Agenda Internacional
Año XXXIII N° 46, 2026, pp. 172-198
e-ISSN 2311-5718

# Responsabilidad internacional del Estado ante armas autónomas y ciberoperaciones. Atribución, trazabilidad y desafíos para el derecho internacional humanitario

José Roberto Rodríguez Bustamante*
*Investigador independiente*

**Resumen**

Este artículo examina los desafíos que plantean las armas autónomas y las ciberoperaciones para el régimen de responsabilidad internacional del Estado, con énfasis en la atribución, la trazabilidad y el cumplimiento de las obligaciones del derecho internacional humanitario (DIH). Partiendo de los Artículos sobre Responsabilidad del Estado por Hechos Internacionalmente Ilícitos (ARSIWA) y de los principios de distinción, proporcionalidad y precaución, el estudio sostiene que el problema contemporáneo no es un "vacío normativo", sino un déficit operativo: la opacidad tecnológica, la cadena distribuida de actores y la arquitectura transnacional del ciberespacio dificultan probar la imputación y asegurar rendición de cuentas. A partir de un enfoque analítico y comparativo, se propone una reinterpretación funcional de los criterios tradicionales de atribución y control, incorporando estándares complementarios: trazabilidad técnica, previsibilidad del sistema, obligaciones reforzadas de prevención

* Diplomático y magíster en Relaciones Internacionales por la Academia Diplomática del Perú. Abogado por la Pontificia Universidad Católica del Perú y magíster en Literatura por la Universidad de Barcelona. Estudios de posgrado en Asuntos Europeos y Geopolítica (Istituto per gli Studi di Politica Internazionale de Milán); Política Exterior (Academia Diplomática de Rusia); Inteligencia Estratégica (CAEN); Strategy Execution for Public Leadership (Harvard Business School); Public Policy Analysis (London School of Economics and Political Science); International Humanitarian Law (International Institute of Humanitarian Law), entre otros. Ministro consejero con más de 23 años en el Servicio Diplomático. Ha servido en Italia, Arabia Saudita y Rusia. Actual Representante Permanente Alterno del Perú ante la OEA en Washington, D.C., y exvicepresidente de la Comisión de Asuntos Jurídicos y Políticos de dicha organización. Correo electrónico: jrodriguezb@rree.gob.pe
https://orcid.org/0009-0009-6474-351X

https://doi.org/10.18800/agenda.202601.009

y diligencia debida, y control humano significativo en decisiones críticas. Finalmente, se argumenta que una aproximación basada en el ciclo de vida (diseño–despliegue–uso) puede fortalecer la atribución y la responsabilidad por omisión, y se ofrecen recomendaciones normativas y prácticas para orientar políticas y marcos regionales.

**Palabras clave:** responsabilidad del Estado, atribución, diligencia debida, trazabilidad, armas autónomas, ciberoperaciones, derecho internacional humanitario

## State International Responsibility for Autonomous Weapons and Cyber Operations: Attribution, Traceability, and Challenges for International Humanitarian Law

**Abstract**

This article examines how autonomous weapons systems and cyber operations challenge the law of State responsibility, focusing on attribution, traceability, and compliance with International Humanitarian Law (IHL). Building on the Articles on Responsibility of States for Internationally Wrongful Acts (ARSIWA) and the core IHL principles of distinction, proportionality, and precautions, the paper argues that today's problem is not a legal "gap" but an operational accountability deficit: technological opacity, distributed supply chains, and the transnational architecture of cyberspace make attribution harder to evidence and weaken pathways to responsibility. Using an analytical and comparative approach, the study advances a functional reinterpretation of traditional attribution and control tests by integrating complementary standards: technical traceability, system foreseeability, enhanced duties of prevention and due diligence, and meaningful human control over critical decisions. It further proposes a life-cycle framework (design–deployment–use) to strengthen attribution and responsibility for omissions, and concludes with targeted normative and policy recommendations to support clearer accountability in algorithmic and cyber-enabled warfare, including avenues for regional standard-setting.

**Keywords:** State responsibility, attribution, due diligence, traceability, autonomous weapons, cyber operations, International Humanitarian Law

## 1. Introducción

Los avances recientes en inteligencia artificial, robótica militar y cibertecnología han modificado de manera estructural la forma en que se desarrollan los conflictos armados. La progresiva autonomía de los sistemas de armas y la expansión de las operaciones en el ciberespacio plantean interrogantes sobre los límites de la acción estatal, la imputación de conductas y la vigencia de los principios del derecho internacional humanitario (DIH). A diferencia de los medios tradicionales de combate, las nuevas tecnologías incorporan procesos de decisión automatizados que, en muchos casos, reducen o eliminan la intervención humana directa en la selección y ataque de objetivos (Comité Internacional de la Cruz Roja [CICR], 2023).

Este desplazamiento de la agencia humana hacia sistemas algorítmicos introduce una tensión central en el derecho internacional público: la posibilidad de que existan actos internacionalmente ilícitos sin un vínculo claro entre la conducta material y la voluntad del Estado. La atribución —elemento esencial de la responsabilidad estatal— se ve desafiada por la opacidad tecnológica, la trazabilidad limitada y la multiplicidad de actores que intervienen en la creación, despliegue y operación de estos sistemas (Roscini, 2014).

Desde la perspectiva del DIH, los principios de distinción, proporcionalidad y precaución constituyen el núcleo normativo de la limitación del conflicto armado. No obstante, el uso de sistemas autónomos introduce una incertidumbre operacional respecto de la capacidad del Estado de garantizar el cumplimiento de dichos principios. El control humano significativo (meaningful human control) ha sido propuesto por el Comité Internacional de la Cruz Roja como requisito indispensable para preservar la responsabilidad estatal en decisiones que afectan la vida humana (ICRC, 2021).

Este trabajo pretende contribuir al debate doctrinal y a la formulación de propuestas concretas de adaptación del derecho internacional. Su propósito último es fortalecer el principio de responsabilidad estatal, asegurando que el progreso tecnológico no erosione la base ética y jurídica sobre la cual se funda la protección de la persona humana en los conflictos armados.

Este artículo sostiene que el problema contemporáneo no es principalmente un "vacío normativo", sino un déficit operativo de rendición de cuentas. En la práctica, la atribución —condición de posibilidad de la responsabilidad internacional del Estado— se vuelve más difícil de probar cuando la conducta dañosa emerge de sistemas algorítmicos, de infraestructuras distribuidas o de la intervención de múltiples actores (estatales y privados) en el diseño, entrenamiento, despliegue y uso de tecnologías bélicas (Roscini, 2014). En paralelo, en el ámbito cibernético, el anonimato, el enmascaramiento técnico y el uso de infraestructura transfronteriza erosionan los estándares probatorios y complejizan la identificación del vínculo jurídico entre la operación y el Estado (Schmitt, 2019; Dinstein, 2022).

Sobre esa base, el trabajo examina si el marco general de la responsabilidad internacional del Estado ofrece respuestas suficientemente robustas para estos escenarios, o si requiere una reinterpretación funcional de los criterios tradicionales de atribución y control. Se sostiene que, sin sustituir el régimen existente, es posible fortalecer su eficacia mediante estándares complementarios: trazabilidad técnica, previsibilidad del sistema, obligaciones reforzadas de prevención y diligencia debida, y la exigencia de un control humano significativo en decisiones críticas que afectan la vida y la integridad de las personas (CICR, 2021; Crawford, 2013).

El argumento se desarrolla a partir de un enfoque analítico y comparativo, apoyado en fuentes primarias (en particular, Convenios de Ginebra de 1949, Protocolo Adicional I de 1977 y jurisprudencia relevante), así como en instrumentos de codificación y referencia (Draft Articles on Responsibility of States for Internationally Wrongful Acts —ARSIWA— de la Comisión de Derecho Internacional) y fuentes secundarias (doctrina especializada, informes institucionales y el Tallinn Manual 2.0). Asimismo, se incorpora una perspectiva de "ciclo de vida" (diseño–despliegue–uso), útil para identificar puntos de control, deberes de supervisión y escenarios de responsabilidad por omisión cuando la arquitectura tecnológica limita la atribución clásica. Los *Draft Articles on Responsibility of States for Internationally Wrongful Acts* (Comisión de Derecho Internacional [CDI], 2001), fueron acogidos con beneplácito por la Asamblea General de las Naciones Unidas (res. 56/83), no constituyen un instrumento convencional vinculante, pero se emplean aquí como marco de codificación y referencia ampliamente utilizado para sistematizar el régimen de responsabilidad del Estado, en la medida en que varias de sus disposiciones reflejan reglas consuetudinarias.

Finalmente, el artículo subraya la pertinencia regional del debate. En el sistema interamericano, las discusiones sobre ciberseguridad y gobernanza tecnológica —incluyendo la Estrategia Interamericana Integral de Ciberseguridad y desarrollos recientes en instancias especializadas— reflejan la necesidad de criterios comunes sobre obligaciones estatales y cooperación, especialmente en un entorno de interdependencia tecnológica y riesgos transfronterizos. En ese marco, el propósito último de este trabajo es contribuir a la clarificación doctrinal y a la formulación de criterios operativos que preserven la eficacia del principio de responsabilidad estatal y, con ello, la función protectora del DIH frente a modalidades emergentes de conducción de los conflictos armados.

## 2. Responsabilidad del Estado y atribución en el derecho internacional

La responsabilidad internacional del Estado constituye un pilar estructural del orden jurídico internacional. Así pues, toda acción u omisión atribuible a un Estado y constitutiva de la violación de una obligación internacional genera consecuencias jurídicas en el plano internacional. En el derecho contemporáneo, esta arquitectura se expresa con particular claridad en los Draft Articles on Responsibility of States for Internationally Wrongful Acts, adoptados por la Comisión de Derecho Internacional en 2001 (ARSIWA, por sus siglas en inglés), que son considerados como reflejo del derecho consuetudinario en materia de responsabilidad (Crawford, 2013).

El régimen de responsabilidad no crea nuevas obligaciones sustantivas, sino que determina las consecuencias jurídicas del incumplimiento de obligaciones ya existentes. En ese sentido, la responsabilidad opera como mecanismo secundario que garantiza la efectividad del orden jurídico internacional (Cassese, 2005).

La atribución es el eje que conecta la conducta material con la personalidad jurídica del Estado. Los artículos 4 a 11 del proyecto de la CDI (ARSIWA por sus siglas en inglés) establecen las reglas que determinan cuándo una acción u omisión puede considerarse un "acto del Estado". Esta relación puede basarse en tres criterios principales: (a) el ejercicio de funciones oficiales por órganos estatales; (b) la actuación de personas o entidades que ejercen autoridad gubernamental; y (c) la conducta de actores privados bajo control o dirección estatal.

El artículo 4 establece que la conducta de cualquier órgano del Estado se considera acto estatal, cualquiera sea su posición o función. En un contexto de guerra tecnológica, esto implica que los actos realizados por unidades de ciberdefensa o inteligencia artificial militar caen *prima facie* bajo la responsabilidad del Estado respectivo, aun cuando utilicen algoritmos o sistemas automatizados. El artículo 8 resulta especialmente problemático para las ciberoperaciones, pues regula la atribución de conductas de particulares que actúan bajo "dirección o control" del Estado. La Corte Internacional de Justicia, en el caso Nicaragua v. Estados Unidos (1986), estableció el estándar de "control efectivo" para determinar dicha atribución. Este criterio exige que el Estado haya dirigido o controlado las operaciones específicas en cuestión.

Conviene subrayar, además, una distinción metodológica fundamental entre reglas "primarias" y "secundarias". Mientras las primeras establecen obligaciones sustantivas —por ejemplo, las del derecho internacional humanitario (DIH) relativas a la conducción de hostilidades—, las segundas determinan las consecuencias jurídicas del incumplimiento. Los ARSIWA se ubican en este segundo plano: no "crean" el deber de distinguir o de adoptar precauciones, pero sí fijan cuándo una conducta que vulnera esos deberes se considera atribuible al Estado y qué consecuencias se derivan de ello (Crawford, 2013; Cassese, 2005). Esta distinción resulta clave para evitar la idea de que la aparición de nuevas tecnologías exigiría, por sí misma, "nuevas reglas" para que la responsabilidad estatal pueda operar.

Sin embargo, el Tribunal Penal Internacional para la ex Yugoslavia, en Tadić (1999), adoptó un estándar más amplio de "control general", aplicable cuando existe una relación de subordinación o apoyo sostenido. La coexistencia de ambos criterios ha generado un debate doctrinal sobre la flexibilidad del concepto de control (Negro, 2023). Como advierte Negro, el régimen de la responsabilidad internacional del Estado ha evolucionado desde una concepción restringida, centrada en los daños

a los nacionales extranjeros, hacia un sistema integral que abarca toda violación de normas internacionales, incluidas las de derechos humanos y del DIH. Ello implica que actualmente, el régimen de la responsabilidad internacional se extiende a todos los hechos ilícitos cometidos por el Estado y no solo los Estados pueden invocar dicha responsabilidad, sino también los individuos.

Esta expansión subjetiva del régimen internacional refuerza la idea de que la responsabilidad del Estado no se limita a la acción directa de sus órganos, sino que también abarca su deber de garantizar los derechos humanos y de prevenir la comisión de actos ilícitos por omisión o falta de control. En consecuencia, el análisis de atribución en contextos tecnológicos no solo debe atender al vínculo formal con el órgano estatal, sino también al cumplimiento de las obligaciones de garantía derivadas de los tratados internacionales.

En escenarios tecnológicos, el control efectivo puede ser imposible de demostrar cuando los sistemas operan de manera autónoma o semiautónoma. Por ello, algunos autores proponen una reinterpretación "funcional" del control, que abarque no solo la dirección operacional, sino también el diseño, entrenamiento y supervisión de las tecnologías utilizadas (Roscini, 2014).

El artículo 11 de los ARSIWA añade otra vía: la adopción o reconocimiento de actos de particulares como propios del Estado. En el contexto de ciberataques, esto podría ocurrir si un Estado reconoce explícitamente haber empleado un grupo de hackers o haber validado sus acciones *ex post*.

La noción de diligencia debida complementa los criterios de atribución. Como señaló la Corte Internacional de Justicia (CIJ) en el caso Canal de Corfú (1949), los Estados tienen la obligación de no permitir que su territorio sea utilizado para actos contrarios al derecho internacional. En el ámbito cibernético, esto implica deberes de prevención, monitoreo y cooperación para evitar que infraestructuras nacionales sean empleadas en ciberataques (Dinstein, 2022).

Asimismo, los artículos 16 a 18 de los ARSIWA regulan la responsabilidad derivada de la ayuda, asistencia o coacción en la comisión de actos ilícitos. En un entorno interconectado, la transferencia de software, inteligencia o infraestructura a otro Estado que los use ilícitamente podría generar corresponsabilidad, incluso sin participación directa en el ataque.

El desarrollo tecnológico ha diluido las fronteras entre órgano estatal y actor privado. Las alianzas público-privadas en materia de defensa digital complican la imputación, pues la ejecución de un ciberataque puede involucrar empresas, contratistas y desarrolladores civiles. Ello demanda una revisión de la noción clásica de órgano estatal (Cassese, 2005).

Desde el punto de vista del DIH, la responsabilidad estatal no se limita a actos comisivos. También incluye la omisión en el deber de controlar, entrenar y supervisar las armas y sistemas empleados. En ese sentido, la falta de pruebas de fiabilidad o la ausencia de mecanismos de intervención humana en un sistema autónomo podrían configurar una violación del deber de diligencia (CICR, 2023).

La tendencia doctrinal contemporánea se orienta hacia una responsabilidad funcional, que reconoce la complejidad de las cadenas causales tecnológicas. Este enfoque no busca reemplazar el régimen existente, sino expandir sus criterios interpretativos para asegurar su eficacia frente a nuevas modalidades de conflicto.

En la práctica interestatal, varios países —como Francia, los Países Bajos y Brasil— han emitido declaraciones sobre la aplicación del derecho internacional en el ciberespacio. Estas posiciones evidencian un consenso emergente sobre la validez de los ARSIWA, pero también una demanda de clarificación de los estándares de atribución y diligencia debida en contextos digitales.

En el sistema interamericano, la OEA, a través del Comité Interamericano contra el Terrorismo (CICTE), ha promovido la cooperación en materia de ciberseguridad y la adopción de medidas de fomento de la confianza, incluida la implementación de la Estrategia Interamericana Integral de Ciberseguridad aprobada por la Asamblea General (OEA, 2004). Estos desarrollos son relevantes para la región, donde la interdependencia tecnológica exige reglas comunes para la imputación y la prevención de incidentes transfronterizos.

En suma, la atribución y la responsabilidad estatal siguen siendo pilares del orden jurídico internacional, pero enfrentan tensiones crecientes ante la ejecución de conductas mediante algoritmos o actores descentralizados. Las consideraciones anteriores permiten introducir el análisis del DIH frente a medios y métodos de conducción de hostilidades tecnológicamente mediadas, así como su interacción con la rendición de cuentas y la trazabilidad.

## 3. El derecho internacional humanitario en la era del conflicto armado tecnológico

La progresiva incorporación de tecnologías avanzadas en la conducción de las hostilidades —particularmente los sistemas de armas con creciente autonomía, la inteligencia artificial y las capacidades cibernéticas— está transformando de manera sustantiva la fisonomía contemporánea de los conflictos armados y plantea desafíos inéditos para la aplicación del DIH. Estas innovaciones tensionan algunos de los principios estructurantes del DIH, en especial la distinción entre combatientes y

civiles, la proporcionalidad en el uso de la fuerza, la previsibilidad de los efectos de los ataques y la exigencia de un control humano significativo en las decisiones que implican el empleo letal de la fuerza. En este contexto, la pregunta ya no es únicamente si las normas existentes del DIH resultan formalmente aplicables a estas nuevas modalidades de conducción de las hostilidades, sino si los marcos normativos actuales ofrecen respuestas suficientemente robustas para garantizar una protección efectiva de las personas y bienes civiles frente a sistemas cuyo grado de autonomía, opacidad algorítmica y capacidad de aprendizaje plantea serios retos en términos de responsabilidad, control y rendición de cuentas.

A efectos de este artículo, se distingue entre sistemas automatizados —que ejecutan reglas predeterminadas o respuestas preprogramadas— y sistemas autónomos, que pueden seleccionar cursos de acción o objetivos dentro de un margen de discrecionalidad operacional (a veces apoyados en aprendizaje automático). Esta distinción resulta relevante porque los desafíos jurídicos sobre previsibilidad, control humano y rendición de cuentas se intensifican a medida que aumenta el grado de autonomía del sistema.

En este sentido, Sassòli advierte que el empleo de sistemas de armas con alto grado de autonomía no solo plantea problemas técnicos, sino que incide directamente en la posibilidad real de aplicar los principios clásicos del DIH, en particular cuando la toma de decisiones letales se desplaza del juicio humano hacia procesos automatizados, con efectos difíciles de anticipar y controlar (Sassòli, 2024, p. 569).

Antes de desarrollar el alcance de los principios de distinción, proporcionalidad y precaución, conviene precisar el umbral de aplicabilidad del DIH en el dominio cibernético. En términos generales, el DIH se activa cuando existe un conflicto armado —internacional o no internacional— y regula la conducción de las hostilidades en ese contexto. Por ello, no toda operación en el ciberespacio queda automáticamente comprendida en el DIH: la cuestión decisiva es su vinculación con un conflicto armado y, especialmente, si la operación forma parte de las hostilidades o incide directamente en ellas. Este punto es esencial para evitar extrapolaciones y para delimitar con rigor qué normas son aplicables en cada escenario.

En esa misma línea, una ciberoperación plantea un desafío conceptual adicional: determinar cuándo constituye un "ataque" a efectos del DIH. Sin necesidad de entrar aquí en un debate técnico, resulta suficiente advertir que esta calificación suele depender de sus efectos, particularmente cuando produce —o está diseñada para producir— consecuencias comparables a las de un ataque cinético, como daño a personas, destrucción o inutilización de bienes, o interferencias graves en infraestructura esencial. En tales supuestos, cobra plena relevancia el marco normativo del

DIH y, con él, la exigencia de que el empleo de capacidades cibernéticas respete los principios de distinción, proporcionalidad y precaución, cuya adaptación operativa a entornos de opacidad y automatización constituye el núcleo del análisis que sigue. A continuación, abordaremos algunos de los retos que enfrentan los principios centrales del DIH en relación al uso de tecnología avanzada en los conflictos armados.

### 3.1. Principio de distinción

El principio de distinción constituye el núcleo del DIH contemporáneo y cumple la función de limitar los efectos de la guerra protegiendo a quienes no participan directamente en las hostilidades. En la medida en que una operación —incluida una ciberoperación— alcance el umbral de "ataque" a efectos del DIH, este principio adquiere plena centralidad: la conducción de las hostilidades debe orientarse, en todo momento, a distinguir entre combatientes y civiles, así como entre objetivos militares y bienes de carácter civil. Recogido en el artículo 48 del Protocolo Adicional I a los Convenios de Ginebra (1977), obliga a las partes en conflicto a mantener esa distinción como condición de licitud del empleo de la fuerza. Sin embargo, la aparición de sistemas autónomos de armas y la realización de ciberoperaciones pone a prueba la capacidad de los Estados para cumplir efectivamente este principio.

En los sistemas bélicos tradicionales, la identificación del objetivo era una tarea eminentemente humana. En cambio, en las armas autónomas, la decisión de seleccionar y atacar puede ser ejecutada por algoritmos basados en aprendizaje automático, lo que introduce grados de opacidad y de incertidumbre operacional incompatibles con un control humano directo y verificable. El CICR (2023) advierte que dicha opacidad algorítmica puede incrementar el riesgo de ataques indiscriminados o desproporcionados, especialmente en entornos densamente poblados o con presencia de infraestructura dual (civil y militar).

La cuestión central es si un sistema autónomo puede cumplir, en términos jurídicamente exigibles, el deber de distinguir impuesto a las partes en conflicto. Como señalan Galliott y Scholz (2021), la respuesta depende de la capacidad de ese sistema para procesar información contextual con precisión y bajo supervisión humana significativa. Sin embargo, incluso los sistemas más avanzados presentan limitaciones —técnicas y contextuales— que dificultan garantizar la fiabilidad del juicio normativo requerido por el DIH, lo que repercute directamente en la determinación de responsabilidad por hechos internacionalmente ilícitos cuando se producen daños a civiles o bienes de carácter civil.

En el ámbito de las ciberoperaciones, el principio de distinción enfrenta un desafío análogo. Ciberataques dirigidos contra infraestructuras críticas —como redes

eléctricas o sistemas hospitalarios— pueden traducirse en afectaciones extensas a la población civil, incluso cuando el objetivo militar invocado sea legítimo. Como subraya el profesor Schmitt (2019), la dificultad de determinar con certeza la naturaleza —civil, militar o dual— de numerosos sistemas informáticos convierte la distinción en un proceso altamente incierto. En ese contexto, la mera posibilidad de daño colateral significativo refuerza el deber estatal de diligencia y verificación antes, durante y después de la operación, como presupuesto para sostener que la conducta se mantuvo dentro de los límites impuestos por el DIH

### 3.2. Principio de proporcionalidad

El principio de proporcionalidad, consagrado en el artículo 51(5)(b) del Protocolo Adicional I, prohíbe los ataques en los que el daño incidental esperado a civiles o bienes de carácter civil sea excesivo en relación con la ventaja militar concreta y directa que se busca alcanzar. En la práctica, este principio exige un juicio ponderativo que combina apreciaciones jurídicas, técnicas y éticas. En consecuencia, la proporcionalidad depende de una valoración humana informada, lo cual adquiere especial complejidad cuando ciertas fases del proceso decisorio militar se automatizan o se delegan a sistemas algorítmicos.

Como sostiene Dinstein (2022), la proporcionalidad no puede reducirse a un cálculo estadístico, porque involucra la apreciación contextual de la necesidad militar y el respeto debido a la vida humana. La sustitución de este juicio por algoritmos plantea un riesgo jurídico evidente: transformar un principio humanitario en una operación de optimización de daños, con márgenes de error difíciles de justificar en escenarios donde la información es incompleta, cambiante o deliberadamente ambigua.

En el caso de las ciberoperaciones, la proporcionalidad se ve afectada además por la imprevisibilidad de los efectos secundarios. Un ataque dirigido contra un servidor o un sistema militar puede propagarse, de manera no intencionada, a redes civiles interconectadas, generando daños desproporcionados o interrupciones graves de servicios esenciales. La labor de la Corte Internacional de Justicia en casos como Nuclear Weapons Advisory Opinion (1996) recuerda que el uso de cualquier medio de guerra debe respetar las limitaciones inherentes al DIH y al principio de humanidad. Por analogía, cuando los efectos cibernéticos no pueden ser razonablemente previstos o contenidos, la evaluación de proporcionalidad se vuelve frágil incluso en ausencia de intención dolosa.

Por ello, la proporcionalidad en la era digital exige una aproximación intensamente preventiva. Los Estados tienen la obligación de realizar evaluaciones jurídicas *ex ante* sobre los riesgos de daños colaterales derivados del uso de sistemas autónomos o de la ejecución de ciberataques. El artículo 36 del Protocolo Adicional I impone el

deber de revisar la legalidad de nuevas armas, medios o métodos de guerra; en consecuencia, esta exigencia aplica a sistemas algorítmicos de decisión y a herramientas digitales ofensivas (CICR, 2021), particularmente cuando su funcionamiento —por diseño o por el entorno operativo— pueda generar efectos extendidos o difíciles de controlar.

### 3.3. Precaución y control humano

El principio de precaución, estrechamente vinculado a la proporcionalidad, establece la obligación de adoptar todas las medidas factibles para evitar o, al menos, minimizar el daño a la población civil y a los bienes civiles. En los conflictos contemporáneos, su aplicación se vuelve especialmente exigente cuando la conducción de las hostilidades incorpora automatización, mediación algorítmica o capacidades cibernéticas, pues aumenta la dificultad de anticipar, contener y verificar efectos.

En ese marco, la noción de "control humano significativo" (meaningful human control), promovida por el CICR y respaldada por diversos Estados, alude al nivel de intervención y supervisión que debe conservar la persona humana sobre el funcionamiento del sistema y, en particular, sobre funciones críticas como la selección de objetivos y la decisión de atacar. De acuerdo con el CICR (2023), este control debe reflejarse en varias fases: diseño y programación, parámetros de empleo, selección/validación del objetivo y posibilidad de intervención o cancelación. La ausencia de un control humano adecuado incrementa el riesgo de incumplir el deber de adoptar precauciones y, con ello, de incurrir en violaciones del DIH.

En el ámbito de las ciberoperaciones, el deber de precaución se manifiesta en obligaciones concretas de verificación y evaluación *ex ante:* identificar el objetivo, estimar razonablemente los efectos previsibles —incluidos efectos secundarios— y adoptar medidas para limitar el alcance y duración del daño. En este dominio, la previsibilidad técnica y la capacidad de detener, contener o revertir una operación adquieren un valor decisivo para evaluar el cumplimiento del principio de precaución (Roscini, 2014). Cuando tales salvaguardas faltan, el riesgo de afectación a bienes civiles e infraestructura esencial se incrementa de manera significativa.

Finalmente, este punto enlaza con un estándar de prevención propio del régimen general de responsabilidad internacional: la debida diligencia. En situaciones donde la atribución directa puede ser difícil de acreditar, la debida diligencia opera como un puente analítico entre el DIH y las obligaciones estatales de prevención de daños previsibles. Como enfatiza Negro (2023), la responsabilidad estatal no se agota en la imputación inmediata de la conducta, sino que puede comprender omisiones relevantes —por ejemplo, la falta de medidas razonables para prevenir consecuencias

previsibles del despliegue de tecnologías de alto riesgo—. Así, un Estado que emplea un sistema autónomo sin garantías suficientes de control humano, o que ejecuta una operación cibernética sin verificar razonablemente sus efectos previsibles, podría incurrir en incumplimientos que comprometan su responsabilidad internacional.

### 3.4. El DIH y los nuevos métodos y medios de guerra

El DIH se ha caracterizado históricamente por su capacidad para aplicarse a innovaciones en los medios y métodos de guerra. La Cláusula de Martens, incluida en el Preámbulo de los Convenios de La Haya (1899) y reiterada en el Protocolo Adicional I, dispone que incluso en los casos no previstos por los tratados, las personas civiles y combatientes quedan bajo la protección de los principios del derecho de gentes, las leyes de la humanidad y las exigencias de la conciencia pública. A la luz de esta cláusula y del carácter principialista del DIH, la ausencia de una regulación específica sobre el diseño, fabricación, despliegue y uso de sistemas de armas autónomas letales, así como sobre la conducción de ciberoperaciones con fines militares, no implica necesariamente un vacío normativo: desplaza, más bien, el problema hacia la aplicación rigurosa de principios existentes y hacia la garantía de mecanismos efectivos de control y rendición de cuentas.

En esa lógica, las nuevas tecnologías no crean un "nuevo derecho de la guerra", pero sí exigen una aplicación particularmente cuidadosa —y, en ocasiones, una actualización interpretativa— de principios como distinción, proporcionalidad y precaución, atendiendo a los rasgos del entorno operativo contemporáneo: opacidad algorítmica, velocidad de ejecución, dependencia de datos y riesgos de efectos secundarios difíciles de predecir o contener. La legalidad del uso de sistemas autónomos o del empleo de ciberataques debe evaluarse, por tanto, conforme a los principios fundamentales del DIH, aplicados mutatis mutandis a los medios digitales y cibernéticos (Dinstein, 2022; CICR, 2021).

Los debates recientes en el marco de la Convención sobre Ciertas Armas Convencionales (CCAC) y las resoluciones de la Asamblea General de las Naciones Unidas sobre sistemas de armas autónomas reflejan una preocupación creciente por los riesgos éticos y jurídicos de la autonomía bélica. En este contexto, el CICR ha propuesto un marco regulatorio orientado a prohibir determinados sistemas completamente autónomos y a regular estrictamente los sistemas parcialmente autónomos, con especial énfasis en límites al diseño, condiciones de empleo y exigencias de control humano significativo (CICR, 2021).

En su posición institucional, el CICR ha llamado a establecer límites claros al desarrollo y uso de sistemas de armas autónomas, incluyendo la prohibición de determinados

sistemas que no puedan operar conforme al DIH y la regulación estricta de aquellos que mantengan funciones críticas bajo control humano significativo, con el fin de preservar la protección de la población civil y la rendición de cuentas.

### 3.5. Hacia una actualización interpretativa del DIH

La convergencia entre autonomía tecnológica y acción militar exige replantear el equilibrio entre eficacia operativa y humanidad. En palabras de Dinstein (2022), el DIH debe seguir siendo "el derecho de la moderación en la guerra", incluso frente a la sofisticación de las armas inteligentes. Preservar la centralidad del juicio humano —y, con ello, la posibilidad de control, explicación y verificación— se vuelve una condición esencial para mantener el valor protector del derecho en escenarios donde decisiones críticas pueden quedar mediadas por sistemas tecnológicos.

En el plano regional, la Organización de los Estados Americanos podría desempeñar un papel relevante promoviendo directrices hemisféricas sobre el uso responsable de tecnologías militares autónomas y el cumplimiento del DIH en el ciberespacio, a partir de los trabajos del Comité Jurídico Interamericano (CJI), como el informe "El derecho internacional aplicable al ciberespacio" (CJI, 2022). Dicho informe reafirma la aplicabilidad del marco vigente del derecho internacional a las actividades desarrolladas en el ciberespacio, incluyendo principios como soberanía y prohibición del uso de la fuerza; advierte que determinadas ciberoperaciones pueden vulnerar la soberanía y el principio de no intervención; y subraya la centralidad de la responsabilidad internacional del Estado, destacando que los Estados no solo deben abstenerse de conductas internacionalmente ilícitas, sino también adoptar medidas razonables para evitar que su territorio sea utilizado por actores no estatales como plataforma para ciberataques contra otros Estados. La experiencia hemisférica en medidas de fomento de la confianza y cooperación cibernética ofrece, en esa línea, una base para avanzar hacia estándares compartidos. En particular, la Estrategia Interamericana Integral de Ciberseguridad (OEA, 2004) adopta un enfoque multidimensional que combina prevención, cooperación, creación de capacidades y promoción de una cultura de ciberseguridad, elementos que pueden contribuir a operacionalizar obligaciones de diligencia debida, intercambio de información y medidas de fomento de la confianza.

El DIH conserva plenamente su vigencia frente a la aparición de nuevos métodos y medios de conducción de las hostilidades; sin embargo, su eficacia depende cada vez más de una aplicación particularmente rigurosa de sus principios estructurantes. Distinción, proporcionalidad y precaución no pierden relevancia ante la innovación tecnológica; por el contrario, se tornan más exigentes en contextos caracterizados por automatización, opacidad algorítmica y creciente distancia entre el decisor humano y los efectos materiales de la violencia.

Esta constatación prepara el terreno para abordar, en el capítulo siguiente, uno de los desafíos más críticos del DIH en la era de la guerra tecnológica: la brecha de rendición de cuentas. La progresiva delegación de funciones decisorias en sistemas autónomos y la complejidad técnica de las operaciones cibernéticas dificultan mecanismos tradicionales de atribución de conductas, determinación de responsabilidades y trazabilidad de decisiones que producen efectos letales o daños humanitarios significativos. En este contexto, la efectividad del régimen jurídico internacional dependerá, en buena medida, de la capacidad de los Estados y de la comunidad internacional para asegurar que el desarrollo y empleo de estas tecnologías no diluyan los umbrales de responsabilidad ni erosionen los estándares de control y rendición de cuentas exigidos por el DIH.

## 4. La brecha de rendición de cuentas. Atribución y trazabilidad

La progresiva incorporación de sistemas tecnológicos complejos en la conducción de las hostilidades —incluidos los sistemas de armas con creciente autonomía y las operaciones en el ciberespacio— ha puesto en evidencia una brecha emergente en los mecanismos tradicionales de rendición de cuentas del DIH. En escenarios donde la cadena de decisiones se fragmenta entre diseñadores, programadores, operadores, mandos militares y sistemas algorítmicos que median o ejecutan la acción, los esquemas clásicos de atribución de conductas y determinación de responsabilidades enfrentan tensiones significativas. La opacidad técnica de los sistemas, la dificultad para reconstruir *ex post* la secuencia decisoria y la posible dilución del control humano efectivo dificultan, a la vez, la identificación de los sujetos responsables y la trazabilidad de las decisiones que derivan en efectos humanitarios adversos.

Este capítulo examina, desde esa perspectiva, (i) los desafíos técnicos y operativos para la atribución, (ii) el papel del control humano en la cadena de mando como criterio funcional de imputación, y (iii) las implicancias de la responsabilidad por omisión y de la diligencia debida estatal; para finalmente explorar la necesidad de avanzar hacia enfoques de responsabilidad adaptativa capaces de preservar la efectividad del régimen jurídico internacional frente a las transformaciones tecnológicas de los conflictos armados.

### 4.1. Desafíos técnicos y operativos para la atribución

Existe un consenso creciente en el sentido de que los sistemas de armas autónomos deben emplearse únicamente bajo un control humano significativo; sin embargo, aun bajo ese marco, persiste el riesgo de que decisiones materialmente determinantes estén condicionadas por sistemas computarizados y aplicaciones de inteligencia artifi-

cial que inciden —directa o indirectamente— en el uso de la fuerza. En ese contexto, uno de los mayores desafíos del derecho internacional contemporáneo frente a las nuevas tecnologías militares es la dificultad de atribuir actos ilícitos a un Estado de manera precisa, verificable y suficientemente probada.

En el caso de las ciberoperaciones, el anonimato estructural del ciberespacio —caracterizado por la posibilidad de enmascarar direcciones, suplantar identidades y utilizar infraestructura ubicada en terceros países— impide en muchos casos establecer con certeza el origen de un ataque. Como señala Roscini (2014), la atribución técnica (identificación del atacante mediante evidencias digitales) y la atribución jurídica (vinculación del acto al Estado) constituyen niveles distintos que rara vez coinciden. Esta disociación debilita los mecanismos clásicos de rendición de cuentas en la medida en que las víctimas de un ataque cibernético no pueden identificar claramente al responsable en términos jurídicos, aun cuando existan indicios técnicos plausibles.

Los artículos del proyecto de la CDI presuponen que la existencia del hecho ilícito y su autoría estatal pueden determinarse a partir de pruebas objetivas. Sin embargo, en el entorno digital esas pruebas suelen ser frágiles, efímeras y, en ocasiones, fácilmente manipulables. La volatilidad de los datos y la ausencia de trazabilidad integral dificultan alcanzar el estándar probatorio requerido para una imputación de responsabilidad que sea aceptable a nivel internacional.

En el ámbito de las armas autónomas, la dificultad se manifiesta de otra forma: determinar el nexo causal entre diseño, programación, parámetros de empleo y daño causado. La cadena de responsabilidad tiende a extenderse desde el programador hasta el operador, pasando por el comandante y, en última instancia, por el Estado. No obstante, la autonomía parcial o total de estos sistemas puede interrumpir la relación causal tradicional y generar zonas de incertidumbre respecto de qué decisión humana específica produjo el resultado. Galliott y Scholz (2021) advierten que, en estos casos, la falta de transparencia algorítmica —el llamado "problema de la caja negra"— puede hacer impracticable reconstruir a posteriori las decisiones tomadas por el sistema.

La trazabilidad tecnológica se convierte entonces en un elemento esencial para la responsabilidad internacional. Sin mecanismos que permitan auditar el comportamiento de los sistemas autónomos, verificar registros de operación y analizar parámetros de decisión, el derecho internacional enfrenta un déficit estructural de evidencia. En esa línea, se ha propuesto incorporar el principio de "accountability by design", entendido como la obligación de garantizar desde la etapa de desarrollo que las tecnologías bélicas incluyan registros verificables de sus acciones (CICR, 2023).

A nivel operativo, la cooperación internacional es también un factor determinante. La atribución de un ciberataque exige, con frecuencia, colaboración entre Estados y provisión de información técnica e inteligencia. Sin embargo, intereses estratégicos y políticas de confidencialidad en materia de defensa limitan dicha cooperación. Como resultado, la brecha de rendición de cuentas no es solo técnica, sino también política, al depender de la voluntad estatal de compartir evidencias y reconocer responsabilidades (NATO CCDCOE, 2017).

### 4.2. El papel del control humano en la cadena de mando

Para los efectos del presente trabajo, la noción de control humano significativo adquiere una dimensión transversal: no solo opera como garantía ética y operacional, sino también como criterio funcional de imputación jurídica. En la medida en que el Estado, a través de sus agentes políticos o militares, conserve un control suficiente sobre el sistema, la conducta desplegada puede atribuirse al Estado conforme al artículo 4 de los Artículos sobre Responsabilidad del Estado por Hechos Internacionalmente Ilícitos (ARSIWA). Por el contrario, cuando dicho control se pierde o se delega de manera sustancial en sistemas autónomos, la cadena de responsabilidad tiende a fragmentarse, dificultando la determinación de los sujetos responsables.

Desde la perspectiva del DIH, el control humano resulta igualmente esencial para asegurar la vigencia del principio de responsabilidad de mando y de los deberes de supervisión. Los comandantes militares tienen la obligación de prevenir y sancionar las violaciones al DIH cometidas por sus subordinados. Esta obligación, consagrada en el artículo 86 del Protocolo Adicional I, debe ser leída a la luz de las transformaciones tecnológicas para abarcar, en lo pertinente, el empleo de sistemas de armas autónomos que operan bajo su autoridad o dentro de su esfera de mando. En tal sentido, el comandante no puede ampararse en el desconocimiento de decisiones adoptadas por un sistema técnico que opera sin monitoreo humano efectivo. Una intervención humana meramente formal en la activación de estos sistemas —sin supervisión sustantiva de parámetros, contexto y efectos— puede comprometer la responsabilidad internacional del Estado por incumplimiento de obligaciones de prevención y control.

En la misma línea, Sassòli propone una aproximación desagregada del control humano significativo, distinguiendo: (i) el control sobre la decisión de emplear un sistema de armas autónomo en una misión concreta; (ii) el control contextual ejercido mediante la configuración activa de parámetros internos y la imposición de restricciones espaciales y temporales; y (iii) —de manera más controvertida— la posibilidad de intervención humana durante la ejecución de la operación. Esta última

dimensión presupone, además, no solo la existencia de un vínculo de comunicación entre el operador y el sistema, sino también la capacidad del sistema para ofrecer explicaciones comprensibles sobre sus decisiones o recomendaciones. La ausencia de estos niveles de control y de trazabilidad decisional no solo debilita la efectividad operativa del DIH, sino que compromete la posibilidad de atribuir responsabilidades jurídicas claras ante resultados contrarios al derecho (Sassòli, 2024).

En el ámbito de las ciberoperaciones, el control humano se expresa principalmente a través de la supervisión, validación y parametrización de los algoritmos utilizados para ejecutar acciones ofensivas o defensivas automatizadas. La creciente tendencia hacia la automatización de respuestas en el ciberespacio —por ejemplo, mediante sistemas de "respuesta automática" que operan sin intervención humana directa— plantea un riesgo jurídico análogo al de las armas autónomas, en la medida en que debilita el vínculo entre la decisión estatal y sus consecuencias fácticas. Como advierte Schmitt (2019), aun cuando ciberoperaciones sean ejecutadas por programas automatizados, el Estado conserva la obligación de garantizar que tales acciones respeten el principio de proporcionalidad y los derechos soberanos de otros Estados.

Desde el punto de vista institucional, la incorporación de mecanismos de revisión humana puede ser entendida como parte del deber de diligencia debida. La Corte Internacional de Justicia, en el caso Canal de Corfú (1949), estableció que los Estados tienen la obligación de no permitir que su territorio sea utilizado para actos contrarios a los derechos de otros Estados. En el contexto tecnológico, esta obligación se proyecta sobre la necesidad de prever y controlar riesgos derivados de la autonomía de los sistemas, evitando que operen sin supervisión humana razonable cuando sus efectos puedan trascender fronteras o afectar intereses jurídicamente protegidos.

La cadena de mando, en consecuencia, ya no puede limitarse a la jerarquía militar tradicional. Debe ampliarse para incluir, de manera funcional, a actores técnicos —ingenieros, desarrolladores, programadores— que diseñan, entrenan y mantienen sistemas cuyo comportamiento incide en decisiones operacionales. La omisión de control sobre estas etapas puede constituir una forma de responsabilidad por omisión.

En esta misma línea, Negro (2023) sostiene que la transformación tecnológica de los conflictos exige superar lecturas excesivamente rígidas del control efectivo y avanzar hacia una interpretación más funcional de la atribución, que tome en cuenta contextos híbridos y la actuación concurrente de actores estatales y privados. Para el autor, la creciente participación de entidades no estatales —empresas tecnológicas, contratistas o desarrolladores— refuerza la necesidad de reinterpretar reglas de imputación a fin de evitar que el vacío de control se traduzca en vacío de responsabilidad.

Esta observación se integra con la tesis del presente trabajo: la brecha de rendición de cuentas en el ámbito digital y autónomo no es meramente técnica, sino estructural, pues nace de la dificultad de mantener trazabilidad entre decisión humana y acción tecnológica. La función del derecho, en este sentido, es reconstituir ese vínculo mediante estándares de control, previsibilidad, supervisión y diligencia debida.

### 4.3. La responsabilidad por omisión y la diligencia debida estatal

La responsabilidad por omisión constituye una categoría clave para abordar rendición de cuentas en entornos donde funciones decisorias humanas pueden ser sustituidas o mediadas por decisiones algorítmicas. Conforme al artículo 2 de los ARSIWA, la responsabilidad del Estado puede surgir tanto por actos como por omisiones contrarias al derecho internacional. En consecuencia, la falta de supervisión, de pruebas de fiabilidad o de medidas preventivas frente a sistemas autónomos puede constituir una omisión ilícita cuando exista una obligación internacional aplicable que exija actuar.

La debida diligencia, entendida como el deber de actuar con prudencia y previsión razonable, se ha consolidado como un principio general del derecho internacional. Aplicada al ámbito tecnológico, implica que los Estados deben adoptar medidas razonables para evitar que los sistemas que emplean con fines militares —o los sistemas bajo su jurisdicción— causen daño a otros Estados o a la población civil (Dinstein, 2022). Esta obligación se proyecta tanto en la fase de desarrollo y prueba como en la de despliegue operacional.

La doctrina reciente ha propuesto reconocer una diligencia debida tecnológica como estándar emergente. Bajo esta aproximación, los Estados deberían asegurar mecanismos de transparencia, auditoría y control sobre tecnologías empleadas con fines militares. La ausencia de tales mecanismos podría constituir una violación autónoma del deber de prevenir daños previsibles (ICRC, 2023).

En el caso de las ciberoperaciones, la debida diligencia puede implicar monitorear infraestructura nacional, establecer canales de cooperación internacional y responder ante actividades ilícitas detectadas. El Tallinn Manual 2.0 (NATO CCDCOE, 2017) reconoce este deber como una obligación tanto negativa como positiva: los Estados no deben permitir actos ilícitos desde su territorio y deben adoptar medidas activas, razonables y proporcionadas para impedirlos cuando tengan conocimiento de ellos y capacidad para actuar.

Desde una perspectiva política e institucional, existe un riesgo alto de que la rendición de cuentas en el ciberespacio se vea obstaculizada por la falta de transparencia estatal. La ausencia de mecanismos internacionales de investigación comparables a las comisiones de investigación empleadas en otros ámbitos limita la verificación

independiente de ciberoperaciones. En ese sentido, la creación de un marco multilateral de reporte y revisión —eventualmente bajo auspicios regionales o universales— podría contribuir a reducir la brecha, fortalecer la confianza entre Estados y mejorar la disponibilidad de información para evaluaciones jurídicas.

Asimismo, en los debates internacionales sobre normas de conducta estatal en el ciberespacio han emergido enfoques divergentes. Un ejemplo es la Organización de Cooperación de Shanghái (OCS), que ha promovido el concepto de "seguridad de la información" y propuestas orientadas a reforzar el control soberano del entorno informacional, incluida la presentación ante la Asamblea General de la ONU del *International Code of Conduct for Information Security* (A/66/359) y una versión actualizada (A/69/723). Considerar estas iniciativas permite apreciar que la atribución, la prevención y la cooperación se discuten en un contexto normativo plural, con implicancias para la construcción de estándares comunes.

En el ámbito interamericano, el debate sobre atribución, diligencia debida y conducta estatal en el ciberespacio cuenta con un insumo jurídico-regional relevante: el Segundo Informe (2022) del Comité Jurídico Interamericano (CJI) sobre el derecho internacional aplicable al ciberespacio. Dicho informe reafirma la aplicabilidad del derecho internacional vigente a las actividades cibernéticas, destaca la centralidad de la responsabilidad internacional del Estado y subraya deberes de prevención para evitar que el territorio sea utilizado por actores no estatales como plataforma de ataques. Asimismo, constata que la formulación de posiciones nacionales más completas en la región sigue siendo incipiente, al tiempo que resalta la dimensión institucional de apoyo técnico —incluidos los cursos y actividades del Programa de Ciberseguridad— en el marco del Comité Interamericano contra el Terrorismo (CICTE), como mecanismo de fortalecimiento de capacidades estatales y de desarrollo de políticas públicas en ciberseguridad y ciberoperaciones.

Complementariamente, por sucesivos mandatos de la Asamblea General de la OEA, la Comisión de Seguridad Hemisférica y el Comité Interamericano contra el terrorismo han venido abordando de manera continuada los desafíos de la ciberseguridad y la aplicación del derecho internacional en el ciberespacio, incluyendo el fortalecimiento de capacidades, la cooperación técnica y el intercambio de buenas prácticas. Esta línea de trabajo, junto con la difusión de los aportes del Comité Jurídico Interamericano, proporciona una base regional para reforzar prevención, diligencia debida y trazabilidad en el contexto regional.

En suma, la brecha de rendición de cuentas en la era digital y autónoma no responde a un vacío normativo propiamente dicho, sino a una creciente desincronización entre el ritmo del desarrollo tecnológico y la operatividad de los marcos de atribución,

control y diligencia debida. Los principios estructurantes del derecho internacional conservan su vigencia, pero requieren una lectura funcional capaz de integrar desafíos derivados de opacidad técnica, fragmentación de la cadena decisoria y autonomía creciente de los sistemas.

En este sentido, lo que podría denominarse una responsabilidad adaptativa remite a la capacidad del derecho internacional para absorber complejidad tecnológica sin perder fuerza normativa ni función garantista. Ello supone reforzar, de manera articulada, exigencias de transparencia, control humano efectivo y cooperación internacional, como condiciones mínimas para evitar que los avances tecnológicos se traduzcan en zonas de impunidad. Sobre esta base, el capítulo siguiente examina criterios normativos y funcionales que podrían orientar una actualización coherente de los esquemas de imputación y responsabilidad estatal a lo largo del ciclo de vida de los sistemas de armas autónomos y de las ciberoperaciones.

## 5. Hacia una reinterpretación funcional de la responsabilidad estatal. Criterios para la imputación en el ciclo de vida de las armas autónomas

### 5.1. La imputación a lo largo del ciclo de vida tecnológico

La evolución de los sistemas de armas autónomas y de las ciberoperaciones exige repensar los criterios clásicos de imputación en el derecho internacional, tradicionalmente centrados en la acción u omisión de órganos humanos del Estado. En el contexto tecnológico contemporáneo, la voluntad estatal puede manifestarse a lo largo de un ciclo de vida que comprende el diseño, el desarrollo, el despliegue y el uso operativo de los sistemas; y cada una de estas fases puede generar consecuencias jurídicas en caso de violación del derecho internacional humanitario u otras obligaciones internacionales.

La imputación no debe concebirse, por tanto, como un acto aislado vinculado exclusivamente al resultado material de una operación concreta, sino como una cadena continua de responsabilidad que acompaña el ciclo de vida de la tecnología militar. El control estatal se expresa en la decisión de desarrollar determinados sistemas, en las condiciones bajo las cuales se autoriza su empleo y en las circunstancias específicas de su activación. En cada una de estas etapas, el Estado conserva un grado diferenciable —pero jurídicamente relevante— de dirección, supervisión y responsabilidad.

Durante la fase de diseño y desarrollo, el Estado tiene la obligación de garantizar que las armas, los algoritmos y los sistemas que incorpora a su aparato militar sean

conformes con el DIH. El artículo 36 del Protocolo Adicional I de 1977 impone el deber de someter toda nueva arma o método de guerra a una revisión jurídica previa, obligación que se extiende a sistemas basados en inteligencia artificial y a capacidades cibernéticas con potencial impacto cinético o humanitario. Esta revisión no puede ser meramente formal: debe evaluar, de manera sustantiva, la compatibilidad del sistema con los principios de distinción, proporcionalidad y precaución, así como con obligaciones pertinentes del derecho internacional de los derechos humanos en la medida aplicable. En este sentido, la responsabilidad estatal puede comenzar a configurarse ya en la fase de concepción tecnológica y no únicamente en el momento de su empleo en el terreno.

En la etapa de despliegue operativo, el control estatal se proyecta a través de la estructura de mando, los procedimientos de autorización y los protocolos de supervisión. El principio de responsabilidad de mando, consagrado en el artículo 86 del Protocolo Adicional I, mantiene plena vigencia en entornos caracterizados por altos niveles de automatización: los comandantes son responsables no solo de las conductas de sus subordinados, sino también —en lo pertinente— de los medios y sistemas que deciden emplear bajo su autoridad. En consecuencia, la introducción de sistemas autónomos no exonera al Estado del deber de garantizar la existencia de un control humano significativo ni de adoptar medidas razonables para prevenir o corregir fallas previsibles en el funcionamiento de tales sistemas.

Finalmente, en la fase de uso o ejecución, la autorización o activación de un sistema autónomo constituye un punto crítico del vínculo de imputación. En términos del régimen general de responsabilidad del Estado, la imputación se fundamenta principalmente en los artículos 4 a 8 de los ARSIWA: conducta de órganos del Estado (art. 4), entidades que ejercen funciones gubernamentales (art. 5) y, en su caso, actores que actúan bajo dirección o control del Estado (art. 8). En contextos tecnológicos complejos, la automatización del proceso decisorio no rompe el nexo entre la acción desplegada y la voluntad estatal en la medida en que el Estado conserva la capacidad de configurar, habilitar y emplear el sistema bajo su autoridad.

### 5.2. Trazabilidad, supervisión humana y diligencia debida

El eje central de una lectura funcional de la imputación en contextos tecnológicos reside en la trazabilidad, entendida como la capacidad de reconstruir de manera verificable el proceso decisorio que conduce a un daño o a una violación del derecho internacional. En el ámbito de los sistemas de armas autónomos, la trazabilidad no constituye únicamente una exigencia técnica, sino una garantía jurídica indispensable para la efectividad del régimen de responsabilidad. Un sistema cuyo funcionamiento

no permite identificar las razones de su comportamiento o reconstruir la secuencia de decisiones resulta difícilmente compatible con exigencias del DIH y con el principio de responsabilidad internacional del Estado.

La opacidad algorítmica —frecuentemente descrita como el fenómeno de la "caja negra" de la inteligencia artificial— introduce un riesgo estructural de debilitamiento de la rendición de cuentas. Cuando ni los operadores ni los desarrolladores pueden explicar de manera comprensible por qué un sistema actuó de determinada forma, se compromete el vínculo entre causalidad y atribución jurídica y, con ello, se erosionan los presupuestos de la responsabilidad internacional. En este escenario, la exigencia de trazabilidad adquiere un carácter estructural como condición para evitar zonas de impunidad.

La trazabilidad se vincula estrechamente con la supervisión humana significativa, entendida como la capacidad real —y no meramente formal— de intervenir, modificar o interrumpir el funcionamiento de un sistema autónomo en caso de riesgo de violación del DIH. Este estándar se articula con el deber de diligencia debida del Estado, que impone la adopción de medidas razonables para prevenir daños previsibles derivados del uso de tecnologías militares avanzadas. La utilización de sistemas que no admiten control humano efectivo o que carecen de mecanismos funcionales de interrupción puede configurar una omisión internacionalmente ilícita por falta de previsión y control.

Desde esta perspectiva, la diligencia debida se traduce, al menos, en tres obligaciones complementarias: (i) la identificación y evaluación *ex ante* de los riesgos asociados a la autonomía de los sistemas, así como la imposición de límites operativos claros; (ii) la realización de auditorías técnicas y simulaciones periódicas, en conexión con el deber de revisión jurídica previsto en el artículo 36 del Protocolo Adicional I; y (iii) el mantenimiento de registros verificables que permitan la reconstrucción *ex post* de los procesos decisorios y faciliten la rendición de cuentas. El incumplimiento de estas obligaciones puede generar responsabilidad internacional del Estado aun cuando el daño no haya sido intencional, en la medida en que se trate de consecuencias razonablemente previsibles.

En el ámbito específico de las ciberoperaciones, la diligencia debida implica, además, el monitoreo de infraestructura bajo jurisdicción estatal, el establecimiento de capacidades de detección y respuesta frente a actividades ilícitas, y la cooperación internacional para prevenir el uso del ciberespacio como plataforma para ataques contra otros Estados. En este sentido, la diligencia debida opera tanto como obligación negativa de no permitir la realización de actos ilícitos desde el propio territorio, como obligación positiva de adoptar medidas razonables para impedirlos cuando existan indicios suficientes y capacidad de actuación.

### 5.3. Imputación tecnológica y responsabilidad cooperativa en entornos complejos

La creciente complejidad de los entornos digitales y la fragmentación de las cadenas tecnológicas de diseño, desarrollo y operación de sistemas militares plantean la necesidad de avanzar hacia una lectura funcional de la imputación estatal en contextos caracterizados por altos niveles de interdependencia técnica. Sin sustituir el régimen general de los ARSIWA, esta aproximación permite reforzar criterios de atribución existentes a la luz de los desafíos que plantean la autonomía, la opacidad técnica y la externalización de funciones a actores privados.

En este marco, la responsabilidad estatal tiende a desplazarse desde una lógica centrada exclusivamente en el acto consumado hacia una lógica de gestión del riesgo tecnológico, en la que el énfasis recae en obligaciones de control, previsión y prevención. La imputación se articula así no solo en función del resultado final de una operación, sino también de la conducta estatal a lo largo del proceso de concepción, habilitación y empleo de la tecnología. Este enfoque no debilita el régimen de responsabilidad internacional: lo refuerza al adaptarlo a escenarios donde causalidad y autoría se encuentran mediadas por sistemas técnicos complejos.

Asimismo, el carácter transnacional del desarrollo de software militar, de las infraestructuras digitales y de las ciberoperaciones puede plantear escenarios de pluralidad de Estados potencialmente responsables. Sin embargo, ello no se deriva de "riesgos compartidos" en abstracto, sino de la concurrencia de elementos jurídicamente relevantes, tales como la existencia de una contribución atribuible (por acción u omisión) al hecho ilícito, un nexo suficientemente cercano entre dicha contribución y la violación, y conocimiento o previsibilidad razonable de las consecuencias ilícitas. Los ARSIWA contemplan expresamente la pluralidad de Estados responsables (arts. 47 y 48), lo que ofrece un marco para analizar supuestos de cooperación tecnológica cuando existan vínculos fácticos y normativos suficientes. En cambio, la mera provisión de tecnología, sin control, dirección o asistencia determinante, no debería conducir automáticamente a responsabilidad concurrente.

Desde una perspectiva prospectiva, la consolidación progresiva de estándares interpretativos en materia de trazabilidad, control humano significativo, diligencia debida tecnológica y cooperación internacional podría contribuir a reducir la brecha de rendición de cuentas generada por la autonomía y el ciberespacio. Garantizar que toda tecnología de guerra permanezca bajo un control humano verificable constituye, en este siglo, una de las vías más concretas para preservar la función protectora del derecho internacional humanitario y mantener la vigencia del principio de responsabilidad como piedra angular del orden jurídico internacional.

## Conclusiones

El presente estudio ha examinado los desafíos que las armas autónomas y las ciberoperaciones plantean al régimen de responsabilidad internacional del Estado y a la aplicación efectiva del derecho internacional humanitario. El análisis permite constatar que el marco jurídico vigente —integrado por normas convencionales del DIH y por principios generales del derecho internacional, tal como han sido sistematizados en instrumentos de referencia como los Artículos sobre Responsabilidad del Estado por Hechos Internacionalmente Ilícitos— conserva una capacidad significativa de adaptación. No obstante, dicho marco enfrenta tensiones crecientes derivadas de la progresiva autonomización tecnológica de la acción estatal y de la participación cada vez más intensa de actores privados en el diseño, desarrollo y empleo de medios bélicos.

En primer lugar, se evidencia que la atribución de conductas constituye el núcleo problemático del régimen de responsabilidad en contextos tecnológicos complejos. Los criterios tradicionales de imputación —fundados en la actuación de órganos del Estado, agentes de facto o personas que actúan bajo dirección o control— se ven tensionados por sistemas capaces de operar con grados variables de autonomía. Sin embargo, la ausencia de intervención humana directa en la ejecución de una operación no excluye la responsabilidad estatal: desplaza el análisis hacia las fases previas de decisión, autorización, programación, despliegue y supervisión. En este sentido, la imputación debe concebirse como un proceso dinámico que acompaña el ciclo de vida de la tecnología, y no como un juicio aislado circunscrito al momento del daño.

En segundo lugar, el principio de diligencia debida emerge como un puente normativo indispensable entre el régimen clásico de responsabilidad y las realidades de los conflictos armados en entornos tecnológicamente avanzados. La obligación estatal de prevenir actos ilícitos, de garantizar control humano significativo y de asegurar la trazabilidad de procesos decisorios constituye una manifestación contemporánea de los deberes de control y previsión exigidos por el derecho internacional. La omisión de medidas razonables en las fases de diseño, prueba, despliegue o monitoreo de sistemas autónomos o ciberoperaciones puede dar lugar a responsabilidad internacional por omisión, aun en ausencia de intención o conocimiento directo del daño producido.

En tercer lugar, el DIH continúa ofreciendo los criterios sustantivos más robustos para evaluar la licitud de conductas en conflictos armados tecnológicamente avanzados. Los principios de distinción, proporcionalidad y precaución mantienen plena vigencia y operan como límites objetivos al desarrollo y empleo de sistemas autónomos y capacidades cibernéticas con impacto humanitario. Su aplicación efectiva

exige, no obstante, una lectura funcional que integre consideraciones técnicas, operativas y éticas, a fin de evitar que la automatización decisional vacíe de contenido normas protectoras o diluya umbrales de responsabilidad.

El análisis realizado permite concluir que el problema contemporáneo no es la ausencia de normas aplicables, sino las dificultades de implementación en entornos caracterizados por opacidad algorítmica, fragmentación de la cadena decisoria y reducción del control humano efectivo. En este contexto, la responsabilidad internacional debe proyectarse también sobre la gestión del riesgo tecnológico, incorporando estándares de revisión jurídica, transparencia técnica y rendición de cuentas que refuercen la trazabilidad de decisiones con efectos humanitarios adversos.

Asimismo, el trabajo propone avanzar hacia una lectura funcional de la imputación en contextos tecnológicos, capaz de complementar esquemas clásicos de responsabilidad estatal a la luz de transformaciones introducidas por la autonomía y el ciberespacio. Esta aproximación no implica crear un nuevo régimen jurídico, sino una adaptación interpretativa del marco existente orientada a preservar coherencia sistémica y función garantista. En esa línea, la noción de responsabilidad cooperativa o concurrente —en supuestos de participación conjunta o complementaria en el desarrollo o uso de un mismo sistema tecnológico— se perfila como una proyección posible en escenarios de cooperación tecnológica transnacional, siempre sobre la base de los criterios y requisitos del derecho internacional aplicable.

Desde la perspectiva del sistema interamericano, este debate adquiere una relevancia particular. La Organización de los Estados Americanos, a través de los aportes doctrinales del Comité Jurídico Interamericano y del trabajo continuado de sus instancias técnicas y políticas en materia de ciberseguridad, incluida la cooperación y el fortalecimiento de capacidades impulsados en el marco del Comité Interamericano contra el Terrorismo (CICTE) y la Comisión de Seguridad Hemisférica, se encuentra en una posición propicia para promover una reflexión colectiva sobre la responsabilidad estatal en escenarios de conflictos armados en entornos tecnológicamente avanzados e incentivar orientaciones interpretativas no vinculantes, jurídicamente informadas, sobre el uso y control de tecnologías basadas en inteligencia artificial en ámbitos de seguridad y defensa. Ello contribuiría no solo al fortalecimiento del derecho internacional en la región, sino también a preservar el valor central del ser humano en la conducción de los conflictos armados y en la protección de la paz.

En suma, la responsabilidad internacional del Estado frente a las armas autónomas y las ciberoperaciones no requiere la creación de un "nuevo derecho", sino una relectura funcional del marco jurídico existente, fundada en los pilares de la trazabilidad, el control humano significativo y la diligencia debida. Reforzar estos ejes constituye

una vía eficaz para asegurar que el progreso tecnológico no erosione la esencia ética y jurídica del DIH: la protección de la dignidad humana en los conflictos armados.

## Referencias bibliográficas


Cassese, A. (2005). *International law.* (2nd ed.). Oxford University Press.

Comisión de Derecho Internacional. (2001). *Draft articles on responsibility of states for internationally wrongful acts, with commentaries*.

Comité Internacional de la Cruz Roja. (2006). *A guide to the legal review of new weapons, means and methods of warfare (Article 36 of Additional Protocol I)*.

Comité Internacional de la Cruz Roja. (2021). *Autonomy, artificial intelligence and robotics: Technical, military, legal and humanitarian aspects*.

Comité Internacional de la Cruz Roja. (2021a). ICRC position on autonomous weapon systems.

Comité Internacional de la Cruz Roja. (2023). *Autonomy, artificial intelligence and robotics: Technical and legal challenges*.

Comité Jurídico Interamericano. (2022). Segundo informe: El derecho internacional aplicable al ciberespacio (CJI/doc.671/22 rev.2 corr.1). En *Informe anual del Comité Jurídico Interamericano a la Asamblea General 2022* (pp. 311–334). Organización de los Estados Americanos.

Convenio de La Haya (II). (1899). Convenio relativo a las leyes y costumbres de la guerra terrestre (Convenio II) y su Preámbulo (Cláusula de Martens).

Convenios de Ginebra. (1949). Convenios de Ginebra de 12 de agosto de 1949.

Corte Internacional de Justicia. (1949). *Corfu Channel (United Kingdom v. Albania)*, Judgment of April 9, 1949. *ICJ Reports*.

Corte Internacional de Justicia. (1986). *Military and Paramilitary Activities in and against Nicaragua (Nicaragua v. United States of America)*, Judgment. *ICJ Reports*.

Corte Internacional de Justicia. (1996). *Legality of the Threat or Use of Nuclear Weapons*, Advisory Opinion. *ICJ Reports*.

Crawford, J. (2013). *State responsibility: The general part*. Cambridge University Press.

Crootof, R. (2015). The killer robots are here: Legal and policy implications. *Cardozo Law Review, 36*(6), 1837–1915.

Dinstein, Y. (2022). *The conduct of hostilities under the law of international armed conflict* (3rd ed.). Cambridge University Press.

Galliott, J. y Scholz, J. (2021). *Lethal autonomous weapons: Re-examining the law of armed conflict*. Routledge.

Hollis, D. B. (Ed.). (2023). *The Oxford handbook of cybersecurity and international law*. Oxford University Press.

Naciones Unidas, Asamblea General. (2001). Responsabilidad del Estado por hechos internacionalmente ilícitos (A/RES/56/83).

Naciones Unidas, Asamblea General. (2011). Letter dated 12 September 2011 from the Permanent Representatives of China, the Russian Federation, Tajikistan and Uzbekistan to the United Nations addressed to the Secretary-General (A/66/359).

Naciones Unidas, Asamblea General. (2015). Letter dated 9 January 2015 from the Permanent Representatives of China, Kazakhstan, Kyrgyzstan, the Russian Federation, Tajikistan and Uzbekistan to the United Nations addressed to the Secretary-General (A/69/723).

NATO Cooperative Cyber Defence Centre of Excellence. (2015). An updated draft of the Code of Conduct distributed in the United Nations: What's new? CCDCOE.

NATO Cooperative Cyber Defence Centre of Excellence. (2017). *Tallinn manual 2.0 on the international law applicable to cyber operations*. Cambridge University Press.

Negro, D. (2023). La atribución de las conductas bajo el régimen de la responsabilidad internacional de los Estados: Nuevos desafíos para el derecho internacional contemporáneo. En Comité Jurídico Interamericano (Ed.), XLVIII Curso de derecho internacional (pp. 161–204). Organización de los Estados Americanos.

Organización de los Estados Americanos. (2004). *Estrategia Interamericana Integral de Seguridad Cibernética: Un enfoque multidimensional y multidisciplinario para la creación de una Cultura de Seguridad Cibernética*. AG/RES. 2004 (XXXIV-O/04).

Protocolo adicional de 1977 a los Convenios de Ginebra del 12 de agosto de 1949 relativo a la protección de las víctimas de los conflictos armados internacionales (Protocolo I de 1977).

Roscini, M. (2014). *Cyber operations and the use of force in international law*. Oxford University Press.

Sassòli, M. (2024). *International humanitarian law: Rules, controversies, and solutions to problems arising in warfare* (2nd ed.). Edward Elgar Publishing.

Schmitt, M. N. (2019). The principle of distinction in 21st century warfare. *Yale Journal of International Law, 44*(2), 317–350.

Tribunal Penal Internacional para la ex-Yugoslavia. (1999, 15 de julio). Prosecutor v. Duško Tadić, Appeals Judgement (Case No. IT-94-1-A). International Criminal Tribunal for the former Yugoslavia.

Fecha de recepción: 9 de marzo de 2026
Fecha de aprobación: 20 de mayo de 2026