https://doi.org/10.18800/iusetveritas.202102.009

Sistemas de armas autónomas y DIH. Advertencia de un futuro cercano(*)

Autonomous weapons systems and IHL. Warning of the near future

Melissa Macher Reyes(**)
Universidad de Leiden (Leiden, Países Bajos)

Resumen: La Inteligencia Artificial y los sistemas de aprendizaje automático están revolucionando no solamente la forma en la que vivimos y nos relacionamos, sino también la manera en la que aplicamos este tipo de tecnología en conflictos bélicos. Si bien es cierto que, en muchos casos la idea de Inteligencia Artificial puede presentarse como un concepto muy remoto, sistemas de aprendizaje automático incluso en sus niveles de desarrollo actual, pueden presentar una amenaza más real que toda esta tecnología desarrollada y potencial. Este artículo busca socializar la información relacionada al potencial actual de los sistemas de aprendizaje automático y analizar en qué medida se encuentran estos avances alineados con las normas del Derecho Internacional Humanitario. A pesar de que el desarrollo de esta tecnología ocurre en mayor escala en ámbitos corporativos, la carrera por militarizar esos hallazgos existe, no es ajena a nuestra realidad y está más presente de lo que imaginamos.

Palabras clave: Derecho Internacional Humanitario - Inteligencia Artificial - Aprendizaje Automático - Información como arma de guerra

Abstract: Artificial Intelligence and machine learning systems are revolutionizing not only the way we live and interact but also the way we apply this type of technology in war conflicts. While it is true that in many cases the idea of Artificial Intelligence can be presented as a very remote concept, machine learning systems, even at their current levels of development, can present a more real threat than all this developed and potential technology. This article seeks to socialize the information related to the current potential of machine learning systems and analyze to what extent these advances are aligned with the norms of International Humanitarian Law. Although the development of this technology occurs on a larger scale in corporate settings, the race to militarize these findings exists is not alien to our reality and is more present than we imagine.

Keywords: International Humanitarian Law - Artificial Intelligence - Machine Learning - Information as a weapon of war

(*) Nota del Equipo Editorial: Este artículo fue recibido el 4 de septiembre de 2021 y su publicación fue aprobada el 21 de diciembre de 2021.

(**) LLM en Derecho Internacional Público por la Universidad de Leiden, en Holanda (Países Bajos). Bachiller en Derecho por la Pontificia Universidad Católica del Perú. ORCID: https://orcid.org/0000-0001-6705-4220. Correo electrónico: machermelissa@gmail.com.

1. Introducción

1.1. Antecedentes

Cuando hablamos de armas e Inteligencia Artificial (IA) se nos vienen a la cabeza una serie de imágenes de las cuales nos sentimos alejados temporal y espacialmente. Pensamos en drones capaces de identificar a sus objetivos de manera automática e independiente, o figuras humanoides capaces de procesar grandes cantidades de información en un corto periodo de tiempo como las sacadas de películas futuristas. Estas preconcepciones, cada una intimidante en su propio contexto, tienen en común que la Inteligencia Artificial, es sin duda capaz de superar infinitamente el proceso cognitivo de los seres humanos.

¿Pero qué pasa si este tipo de arma no es en la que debería atemorizarnos más? La Inteligencia Artificial está mucho más cercana y accesible a nosotros de lo que imaginamos cuando evocamos drones o figuras humanoides. Actualmente usamos Inteligencia Artificial en muchos ámbitos cotidianos que seguramente han pasado desapercibidos como un motor de búsqueda en internet o el asistente personal en nuestros celulares. De igual forma, se viene desarrollando esta tecnología en medicina, finanzas, movilidad, traducciones y como veremos detenidamente, en defensa estatal.

Es una realidad que esta tecnología viene investigándose, desarrollándose y aplicándose por distintos estados y en contextos militarizados. Este escenario toma relevancia cuando estas herramientas son aplicadas por estos mismos estados con la finalidad garantizar su labor de defensa y seguridad, principalmente en el ámbito regulado por el Derecho Internacional Humanitario (DIH). Es el caso por ejemplo en Sudamérica, en donde venimos experimentado un movimiento social y gubernamental durante los últimos años, presenciando una decodificación de la realidad que da paso a grandes cambios de paradigma, normalmente por motivos políticos, sociales, étnicos y/o religiosos.

Es el caso de las revueltas de Chile, las protestas en Perú, Colombia y Bolivia y los casos de violencia en Cuba durante el 2020 y el 2021. En un momento en que el apoyo público a la democracia en Latinoamérica está flaqueando, la IA podría ayudar al rescate. Los órganos del Congreso de toda la región podrían utilizar la inteligencia artificial para impulsar la transparencia y los aportes del proceso legislativo. De hecho, Hacker Laboratory, un laboratorio de innovación dentro de la Cámara de Diputados de Brasil, está utilizando plataformas de inteligencia artificial para facilitar las interacciones entre legisladores y ciudadanos.

> La esfera digital y el desarrollo de un espacio público digital y animado por inmensas plataformas de redes sociales, a su vez dirigidas por algoritmos que unen a los individuos por enlaces de afinidad y centros de intereses comunes, impactan en los movimientos sociales en varios niveles. El espacio digital y las redes sociales en particular (Facebook, Instagram, Twitter, WhatsApp, TikTok) multiplican las interacciones de los movimientos y sus miembros sin intermediarios, en todos los niveles de la sociedad y los territorios (local, nacional, global), y los conectan con una serie de individuos no conocidos, comprometidos y organizados anteriormente, pero que pueden encontrarse de acuerdo con sus valores, demandas y resistencias. Al permitir que los movimientos difundan su mensaje a millones de personas, que son tanto medios como actores en sí mismos, y que convoquen a movilizaciones callejeras sin la necesidad de recursos financieros o de medios tradicionales (Ventura & Billion, 2020).

Lo antes señalado ha generado una infinidad de situaciones de violencia, en las cuales cada Estado tiene un rol, que incluye velar por la seguridad respetando lo establecido en los Convenios de Ginebra (CG) y las disposiciones inderogables de la Convención Americana sobre Derechos Humanos (CADH) y el Pacto Internacional de Derechos Civiles y Políticos (PIDCP)[1]. En este escenario, utilizar tecnologías que aplican Inteligencia Artificial podría contribuir positivamente a que se tenga un nivel de prevención y control mucho más preciso. Así, se generaría una reducción en los muy comunes excesos en el uso de la fuerza por parte del Estado y/o participantes en las hostilidades, así como a la capacidad de brindar a la población civil una sensación de seguridad más sostenible en el tiempo.

Sin embargo, podemos estar en un supuesto en el que estos objetivos se alcancen en detrimento de no combatientes, como consecuencia de la falta de control que tiene un civil en el uso y manejo de su propia información. El establecimiento de un marco normativo claro y ajustado a la realidad sobre los medios de guerra que funcionan aplicando la Inteligencia Artificial es poco, por no decir nulo. En un contexto de violencia, la línea que divide un uso ‘‘ético’’ y el ‘‘práctico’’ puede no ser tan clara.

La falta de un sistema normativo preciso como consecuencia de su novedad, tanto

(1) Especialmente en lo que respecta al Art. 3 común a todos los Convenios de Ginebra, articulo 27 de CADH y artículo 4 del PIDCP.

en el ámbito internacional como en el nacional, significa que existe un vacío normativo en la aplicación de dicha tecnología a la realidad y, por ende, una oportunidad para abusar de ella ya sea de parte del Estado o de cualquier grupo que participe activamente en las hostilidades. Es más, como veremos más adelante, incluso algunos de los supuestos básicos del DIH se ponen en tela de juicio y las categorías y distinciones fundamentales como ‘‘conflicto armado’’, ‘‘civil’’ y ‘‘objetivo militar’’, no se distinguen con tanta facilidad en este ámbito (Diamond, 2014, p. 70).

1.2. Metodología

Este artículo presentará una perspectiva con respecto a los sistemas de armas autónomas que pueden generarse del uso de la información emitida por civiles no combatientes y a su uso hipotético como arma de guerra en el contexto de un conflicto armado. El objetivo es brindar un entendimiento básico pero exhaustivo de las posibles amenazas que existen en torno a la utilización de esta nueva tecnología.

El artículo está estructurado alrededor de las siguientes nociones: (i) ¿Cuál es el tipo de inteligencia artificial que nos concierne? (ii) ¿Cuáles son las posibles amenazas a las que nos enfrentamos?, y (iii) ¿Cuáles son las violaciones al DIH que se desprenden de este análisis?

El análisis presentado se basa es una revisión multidisciplinaria acerca del desarrollo de tecnologías de Inteligencia Artificial y sistemas de armas autónomos proveniente de fuentes civiles y militares. Asimismo, se sustenta en la recolección de datos relacionados con esta tecnología de fuentes digitales e impresas y su aplicación con la normativa positiva y consuetudinaria del DIH.

1.3. Objetivo

Como explicaremos en detalle más adelante, la Inteligencia Artificial se alimenta y se crea principalmente a base de información. Mientras más información volquemos nosotros sobre estas herramientas, más precisas se vuelven. Muchas de estas tecnologías son capaces de aprender por sí mismas, a través de, entrenamientos, el uso habitual, el *feedback* del usuario y toda la información que se les proporciona en conjunto. Este atributo inherente a la Inteligencia Artificial se conoce como ‘‘aprendizaje automático’’ (De Spiegeleire, 2017, pp. 33-35).

En este articulo analizaremos qué podemos entender realmente por Inteligencia Artificial y cómo el Aprendizaje Automático facilita que esta tecnología se encuentre disponible para ser aplicada en los sistemas de armas autónomas para fines bélicos. Asimismo, analizaremos las potenciales amenazas y riesgos a los que estarían expuestos quienes no participan directamente en las hostilidades, ya sean estas de carácter internacional (CAI) o internas (CANI).

2. La inteligencia artificial como arma de guerra

Para entender realmente qué es la Inteligencia Artificial y cómo puede ser aplicada actualmente en contextos de conflictos armados, es necesario entender su razón de ser y el posible potencial de aplicación en el futuro. Asimismo, es necesario comprender como el Aprendizaje Automático es el hilo conductor de los sistemas de armas autónomas y como se ha convertido en un tema de interés para el DIH.

2.1. ¿Qué es la Inteligencia Artificial?

Todas las herramientas que funcionan luego de haber sido programadas con un algoritmo tienen un grado de Inteligencia Artificial. John McCarthy acuño el término en 1955 y señalaba que esta puede definirse como ‘‘la ciencia y la ingeniería de hacer máquinas inteligentes’’ (Boulanin, 2017, p. 89). Esta puede definirse como la inteligencia no humana que se mide por la habilidad de ‘‘replicar las habilidades mentales propias del ser humano’’. Estas habilidades pueden ser, por ejemplo: el reconocimiento de patrones, el entendimiento del lenguaje natural, el aprendizaje basado en la experiencia, la capacidad de crear estrategias, el razonamiento sobre ciertos preceptos, entre otros (De Spiegeleire, 2017, p. 28).

Muchos autores, reconocen a la Inteligencia Artificial, como la definición paraguas que cubre todo tipo de investigación que busca hacer que las máquinas resuelvan problemas que los humanos calificamos de inteligentes (Boulanin, 2017, p. 90). Dentro de esta definición general, existen distintos tipos de tecnologías que componen varias capas o ‘‘generaciones’’ de lo que la Inteligencia Artificial comprende actualmente. Esta primera clasificación nos va a ayudar a tener claro los tipos de tecnología de los que estamos hablando y cuál es realmente el nivel de aplicación de estos en la realidad (De Spiegeleire, 2017, pp. 12-13).

- Artificial Super Inteligence (ASI) Es la tecnología que supera la inteligencia humana en cualquier ámbito.
- Artificial General Inteligence (AGI) Es la tecnología que equipara en su total capacidad a los humanos en cualquier tipo de ámbito.

- Artificial Narrow Inteligence: (ANI) Es la tecnología que equipara o excede la inteligencia humana para una tarea en específico.

Estos tres tipos de Inteligencia Artificial se encuentran en distintos niveles de desarrollo y cada uno viene con una implicación ética, moral y legal independiente de la otra. Aún nos encontramos lejos de lo que sería entendido como un ASI y un AGI. Esta tecnología es muy parecida a una entidad que -bajo cualquier precepto- piensa de manera superior a la mente humana. Una especie de deidad que nos superaría en todos los ámbitos y que por el momento en encuentra dentro del ámbito de la ciencia ficción (Boulanin, 2017, p. 92). Si bien es cierto que esto parece una realidad utópica, es necesario empezar a pensar en las implicancias legales y éticas, así como considerar medidas para la supervisión, regulación y conducción responsable por parte de quienes la controlen (De Spiegeleire, 2017, p. 16). Por otro lado la tecnología ANI, sobre la cual nos enfocaremos durante este artículo, tiene la potencialidad de mejorar nuestra experiencia humana como consecuencia de sus efectos prácticos, por lo que es muy posible que la tomemos con los brazos abiertos ya que probablemente facilite la vida cotidiana.

Entre todas estas formas de Inteligencia Artificial, existe un factor común que es la capacidad de aprender autónomamente desde la información que se le proporciona. Este tipo de tecnología se llama *Machine Learning* o Aprendizaje Automático y es la característica que permite que este tipo de tecnología evolucione y ‘‘aprenda’’ a procesar la información de forma superior a la humana, basándose en patrones y algoritmos.

2.2. ¿Qué es el Aprendizaje Automático?

El Aprendizaje Automático, como su nombre lo dice, es la capacidad que tiene una máquina para aprender y evolucionar por sí misma. El Aprendizaje Automático, una de las características de la Inteligencia Artificial, hace referencia a una serie de códigos para detectar patrones, aprender y hacer predicciones de acuerdo con la información proporcionada al momento de su programación. Esta particularidad es la que permite que las máquinas aprendan y realicen tareas sin necesidad de programarlas explícitamente (Lewis, 2018, p. 4).

El reciente éxito y desarrollo acelerado de este tipo de tecnología es consecuencia del aumento de la capacidad de las computadoras y el crecimiento en la cantidad de información que tenemos disponible para su alimentación. Conjuntamente estos contribuyen al entrenamiento de los algoritmos de la Inteligencia Artificial. Su efectividad, como ya mencionamos, depende no solo de un algoritmo correcto sino de la cantidad y la calidad de la información proporcionada para su entrenamiento (Lewis, 2018, p. 4).

En mucha mayor escala, el Aprendizaje Automático es cada día más indispensable en las actividades que realizamos, por ejemplo:

- El uso de los asistentes personales en los teléfonos. Estos siguen instrucciones a través del reconocimiento de voz y aprenden a través de conversaciones. De manera más relevante, estos toman muestras de nuestros datos y de muchas más conversaciones para ‘‘aprender’’ a reconocer palabras o distintas formas de pronunciación.
- Los servicios de *streaming* de películas, series o música, usan el aprendizaje automático para hacer recomendaciones, utilizando como herramientas los hábitos de visualización, detectando patrones y gustos.
- En las redes sociales se usan algoritmos para reconocer rostros e imágenes, además de mostrar el tipo de información más a fin con el usuario.
- Con relación a temas médicos, las computadoras pueden analizar imágenes en búsqueda de pistas visuales que indiquen la presencia de afecciones médicas utilizadas para el diagnóstico sanitario (Hauert, 2017).

Las muchas aplicaciones comerciales de la Inteligencia Artificial derivan de la habilidad de la tecnología para encontrar patrones, valores atípicos y soluciones óptimas para grandes cantidades de información. De acuerdo con Lewis, las más comunes son las siguientes:

> Es en estas tecnologías, las que se encuentran incorporadas en el día a día, en las que haremos hincapié. Ya que, por su amplio uso, el cual pasa inadvertido, podrían ser usadas en contra de quienes no participan directamente en las hostilidades (Lewis, 2018, p. 6).

Tabla 1: Funciones comunes de los AWS y sus descripciones

Función	Descripción
Automatización de tareas	Automatizar funciones de rutina y emplear la formación de equipos hombre-máquina para reducir el tiempo y carga del personal (por ejemplo: hacerse cargo funciones administrativas).
Procesar conjuntos de datos grandes o complejos	Permitir el análisis de gran cantidad de datos y fuentes de datos (por ejemplo: reconocimiento de voz e imagen, mapas y vigilancia).
Predecir el comportamiento	Aprender de los datos pasados para anticipar un posible comportamiento futuro (por ejemplo: predicción en tiempos de tráfico).

Función	Descripción
Marcar anomalías o eventos de interés	Identificación de indicadores de problemas potenciales o eventos de interés para crear alertas (por ejemplo: bancos que predicen transacciones fraudulentas).
Etiquetado de datos y corrección de errores	Reconocimiento de contenido y creación de etiquetas para explotar datos eficientemente (por ejemplo: taggeado automático en Facebook).

(Lewis, 2018, p. 6)

2.3. Sistemas de Armas Autónomos

Cuando aplicamos la Inteligencia Artificial (ANI) al ámbito bélico, estamos hablando de Sistemas de Armas Autónomas (AWS[2]). En ese sentido, la clave para entender este tipo de armamento es la palabra "autónomo", la cual hace referencia a un sistema, ya sea de hardware o software capaz de ejecutar una tarea sin intervención humana (Boulanin, 2017, p. 6). En ese sentido, de forma técnica la autonomía de la que hablaremos se entiende como la capacidad de "transformar información del entorno en un plan/acción con propósito" (Boulanin, 2017, p. 7).

El Comité International de la Cruz Roja (CICR) basándose en la naturaleza de las tareas que desempeñan las AWS las define como cualquier tipo de armamento con autonomía en sus funciones criticas; es decir, un arma que sin intervención humana puede seleccionar (buscar o detectar, identificar y rastrear) y atacar (interceptar, aplicar fuerza, neutralizar, dañar o destruir) objetivos (CICR, 2015).

Actualmente, los sistemas militares ya incluyen muchas actividades ''autónomas'', específicamente en las áreas de: movilidad, focalización, inteligencia, interoperabilidad y manejo de la salud (Boulanin, 2017, p. 22). Asimismo, el CICR señala que las AWS en el marco del DIH puede emplearse en tres escenarios (CICR, 2021, pp. 465-468):

A) Robots y Armas Físicas

La primera aplicación implica aumentar el nivel de Inteligencia Artificial en robots y armas físicas; es decir, drones. La principal característica de estas armas es que son capaces de identificar, de manera independiente, sus objetivos sin contar con la participación de un humano en la tarea. Esta tecnología, si bien es una de las más peligrosas puesto que genera muchas implicancias y cuestionamientos con respecto a su aplicación, se encuentra en proceso de desarrollo y en etapas iniciales. Asimismo, en contextos como el de Sudamérica, se encuentra lejos de poder ser desarrollada a un nivel ''confiable'' por alguno de esos gobiernos.

B) Fuente de información

El segundo tipo, sobre el cual vamos a detenernos y desarrollar durante este artículo, hace referencia a los usos de la Inteligencia Artificial como fuente de nuevos medios información y/o en lo que podría considerarse una guerra cibernética[3].

Este tipo de tecnología aplicará los algoritmos diseñados y la información de patrones que estos arrojen a ayudar a los humanos a tomar decisiones más precisas basadas en el análisis de miles de escenarios preprogramados en su algoritmo. Su relación con la información que manejamos de forma diaria la convierte en el mayor factor de riesgo dentro de la aplicación de la Inteligencia Artificial.

En relación con acciones de inteligencia, los AWS son capaces de recopilar datos, a modo de ejemplo:

- Generación de mapas: Pueden ser usados sistemas subacuáticos, vigilancia y reconocimiento aéreo capaces de generar de forma autónoma detalles sobre el entorno. Un ejemplo notable es Shield AI, un Sistema aéreo no tripulado táctico actualmente en desarrollo por Estados Unidos. El Shield AI puede generar mapas tridimensionales (3-D) usando cámaras, láseres y sensores inerciales y ultrasónicos. No requiere pilotaje humano ni GPS (Tucker, 2016).
- Evaluación de amenazas: En este caso, el sistema está programado para evaluar el nivel de riesgo basado en criterios predefinidos. Por ejemplo: el sistema de defensa de antimisiles ''Iron Dome'' de Israel, puede evaluar dónde detonará un misil entrante y sugerir contramedidas dependiendo de la evaluación de la amenaza.
- Analítica de Big Data: el uso de análisis de big data para el reconocimiento de patrones en datos de inteligencia. Los

(2) Autonomous Weapons Systems (AWS), por sus siglas en inglés.
(3) Cabe mencionar que, durante este artículo no se hará referencia a la ''guerra cibernética'. Con relación a este tema, se puede consultar el Manual Tallin creado en 2009 por un grupo de expertos de la OTAN buscando crear un documento de referencia de carácter no obligatorio.

> avances en los algoritmos de aprendizaje automático permiten encontrar correlaciones en conjuntos de inteligencia grandes y potencialmente heterogéneos. Una ilustración reciente de esta capacidad es el supuesto uso de aprendizaje automático algoritmos de EE. UU. para buscar en el sistema global para comunicaciones móviles (GSM) metadatos de 55 millones de usuarios de teléfonos móviles en Pakistán. El algoritmo fue entrenado para localizar a los mensajeros que llevan mensajes entre miembros de al-Qaeda, finalmente ayudó a Estados Unidos a localizar la residencia de Osama Bin Laden (Boulanin, 2017, p. 29).

C) Toma de Decisiones

Este acápite hace referencia al proceso de toma de decisión per se; es decir, dejar que el algoritmo programado tome la decisión en lugar de que sea discreción de un humano. De los tres tipos de aplicaciones de la Inteligencia Artificial en AWS esta es la más lejana puesto que, al no existir un proceso claro y definido que nos ilustre del razonamiento detrás de la ‘‘toma de decisiones’’ de modo que pueda ser interiorizado por un humano, este tipo de tecnología siempre pasará por un filtro previo. Asimismo, muchos expertos señalan que podríamos ver aplicada este tipo de tecnología en no menos de 30 años y que es muy probable que estemos lidiando con los efectos del calentamiento global antes que con AWS (Deeks, 2020).

3. Amenazas

Como es el caso con muchas de las armas que existen, la Inteligencia Artificial, en este momento de su desarrollo, no es una tecnología del todo precisa. Las armas no son del todo predecibles. Un cambio de viento puede afectar la trayectoria de una flecha, una granada puede explotar en cualquier segundo y una bomba lacrimógena puede caer en cualquier parte. La aplicación de la Inteligencia Artificial para garantizar la seguridad del Estado o su defensa, no nos garantiza que el resultado vaya a ser el esperado. Sin importar la cantidad y calidad de información con la que se programe este algoritmo, con el objetivo de cubrir todos los escenarios posibles, no se podrían replicar todos los factores que pueden presentarse de imprevisto en la vida real. Es decir, es difícil predecir como va a funcionar la Inteligencia Artificial en la práctica (Schuller, 2017, p. 410).

A continuación, pasaremos a analizar con exactitud cuáles son los puntos que deberíamos tener en cuenta al momento de analizar los Sistemas de Armas Autónomos y su relación con las normas establecidas por el DIH.

3.1. La información como arma de guerra

La Inteligencia Artificial tiene un gran potencial para transformar nuestras vidas y la forma en la que nos sentimos seguros. En Estados Unidos se dio el caso en el que un policía de Texas durante su turno se presentó a atender dos llamados consecuencia de suicidios. Más tarde, en el mismo turno, fue llamado para poner orden en una fiesta de jóvenes. En ella, terminó sacando su arma de manera innecesaria. Existe ya un sistema creado por Rayid Ghani, de la Universidad de Chicago, que ha mejorado en 12 por ciento las predicciones sobre los oficiales de policía que podrían ser considerados ‘‘riesgosos’’. De esta manera se pueden tomar decisiones más acertadas con relación al personal de seguridad estatal. Idealmente, un sistema de Inteligencia Artificial hubiese sido capaz de prever la situación del policía y recomendar que se elija un policía menos expuesto al estrés para atender la llamada de los jóvenes (The Economist, 2016).

Cuando extrapolamos este ejemplo y pensamos en el potencial militar, la gran cantidad de información que se crea diariamente en diversos sectores, podemos imaginarnos los múltiples usos en los que se podría implementar la Inteligencia Artificial. Entre ellos se encuentra el procesamiento de informaciones complejas para encontrar soluciones óptimas, la predicción de un comportamiento, la identificación de anomalías y eventos de interés, la corrección de errores, entre otros.

Las oportunidades en la aplicación de este tipo de tecnología no han pasado desapercibidas para los países que se encuentran siempre a la vanguardia de la tecnología de guerra. El Departamento de Defensa de Estados Unidos busca capitalizar el rápido desarrollo de la Inteligencia Artificial en aplicaciones militares. Por su lado, el presidente ruso ha declarado que el país que domine la Inteligencia Artificial será quien mantenga el poder mundial. Es el mismo caso con China, que quiere crear una industria de Inteligencia Artificial de 150 mil millones de dólares. Dicha industria será la más avanzada del mundo y pretende obtener una ventaja militar en un ejemplo sin precedentes de cooperación civil-militar (Lewis, 2018, p. 1).

> En Argentina, por ejemplo, la inteligencia artificial se está utilizando para predecir y prepararse para el embarazo adolescente y la deserción escolar, así como para delinear oportunidades comerciales invisibles en los barrios de la ciudad. En Colombia y Uruguay,

> se ha desarrollado software para predecir dónde es probable que ocurran los delitos. En Brasil, la Universidad de São Paulo está desarrollando tecnología de aprendizaje automático que evaluará rápidamente la probabilidad de que los pacientes tengan dengue, zika o chikungunya cuando lleguen a un centro médico (MarshMcLennan, s. f.).

Esta lucha de poder nos lleva a analizar el hipotético caso en el que estas tecnologías sirvan para objetivos más perversos y efectivamente sean usadas en contextos militarizados sin ningún tipo de restricción. La cantidad de información que hemos ido volcando como usuarios aumenta diariamente de manera exponencial. Los últimos datos de la Unión Internacional de Telecomunicaciones (UIT), el organismo especializado de las Naciones Unidas para las tecnologías de la información y la comunicación (TIC), revelan un incremento en el uso de Internet a escala mundial causado por el Covid-19 de 4,100 millones usuarios de Internet en 2019 a 4,900 millones en 2021 (ITU, 2021).

Ahora bien, en un contexto de uso de la fuerza, existe la posibilidad de tomar ‘‘ventaja’’ frente a la contraparte como consecuencia de la calidad y facilidad de acceso a información que se generaría al introducir un sistema de armas autónomas a los usos de la guerra. No es insensato pensar en las consecuencias que podrían generarse si toneladas de información -incluyendo no solamente a la personal sino a la que hace referencia a nuestros procesos mentales y la manera en la que tomamos decisiones- cayera en las manos equivocadas, un gobierno autoritario, un grupo armado, etc.

Se han visto casos en los que herramientas que funcionan con Inteligencia Artificial ha sido utilizada para la manipulación de la población civil con resultados comprobables. Es el caso del famoso escándalo de *Cambridge Analytica*, en donde se desarrolló un algoritmo producto de la información de los perfiles de algunos usuarios de Facebook, sus conexiones de amistad y los resultados de un cuestionario de personalidad. Dicho algoritmo pudo identificar 253 predicciones (la capacidad de un humano para procesar información es de 5 a 7 piezas) relacionadas a las afiliaciones políticas de las personas y a las maneras de poder ser influenciadas más fácilmente a través de las noticias que se mostraban en sus ordenadores (Weller, 2019).

El problema con *Cambridge Analytica* es la manera en la que se usó la información proporcionada por este algoritmo que a primera vista no parece negativa. Un grupo de personas identificadas como ‘‘influenciables’’ fueron alimentadas de información, no necesariamente veraz, que soportaban las aseveraciones del entonces candidato Donald Trump. Si bien no es posible establecer con certeza que los votantes expuestos al algoritmo de *Cambridge Analytica* basaron su decisión de voto en esta información, es preocupante saber que esta esta forma de manipulación fue convincente ya que este grupo de personas compartía un perfil, percibían de manera parecida la realidad y no cuestionaron la información que vieron (Weller, 2019).

Este ejemplo ilustra perfectamente lo que puede pasar cuando se tiene acceso a este tipo de información. Tenemos que ponernos en el caso en el que un Estado, o cualquiera que participe directamente de las hostilidades, puedan ‘‘manipular’’ nuestro pensamiento y alimentarnos de un tipo de información que se inclina para un lado u otro en específico.

Sin ir muy lejos:

> El Gobierno peruano, a través de la secretaria de Gobierno Digital de la Presidencia del Consejo de Ministros, ha puesto a disposición de la ciudadanía una Estrategia Nacional de Inteligencia Artificial (ENIA) correspondiente al periodo 2021-2026 que tiene como finalidad impulsar la investigación, desarrollo y adopción de la IA en el país. Sin embargo, aún no hay una disposición específica sobre IA que regule sus riesgos, responsabilidades o su supervisión. Al no contar con una regulación que limite el desarrollo de la Inteligencia Artificial mientras su uso va en incremento, en un corto o mediano plazo hablaremos de un gran riesgo para la sociedad en general. Hay que tomar en cuenta que la IA en ocasiones funciona mal, es una tecnología en expansión y crecimiento cuyos errores pueden tener consecuencias (Linares, 2021).

3.2. Sesgo en el Aprendizaje Automático

Lo avances en la autonomía de las computadoras las hacen cada vez más capaces de reconocer y entender su entorno; sin embargo, es muy difícil para los programadores representar en un código cual es la verdadera relación entre objetos, personas y situaciones del mundo real. Debemos tener en cuenta que los algoritmos que crean la Inteligencia Artificial son diseñados por humanos, es decir, son factibles de ser creados con errores. De acuerdo con lo señalado por Pratt, existen casos en los que el algoritmo produce resultados sistémicamente dañados debido a suposiciones erróneas en el proceso de Aprendizaje Automático. A esto se le denomina ‘‘Sesgo del Aprendizaje Automático’’.

Los algoritmos pueden tener sesgos incorporados dentro de su programación, porque son creados por individuos (ya sean los programadores o quienes tengan el

interés) que tienen preferencias conscientes o inconscientes que pueden pasar desapercibidas hasta que los algoritmos se utilicen. Incluso cuando los diseñadores no tienen un motivo nefasto, un sesgo puede infiltrarse en la programación del sistema por una situación no prevista con anterioridad (Pratt, 2020).

Existen muchas decisiones en las que inclusive la forma en la que se ha establecido el código no es clara. En Estados Unidos algunos jueces utilizan los *risk assessments*, generados por un software, que indican cuál es la posibilidad de reincidencia de una persona. Estos exámenes son utilizados para el pago de fianzas y muchas veces -controversialmente- para definir las condenas. Sin embargo, se comprobó por un grupo de investigadores que, en Florida, el algoritmo etiquetaba a la gente de raza negra como futuros criminales casi el doble que a la gente de raza blanca (The Economist, 2016).

Es por esto que es de gran una importancia que los ciudadanos conozcan cómo se han tomado las decisiones que los afectan, especialmente si son producto de un algoritmo que puede tener un sinnúmero de factores determinantes en su programación. Estos sesgos pueden inclusive considerarse como ‘‘perfiles raciales’’. Tanto es así que la Unión Europea está considerando otorgar a los ciudadanos afectados por decisiones tomadas por algoritmos, el derecho de proveer una explicación en pro de la transparencia (The Economist, 2016).

3.3. Falta de explicabilidad e imprevisibilidad

Es por esto que una de las barreras más grandes del uso de las AWS es la ‘‘explicabilidad’’ de las decisiones que toma el algoritmo. Esta tecnología procesa datos y es capaz de llegar a conclusiones mucho más rápido que la inteligencia humana, sin embargo, una máquina solo puede superarnos en análisis deductivos mas no en análisis inductivos. Es decir, el resultado arrojado por una máquina en aplicación del algoritmo, el cual ha sido procesado con mucha más velocidad y efectividad que un proceso lógico humano, no puede brindarnos una razón fundamentada del proceso lógico que se ha seguido para tomar esta decisión. Incluso en el caso de que esto sea posible, el proceso de aprendizaje humano va a tomar un tiempo más prolongado.

La Agencia de Proyectos de Investigación Avanzados (DARPA)[4] de los Estados Unidos de Norteamérica tiene un presupuesto para el año 2021 de 3,5 billones de dólares americanos. Actualmente, se encuentra en pleno desarrollo de un programa para generar un sistema de explicabilidad de la Inteligencia Artificial (DARPA, 2021). Esto, dada la necesidad imperante que existe de entender, confiar y manejar de manera efectiva una generación de herramientas artificialmente inteligentes (Turek, 2021).

La incapacidad de explicar las decisiones que se toman y, por lo tanto, convertirlas en predecibles, es una característica que bloquea la aplicación de esta tecnología en la vida real. Un ejemplo que podemos ver actualmente son los automóviles autónomos. Estos automóviles no requieren ningún tipo de instrucción especifica acerca de qué maniobras utilizar. Sin embargo, el problema es que no es claro como este automóvil toma las decisiones (Boulanin, 2017, pp. 27-29). Inclusive para los propios desarrolladores de estas tecnologías es difícil entender cómo se ha tomado la decisión de despistarse o chocarse con una superficie.

Con relación a la Inteligencia Artificial y el software de Aprendizaje Automático, específicamente del tipo desarrollado para el ‘‘reconocimiento automático de objetivos’’ que se viene desarrollando para ser aplicados en sistemas de armas autónomas, la dimensión de imprevisibilidad eleva el nivel de preocupación sobre la falta de explicabilidad en su toma de decisiones y el posible sesgo al que han sido sometidas durante su programación (CICR, 2021, p. 466).

Es preciso preguntarnos: ¿Qué pasaría en el caso de que el sistema señale que es necesario eliminar a cierta persona? Puede ser porque ha identificado que es un combatiente o porque ha identificado a una persona potencialmente riesgosa para los intereses de la contraparte. Debe existir siempre un elemento de ‘‘humanidad’’ que permita tomar la decisión de si se debe actuar o no luego de recibir este tipo de información, sobre todo teniendo en cuenta que no conocemos cuál ha sido el proceso de análisis que ha dado este resultado.

Schuller menciona que quien programa debe tener en cuenta la incertidumbre en la programación del sistema de manera que podamos predecir razonablemente que cumplirá con el DIH a pesar de las complejidades inherentes al combate. Dependiendo del específico características del sistema, esta predicción podría ser simple, imposible o en algún lugar entremedio

(4) Defense Advanced Research Projects Agency, siglas en inglés.

(Schuller, 2017, pp. 411-412). La prueba, debe ser si un humano puede predecir razonablemente que la acción que tomará el AWS cumplirá con el DIH. Por lo tanto:

> Si podemos predecir razonablemente el cumplimiento del DIH, entonces mantenemos el control sin importar el tipo de interacción necesaria en el instante de la acción letal. Pero si no se puede predecir razonablemente que la máquina cumplirá con el DIH, este AWS puede ser ilegalmente autónomo (Schuller, 2017, p. 408).

4. Análisis jurídico

El Derecho Internacional Humanitario es el cuerpo normativo a quien le concierne regular de los conflictos armados, el tipo de armas que se usa y la protección de víctimas tanto en conflictos internos como internacionales. Esta muy relacionado al Derecho Internacional de los Derechos Humanos, puesto que ambos tienen los mismos valores, principalmente el de salvaguardar la dignidad y el respeto de la persona humana pero son aplicados en distintas realidades.

La mayoría de los conflictos modernos son de carácter no internacional, es decir que muchas veces el ‘‘total’’ del aparato legislativo con relación al DIH no es aplicable sino únicamente lo que se entiende como el núcleo duro y las normas de *Ius Cogens*. Una de las propias dificultades de la aplicación del DIH en sí, es que muchas veces los propios estados son quienes niegan la existencia de un conflicto que active la protección ofrecida por el DIH. Adicionalmente, muchas veces las organizaciones internacionales pueden influenciar en la categorización de un conflicto, pero no existe una definición exacta de lo que es un ‘‘conflicto armado’’ ni mucho menos una autoridad encargada de esta labor (Sivakumaran, 2018, p. 508).

Ahora bien, las normas de DIH han sido aplicadas a muchas situaciones y desde que se codificaron han sido capaces de ‘‘adaptarse’’ a las nuevas tecnologías de guerra que han podido desarrollarse durante el siglo pasado. El Protocolo Adicional I a los Convenios de Ginebra, en su artículo 36 establece que los estados tienen la obligación de adecuar cualquier nueva arma o método de guerra a las normas del DIH:

> Artículo 36. Armas nuevas Cuando una Alta Parte contratante estudie, desarrolle, adquiera o adopte una nueva arma, o nuevos medios o métodos de guerra, tendrá la obligación de determinar si su empleo, en ciertas condiciones o en todas las circunstancias, estaría prohibido por el presente Protocolo o por cualquier otra norma de derecho internacional aplicable a esa Alta Parte contratante(5).

Si bien es cierto que esta obligación establece un buen punto de partida para los estados, la realidad es que los métodos de guerra están en una evolución continua y los escenarios beligerantes cambian constantemente. La regulación que brinda el DIH, fue creada teniendo en mente para métodos y medios de guerra que implican el uso del mundo físico, por lo que muchas de sus estipulaciones no encajan perfectamente con el mundo cibernético y/o de codificación. Por ello, es necesario recurrir a las principales regulaciones del DIH, incluyendo las normas consuetudinarias y principios rectores para analizar las posibles violaciones que pueden darse en el marco de un conflicto armado en el cual se apliquen un AWS de manera irrestricta, sin detenernos en la clasificación o naturaleza del conflicto en sí.

4.1. Violación al principio de distinción

El principio de distinción establece que únicamente quienes participan directamente en las hostilidades y los objetivos militares pueden ser objeto de ataques. De esta manera se estableció la prohibición de dirigir acciones bélicas a la población civil(6). En este sentido, el CICR ha establecido que para definir quienes son los que participan directamente en las hostilidades, se debe tomar en cuenta la aplicación combinada de tres requisitos circunstanciales y particulares de cada situación. Estos son en cada acción se debería analizar: el umbral de daño, causalidad directa y nexo beligerante, para hacer una distinción entre las actividades que son consideradas participación directa en las hostilidades y que, como consecuencia, conllevan a la pérdida de protección contra ataques directos (Melzer, 2010, p. 64).

Una de las más grandes preocupaciones que existe en torno a la aplicación de la Inteligencia Artificial en un contexto de conflicto armado es la capacidad para discernir adecuadamente entre quién es un objetivo militar legal y quiénes se encuentran protegidos de cualquier tipo de ataque por no participar activamente en las hostilidades. Es decir, esta AWS debe estar programada para discriminar entre un combatiente y un no combatiente de manera efectiva y

(5) Artículo 49, Protocolo I Adicional a los Convenios de Ginebra de 1949 Relativos a la Protección de las Víctimas de los Conflictos Armados Internacionales.

(6) Principio contenido en el Artículo 3 común a los Convenios de Ginebra, artículos 48 y 51 del Protocolo Adicional I, artículos 4, 13.2 y 14 del Protocolo Adicional II.

en concordancia con el principio de distinción establecido por el DIH. Inclusive, la capacidad de entendimiento de este sistema debería llegar hasta el momento exacto en el que un combatiente se rinde y por lo tanto ya no es considerado un objetivo militar. Como plantea Lewis, es válido cuestionarnos si una máquina va a ser capaz de responder adecuadamente en esta situación y reconocer ciertamente el momento en que un combatiente deja de serlo (Lewis, 2018, p. 14).

No es tan descabellado pensar que una máquina podría representar un riesgo menor para la población civil en comparación a un combatiente humano y que, en algunos contextos, este estándar podría ser relativamente fácil de cumplir. En concordancia con lo mencionado por Schuller, el sistema podría tener una ‘‘predictibilidad’’ suficiente para no caer en ninguna irregularidad en el marco del DIH. Sin embargo, muchas veces es difícil ‘‘programar’’ los múltiples escenarios que pueden darse en la vida real y este estándar de distinción puede verse afectado por el contexto operativo o el entorno en sí.

> El rendimiento técnico y operativo de los sistemas que emplean autonomía e inteligencia artificial podría capturarse y compararse con los sistemas heredados existentes utilizando la toma de decisiones humana para garantizar que se cumpla este criterio ético (Lewis, 2018, p. 15).

A manera de ejemplo, podríamos referirnos a la tecnología de reconocimiento facial, la cual tiene el potencial de ser de mucha ayuda en temas de seguridad pública y establecimiento del orden. Principalmente en casos en los que es posible rastrear e identificar a personas en espacios públicos, lo cual haría muy fácil distinguir entre un objetivo militar y un no combatiente. Sin embargo, no siempre nos encontramos en el escenario ideal. Por ejemplo, un individuo que participa en las hostilidades podría usar la información del rostro de un civil a través de un AWS y utilizarla para acceder a la computadora de esta persona de forma remota, bloquear el acceso a la información o inclusive implantar información falsa. En caso de una persona pública como un político, un alto mando militar, líder de algún movimiento y/o celebridad, esto puede traer consecuencias muy negativas (Walch, 2019).

4.2. Violación al principio de proporcionalidad

El principio de distinción mencionado anteriormente es moderado por el principio de proporcionalidad, ya que la ley reconoce que puedan existir daños o víctimas civiles (‘‘daño colateral’’) en el contexto de un ataque militar siempre y cuando este resultado no sea excesivo en relación con la ventaja militar anticipada del ataque (Turns, 2014, p. 837). Es importante enfatizar que no existe una fórmula matemática para decidir cuál sería, o no, un nivel proporcional de daño colateral. Al igual que en el caso de quienes participan directamente de las hostilidades, este estándar debe analizarse en luz de las circunstancias y del contexto operativo. La decisión de atacar o no a un blanco determinado es tomada por la persona a cargo, quien debe basar su evaluación en la inteligencia e información que razonablemente esté disponible para ella. Por lo tanto, es incorrecto suponer que cualquier daño a civiles en una operación militar constituye *ipso facto* un crimen de guerra (Turns, 2014, p. 838).

De acuerdo con lo señalado por Lewis, el Aprendizaje Automático y la Inteligencia Artificial podrían usarse para disminuir víctimas civiles. Por ejemplo, reduciendo el número de civiles identificados como combatientes, que es una de las mayores causas de violaciones al DIH. La tecnología permite monitorear áreas para definir un momento donde exista un menor daño colateral y así reducir los daños a infraestructura civil en áreas de conflicto o evitar efectos negativos de largo plazo como pérdida de electricidad, agua o suministro de alimentos (Lewis, 2018, p. 32).

Sin embargo, con relación a la proporcionalidad y basándonos nuevamente en el estándar de predictibilidad razonable de Schuller, cuando estamos ante un AWS capaz de procesar y tomar decisiones con una rapidez superior a la humana, ¿Cómo podemos medir anticipadamente la ventaja militar del ataque? Google recientemente retiró su apoyo al ‘‘Proyecto Maven’’ del Departamento de Defensa de Estados Unidos, el cual se encuentra aplicando algoritmos de Inteligencia Artificial para escanear imágenes tomadas por drones y hacer sugerencias para clasificar objetos como personas, edificios o vehículos (Lewis, 2018, p. 1). Esto quiere decir que esta tecnología puede analizar los vídeos y fotos proporcionadas para detectar de manera automática cualquier elemento ‘‘relevante’’ de acuerdo con su entrenamiento. En una carta abierta, los empleados de Google le solicitaron a su empleador que se inmiscuya en el ‘‘negocio de la guerra’’ y que este desarrollo podría generar resultados ‘‘potencialmente letales’’ para la humanidad.

Google, una de las principales desarrolladoras e inversoras en el desarrollo de estas tecnologías nos ha demostrado que existe un lugar para explorar las muchas aplicaciones que se le puede dar a la Inteligencia Artificial. Estas podrían tener

un efecto en la creación de políticas para muchos sectores, desde la educación hasta la medicina. Siempre y cuando sean aplicadas en respeto de las normas y principios internacionales pertinentes. Estos programas, como Proyecto Maven, pueden significar una intervención proactiva mientras se ahorra una gran cantidad de los fondos que son tan escasos, sobre todo en Latinoamérica (De Spiegeleire, 2017, p. 47). Ahora bien, es necesario preguntarnos si es que las posibles consecuencias del uso de esta tecnología son proporcionales o no con la ventaja militar que el Departamento de Defensa de Estados Unidos obtendría.

4.3. Prohibición de ataques indiscriminados

El Protocolo Adicional I, en su artículo 51.4 establece la obligación de evitar ataques indiscriminados, es decir aquellos que no están dirigidos contra un objetivo militar concreto y/o cuyos efectos no sean posibles de limitarse de acuerdo con lo establecido en las normas de DIH. Si bien es cierto esta estipulación fue pensada para restringir ataques físicos (como la posible destrucción causada por una bomba arrojada en una ciudad), puede extrapolarse para incluir ataques que se den en el marco de una AWS.

Hablemos por ejemplo de la desinformación que puede ser usada como un método de guerra, es decir: la creación y difusión de información falsa con la intención de engañar o manipular como una estrategia de acción(7).

Con relación a la ‘‘predictibilidad’’ suficiente introducida por Schuller en estos contextos, en los que es prácticamente imposible hacer una distinción entre combatientes y no combatientes, la producción de información falsa; ya sea en textos, audios, fotos o video; se hace cada vez más difícil de distinguir con la información real (CICR, 2021, p. 467). Este escenario se agrava en casos como el sudamericano por el analfabetismo y el hecho de que el pensamiento crítico hacia el contenido digital es prácticamente inexistente. De esta manera, es imposible predecir cual será el alcance que podrá tener esta acción dentro de la población civil y los efectos en podrían ser devastadores.

El uso de estas estrategias en una situación de conflicto interno o externo solo amplifica los métodos arcaicos de propaganda con la finalidad de manipular la opinión pública, influenciar las decisiones de no combatientes, aumentar la tensión, promover la violencia y extender los conflictos. En este contexto, para el CICR existe una preocupación latente de que los civiles, como resultado de la desinformación digital, sean sujetos de arrestos ilegales, de tratos inhumanos, de discriminación, sean negados del acceso a los servicios básicos o que sufran de ataques personales o a su propiedad privada (CICR, 2019, p. 9).

4.4. Atribución de Responsabilidad

Lo mencionado anteriormente, además de todas las dudas que se pudieran levantar con respecto a las características intrínsecas de la tecnología, nos deja con otra incógnita: ¿quién es el responsable por la decisión tomada por un algoritmo? En un contexto de guerra, si un combatiente viola lo establecido por el DIH y comete un crimen de guerra, los estándares básicos internacionales dictan que este combatiente sea acusado de ese crimen y procesado de acuerdo con lo establecido por ley. En el caso de una máquina que ha violado el DIH, ¿quién puede ser responsable?

Dependiendo del tipo de tecnología que se utilice, existen distintas posibles respuestas a esta pregunta. En un caso más utópico de aplicación de un AWS, ¿la responsabilidad debería ser asumida por el programador, por las autoridades civiles que decidieron desplegar el sistema o por el comandante que eligió confiar en la máquina en esa operación en particular? Esto genera una ‘‘brecha’’ en la responsabilidad o la rendición de cuentas y podría significar un dilema al momento de tomar la decisión de insertar un sistema de este tipo en las hostilidades (Lewis, 2018, p. 13). Adicionalmente, en el ámbito internacional, se ha establecido un estándar mínimo aplicable a los AWS, las cuales requieren ‘‘control humano significativo’’ en sus funciones críticas, lo cual permitirá que siempre exista un filtro humano quien decida la acción militar, basándose en los principios de proporcionalidad y necesidad militar.

Sin embargo, en un ejemplo que debería considerarse más alarmante, deberíamos detenernos a pensar qué sucede en el caso de que el AWS sea la información. ¿Cómo identificamos a la persona que difunde la desinformación y/o discurso de odio? Asimismo, en un contexto en el que no es fácil distinguir combatientes y no combatientes, ¿cuál es la efectividad de contramedidas o alertas que previenen que la desinformación llegue a un nivel de descontrol generalizado? La pregunta más importante sigue siendo: ¿quién tiene la responsabilidad legal en estos casos? (CICR, 2019, p. 9).

(7) Es importante recalcar que no todo el tipo de información falsa es creada por herramientas de aprendizaje automático o inteligencia artificial.

5. Conclusión

En una realidad en la que los humanos estamos produciendo una cantidad de información nunca vista, no podemos dejar de ser conscientes de que estamos entregándola a una entidad, empresa, organización u gobierno. No es descabellado pensar que puede ser utilizada en nuestra contra, de ser el caso que caiga en manos de malos administradores. Peor aún, que las mismas instituciones en las que hemos depositado nuestra confianza sean las responsables de manipular nuestra forma de pensar con la finalidad de cumplir sus propias metas.

En el contexto del DIH, los múltiples escenarios en los que podemos vernos envueltos y estar expuesto a una máquina o sistema, se multiplican y muchas veces pueden hacernos participes de una ‘‘noción” con el que no estamos realmente alineados. ¿Son los principios del DIH lo suficientemente exactos para proteger a los no combatientes en estos casos? ¿Es necesaria una actualización en el sistema?

Aplicar las normas existentes en el DIH a los nuevos usos y métodos de guerra genera dudas relacionadas en la idoneidad de estos aparatos normativos. Se ha estudiado por varios escolares la falta de definiciones claras en lo que respecta a conflictos armados, como estas reglas puede manipularse para solapar conductas que deberían ser sin duda consideradas crímenes de guerra y que muchas veces han puesto en tela de juicio la efectividad de todo el sistema normativo. Por lo que es pertinente analizar si realmente estas reglas son lo suficientemente claras con relación a los nuevos desarrollos bélicos y lo concerniente a Inteligencia Artificial. De no ser el caso, estaríamos ante una carrera en búsqueda de normar un escenario que viene cambiando constantemente y cuyo límite, es literalmente la imaginación del hombre.

Asimismo, debemos tener la certeza de que las AWS son lo suficientemente confiables como para poder delegar en ellas la toma de decisiones que influyen en el tratamiento de los civiles en tiempos de guerra, por ejemplo, en relación con la calidad o veracidad de la información sobre las que se han desarrollado los algoritmos. Asimismo, esta confianza generada debe circunscribirse al respeto a los principios rectores del DIH y sus normas consuetudinarias. Debemos tomar en cuenta las líneas de pensamiento que se encuentran completamente en contra de la utilización de esta tecnología en las hostilidades y las que, por otro lado, consideran que es posible llegar a un estándar mínimo aceptable para aplicar las AWS a los conflictos bélicos, de manera que sean más sus aspectos positivos que los negativos.

Referencias bibliográficas

Boulanin, V. (2017). *Mapping the Development of Autonomy in Weapon Systems.* SIPRI.

CICR. (2014, 1 de enero). *Los Convenios de Ginebra de 1949 y sus Protocolos adicionales.* https://www.icrc.org/es/document/los-convenios-de-ginebra-de-1949-y-sus-protocolos-adicionales

CICR. (2015). Autonomoys weapon systems: is it morally acceptable for a machine to make life and death decisions?. *CCW Meeting of Experts on Lethal Autonomous Weapons Systems* (pp. 13-17). Geneva.

CICR. (2019). *Symposium Report: Digital Risks in Situations of Armed Conflict.* www.icrc.org/en/event/digital-risks-symposium

CICR. (2021). *Artificial intelligence and machine learning in armed conflict: A human-centred approach.* International Review of the Red Cross. https://international-review.icrc.org/sites/default/files/reviews-pdf/2021-03/ai-and-machine-learning-in-armed-conflict-a-human-centred-approach-913.pdf

DARPA. (2021). Defense Advanced Research Project Agency. https://www.darpa.mil/about-us/budget

De Spiegeleire, S. (2017). Artificial Intelligence and the Future of Defense: Strategic Implications for Small-and-Medium-sized Force Providers. *The Hague Centre for Strategic Studies Report*.

Deeks, A. (2020). Ashley Deeks on AI and the Laws of War [episodio de podcast]. En *Podcast Jib Jab The Laws of War.* https://soundcloud.com/cxm-444054529/jib-jab-podcast-ep-5-ai-and-the-laws-of-war

Diamond, E. (2014). *Applying International Humanitarian Law to Cyber Warfare.* https://www.inss.org.il/he/wp-content/uploads/sites/2/systemfiles/SystemFiles/05%20Applying.pdf

Hauert, S. (2017, 25 de abril). *Eight ways intelligent machines are already in your life.* BBC. https://www.bbc.com/news/uk-39657382

Iron Dome System and SkyHunter Missile. (s. f.). Raytheon Missiles and Defense. https://www.raytheonmissilesanddefense.com/capabilities/products/irondome

ITU. (2021, 30 de noviembre). *Union Internacional de Telecomunicaciones* . Obtenido de 2.900 millones de personas siguen careciendo de conexión. https://www.itu.int/es/mediacentre/Pages/PR-2021-11-29-FactsFigures.aspx

Lewis, L. (2018). *AI and Autonomy in War: Understanding and Mitigating Risks.* Center for Autonomy and AI. https://www.cna.org/CNA_files/PDF/DOP-2018-U-018296-Final.pdf

Linares, E. (2021, 20 de setiembre). *Regular la Inteligencia Artificial en Latinoamérica,*

¿necesario?. Lex Latin. https://lexlatin.com/opinion/regular-inteligencia-artificial-latinoamerica

MarshMcLennan. (s. f.). La creciente ola de inteligencia artificial en Latinoamérica. https://www.marsh.com/co/migrated-articles/aumento-inteligencia-artificial-latinoamerica.html

Melzer, N. (2010). *Guía para interpretar la noción de participación directa en las hostilidades según el derecho internacional humanitario.* CICR. https://www.icrc.org/es/doc/assets/files/other/icrc_003_0990.pdf

Pratt, M. (2020). *Machine Learning Bias (AI bias).* Search Enterprise AI. https://searchenterpriseai.techtarget.com/definition/machine-learning-bias-algorithm-bias-or-AI-bias

Salmón, E. (2012). *Introducción al Derecho Internacional Humanitario.* CICR.

Schuller, A. (2017). At the Crossroads of Control: The Intersection of Artificial Intelligence in Autonomous Weapon Systems with International Humanitarian Law. *Harvard National Security Journal*.

Sivakumaran, S. (2018). International Humanitarian Law (pp. 503-520). Oxford University Press.

The Economist. (2016, 20 de agosto). Of prediction and policy Machine learning. https://www.economist.com/finance-and-economics/2016/08/20/of-prediction-and-policy

Tucker, P. (2016, 11 de setiembre). *Special Operators Are Getting a New Autonomous Tactical Drone*. Defense One. https://www.defenseone.com/technology/2016/09/special-operators-are-getting-new-autonomous-tactical-drone/131431/

Turek, M. (2021). *Explainable Artificial Intelligence (XAI)*. DARPA. https://www.darpa.mil/program/explainable-artificial-intelligence

Turns, D. (2014). The Law of Armed Conflict (International Humanitarian Law). En M. Evans, *International Law.* Oxford University Press.

Ventura, C. & Billion, D. (2020). ¿Por qué protesta tanta gente a la vez?. *Nueva Sociedad*. https://nuso.org/articulo/por-que-protesta-tanta-gente-la-vez/

Walch, K. (2019, 8 de noviembre). *Training data in facial recognition use cases reveals bias*. Search Enterprise AI. https://searchenterpriseai.techtarget.com/feature/Training-data-in-facial-recognition-use-cases-reveals-bias

Weller, A. (2019). *Design Thinking for a User-Centered Approach to Artificial Intelligence*. She Ji: The Journal of Design, Economics, and Innovation. https://www.sciencedirect.com/science/article/pii/S2405872619300887