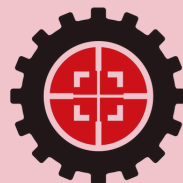


ELEMENTOS CLAVE DE UN TRATADO

SOBRE ARMAS TOTALMENTE
AUTÓNOMAS



CAMPAIGN TO **STOP**
KILLER ROBOTS

ELEMENTOS CLAVE DE UN TRATADO SOBRE ARMAS TOTALMENTE AUTÓNOMAS

La creciente capacidad tecnológica para la autonomía en sistemas de armas plantea una serie de preocupaciones morales, jurídicas, tecnológicas, en la rendición de cuenta y de seguro. Los sistemas de armas que seleccionan y atacan objetivos sin un control humano significativo, conocidos como armas totalmente autónomas, sistemas de armas autónomas letales o robots asesinos, traspasarían el umbral de aceptabilidad y se deberían impedir y prohibir por medio de nuevo derecho internacional.

La Campaña contra los robots asesinos exige un instrumento legalmente vinculante para abordar esta tecnología emergente mediante la preservación de un control humano significativo sobre el uso de la fuerza. El instrumento debería aplicarse a la gama de sistemas de armas que seleccionan y atacan objetivos sobre la base de datos de entrada de sensores, es decir, sistemas en los que el objeto a atacar se determina mediante el procesamiento de sensores, no por seres humanos.[1] Este amplio alcance está diseñado para asegurar que la tecnología problemática no se escape a la regulación.

Sin embargo, las restricciones del tratado se centrarían en los sistemas que contravienen el requisito de un control humano significativo. Utilizaría una combinación de prohibiciones y obligaciones positivas para de hecho prohibir los sistemas que equivalen, o que se utilizan como, armas totalmente autónomas. Si bien durante los debates multilaterales y las negociaciones de los tratados habría que elaborar un lenguaje y un contenido específicos, el instrumento final debería incorporar los elementos clave señalados en el presente documento.

[1] Para obtener más información sobre esta categorización, consulte Richard Moyes, Article 36, “Target Profiles”, agosto de 2019, <http://www.article36.org/wp-content/uploads/2019/08/Target-profiles.pdf>, p. 3.

Este documento examina el concepto de control humano significativo, que sería fundamental para el nuevo tratado o protocolo. Luego, propone tres tipos de obligaciones fundamentales:

- Una obligación general de mantener un control humano significativo sobre el uso de la fuerza.
- Prohibiciones (es decir, obligaciones negativas) de los sistemas de armas que seleccionan y atacan objetivos y que, por su naturaleza, plantean problemas morales o jurídicos fundamentales.
- Obligaciones positivas específicas para ayudar a asegurar que se mantenga un control humano significativo en el uso de todos los demás sistemas que seleccionan y atacan objetivos.

EL CONCEPTO DE CONTROL HUMANO SIGNIFICATIVO



El instrumento jurídicamente vinculante propuesto debería centrarse sobre el control humano significativo, porque muchas de las preocupaciones que plantean las armas totalmente autónomas pueden atribuirse a la ausencia de dicho control. Sería perjudicial para la dignidad humana delegar las decisiones de vida o muerte en máquinas inanimadas que reducen a los seres humanos a puntos de datos y que no pueden comprender el valor de la vida humana. Esos sistemas de armas carecerían también de la capacidad de discernimiento humano necesaria, por ejemplo, para respetar el principio de proporcionalidad de un ataque, como se exige en el derecho internacional. Además, sería jurídicamente difícil y discutiblemente injusto considerar a un ser humano responsable de las acciones de un sistema que opera fuera de su control.[2]

Por estas y otras razones, los Estados, así como las organizaciones internacionales y no gubernamentales, han expresado un amplio acuerdo sobre la necesidad de alguna forma de control humano sobre el uso de la fuerza. Su elección de terminología y sus puntos de vista específicos sobre el rol del ser humano pueden diferir, pero han identificado muchos de los mismos factores. En base a debates internacionales y numerosas publicaciones, este documento destila el concepto de control humano significativo en los componentes de tomas de decisiones, componentes tecnológicos y componentes operacionales.[3]

[2] Para obtener más información sobre los problemas de las armas totalmente autónomas, consulte el documento de Human Rights Watch y la Clínica Internacional de Derechos Humanos de la Facultad de Derecho de Harvard, *Making the Case: The Dangers of Killer Robots and the Need for a Preemptive Ban* (2016), <https://www.hrw.org/report/2016/12/09/making-case/dangers-killer-robots-and-need-preemptive-ban>.

[3] Si bien existen maneras diferentes de enmarcar este concepto, la expresión "control humano significativo" tiene muchas ventajas. "Control" es un término ampliamente utilizado en el derecho internacional y es más fuerte y amplio que las alternativas propuestas por algunos Estados, como "intervención" y "juicio". El calificativo "significativo" tiene la función de asegurar que el control sea sustancial en lugar de superficial y que sea menos específico del contexto o impulsado por los resultados que las alternativas como "apropiado" y "efectivo".

COMPONENTES DE TOMA DE DECISIONES

Los componentes de toma de decisiones del control humano significativo proporcionan a los seres humanos la información y la capacidad de tomar decisiones sobre si el uso de la fuerza cumple con las normas jurídicas y los principios éticos. En particular, el operador humano de un sistema de armas debería tener: una comprensión del entorno operacional; una comprensión de cómo funciona el sistema, incluido lo que podría identificar como objetivo; y tiempo suficiente para la deliberación.

COMPONENTES TECNOLÓGICOS

Los componentes tecnológicos son elementos incorporados a un sistema de armas que pueden mejorar el control humano significativo. Entre ellos se incluye: la previsibilidad y la fiabilidad; [4] la capacidad del sistema para transmitir la información pertinente al operador humano; y la capacidad de un ser humano para intervenir después de la activación del sistema.

COMPONENTES OPERACIONALES

Los componentes operacionales hacen que el control humano sea más significativo al limitar cuándo y dónde puede operar un sistema de armas y a qué puede apuntar. Los factores que podrían limitarse incluyen: el tiempo que transcurre entre la evaluación jurídica de un ser humano y la aplicación de la fuerza por parte del sistema; la duración del funcionamiento del sistema; la naturaleza y el tamaño de la zona geográfica de operación; y los tipos permisibles de objetivos (por ejemplo, personal o material).

Si bien ninguno de estos componentes es suficiente por sí solo para constituir un control humano significativo, todos tienen el potencial de mejorar el control de alguna manera. Además, los componentes generalmente funcionan conjuntamente. Un análisis más profundo de la tecnología existente y emergente podría ayudar a determinar cuáles de estos u otros componentes deberían codificarse en un instrumento jurídico como requisitos previos para un control humano significativo.

[4] En general, la previsibilidad se refiere al grado en que un sistema de armas funciona como prevén los seres humanos, y la fiabilidad se refiere al grado en que el sistema funcionará de manera coherente. Declaración del CICR en relación con el Punto 5(b) del orden del día, el Grupo de Expertos Gubernamentales (GEG) del Convenio sobre Ciertas Armas Convencionales (CCAC) sobre sistemas de armas autónomas letales, Ginebra, marzo de 2019.

OBLIGACIONES FUNDAMENTALES DEL TRATADO

La parte central del instrumento jurídicamente vinculante debería consistir en tres tipos de obligaciones básicas: una obligación general, junto con prohibiciones y obligaciones positivas para implementarlo.

UNA OBLIGACIÓN GENERAL DE MANTENER UN CONTROL HUMANO SIGNIFICATIVO SOBRE EL USO DE LA FUERZA

Esta disposición general facilitaría el cumplimiento de las normas jurídicas y éticas aplicables al obligar a los Estados Partes a mantener un control humano significativo sobre el uso de la fuerza. La generalidad de la obligación ayudaría a evitar lagunas, y el principio que representa podría servir de base para la interpretación de las demás disposiciones del tratado. Como se ha señalado anteriormente, la mayoría de los Estados ya han expresado su apoyo a un requisito de control humano.

La obligación general debería centrarse en el control de la conducta ("uso de la fuerza") más que en la tecnología específica. Este enfoque ayudaría a asegurar el futuro del tratado al evitar la necesidad de predecir cómo se desarrollará la tecnología. El término "uso de la fuerza" también hace que la obligación general se aplique a las situaciones de conflicto armado y a la aplicación de la ley de las agencias de seguridad.[5]

[5] Si bien el término "uso de la fuerza" aparece con frecuencia en los debates y documentos del derecho internacional humanitario y del derecho internacional de los derechos humanos, los dos conjuntos de normas lo rigen de manera algo diferente. Es posible que el nuevo tratado deba tener en cuenta dichas diferencias.

PROHIBICIONES DE LOS SISTEMAS DE ARMAS QUE SELECCIONAN Y ATACAN OBJETIVOS Y QUE, POR SU NATURALEZA, PLANTEAN PROBLEMAS MORALES O JURÍDICOS FUNDAMENTALES

El tratado debería prohibir el desarrollo, la producción y el uso de sistemas de armas que seleccionen y ataquen objetivos y que sean intrínsecamente inaceptables por razones éticas o legales. La claridad de las prohibiciones facilitaría supervisión y ejecución, y su carácter absoluto crearía un fuerte estigma contra los sistemas prohibidos.

El nuevo instrumento debería prohibir los sistemas de armas que, por su naturaleza, seleccionan y atacan objetivos sin un control humano significativo. La prohibición debería abarcar, por ejemplo, los sistemas que se vuelven demasiado complejos para que los usuarios humanos los comprendan y, por lo tanto, producen efectos impredecibles e inexplicables. Estos sistemas complejos pueden aplicar la fuerza en base al aprendizaje automático o permitir que los parámetros críticos del sistema cambien sin autorización humana. Dichos sistemas de armas incumplirían la obligación general del nuevo instrumento antes mencionada.

Las prohibiciones también podrían extenderse a otros sistemas de armas específicos que seleccionan y atacan objetivos y son problemáticas por su naturaleza, y no por su forma de uso. En particular, el tratado podría prohibir los sistemas de armas que seleccionan y atacan a seres humanos como objetivos, independientemente de si operan bajo un control humano significativo.[6] Dichos sistemas se basarían en ciertos tipos de datos, como el peso, el calor o el sonido, para representar a personas o categorías de personas. Al matar o herir a personas en base a dichos datos, esos sistemas contravendrían el principio de la dignidad humana y deshumanizarían la violencia. La prohibición de esta categoría de sistemas también abarcaría los sistemas que, deliberada o involuntariamente, se dirigen a grupos de personas sobre la base de indicadores discriminatorios relacionados con la edad, el género u otras identidades sociales.

[6] Para obtener más información sobre estos sistemas y la propuesta de prohibirlos, véase, en general, Moyes, "Target Profiles."

OBLIGACIONES POSITIVAS ESPECÍFICAS PARA ASEGURAR QUE SE MANTENGA UN CONTROL HUMANO SIGNIFICATIVO EN EL USO DE TODOS LOS DEMÁS SISTEMAS QUE SELECCIONAN Y ATACAN OBJETIVOS

Las obligaciones positivas del nuevo instrumento deberían abarcar los sistemas de armas que no son intrínsecamente inaceptables, pero que todavía podrían tener el potencial de seleccionar y atacar objetivos sin un control humano significativo. Las obligaciones requerirían que los Estados Partes se aseguren de que los sistemas de armas que seleccionan y atacan objetivos sólo se utilicen con un control humano significativo.

El contenido de las obligaciones positivas debería basarse en los componentes de un control humano significativo mencionado anteriormente. Por ejemplo, el tratado podría exigir que los operadores comprendan cómo funciona un sistema de armas antes de activarlo. Podría establecer normas mínimas de previsibilidad y fiabilidad. Además, o alternativamente, el tratado podría limitar los sistemas permisibles a los que operan dentro de ciertos parámetros temporales o geográficos. Al hacerlo, las obligaciones positivas ayudarían a preservar un control humano significativo sobre el uso de la fuerza y establecerían requisitos que, en efecto, harían que el uso de sistemas que funcionan como armas totalmente autónomas sea ilegal.

OTROS ELEMENTOS



Si bien los elementos clave señalados anteriormente son fundamentales para lograr los objetivos del nuevo instrumento, otros elementos deberían complementarlos. Por ejemplo, un preámbulo debería articular el propósito del tratado y situarlo en el contexto del derecho pertinente. Los requisitos de presentación de informes promoverían la transparencia y facilitarían la supervisión independiente. Las medidas de verificación detalladas o los mecanismos de cumplimiento cooperativo ayudarían a prevenir las violaciones del tratado. Las reuniones periódicas de los Estados Partes brindarán la oportunidad de examinar la situación y el funcionamiento del tratado, determinar las brechas en su aplicación y establecer objetivos para el futuro. Otros elementos importantes incluirían el requisito de adoptar medidas nacionales de aplicación y un umbral para la entrada en vigencia.

Este documento informativo de la Campaña para detener a los robots asesinos fue elaborado por Bonnie Docherty de Human Rights Watch y la Clínica Internacional de Derechos Humanos de la Facultad de Derecho de Harvard, con el apoyo de sus estudiantes de derecho en la Clínica.

Retain meaningful human control over the use of force.
Prohibit fully autonomous weapons.
WWW.STOPKILLERROBOTS.ORG

