- 1. Proliferación:** sin un freno internacional al desarrollo, pruebas y producción de sistemas de armas autónomas, es probable que veamos una proliferación masiva de estas armas y contra-armas, y así sucesivamente. Es importante destacar que no todas las naciones tendrán la capacidad de llevar a cabo las revisiones de los sistemas de armas autónomas que exige el derecho internacional.
- 2. Aceleración del ritmo de batalla:** la aceleración del ritmo de batalla es un fenómeno que ocurre inclusive actualmente, al punto que los procesos de decisión humanos involucrados no son lo suficientemente rápidos. Nuevos prototipos de armas autónomas letales aéreas están siendo probados, y estos sistemas alcanzan velocidades de vuelo hipersónicas. No es difícil notar que esto puede disminuir el control que los humanos tienen sobre el campo de batalla.
- 3. Conflicto accidental:** si el desarrollo y proliferación de armas continúa, los sistemas de defensa de un Estado podrían interactuar con sistemas de ofensiva igualmente rápidos de otro Estado. La velocidad de esta interacción impredecible podría desencadenar en un conflicto armado inintencionado antes de que los humanos tengan tiempo de reaccionar.
- 4. Militarización del mundo civil:** ya estamos viendo el uso de las nuevas tecnologías bélicas no tripuladas en entornos civiles. Las fuerzas represivas y las agencias de control de fronteras están utilizando sistemas no tripulados para la vigilancia. Algunas empresas incluso los están armando con pistolas eléctricas, sprays de pimienta y otras municiones denominadas “menos letales”. Con la tecnología de objetivos autónomos, esto podría dar lugar a violaciones de los derechos humanos y civiles por parte de la policía y las fuerzas de seguridad privadas con poca posibilidad de rendir cuentas.

5. **Opresión automatizada:** los sistemas de armas autónomas podrían ser una herramienta atractiva para la opresión de poblaciones y la supresión de protestas pacíficas y cambio político. Mientras que los soldados y oficiales policiales pueden, en principio, negarse a dirigir sus armas contra su propia gente, los sistemas de armas autónomas serían programados por personas lejos de los enfrentamientos y entonces podrían matar sin piedad en base a sus instrucciones codificadas.
6. **Actores no estatales:** actualmente estamos presenciando una difusión tecnológica sin precedentes. El coste del desarrollo de la robótica está disminuyendo y el hardware necesario está ya ampliamente disponible. Si se permite que el desarrollo de armas autónomas continúe, no pasará mucho tiempo antes de que veamos copias burdas o exportaciones del mercado gris en manos de actores armados no estatales.
7. **Vulnerabilidades cibernéticas:** el riesgo de errores de codificación del software, el mal funcionamiento, la degradación de las comunicaciones y, sobre todo, los ciberataques del enemigo, las infiltraciones en la cadena de suministro industrial, las interferencias y la suplantación de identidad hacen que los sistemas de armas autónomas sean intrínsecamente inseguros.

¿Qué podemos hacer para apoyar la campaña “Stop Killer Robots”?

La campaña “Stop Killer Robots” fue fundada por diversos organismos internacionales y de la sociedad civil para dar una respuesta coordinada ante esta problemática. La iniciativa comenzó a partir de un llamado de atención realizado por científicos expertos en robótica e inteligencia artificial, donde dieron alarma ante la posibilidad de la creación de sistemas de armas que podrían seleccionar y atacar objetivos sin intervención humana.

El rol de la comunidad científica ha sido central en el debate y activismo alrededor de los robots asesinos desde el inicio. Estas son las personas que tienen un mayor entendimiento de las tecnologías de inteligencia artificial y las implicancias de utilizarla para determinar objetivos en las armas autónomas. Estos conocimientos son vitales para dar alerta temprana ante situaciones en los ámbitos industriales, académicos y de investigación que podrían dar pie a la generación de tecnologías de ataque automático.

En 2015 en la Conferencia Internacional de Inteligencia Artificial que se realizó en Buenos Aires, se hizo pública una carta¹⁷ firmada por miles de científicos y expertos en inteligencia artificial y tecnología, contra el desarrollo de armas que sean autónomas y que operen sin intervención humana; entre los firmantes se encuentran Elon Musk, Stephen Hawking y Steve Wosniak.

Uno de los ejemplos más recientes del poder que puede tener la comunidad técnica como agente de reflexión y cambio se dio en el ámbito industrial en 2018, con el inicio de colaboraciones entre Google y el Departamento de Defensa de los Estados Unidos: el Proyecto Maven, pretende generar e integrar algoritmos de visión artificial para apoyar a analistas militares y civiles en la clasificación de enormes volúmenes de datos de vídeo¹⁸. Estos datos son recogidos por el Departamento de Defensa en apoyo de las operaciones de contraintersurveillance y contraterrorismo utilizando drones.

En respuesta a esta alianza, más de 3 mil empleados de Google firmaron una carta de protesta, en la que expresaron su preocupación con el involucramiento de la empresa en temas de guerra, y determinaron que no podían delegar la responsabilidad moral de sus tecnologías a terceros. Gracias a esta iniciativa, Google no renovó su contrato con el Proyecto Maven para el siguiente año¹⁹. Esta protesta además propició la reformulación de los principios del uso de inteligencia artificial de la compañía²⁰.

Este tipo de iniciativa no se ha dado solamente en el ámbito industrial y corporativo. Las universidades tienen un papel extremadamente importante en la construcción de la sociedad. Muchas innovaciones utilizadas en el día a día, como los cinturones de seguridad o las pantallas táctiles, vienen de la investigación universitaria, ilustrando los impactos positivos y aplicaciones que la investigación universitaria puede brindar a la sociedad. Es importante destacar que la financiación que tienen estas investigaciones viene tanto del sector industrial, en específico la industria armamentista, como de entes estatales (ministerios de defensa).

Aunque la colaboración con el sector militar no es necesariamente problemática de por sí, es crucial comprender que las investigaciones en ciencia de la computación, inteligencia artificial y robótica tienen una influencia clave en el desarrollo de la industria armamentista, que está cambiando rápidamente la naturaleza de la guerra y puede presentar amenazas a la paz y seguridad internacional.

El diseño del futuro de la guerra ha sido liderado en varias instancias por los centros de investigación universitarios a lo largo de la historia. Un ejemplo emblemático de este fenómeno es el Proyecto Manhattan, donde el trabajo de numerosas universidades estadounidenses fue clave para el desarrollo de la primera arma nuclear. Hasta el 2011, en Estados Unidos, el Pentágono fue el tercer patrocinador más grande de toda la investigación académica del país²¹, y el mayor patrocinador de las investigaciones relacionadas a las ciencias físicas, matemáticas, ingeniería y ciencias de la computación. Instancias similares se reproducen en dos de los países más grandes en términos de producción de armas: Reino Unido²² y China²³.

La colaboración entre universidades y el sector militar es un tema polémico que ha generado muchas dudas y reclamos por parte de la comunidad universitaria a lo largo de los años. La resistencia estudiantil y del sector académico históricamente han sido agentes de cambio en cuanto a las condiciones de estas colaboraciones, y específicamente en los últimos años con respecto al desarrollo de armas autónomas:

- En el año 2018, el Instituto Avanzado de Ciencia y Tecnología de Corea abrió un centro de investigación junto a la empresa Hanwha Systems, con el objetivo de llevar adelante estudios acerca de cómo las tecnologías de la cuarta revolución industrial podrían ser utilizadas para el futuro del campo de batalla²⁴. Este anuncio dio lugar a un boicot organizado donde más de 50 investigadores de IA y robótica declararon a través de una carta científica que “boicotearían a la universidad KAIST de Corea del Sur por los planes del instituto de ayudar a desarrollar armas impulsadas por IA”²⁵. Como resultado de esta iniciativa, la universidad de KAIST cuenta hoy en día con un código de ética para la implementación de sistemas de Inteligencia Artificial²⁶.
- En Bélgica, 116 científicos del área de IA y robótica, en su mayoría profesores e investigadores universitarios, publicaron una carta abierta expresando una gran preocupación por el desarrollo de sistemas de armas letales que carecen de control humano significativo sobre las funciones de selección de objetivos y lanzamiento de ataques²⁷. La carta fue publicada en el mismo día en que el parlamento belga mantuvo su primera audiencia sobre armas autónomas, y la atención mediática fue fundamental para la resolución unánime de prohibir el desarrollo de este tipo de tecnología²⁸.

Cartas científicas abiertas y el rol de la prensa

Esta sección está basada en el artículo "How to do a scientist letter" de Daan Kayser (PAX), publicado en el "Campaign to Stop Killer Robots - Campaigner's Kit".

A partir de lo visto en la sección anterior, es fácil destacar que el activismo por parte de la comunidad científica y universitaria ha tenido un rol central en el esfuerzo colectivo para el desarrollo de tecnología alineada con los principios básicos de humanidad. Las cartas científicas abiertas son una de las herramientas principales con las que cuenta este sector: han sido útiles para detener instancias de desarrollo, implementación y adquisición de tecnología de guerra autónoma. Además, cuentan con un potencial catalizador para la creación de códigos de ética sobre el uso de las tecnologías de inteligencia artificial para universidades y centros de desarrollo.

Estas cartas pueden además influenciar a los tomadores de decisiones de instancias estatales para generar acciones hacia la prohibición, no sólo del desarrollo de armas autónomas, sino también de la adquisición o implementación de tecnologías de vigilancia automatizadas, a partir de las preocupaciones de la comunidad científica y la atracción mediática que pueden generar.

En su artículo "How to do a scientist letter", Daan Kayser (PAX) cita los siguientes pasos para la publicación de una carta científica:

1. Encontrar un buen momento para publicar la carta: es importante poder enlazar la carta con el contexto local y temporal. Eventos como la discusión en el gobierno sobre el desarrollo y adquisición de armas autónomas letales y tecnologías de vigilancia, o alianzas entre el sector académico y militar con el fin de desarrollar este tipo de tecnologías pueden ser buenos momentos. Esto es además importante ya que puede facilitar alianzas con medios de prensa.
2. Desarrollar el texto de la carta: debe poner en contexto al lector acerca de la problemática, especificando claramente la definición de la tecnología que se busca cuestionar, las consecuencias negativas de implementación y un claro llamado a la acción hacia el debate informado. Es interesante además referir a otras cartas científicas de este estilo, demostrando así el esfuerzo como parte de un movimiento internacional. Además, hay que hacer referencia las implementaciones positivas de la Inteligencia Artificial, teniendo en cuenta el posible rechazo de la comunidad científica a firmar una carta que sólo mencione aspectos negativos de su campo de investigación.
3. Crear un formulario de firmas: el uso de formularios online para la firma de esta carta facilita la adhesión tanto de miembros de la comunidad científica (firmatarios de la "primera ola") como del público general.
4. Identificar y contactar a los representantes de la comunidad científica: es interesante identificar y contactar personas reconocidas dentro de la comunidad científica local. Esto tanto para la firma de la carta como para que puedan contactar a otros miembros de la comunidad.

5. Abrir la carta para firmas: es importante contar con la mayor cantidad de firmas posibles de miembros representantes de la comunidad científica y expertos tanto nacionales como internacionales. Tanto la Campaña “Stop Killer Robots”²⁹ como ICRC³⁰ reciben solicitudes de apoyo para la firma de este tipo de documento, y ambas organizaciones cuentan con una importante base de datos de expertos aliados a la campaña a nivel internacional.
6. Publicación de la carta y alianza con medios: la publicación debe realizarse con el objetivo de llegar a la mayor cantidad de personas y buscar la atención de los medios de prensa. El uso de comunicados de prensa, mencionando el momento político que se está intentando influenciar, y la generación de eventos públicos de protesta son cruciales para lograr este cometido. Es importante asegurar el envío de copias de esta carta a los tomadores de decisiones que se está intentando influenciar.

Referencias

1. Lawson, Kimberly. (2019). Robotics within Military Applications. AZoRobotics. <https://www.azorobotics.com/Article.aspx?ArticleID=199>
2. Rodríguez Joaquín, Xavi Mojar, Tica Font, Pere Brunet (2019). Nuevas armas contra la ética y las personas. Drones armados y autónomos. Centre Delàs d'Estudis per la Pau. http://arxiu.centredelas.org/images/INFORMES_i_altres_PDF/informe39_DronesArmados_CAST_web_DEF.pdf
3. Reimann, Matt (2015). A brief history of robotics in literature: <https://blog.bookstell-youwhy.com/a-brief-history-of-robots-in-literature>
4. Qureshi, Asim (2017). 10 Ancient robots that were built without modern technology: <https://wonderfulengineering.com/10-amazing-robots-no-modern-technology/>
5. Bradford, Alina (2017) - The Five (and More) Senses: <https://www.livescience.com/60752-human-senses.html>
6. Miller, Sara (2016) - A sixth sense? it's in your genes: <https://www.livescience.com/56223-sixth-sense-genes.html>
7. Página web oficial - Boston Dynamics: <https://www.bostondynamics.com/atlas>
8. (2007) "KUKA Robotics introduces world's largest and strongest robot", Industrial Robot, Vol. 34 No. 6. <https://doi.org/10.1108/ir.2007.04934fad.001>
9. Tamborini, Marco (2021), The Material Turn in the Study of Form: From Bio-Inspired Robots to Robotics-Inspired Morphology. Perspectives on Science; 29 (5): 643–665. doi: https://doi.org/10.1162/posc_a_00388
10. Lanfranco, A. R., Castellanos, A. E., Desai, J. P., & Meyers, W. C (2004).. Robotic surgery: a current perspective. Annals of surgery, 239(1), 14–21. <https://doi.org/10.1097/O1.sla.0000103020.19595.7d>
11. HWY18MH010 - Preliminary Report - National Transportation Safety Board: <https://www.technischweekblad.nl/files/46e4a0ca1a0bf2cb04f4751doaf26913.pdf>
12. Ross, Casey. Swetlitz, Ike (2017). IBM's Watson supercomputer recommended 'unsafe and incorrect' cancer treatments, internal documents show: <https://www.statnews.com/2018/07/25/ibm-watson-recommended-unsafe-incorrect-treatments/>
13. Página web oficial - Faception: <https://www.faception.com/>
14. Lor, Steve (2018). Facial Recognition Works, if you are a white guy. Publicado en The New York Times: <https://www.nytimes.com/2018/02/09/technology/facial-recognition-race-artificial-intelligence.html>

15. Simonite, Tom (2018). AI Is the Future—But Where Are the Women?. WIRED. <https://www.wired.com/story/artificial-intelligence-researchers-gender-imbalance/>
16. Campaign to Stop Killer Robots, “Resultados de una encuesta mundial muestran que el 61% se opone a los robots asesinos,” Enero 2019, <https://www.stop-killerrobots.org/2019/01/global-poll-61-oppose-killer-robots/>
17. (2015) Autonomous Weapons: an Open Letter from AI & Robotics Researchers <https://futureoflife.org/2016/02/09/open-letter-autonomous-weapons-ai-robotics/#signatories>
18. Pellerin, Cheryl (2017). Project Maven to Deploy Computer Algorithms to War Zone by Year’s End <https://www.defense.gov/News/News-Stories/Article/Article/1254719/project-maven-to-deploy-computer-algorithms-to-war-zone-by-years-end/>
19. Cogner, Kate (2017). Google Plans Not to Renew Its Contract for Project Maven, a Controversial Pentagon Drone AI Imaging Program: <https://gizmodo.com/google-plans-not-to-renew-its-contract-for-project-mave-1826488620>
20. D’Onfro, J. (2018). Google promises not to use A.I. for weapons or surveillance, for the most part. <https://www.cnbc.com/2018/06/07/google-ai-ethical-principles.html>
21. Ghoshroy, Subrata (2011). Fact Sheet: The Pentagon and the Universities. <http://archive.demilitarize.org/enfact-sheet-pentagon-universities/>
22. Burt, Peter (2018). “Off the Leash. The development of autonomous military drones in the UK”, Drone Wars: <https://dronewarsuk.files.wordpress.com/2018/11/dw-leash-web.pdf>
23. Kania, Elsa (2018). “Tsinghua’s Approach to Military-Civil Fusion in Artificial Intelligence”, Centre for New American Security: <https://www.cnas.org/publications/commentary/tsinghuas-approach-to-military-civil-fusion-in-artificial-intelligence>
24. Jun Ji-hye (2018), “Hanwha, KAIST to develop AI weapons”, Korea Times: https://www.koreatimes.co.kr/www/tech/2018/12/133_244641.html
25. Vincent, James (2018), “Leading AI researchers threaten Korean university with boycott over its work on ‘killer robots’”, The Verge: <https://www.theverge.com/2018/4/4/17196818/ai-boycot-killer-robots-kaist-university-hanwha>
26. Página oficial - Kaist Institute for Artificial Intelligence: https://kis.kaist.ac.kr/index.php?mid=KIAI_O
27. Carta científica Belga sobre armas autónomas: <https://docs.google.com/document/d/e/2PACX-1vQU8W-mpdjBqLHIA4Xgbe1BhKI4sc-m2UyQg3cPpylpjnOVF81OmPSE7QmzaXNDfqBeLGrNFS4ozRL8-/pub>

28. PAX (2018), "Belgium votes to ban killer robots": <https://paxforpeace.nl/news/overview/belgium-votes-to-ban-killer-robots>
29. Página oficial de la Campaña Stop Killer Robots: <https://www.stopkillerrobots.org/>
30. Página oficial del "International Committee for Robot Arms Control": <https://www.icrac.net/>



Esta obra está bajo una
Licencia Creative Commons
Atribución-CompartirIgual 4.0
Internacional.

