

ARMAS AUTÔNOMAS E CONTROLE HUMANO SIGNIFICATIVO: DESAFIOS DO DIREITO INTERNACIONAL HUMANITÁRIO

**AUTONOMOUS WEAPONS AND MEANINGFUL HUMAN CONTROL:
CHALLENGES OF INTERNATIONAL HUMANITARIAN LAW**

Ciências Sociais Aplicadas • 30/09/2026

REGISTRO DOI: [10.70773/revistatopicos/790654280](https://doi.org/10.70773/revistatopicos/790654280)

Sérgio Ricardo Ayres Rocha¹

Anatólio Rocha dos Santos Junior²

RESUMO

Este artigo examina se o Direito Internacional Humanitário oferece critérios suficientes para proteger civis e atribuir responsabilidade jurídica por decisões de ataque em operações militares com sistemas de armas autônomas. O estudo utiliza como caso instrumental o episódio ocorrido em 6 de julho de 2026, em Zaporizhzhia, no qual, segundo investigação jornalística, um drone guiado por inteligência artificial selecionou computacionalmente um alvo na fase terminal do voo, mas colidiu com um prédio residencial ao tentar contorná-lo, matando três civis. A pesquisa adota abordagem qualitativa e jurídico-dogmática, com procedimentos bibliográfico e documental, e propõe a cadeia decisória humano-algoritmo como unidade de análise, a partir dos elementos que caracterizam o controle humano juridicamente relevante. A análise confronta o episódio com as regras de distinção, proporcionalidade, precaução e revisão jurídica de novas armas, além de examinar a atribuição de responsabilidade ao Estado e aos agentes envolvidos. Os resultados indicam que o Direito Internacional Humanitário permanece aplicável e oferece parâmetros materiais relevantes, mas que sua incidência sobre cadeias decisórias parcialmente automatizadas enfrenta dificuldades de previsibilidade, supervisão e prova, sem configurar vazio normativo. Conclui-se pela insuficiência de garantias de verificação e atribuição nessa delegação de funções, o que sustenta regras que restrinjam sistemas imprevisíveis e imponham mecanismos de supervisão, registro e desativação.

Palavras-chave: Direito Internacional Humanitário; Sistemas de Armas Autônomas; Controle Humano Significativo; Responsabilidade Internacional; Inteligência Artificial Militar.

ABSTRACT

This article examines whether International Humanitarian Law offers sufficient criteria to protect civilians and to attribute legal responsibility for the decision to attack in military operations involving autonomous weapon systems. The study uses as an instrumental case the episode that occurred on 6 July 2026 in Zaporizhzhia, in which, according to journalistic investigation, an artificial-intelligence-guided drone computationally selected a target during the terminal phase of its flight, but collided with an apartment building while attempting to go around it, killing three civilians. The research adopts a qualitative and doctrinal legal approach, based on bibliographic and documentary procedures, and proposes the human-algorithm decision chain as its unit of analysis, based on the elements that characterize legally relevant human control. The analysis confronts the episode with the rules of distinction, proportionality, precaution, and the legal review of new weapons, as well as with the attribution of responsibility to the State and to the agents involved. The results indicate that International Humanitarian Law remains applicable and offers relevant substantive parameters, but its application to partially automated decision chains faces difficulties of predictability, supervision, and evidence, without amounting to a normative vacuum. The article concludes that safeguards for verification and attribution are inadequate in this delegation of functions, which supports rules that restrict unpredictable systems and impose mechanisms of supervision, record-keeping, and deactivation.

Keywords: International Humanitarian Law; Autonomous Weapon Systems; Meaningful Human Control; International Responsibility; Military Artificial Intelligence.

1. INTRODUÇÃO

Funções antes reservadas a agentes humanos, como reconhecimento, seleção e emprego da força contra alvos, foram transferidas a sistemas algorítmicos diante da incorporação da inteligência artificial às operações militares. Essa mudança suscita dúvidas sobre a aplicação de normas estruturadas em torno de decisões e avaliações juridicamente atribuíveis a agentes humanos.

O ponto de partida para a compreensão desta questão origina-se na conceituação do que caracteriza, tecnicamente, um sistema de armas autônomo. Segundo a definição oficial do Comitê Internacional da Cruz Vermelha, trata-se de um sistema que, depois de ativado por uma pessoa, "seleciona e aplica força contra alvos sem intervenção humana", iniciando um ataque "em resposta a informações do ambiente recebidas por sensores e com base em um perfil de alvo generalizado" (CICV, 2021, p. 5, tradução nossa).

Essa capacidade de selecionar e aplicar força sem intervenção humana ganhou repercussão internacional com o episódio ocorrido em 6 de julho de 2026, em Zaporizhzhia, quando um drone russo do tipo Molniya teria identificado computacionalmente, na fase terminal do voo, os tanques de propano de um posto de combustível como alvo, mas, ao tentar contornar sem sucesso um prédio de apartamentos interposto na rota, teria colidido com a parede dessa edificação e matado três civis que se abrigavam junto a ela (Kramer, 2026). A reconstrução pública não autoriza afirmar que a máquina decidiu de modo integralmente independente, pois agentes humanos desenvolveram o sistema, definiram a área da missão e autorizaram previamente o emprego da força.

Independentemente da tecnologia empregada nesse episódio, permanecem exigíveis as obrigações de distinção,

proporcionalidade e precaução, bem como o dever de revisão jurídica de novas armas (CICV, 1977, arts. 36, 48, 51(5)(b) e 57). Essa aplicação é dificultada, contudo, quando a autorização humana antecede o contato com o alvo e inexiste possibilidade de intervenção na fase terminal. Parte da literatura ainda opõe abstratamente decisão humana e decisão da máquina, representação que não descreve adequadamente operações em que funções são distribuídas entre desenvolvedores, planejadores, comandantes e sistemas algorítmicos. Essa interpretação é encontrada tanto em Amoroso e Tamburrini (2021) quanto em Schwarz (2021).

Essa distribuição de funções impõe delimitar o problema de pesquisa deste artigo. Em que medida as normas do Direito Internacional Humanitário, aplicadas a uma cadeia decisória parcialmente automatizada, oferecem critérios suficientes para proteger civis e atribuir juridicamente a decisão de ataque no emprego de sistemas de armas autônomas, à luz do episódio de Zaporizhzhia? Parte-se da hipótese de que o Direito Internacional Humanitário permanece aplicável, mas suas normas gerais não bastam, isoladamente, para controlar os riscos da delegação de funções críticas a sistemas algorítmicos; a dificuldade reside menos em vazio normativo do que na insuficiência de mecanismos específicos de previsibilidade, supervisão, intervenção, documentação e atribuição. O objetivo geral consiste em analisar essa suficiência normativa, tomando Zaporizhzhia como caso instrumental. Os objetivos específicos são caracterizar a autonomia funcional e a cadeia decisória humano-algoritmo, confrontar o episódio com as regras de distinção, proporcionalidade, precaução e revisão jurídica, e examinar a atribuição da conduta ao Estado e aos indivíduos envolvidos.

Justifica-se o estudo pela urgência prática do tema, intensificada pelos apelos das Nações Unidas e do Comitê Internacional da Cruz Vermelha por regras vinculantes (Nações Unidas, 2024), e por sua relevância para o debate interdisciplinar entre Direito Internacional, governança de tecnologias emergentes e controle da atuação estatal. Pretende-se demonstrar, ao final, que a suficiência do Direito Internacional Humanitário diante de sistemas de armas autônomas depende menos da criação de normas inteiramente novas do que da incorporação, à cadeia decisória humano-algoritmo, de mecanismos específicos de previsibilidade, supervisão, intervenção e rastreabilidade.

2. FUNDAMENTAÇÃO TEÓRICA

Sistemas de armas autônomas caracterizam-se, em termos funcionais, pela capacidade de identificar, selecionar e atacar alvos sem intervenção humana em cada ataque específico após a ativação, o que os distingue de sistemas automatizados ou remotamente operados, nos quais essa decisão permanece sob comando humano (Boulainin *et al.*, 2020; CICV, 2021).

Diante da ausência de definição internacional consensual sobre os limites dessa transferência de funções, o controle humano significativo tornou-se o principal referencial normativo para avaliar a legitimidade dessas armas. Amoroso e Tamburrini (2021) atribuem à intervenção humana três funções: impedir ataques indiscriminados, viabilizar a responsabilização jurídica dos agentes envolvidos e preservar a autoridade humana sobre decisões que afetam a vida das pessoas. Schwarz (2021) alerta, contudo, para o risco de um controle apenas aparente, e Siebert *et al.* (2023)

sustentam que essas condições sociotécnicas devem ser mantidas ao longo de todo o ciclo de vida do sistema.

Rosert e Sauer (2019) deslocam a base normativa do debate para a dignidade humana, argumentando que a decisão de matar em combate deve permanecer exclusivamente humana, ainda que a tecnologia se torne mais precisa. Woodcock (2024) e Dorsey (2025) demonstram que sistemas de apoio à decisão nem sempre avaliam proporcionalidade como comandantes humanos, e Pacholska (2023) situa o problema na perspectiva da responsabilidade estatal.

A literatura converge quanto à necessidade de preservar controle e responsabilidade humanos, mas diverge sobre a suficiência do direito vigente. Gunawan *et al.* (2022) sustentam que os institutos tradicionais de responsabilidade de comando permanecem aptos a responsabilizar o emprego indevido de armas autônomas na maior parte dos casos. Já Fornasier (2021) defende a criação de regras específicas, inclusive proibitivas, por considerar insuficiente a mera extensão do direito vigente. É nesse espaço de divergência que se situa o problema de pesquisa deste artigo, examinado nas seções seguintes a partir da cadeia decisória humano-algoritmo.

3. METODOLOGIA

Adota-se abordagem qualitativa e desenho jurídico-dogmático, de procedimento bibliográfico e documental, com o episódio de Zaporizhzhia (6 jul. 2026) empregado como estudo de caso instrumental para examinar a capacidade explicativa e regulatória das normas de Direito Internacional Humanitário, sem pretensão de investigação pericial autônoma ou de imputação definitiva de responsabilidade.

Para viabilizar essa análise, o corpus reúne quatro categorias de fontes: normativas (Protocolo Adicional I e regras consuetudinárias correlatas); institucionais (Nações Unidas, CICV, SIPRI); literatura científica sobre autonomia, controle humano e responsabilidade; e fontes factuais qualificadas sobre o ataque, com distinção entre a investigação originária e sua mera reprodução por terceiros. Priorizaram-se documentos oficiais e produção científica recente, sem exclusão de referências clássicas indispensáveis.

Com base nesse corpus, a operação foi examinada a partir da cadeia decisória humano-algoritmo, considerando-se, em cada etapa relevante, o agente predominante, a informação disponível, a previsibilidade e a possibilidade de intervenção. O recorte temporal abrange documentos publicados até 11 de setembro de 2026. Constituem limitações a ausência de acesso a planejamento militar, regras de engajamento e laudo pericial independente, razão pela qual as conclusões sobre licitude do ataque são apresentadas de forma condicionada.

4. AUTONOMIA E CONTROLE HUMANO NAS DECISÕES DE ATAQUE

Antes de examinar o episódio de Zaporizhzhia, é necessário construir os critérios com os quais ele será avaliado: o que distingue autonomia funcional de automação; como se organiza a cadeia decisória entre agentes humanos e sistemas algorítmicos; e o que caracteriza um controle humano juridicamente relevante.

4.1. Da Automação à Autonomia Funcional

A automação clássica executa respostas predeterminadas diante de condições previstas, de modo que, se o ambiente apresenta variável

não codificada, o sistema simplesmente não responde. A autonomia funcional surge quando o equipamento, após ativado, processa informações captadas do ambiente e seleciona entre cursos de ação possíveis sem nova intervenção humana, o que não implica consciência ou compreensão jurídica, mas processamento computacional orientado por parâmetros e dados de treinamento (Boulainin *et al.*, 2020). Essa distinção funcional é mais útil, do ponto de vista jurídico, do que a simples presença ou ausência de um operador. Um equipamento remotamente pilotado mantém a decisão de disparo sob apreciação humana em tempo real, ao passo que um sistema autônomo prescinde dessa apreciação na fase terminal.

O risco jurídico varia conforme a função efetivamente transferida à máquina. Um sistema pode navegar autonomamente até coordenadas definidas e depender de autorização humana para atacar. Outro pode receber uma área de operação e, dentro dela, selecionar e atacar objetos correspondentes a determinado perfil, hipótese em que o alvo concreto e o instante da aplicação da força deixam de ser escolhidos individualmente pelo operador.

Não se deve, portanto, classificar uma arma apenas como autônoma ou não autônoma, mas identificar quais funções foram transferidas, por quanto tempo e com quais possibilidades de supervisão, ajustando a exigência de controle ao tipo de alvo e ao ambiente operacional. Amoroso e Tamburrini (2021) sustentam que essa exigência não pode ser uniforme, devendo corresponder à arma e à situação concretas de emprego, como ilustra a diferença entre um sistema defensivo em área isolada e uma arma destinada a selecionar objetos em espaço urbano.

4.2. A Cadeia Decisória Humano-Algoritmo

Identificar quais funções foram transferidas à máquina exige decompor a própria decisão de ataque, que resulta de atos sucessivos de desenvolvimento, treinamento, planejamento, autorização, ativação, reconhecimento e execução, dos quais parte é incorporada antecipadamente ao equipamento. O desenvolvimento e o treinamento não são neutros, pois definem os dados utilizados e o conjunto de objetos que o sistema poderá reconhecer, com seus erros previsíveis. No planejamento, agentes humanos estabelecem finalidade, área e perfis de alvo admissíveis. Ativado o sistema, ele compara o ambiente observado com esses parâmetros e, havendo autonomia de seleção e ataque, os seres humanos autorizam uma classe de resultados possíveis enquanto a máquina determina qual deles ocorrerá no caso concreto.

Essa configuração dificulta atribuir o resultado exclusivamente ao operador, que pode não ter definido os dados de treinamento, o perfil do alvo ou os limites da missão, sem que a dispersão de funções autorize atribuí-lo apenas ao algoritmo, desprovido de personalidade jurídica apta a suportar responsabilidade (Bo; Bruun; Boulanin, 2022). A autonomia também altera a dimensão temporal da decisão. Quando inexiste comunicação na fase terminal, a autorização humana deve antecipar situações supervenientes, tais como o ingresso de civis na área e a perda de função militar do objeto, de modo que, quanto maior a distância entre autorização e ataque, maior deve ser a capacidade de prever essas mudanças.

A expressão "cadeia decisória humano-algoritmo" organiza essa análise sem atribuir personalidade ao sistema, pois localiza onde informações foram produzidas, transformadas ou perdidas, e quais

agentes detinham conhecimento, autoridade e capacidade de intervenção. Kwik (2023) adverte que essa fragmentação gera dificuldades de prova e imputação, mas não implica automaticamente vazio normativo, já que a lacuna de responsabilidade deve ser examinada conforme o sistema e os requisitos aplicáveis, não presumida em abstrato.

Essa proposta distingue-se de um referencial próximo, ao qual poderia ser assimilada sem essa diferenciação. O problema das "muitas mãos", identificado por Kwik (2023) e associado por Bo, Bruun e Boulanin (2022) a situações de responsabilidade distribuída, descreve como múltiplas contribuições causais de diferentes agentes dificultam a responsabilização individual, sem vincular cada contribuição a uma obrigação jurídica específica. A cadeia decisória humano-algoritmo, tal como empregada neste artigo, incorpora esse diagnóstico, mas o reorganiza a partir das obrigações de distinção, proporcionalidade, precaução, revisão jurídica e atribuição de responsabilidade, de modo que cada elo relevante da cadeia corresponda a um requisito jurídico determinado, e não apenas a uma pluralidade causal a ser descrita. Da mesma forma, a cadeia vincula a essas obrigações, como se detalha na seção seguinte, os elementos de controle humano identificados por Boulanin *et al.* (2020) e Siebert *et al.* (2023).

4.3. O Controle Humano Juridicamente Relevante

Cumprir esses requisitos depende de a presença humana na cadeia decisória corresponder a controle real, e não apenas formal. A presença formal de um agente não assegura que ele compreenda o sistema, disponha de informações suficientes ou possua tempo e autoridade para modificar o resultado, e a supervisão pode tornar-se

meramente simbólica quando o operador apenas confirma resultados algorítmicos que não consegue avaliar criticamente, sob pressão temporal e confiança excessiva na automação (Schwarz, 2021).

O controle humano juridicamente relevante consiste, portanto, no controle real, e não meramente formal, exercido sobre a cadeia decisória humano-algoritmo, na medida necessária para permitir o cumprimento das obrigações de distinção, proporcionalidade, precaução, revisão jurídica e atribuição de responsabilidade. Essa expressão operacionaliza, para fins de análise sob o Direito Internacional Humanitário, o debate mais amplo sobre controle humano significativo, a partir dos elementos identificados por Boulanin *et al.* (2020) e Siebert *et al.* (2023). Esse controle é aferido por seis elementos interdependentes: conhecimento das capacidades e limitações do sistema, suficiente para avaliar sua adequação à missão, ainda que não equivalente a domínio técnico integral; informação contextual sobre o alvo, o ambiente e a presença de civis, já que reconhecer tecnicamente um objeto não equivale a determinar seu estatuto jurídico; previsibilidade dos resultados dentro de limites razoáveis; autoridade jurídica e operacional para modificar parâmetros ou suspender a missão; possibilidade material e temporal de intervenção, cancelamento ou redirecionamento; e rastreabilidade dos dados processados, dos parâmetros empregados e dos agentes envolvidos, com função preventiva e retrospectiva (Siebert *et al.*, 2023; Boulanin *et al.*, 2020).

Esses elementos são interdependentes: um agente pode ter autoridade para intervir sem receber informação adequada, ou compreender o sistema sem acompanhar a operação. Além disso, conectam-se diretamente às normas do Direito Internacional

Humanitário: distinção exige informação e qualificação; proporcionalidade depende de previsão de danos; precaução pressupõe autoridade e possibilidade de cancelamento; revisão jurídica requer conhecimento das limitações; responsabilidade demanda rastreabilidade. A legalidade do emprego de um sistema autônomo não pode, assim, ser aferida pela mera existência de autorização humana em algum momento da operação, mas pela combinação efetiva desses elementos diante dos riscos concretos, critério que orienta a análise do episódio de Zaporizhzhia na seção seguinte.

5. O ATAQUE DE ZAPORIZHZHIA PERANTE O DIREITO INTERNACIONAL HUMANITÁRIO

O conflito armado entre a Federação Russa e a Ucrânia qualifica-se como conflito armado internacional, nos termos do artigo 2º comum às Convenções de Genebra de 1949 (CICV, 1949), por opor forças de dois Estados soberanos desde a invasão iniciada em 2022. Ambos os Estados são partes do Protocolo Adicional I de 1977 (CICV, [s. d.]), o que torna os artigos 36, 48, 51(5)(b), 52 e 57 aplicáveis como direito convencional diretamente vinculante entre eles. A declaração russa de 2019, que retirou o reconhecimento da competência da Comissão Internacional de Apuração de Fatos prevista no artigo 90, não afeta a vigência dessas normas substantivas (CICV, [s. d.]).

As regras de distinção, proporcionalidade e precaução refletem, ademais, direito internacional consuetudinário (Henckaerts; Doswald-Beck, 2005), o que reforça sua aplicabilidade independentemente do vínculo convencional. Por sua vez, a obrigação de revisão jurídica de novas armas do artigo 36

permanece fundamentada no vínculo convencional dos Estados-partes do Protocolo.

5.1. Reconstrução Crítica do Episódio

Segundo investigação de Kramer (2026), o drone Molniya empregado em Zaporizhzhia em 6 de julho de 2026 utilizou processamento embarcado, identificado por perícia como módulo Nvidia Jetson Orin, para reconhecer e selecionar, na fase terminal, o alvo, sem comunicação por rádio nessa etapa. O sistema identificou computacionalmente os tanques de propano do posto de combustível como alvo, mas, na aproximação final, tentou contornar um prédio de apartamentos interposto na rota e não completou a manobra, colidindo com a parede dessa edificação antes de alcançá-los. A explosão resultante matou três civis que se abrigavam junto a essa parede, entre eles uma estudante universitária de 19 anos. Agentes humanos desenvolveram o sistema, definiram a área da missão, ativaram o equipamento e autorizaram previamente o emprego da força. Segundo a reconstrução pública disponível, a falha aparente situou-se na execução da manobra terminal de desvio do obstáculo, e não no reconhecimento computacional do alvo previamente definido.

A convergência das matérias posteriores não representa confirmação independente, pois em grande parte reproduzem a investigação originária. Tampouco estão publicamente disponíveis o plano operacional, as regras de engajamento ou laudo pericial internacional. Segundo notícia divulgada por Diallo (2026), o Alto Comissário das Nações Unidas para os Direitos Humanos referiu-se a relatos de drones totalmente autônomos ao comentar o episódio, sem afirmar investigação própria da organização. O caso é tratado,

portanto, como cenário factual publicamente documentado, mas incompleto, sem pretensão de concluir que a inteligência artificial decidiu autonomamente toda a operação nem de realizar imputação definitiva de responsabilidade.

5.2. Distinção e Qualificação do Objetivo

Identificado esse alvo aparente, o posto de combustível e os tanques de propano que o drone teria selecionado, cabe examinar se essa seleção observou o princípio da distinção. O artigo 48 do Protocolo Adicional I exige distinguir a população civil dos combatentes e os bens civis dos objetivos militares. O artigo 52(2) qualifica como objetivo militar o bem cuja natureza, localização, finalidade ou utilização contribua efetivamente para a ação militar, com destruição que ofereça vantagem definida nas circunstâncias do momento. O artigo 52(3) determina que se presuma utilização civil em caso de dúvida sobre bem normalmente destinado a fins civis (CICV, 1977).

Um posto de combustível e seus tanques constituem, em princípio, bens civis, cuja qualificação como objetivo depende de utilização militar efetiva que os dados públicos não esclarecem, pois não se sabe se a instalação abastecia forças em operação, integrava cadeia logística militar ou qual vantagem seria esperada de sua destruição.

O episódio evidencia, sobretudo, a distinção entre classificação computacional e qualificação jurídica, pois um sistema de visão computacional pode reconhecer formas compatíveis com tanques de propano sem que essa capacidade demonstre, por si, contribuição efetiva para a ação militar. Se o sistema foi programado para reconhecer genericamente "infraestrutura de combustível",

sem mecanismo apto a diferenciar sua utilização civil ou militar, há risco elevado de converter uma categoria visual em presunção indevida de objetivo militar (Boulanin *et al.*, 2020). O princípio da distinção não proíbe o emprego de sensores ou classificadores automáticos, mas exige que a escolha final se apoie em informações suficientes para justificar juridicamente o ataque. A autonomia pode auxiliar a identificação, mas não substitui esse juízo.

5.3. Proporcionalidade

Ainda que o alvo fosse corretamente qualificado como objetivo militar, sua destruição permaneceria sujeita a um segundo juízo, o da proporcionalidade entre o dano civil esperado e a vantagem militar concreta. O artigo 51(5)(b) do Protocolo Adicional I proíbe ataques dos quais se possa esperar dano incidental a civis ou bens civis excessivo em relação à vantagem militar concreta e direta prevista (CICV, 1977). Essa formulação prospectiva, ancorada nos próprios termos do dispositivo, é interpretada como exigindo avaliação realizada a partir de informações razoavelmente disponíveis no momento da decisão, e não aferida pelo número de vítimas efetivamente produzido.

A morte de três civis não comprova, isoladamente, que o ataque tenha sido desproporcional. Essa conclusão exigiria conhecer a vantagem militar esperada, a utilização da instalação, a carga explosiva do drone e os efeitos secundários previsíveis, dados não integralmente disponíveis. A configuração relatada, contudo, exigia avaliação rigorosa, dada a proximidade entre instalação inflamável e edificações em ambiente urbano, com possibilidade de explosões secundárias e propagação do fogo.

Isso depende de como cada sistema traduz esse juízo em critérios operacionais. Woodcock (2024) e Dorsey (2025) sustentam que sistemas de apoio à decisão convertem a proporcionalidade em problema de cálculo, produzindo resultados que não correspondem ao juízo qualitativo, contextual e normativo esperado de um comandante razoável. Tais sistemas comprometem, assim, a própria finalidade da avaliação jurídica que deveriam subsidiar. Essa conversão da proporcionalidade em problema de cálculo reforça, no presente caso, que a autonomia terminal não afasta a proporcionalidade, apenas desloca parte de sua avaliação para momento anterior ao lançamento. Quanto maior a distância temporal e informacional entre autorização e o disparo, maior deve ser a cautela na delimitação da missão. Se os efeitos não puderem ser previstos com segurança suficiente, o sistema não deveria ser empregado naquele contexto.

5.4. Precauções e Revisão Jurídica da Arma

Essa mesma incerteza sobre os efeitos previsíveis remete ao dever de precaução, que impõe reavaliar continuamente a adequação do ataque, e não apenas no momento da autorização inicial. O artigo 57 do Protocolo Adicional I obriga quem planeja ou decide um ataque a fazer tudo que for factível para verificar se o objetivo é militar, escolher meios e métodos que evitem ou reduzam danos civis, e cancelar ou suspender a operação quando se torne evidente que o alvo perdeu caráter militar ou que os danos esperados serão excessivos (CICV, 1977).

A ausência de comunicação na fase terminal apresenta efeito ambivalente. De um lado, resiste a interferência eletrônica; de outro, aparentemente impede que o operador redirecione ou suspenda a

missão diante de circunstâncias supervenientes. Quanto menor a possibilidade de intervenção em ambiente dinâmico, mais restritas devem ser duração, área, escala e perfil de alvos autorizados (Boulanin *et al.*, 2020). A colisão com o prédio de apartamentos interposto na rota, ocorrida quando o sistema tentava contorná-lo sem sucesso para alcançar o alvo, ilustra esse risco. A confiabilidade da navegação terminal em ambiente urbano denso, sujeito a obstáculos não previstos, deveria ter integrado a decisão sobre a adequação do próprio meio empregado naquele contexto.

Soma-se a esse dever o de revisão jurídica previsto no artigo 36, que exige aos Estados determinar, no desenvolvimento ou na adoção de nova arma, se seu emprego seria proibido pelo Direito Internacional em algumas ou todas as circunstâncias. Em sistemas autônomos, essa revisão não pode se limitar à estrutura física do equipamento. Deve abranger o uso previsto da capacidade autônoma e o ambiente operacional em que os dados de entrada e saída serão processados, o alcance temporal da revisão diante de eventual aprendizado de máquina ou alteração dos dados após a implantação, e a definição de mecanismo apto a assegurar controle humano e a correspondente responsabilização pelo emprego do sistema (Copeland; Liivoja; Sanders, 2023).

Sistemas que aprendem ou se modificam após a revisão inicial demandam, adicionalmente, acompanhamento contínuo, pois alterações relevantes de software ou de dados podem afastar o funcionamento efetivo daquele originalmente avaliado (Assaad; McFarland, 2023). Não há informação pública suficiente para afirmar se o drone empregado em Zaporizhzhia foi submetido a tal revisão. A opacidade impede verificar se o reconhecimento de infraestrutura civil e a impossibilidade de intervenção terminal foram examinados

previamente. A análise confirma, assim, que as normas de distinção, proporcionalidade, precaução e revisão jurídica permanecem aplicáveis ao caso. O problema não reside em ausência normativa, mas na dificuldade de assegurar que sistemas autônomos disponham de informações, limites e mecanismos de controle compatíveis com essas obrigações.

6. RESPONSABILIDADE JURÍDICA E NECESSIDADE DE REGULAÇÃO ESPECÍFICA

6.1. Responsabilidade Internacional do Estado

A autonomia do sistema não interrompe a atribuição da operação ao Estado, pois, nos termos dos artigos 2º e 4º do Projeto de Artigos sobre Responsabilidade dos Estados, existe ato internacionalmente ilícito quando conduta atribuível a órgão estatal, inclusive militar, viola obrigação internacional vigente, sendo estatais as ações e omissões de seus órgãos (Nações Unidas, 2001). Se o drone integrava operação das forças armadas russas, sua atuação permanece imputável à Federação Russa ainda que funções específicas tenham sido executadas autonomamente. O sistema constitui, nesse caso, meio de execução empregado por seus órgãos, não sujeito independente capaz de interromper a imputação, inclusive quando a atuação estatal excede instruções ou autoridade, conforme o artigo 7º do mesmo Projeto de Artigos.

A atribuição não implica, por si só, ilicitude. É necessário demonstrar a violação de obrigação aplicável, como distinção, proporcionalidade ou precaução, o que os dados públicos ainda não permitem concluir com segurança quanto à instalação atingida em Zaporizhzhia. Comprovado eventual descumprimento, contudo, o Estado não

poderia afastar sua responsabilidade alegando que o resultado foi selecionado pelo algoritmo. Tampouco bastaria a mera alegação de imprevisibilidade algorítmica para caracterizar força maior nos termos do artigo 23 do Projeto de Artigos sobre Responsabilidade dos Estados, que exige força irresistível ou evento imprevisto, fora do controle estatal, capaz de tornar materialmente impossível o cumprimento da obrigação. A excludente também não se aplica quando a situação decorre, isoladamente ou em combinação com outros fatores, da conduta do próprio Estado que a invoca (Nações Unidas, 2001).

Pacholska (2023) argumenta que a introdução de inteligência artificial militar não desloca o fundamento da responsabilidade estatal, pois o vínculo entre o Estado e o resultado do ataque continua a se formar no momento em que o comandante decide autorizar o emprego da arma dentro dos parâmetros do Direito Internacional Humanitário, independentemente do grau de autonomia técnica incorporado ao sistema. A autora propõe, nesse sentido, que sistemas letais plenamente autônomos e baseados em aprendizado de máquina sejam tratados, se um dia empregados em combate real, como órgãos ou agentes do próprio Estado. Esse entendimento não depende da identificação de uma pessoa criminalmente responsável, pois a responsabilidade internacional do Estado e a responsabilidade penal individual constituem regimes distintos, dado que a primeira decorre da atribuição da conduta e da violação de obrigação internacional, ao passo que a segunda exige conduta individual e elementos subjetivos próprios.

6.2. Responsabilização Individual e Fragmentação Decisória

A responsabilidade penal por crimes de guerra é pessoal e exige, via de regra, intenção ou conhecimento. No âmbito da condução das hostilidades, podem configurar crime, preenchidos os requisitos, ataques intencionalmente dirigidos contra civis ou lançados com conhecimento de que causariam dano civil claramente excessivo em relação à vantagem militar esperada (TPI, 1998).

A distribuição de funções entre programadores, planejadores, comandantes e operadores produz o chamado problema das "muitas mãos", pois diversas contribuições causais não correspondem necessariamente a domínio individual sobre o resultado. A presença de autonomia dificulta, além disso, a investigação dos elementos subjetivos exigidos, já que o programador pode desconhecer a missão futura, o operador pode não compreender as limitações do sistema, e o comandante pode receber avaliações técnicas incompletas (Kwik, 2023).

A incerteza quanto a quem detinha o conhecimento relevante repercute nos mecanismos de responsabilidade penal individual, pois pode dificultar a identificação da contribuição humana, do nexo entre a conduta e o resultado e do elemento subjetivo exigido em cada modalidade de responsabilização. Gunawan *et al.* (2022) demonstram que as doutrinas tradicionais de responsabilidade de comando são tensionadas quando parte da decisão letal é incorporada antecipadamente ao sistema e executada sem comando humano imediato. O artigo 28 do Estatuto de Roma exige, quanto ao superior militar, comando e controle efetivos, conhecimento - real ou aferível pelas circunstâncias - e omissão na adoção das medidas necessárias e razoáveis para prevenir ou reprimir a prática dos crimes ou submeter a matéria às autoridades competentes para investigação e persecução penal (TPI, 1998).

Em operações autônomas, esse controle relevante tende a concentrar-se antes do lançamento, na decisão de empregar determinado sistema sob parâmetros específicos, de modo que a responsabilização individual depende da reconstrução da cadeia decisória: quem conhecia as limitações, quem definiu o perfil de alvo, quem avaliou os riscos e quem possuía autoridade para impedir o emprego. No caso de Zaporizhzhia, os dados disponíveis não permitem identificar condutas individuais nem elementos subjetivos, o que impede afirmar a prática de crime de guerra por agente determinado. Isso evidencia, contudo, como a ausência de documentação e de distribuição clara de funções transforma a fragmentação decisória em dificuldade probatória, sem que a mera imprevisibilidade do resultado deva gerar responsabilidade objetiva incompatível com o princípio da culpabilidade.

6.3. Necessidade de Garantias Regulatórias Específicas

Essa mesma fragmentação não interrompe a atribuição da operação militar ao Estado, mas pode dificultar a demonstração da violação internacional e a individualização das condutas humanas envolvidas, o que evidencia a necessidade de garantias regulatórias complementares à aplicação do direito vigente. A aplicabilidade do Direito Internacional Humanitário não encerra, ainda assim, o debate sobre novas regras. O relatório do Secretário-Geral das Nações Unidas registra preocupação ampla dos Estados com imprevisibilidade, perda de controle e dificuldades de responsabilização, embora permaneçam divergências sobre a forma e o conteúdo de eventual instrumento internacional vinculante (Nações Unidas, 2024).

Sistemas cujos efeitos não possam ser suficientemente compreendidos, previstos e explicados deveriam ser proibidos, pois seu emprego compromete a própria capacidade de limitar os efeitos da força. Os demais sistemas deveriam sujeitar-se a restrições cumulativas, tais como perfis de alvo limitados a objetos inequivocamente militares, delimitação geográfica e temporal da missão, exclusão de ambientes com concentração provável de civis, testes em condições operacionais representativas, supervisão com possibilidade efetiva de desativação, manutenção de registros técnicos e operacionais, e revisão jurídica após alterações relevantes de software ou de dados de treinamento (CICV, 2021). Fornasier (2021) acrescenta que essas restrições operacionais devem ser acompanhadas por padrões prospectivos de responsabilidade e *accountability* aplicáveis também a desenvolvedores e fabricantes, bem como por mecanismos de controle da proliferação de armamentos autônomos para Estados ou grupos cujo compromisso com o Direito Internacional Humanitário seja duvidoso, sob pena de garantias técnicas serem neutralizadas pela ausência de controle sobre quem emprega o sistema.

Essas exigências não substituem os princípios de distinção, proporcionalidade e precaução, pois sua função é criar condições para que sejam efetivamente aplicados e verificados. A rastreabilidade, por exemplo, não determina a licitude do ataque, mas permite demonstrar quais informações estavam disponíveis e como o alvo foi selecionado. O episódio de Zaporizhzhia ilustra essa necessidade, pois um sistema que operava sem comunicação, reconhecia genericamente infraestrutura de combustível e não permitia intervenção terminal exigiria, em área urbana, validações e restrições especialmente rigorosas antes de seu emprego. Conclui-se que a autonomia não cria zona livre de responsabilidade. A

conduta militar permanece atribuível ao Estado e pode, presentes os requisitos, fundamentar responsabilidade individual. Todavia, a opacidade, a antecipação e a fragmentação dificultam a aplicação e a prova das normas existentes, o que justifica a adoção de regras internacionais específicas de previsibilidade, restrição operacional, intervenção e rastreabilidade.

7. CONSIDERAÇÕES FINAIS

A pesquisa confirma a hipótese de que as normas do Direito Internacional Humanitário permanecem aplicáveis ao emprego de sistemas de armas autônomas e oferecem parâmetros materiais para a proteção de civis, mas sua aplicação enfrenta dificuldades de previsibilidade, supervisão e prova diante de cadeias decisórias parcialmente automatizadas. Isso não significa vazio normativo, mas insuficiência de garantias específicas de verificação e atribuição. O episódio de Zaporizhzhia ilustra um problema que transcende o caso isolado, pois categorias técnicas, como aquilo que um sensor reconhece ou o momento em que um sistema permite a intervenção humana, não se convertem automaticamente em categorias jurídicas. Essa conversão depende de mecanismos de governança, supervisão e rastreabilidade distribuídos ao longo da cadeia decisória humano-algoritmo investigada neste artigo.

A principal contribuição do estudo é adotar a cadeia decisória humano-algoritmo como unidade de análise, demonstrando que a autonomia não elimina a decisão humana, mas pode antecipá-la, fragmentá-la e afastá-la das circunstâncias concretas da aplicação da força. A conduta militar permanece atribuível ao Estado, e a responsabilidade individual depende da reconstrução do conhecimento, da autoridade e das contribuições de cada agente.

Os resultados sustentam a adoção de regras internacionais que proíbam sistemas imprevisíveis e imponham, aos demais, restrições de alvo, tempo, espaço e contexto, testes, registros e mecanismos de desativação.

Essas recomendações devem ser lidas, contudo, à luz das limitações do próprio estudo, pois o caráter recente do caso e a ausência de documentação militar e pericial integralmente pública impõem cautela às inferências sobre a licitude do ataque específico. Pesquisas futuras poderão comparar sistemas autônomos empregados em diferentes conflitos, examinar práticas nacionais de revisão jurídica e avaliar como os elementos do controle humano juridicamente relevante podem ser convertidos em padrões técnicos verificáveis.

REFERÊNCIAS BIBLIOGRÁFICAS

AMOROSO, Daniele; TAMBURRINI, Guglielmo. Toward a normative model of meaningful human control over weapons systems. **Ethics & International Affairs**, v. 35, n. 2, p. 245-272, 2021. DOI: <https://doi.org/10.1017/S0892679421000241>.

ASSAAD, Zena; MCFARLAND, Tim. Legal reviews of in situ learning in autonomous weapons. **Ethics and Information Technology**, v. 25, n. 1, art. 9, 2023. DOI: <https://doi.org/10.1007/s10676-023-09688-9>.

BO, Marta; BRUUN, Laura; BOULANIN, Vincent. **Retaining human responsibility in the development and use of autonomous weapon systems: on *accountability* for violations of international humanitarian law involving AWS**. Stockholm: SIPRI, 2022. DOI: <https://doi.org/10.55163/AHBC1664>.

BOULANIN, Vincent; DAVISON, Neil; GOUSSAC, Netta; PELDÁN CARLSSON, Moa. **Limits on autonomy in weapon systems: identifying practical elements of human control.** Stockholm: SIPRI; Geneva: ICRC, 2020. Disponível em: https://www.sipri.org/sites/default/files/2020-06/2006_limits_of_autonomy_0.pdf. Acesso em: 11 set. 2026.

COMITÉ INTERNACIONAL DA CRUZ VERMELHA (CICV). **Convention (I) for the Amelioration of the Condition of the Wounded and Sick in Armed Forces in the Field.** Geneva, 12 ago. 1949. Disponível em: <https://ihl-databases.icrc.org/en/ihl-treaties/gci-1949/article-2>. Acesso em: 17 set. 2026.

_____. **Protocol additional to the Geneva Conventions of 12 August 1949, and relating to the protection of victims of international armed conflicts (Protocol I).** Geneva, 8 jun. 1977. Disponível em: https://www.icrc.org/sites/default/files/external/doc/en/assets/files/other/icrc_002_0321.pdf. Acesso em: 11 set. 2026.

_____. **ICRC position on autonomous weapon systems.** Geneva, 12 maio 2021. Disponível em: <https://www.icrc.org/en/document/icrc-position-autonomous-weapon-systems>. Acesso em: 11 set. 2026.

_____. **IHL Treaties: Additional Protocol (I) to the Geneva Conventions, 1977, state parties.** [s. d.]. Disponível em: <https://ihl-databases.icrc.org/en/ihl-treaties/api-1977/state-parties>. Acesso em: 17 set. 2026.

COPELAND, Damian; LIIVOJA, Rain; SANDERS, Lauren. The utility of weapons reviews in addressing concerns raised by autonomous

weapon systems. **Journal of Conflict and Security Law**, v. 28, n. 2, p. 285-316, 2023. DOI: <https://doi.org/10.1093/jcs/krac035>.

DIALLO, Alpha. AI: Türk urges action before it becomes an 'existential risk to humanity'. **UN News**, Geneva, 7 set. 2026. Disponível em: <https://news.un.org/en/story/2026/09/1168288>. Acesso em: 14 set. 2026.

DORSEY, Jessica. The erosion of human(e) judgement in targeting? Quantification logics, AI-enabled decision support systems and proportionality assessments in IHL. **International Review of the Red Cross**, v. 107, p. 1041-1071, 2025. DOI: <https://doi.org/10.1017/s1816383125100969>.

FORNASIER, Mateus de Oliveira. The regulation of the use of artificial intelligence (AI) in warfare: between International Humanitarian Law (IHL) and Meaningful Human Control. **Revista Jurídica da Presidência**, v. 23, n. 129, p. 67-93, 2021. DOI: <https://doi.org/10.20499/2236-3645.rjp2021v23e129-2229>.

GUNAWAN, Yordan; AULAWI, Muhamad Haris; ANGGRIAWAN, Rizaldy; PUTRO, Tri Anggoro. Command responsibility of autonomous weapons under international humanitarian law. **Cogent Social Sciences**, v. 8, n. 1, art. 2139906, 2022. DOI: <https://doi.org/10.1080/23311886.2022.2139906>.

HENCKAERTS, Jean-Marie; DOSWALD-BECK, Louise (ed.). **Customary international humanitarian law: volume I: rules**. Cambridge: Cambridge University Press; Geneva: ICRC, 2005. Disponível em: <https://www.icrc.org/sites/default/files/external/doc/en/assets/files/oth>

[er/customary-international-humanitarian-law-i-icrc-eng.pdf](#). Acesso em: 15 set. 2026.

KRAMER, Andrew E. A drone killed three Ukrainians. It was guided entirely by A.I. **The New York Times**, New York, 24 ago. 2026. Disponível em: <https://www.nytimes.com/2026/08/24/world/europe/russia-drones-autonomous-ai-kill-ukraine-war.html>. Acesso em: 11 set. 2026.

KWIK, Jonathan. The conceptual roots of the criminal responsibility gap in autonomous weapon systems. **Melbourne Journal of International Law**, v. 24, n. 1, p. 1-26, 2023. Disponível em: https://law.unimelb.edu.au/__data/assets/pdf_file/0006/4970067/01-Kwik-unpaginated.pdf. Acesso em: 11 set. 2026.

NAÇÕES UNIDAS. Comissão de Direito Internacional. **Responsibility of States for internationally wrongful acts**. 2001. Anexo à Resolução 56/83 da Assembleia Geral. Disponível em: https://legal.un.org/ilc/texts/instruments/english/draft_articles/9_6_2001.pdf. Acesso em: 11 set. 2026.

_____. Assembleia Geral. **Lethal autonomous weapons systems: report of the Secretary-General**. Documento A/79/88. New York, 2024. Disponível em: <https://digitallibrary.un.org/record/4055753>. Acesso em: 11 set. 2026.

PACHOLSKA, Magdalena. Military artificial intelligence and the principle of distinction: a State responsibility perspective. **Israel Law Review**, v. 56, n. 1, p. 3-23, 2023. DOI: <https://doi.org/10.1017/s0021223722000188>.

ROSERT, Elvira; SAUER, Frank. Prohibiting autonomous weapons: put human dignity first. **Global Policy**, v. 10, n. 3, p. 370-375, 2019. DOI: <https://doi.org/10.1111/1758-5899.12691>.

SCHWARZ, Elke. Autonomous weapons systems, artificial intelligence, and the problem of meaningful human control. **Philosophical Journal of Conflict and Violence**, v. 5, n. 1, p. 53-72, 2021. DOI: <https://doi.org/10.22618/TP.PJCV.20215.1.139004>.

SIEBERT, Luciano Cavalcante *et al.* Meaningful human control: actionable properties for AI system development. **AI and Ethics**, v. 3, n. 1, p. 241-255, 2023. DOI: <https://doi.org/10.1007/s43681-022-00167-3>.

TRIBUNAL PENAL INTERNACIONAL (TPI). **Rome Statute of the International Criminal Court**. Rome, 17 jul. 1998. Disponível em: <https://www.icc-cpi.int/sites/default/files/Publications/Rome-Statute.pdf>. Acesso em: 11 set. 2026.

WOODCOCK, Taylor Kate. Human/machine(-learning) interactions, human agency and the International Humanitarian Law proportionality standard. **Global Society**, v. 38, n. 1, p. 100-121, 2024. DOI: <https://doi.org/10.1080/13600826.2023.2267592>.

¹ Graduação. Especialização. Mestre em Administração. Mestrando em Legal Studies, Emphasis in International Law pela Must University. E-mail: [acesse o artigo original para visualizar o e-mail](#)

² Graduação. Especialista em Direito Público. E-mail: [acesse o artigo original para visualizar o e-mail](#)