

ÉTICA E INTELIGÊNCIA ARTIFICIAL EM ARMAS AUTÔNOMAS: IDENTIFICAÇÃO DE LACUNAS REGULATÓRIAS POR MEIO DE UMA ANÁLISE ESTRUTURADA COM *SOFT SYSTEMS METHODOLOGY*

Renan Vieira de Cerqueira

Instituto Tecnológico de Aeronáutica (ITA), São José dos Campos/SP – Brasil
Praça Marechal Eduardo Gomes, 50 - Vila das Acácias, São José dos Campos - SP, 12228-900
renan_izl@icloud.com

Cláudio Marcos Dias Peixoto

Instituto Tecnológico de Aeronáutica (ITA), São José dos Campos/SP – Brasil
Praça Marechal Eduardo Gomes, 50 - Vila das Acácias, São José dos Campos - SP, 12228-900
claudio82bsb@gmail.com

Mischel Carmen Neyra Belderrain

Instituto Tecnológico de Aeronáutica (ITA), São José dos Campos/SP – Brasil
Praça Marechal Eduardo Gomes, 50 - Vila das Acácias, São José dos Campos - SP, 12228-900
carmen.beld@gmail.com

Níssia Carvalho Rosa Bergiante

Universidade Federal Fluminense (UFF), Niterói/RJ – Brasil
Rua Miguel de Frias, 9, Icaraí. Niterói – RJ – CEP 24220-900
nissiabergiante@id.uff.br

Lamartine Nogueira Frutuoso Guimarães

Instituto Tecnológico de Aeronáutica (ITA), São José dos Campos/SP – Brasil
Praça Marechal Eduardo Gomes, 50 - Vila das Acácias, São José dos Campos - SP, 12228-900
lamar.guima@gmail.com

RESUMO

Este estudo analisa os desafios éticos, jurídicos e geopolíticos associados ao uso de sistemas de armas letais autônomos baseados em Inteligência Artificial. Por meio da *Soft Systems Methodology*, estruturou-se a problemática considerando múltiplos atores e visões de mundo. Foram identificadas lacunas regulatórias, riscos humanitários e limitações técnicas desses sistemas em contextos de conflito. A modelagem permitiu propor transformações desejáveis, tarefas estruturadas e critérios de avaliação, além de destacar a importância da supervisão humana em decisões letais. O estudo enfatiza a necessidade urgente de regulamentações internacionais robustas e de governança ética para mitigar os impactos negativos da automação bélica.

PALAVRAS CHAVE. *Soft Systems Methodology*, Inteligência Artificial, Sistemas Autônomos.

Tópicos: D&SP – PO em Defesa e Segurança; IA – PO e IA.

ABSTRACT

This study analyzes the ethical, legal, and geopolitical challenges associated with the use of lethal autonomous weapon systems based on Artificial Intelligence. Through the application of *Soft Systems Methodology*, the problem was structured by considering multiple stakeholders and worldviews. Regulatory gaps, humanitarian risks, and technical limitations of such systems in

conflict scenarios were identified. The modeling process enabled the proposal of desirable transformations, structured tasks, and evaluation criteria, while highlighting the importance of human supervision in lethal decision-making. The study emphasizes the urgent need for robust international regulations and ethical governance to mitigate the negative impacts of military automation.

KEYWORDS. Soft Systems Methodology, Artificial Intelligence, Autonomous Systems.

Paper topics: D&SP – OR in Defense and Security; AI – OR e AI.

1. Introdução

O avanço da Inteligência Artificial (IA) no domínio militar tem intensificado debates sobre segurança, ética e governança internacional. A incorporação de sistemas autônomos em armamentos e drones, muitas vezes operando com mínima ou nenhuma supervisão humana, potencializa riscos estratégicos, amplia assimetrias de poder e fomenta uma nova corrida armamentista digital. Esse cenário é agravado pela ausência de um marco regulatório global robusto, o que cria lacunas normativas e dificulta a responsabilização por decisões letais tomadas por máquinas inteligentes [Afina e Paoli 2024].

Organizações internacionais têm alertado que o desenvolvimento não regulamentado de Sistemas Letais de Armas Autônomas (LAWS, do termo em inglês *Lethal Autonomous Weapon System*) representa “uma ameaça concreta à estabilidade e à paz internacional”, demandando “mecanismos preventivos, revisões legais e marcos éticos vinculantes” [UNODA 2023]. O Comitê Internacional da Cruz Vermelha reforça que o uso autônomo da força requer salvaguardas mínimas, como o controle humano significativo e o cumprimento irrestrito do Direito Internacional Humanitário (DIH) [ICRC e Geneva Academy 2024]. As discussões recentes no âmbito da ONU têm avançado no reconhecimento de que as obrigações jurídicas e morais devem recair sobre os Estados e indivíduos, não sobre as máquinas [United Nations 2023].

Além da dimensão jurídica, o problema é inerentemente complexo, pois envolve múltiplos *stakeholders*, como governos, empresas de tecnologia, organizações multilaterais, forças armadas, grupos armados e a sociedade civil, que operam com visões de mundo divergentes, interesses conflitantes e níveis assimétricos de poder e conhecimento técnico. Trata-se, portanto, de um problema multifacetado, permeado por incertezas tecnológicas, dilemas morais e disputas geopolíticas [Afina e Paoli 2024].

Casos recentes reforçam a urgência do debate. A utilização de sistemas de IA para selecionar alvos em operações militares já está sendo empregada em conflitos reais, como relatado em reportagens sobre o uso de IA por forças armadas em Gaza, gerando preocupações éticas e acusações de violação de direitos humanos [Apps 2024].

Diante deste cenário, o presente estudo propõe a aplicação da *Soft System Methodology* (SSM) como instrumento de Estruturação de Problemas (*Problem Structuring Method* – PSM) na área de Pesquisa Operacional *Soft*, voltado à análise participativa e crítica do uso bélico da IA. A SSM oferece um arcabouço conceitual adequado para lidar com problemas mal estruturados, onde há multiplicidade de perspectivas, ambiguidade normativa e elevada incerteza. O objetivo é fomentar o debate público, propor caminhos regulatórios e promover a conscientização sobre os riscos e responsabilidades associados ao emprego da IA em contextos de conflito armado.

Esta pesquisa está estruturada em cinco seções: revisão teórica, análise dos desafios éticos e geopolíticos, modelagem do problema com base na SSM, discussão dos resultados e recomendações para pesquisas futuras com abordagens complementares.

2. Panorama dos desafios éticos e geopolíticos da IA e seus impactos

2.1. Inteligência Artificial em aplicações militares

A crescente incorporação de sistemas de IA em aplicações militares, como vigilância, reconhecimento e armamentos letais autônomos, tem suscitado preocupações significativas quanto

à opacidade decisória, confiabilidade técnica e ausência de mecanismos robustos de responsabilização. A opacidade dos algoritmos de IA dificulta a compreensão das decisões tomadas, comprometendo a conformidade com os princípios do DIH, especialmente no que tange à distinção entre combatentes e civis e à proporcionalidade nos ataques. Além disso, a ausência de um marco regulatório internacional robusto para o uso militar da IA contribui para um vácuo normativo perigoso, dificultando a responsabilização por decisões letais tomadas por máquinas inteligentes [West 2024].

A proliferação de tecnologias de IA em armamentos pode acentuar desigualdades entre Estados, fomentando uma nova corrida armamentista e gerando instabilidade geopolítica. A ausência de um consenso internacional sobre a regulamentação de LAWS tem sido agravada por tensões geopolíticas e falta de liderança política, dificultando a implementação de medidas eficazes de controle e supervisão [Blanchard e Bruun 2024].

2.2. Os efeitos na geopolítica

O impacto da IA na geopolítica contemporânea é profundo, influenciando diretamente a projeção de poder estatal e a configuração das relações internacionais. A crescente assimetria tecnológica entre nações tende a acentuar desequilíbrios estratégicos, expondo países com menor capacidade tecnológica a riscos elevados, como ataques cibernéticos e ofensivas conduzidas por sistemas autônomos de caráter militar. Essa assimetria tecnológica, aliada à ausência de mecanismos multilaterais efetivos de controle, favorece o uso indiscriminado de sistemas autônomos e aprofunda a incerteza estratégica no plano global [Afina e Paoli 2024].

Além disso, como demonstram evidências recentes em zonas de conflito real, o uso da IA militar pode induzir líderes políticos a uma falsa percepção de superioridade tática, incentivando decisões precipitadas ou escaladas de conflito com base em avaliações automatizadas [Apps 2024]. Essa percepção equivocada, associada à aceleração da tomada de decisão por sistemas autônomos, compromete a estabilidade internacional e desafia os princípios de contenção e prudência estratégica [ICRC e Geneva Academy 2024].

2.3. Os efeitos militares e sociais

A presença de sistemas autônomos no contexto militar levanta sérias dúvidas quanto à responsabilização por eventuais crimes de guerra, especialmente em cenários urbanos complexos, nos quais limitações técnicas e decisões baseadas em padrões podem resultar em ataques a civis. A possibilidade de erros na identificação de alvos e a dificuldade em atribuir responsabilidade por decisões tomadas por IA configuram ameaças significativas aos Direitos Humanos e às normas das Convenções de Genebra. Como alerta o Comitê Internacional da Cruz Vermelha, é urgente garantir supervisão humana significativa em decisões letais e submeter os sistemas autônomos a critérios normativos estritos [Blanchard 2024].

Casos recentes demonstram que o uso de IA em cenários reais, como nos conflitos em Gaza e na Ucrânia, já evidencia os riscos de sua aplicação descontrolada. A utilização de IA para identificação de alvos tem sido associada a erros sistemáticos e impactos colaterais severos [Apps 2024]. Tais ocorrências apontam para a necessidade de definição clara do regime jurídico aplicável, além da implementação de restrições normativas e operacionais capazes de conter a expansão irrestrita de armamentos autônomos [ICRC e Geneva Academy 2024].

3. Pesquisa Operacional Soft

A problemática abordada neste estudo envolve múltiplos atores com interesses divergentes, cujas interações ocorrem de forma dinâmica e imprevisível. Tal complexidade escapa às abordagens tradicionais da Pesquisa Operacional (*Hard OR*), voltadas a problemas bem estruturados e consensuais [Mingers e Rosenhead 2004]. Em contraste, os PSMs emergiram justamente para lidar com situações mal estruturadas, caracterizadas por ambiguidade, múltiplas perspectivas e objetivos em conflito.

Os PSMs originaram-se da prática e só posteriormente foram sistematizados, incorporando fundamentos teóricos de distintas áreas do conhecimento. Essa gênese pragmática torna a prática uma dimensão essencial da análise metodológica, ainda que dificulte a categorização precisa dos métodos. Com base nisso, o estudo adota a SSM, concebida como abordagem geral para redesenho organizacional, valorizando o pluralismo cultural e a construção participativa de modelos conceituais [Checkland e Poulter 2020].

Conforme apontam [Mingers e Rosenhead 2004], os PSMs são particularmente adequados para lidar com contextos marcados pela presença de múltiplos atores, diversidade de perspectivas, interesses conflitantes ou incomensuráveis, elementos intangíveis relevantes e elevados níveis de incerteza. Dado que a problemática tratada neste estudo exhibe precisamente essas características, justifica-se a adoção da SSM, uma metodologia da Pesquisa Operacional *Soft* especialmente concebida para enfrentar situações complexas, mal estruturadas e ambíguas, permeadas por múltiplas visões de mundo. A estrutura está apresentada na Figura 1.

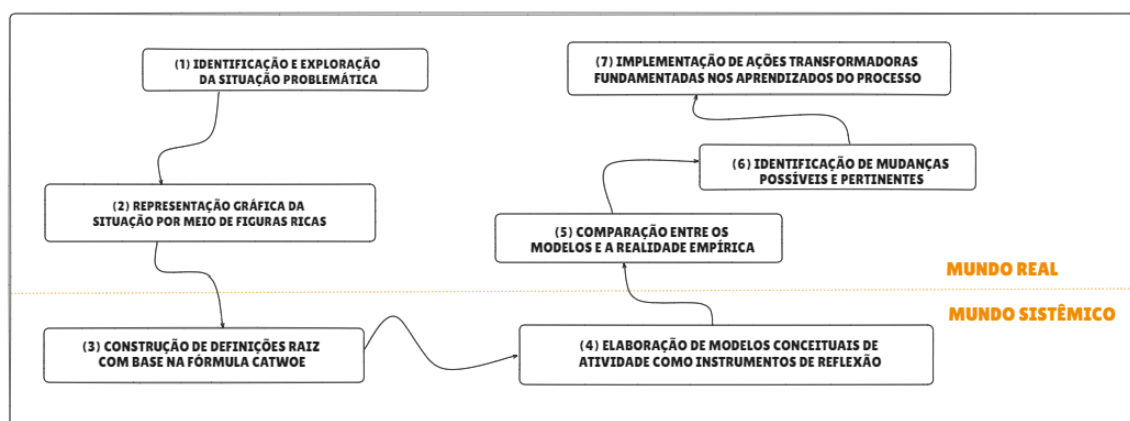


Figura 1 - Etapas de desenvolvimento do SSM

Fonte: adaptado de [Checkland 1981]

A SSM adota uma abordagem qualitativa e interpretativa, voltada à construção colaborativa de entendimentos sobre a realidade social, com o objetivo de promover mudanças que sejam, ao mesmo tempo, sistemicamente desejáveis e culturalmente viáveis. Este trabalho estrutura-se a partir do modelo clássico proposto por [Checkland 1981], que compreende sete estágios interdependentes: (1) identificação e exploração da situação problemática; (2) representação gráfica da situação por meio de figuras ricas; (3) construção de definições raiz com base na fórmula CATWOE; (4) elaboração de modelos conceituais de atividade como instrumentos de reflexão; (5) comparação entre os modelos e a realidade empírica; (6) identificação de mudanças possíveis e pertinentes; e (7) implementação de ações transformadoras fundamentadas nos aprendizados do processo.

4. Modelagem do Problema

4.1. Estágios 1 e 2



Figura 2 – Figura rica da situação problemática abordada no estudo

Na análise realizada no escopo da modelagem do problema, a partir da Figura 2, identificaram-se os principais agentes envolvidos no contexto das armas autônomas letais baseadas em inteligência artificial. Entre os atores mapeados estão os Governos Nacionais, que detêm o poder decisório sobre políticas de segurança e defesa; as Empresas de Tecnologia, enquanto protagonistas no desenvolvimento e comercialização dessas tecnologias; Grupos Terroristas, que podem se apropriar indevidamente das inovações para fins destrutivos; Organizações Mundiais como a ONU, OTAN e outras entidades multilaterais, que atuam como moderadoras e promotoras de consensos normativos; os Acadêmicos, que contribuem para o embasamento crítico e técnico do debate; a População em geral, enquanto sujeito moral e político afetado diretamente pelos riscos dessas tecnologias; e, por fim, a própria Inteligência Artificial atuando de forma autônoma, o que introduz um novo tipo de agente abstrato, não-humano, com potencial para causar danos irreversíveis.

Além da identificação dos atores, a análise também trouxe à tona importantes dinâmicas sociais e políticas que moldam o desenvolvimento, a regulação e o uso dessas tecnologias. Entre elas, destaca-se a percepção pública sobre os riscos e potencialidades da IA militar, muitas vezes permeada por desinformação ou otimismo tecnológico acrítico. Soma-se a isso a dificuldade de quebra de paradigmas e a resistência cultural à aceitação dos riscos inerentes à automação de decisões letais. O avanço tecnológico traz consigo riscos ainda não totalmente compreendidos, o que se agrava frente aos interesses financeiros de *stakeholders* influentes, como corporações tecnológicas e setores militares-industriais. Também foi evidenciada a dificuldade das organizações internacionais em harmonizar interesses e agendas políticas distintas entre os Estados-membros, além da urgência de se estabelecer mecanismos claros de responsabilização por decisões que resultem em perdas de vidas humanas. Outro fator crítico é o uso indevido da tecnologia por atores que atuam à margem da legislação internacional, como grupos terroristas ou regimes autoritários.

Por fim, observou-se que cada agente atua com diferentes capacidades de influência na modelagem do problema. Os Governos Nacionais, por meio da imprensa e da retórica política, podem moldar a opinião pública e direcionar o debate público. As Empresas de tecnologia exercem forte influência por meio do *lobby* junto a políticos e meios de comunicação. Os grupos terroristas utilizam a tecnologia de modo oportunista e destrutivo, desafiando os limites normativos. As Organizações Internacionais possuem a prerrogativa de mediar interesses divergentes e criar

diretrizes globais, ainda que enfrentem limitações políticas e institucionais. Os acadêmicos podem contribuir significativamente para a formação da opinião pública, ao passo que a População exerce papel ativo ao pressionar governos por decisões éticas e transparentes. A IA autônoma emerge como um ator disruptivo, com a capacidade de agir de forma imprevisível e, em última instância, provocar consequências irreversíveis à humanidade, o que impõe desafios éticos e regulatórios inéditos na história da guerra e da tecnologia.

4.1.1. Identificação dos sistemas relevantes

A partir da análise estruturada com base na SSM apresentada na primeira etapa da Figura 1, foram identificados três sistemas relevantes cuja interação constitui o núcleo da problemática associada ao uso de armas autônomas letais com base em inteligência artificial e apresentados na Figura 2. Esses sistemas – **Regulamentador, de Defesa e de Desenvolvimento Tecnológico** – representam conjuntos interdependentes de atores, interesses e funções que influenciam decisivamente o desenvolvimento, a implementação e o controle dessas tecnologias.

O **Sistema de Regulamentadores** é composto por Organizações Mundiais (como a ONU e a OTAN), Governos, Forças Armadas e a população. Este sistema tem como função primária o estabelecimento de normas, protocolos e acordos internacionais que definam os limites éticos, legais e operacionais do uso da IA em contextos bélicos. Organizações multilaterais atuam como fóruns de negociação e harmonização de interesses entre nações, ao passo que os governos detêm a capacidade legislativa e executiva para internalizar tais normativas. As Forças Armadas, como usuárias finais da tecnologia, exercem papel central na aplicação e adequação dessas normas aos cenários reais de operação. A População, por sua vez, figura como agente de pressão política e moral, podendo influenciar diretamente as decisões dos representantes estatais e a agenda das organizações internacionais.

O **Sistema de Defesa** compreende a articulação entre os Governos, as Forças Armadas e a População. Esse sistema é responsável por empregar as tecnologias desenvolvidas em contextos de segurança nacional e internacional, sendo o ponto focal da aplicação prática da IA bélica. Os governos, ao formularem políticas de defesa, determinam os limites de engajamento e uso operacional dessas tecnologias. As Forças Armadas são as executoras diretas, operando os sistemas de armas autônomas e gerenciando seu uso tático e estratégico. Os próprios sistemas de armas, enquanto entidades técnico-autônomas, constituem componentes internos do sistema, com crescente grau de independência decisória. A População permanece como elemento vulnerável e, ao mesmo tempo, influente, pois suas reações sociais, culturais e políticas frente ao uso desses sistemas podem retroalimentar a revisão de políticas e estratégias de defesa.

O **Sistema de Desenvolvimento Tecnológico**, por sua vez, agrega atores envolvidos na pesquisa, desenvolvimento e disseminação de soluções em inteligência artificial para uso militar e civil. Compõem esse sistema os Governos (como financiadores e reguladores), as Empresas de Tecnologia (como principais desenvolvedoras), os Acadêmicos (como geradores de conhecimento e crítica especializada), os Grupos Terroristas (como usuários marginais que exploram brechas legais e tecnológicas para fins ilícitos) e a População (como beneficiária ou vítima dos produtos tecnológicos resultantes). Este sistema é particularmente complexo por abrigar interesses econômicos, científicos, geopolíticos e éticos frequentemente conflitantes. A interação entre esses atores define o ritmo e a direção da inovação, ao mesmo tempo em que revela os riscos da assimetria de poder e da lacuna regulatória.

Portanto, os três sistemas – regulamentador, de defesa e de desenvolvimento tecnológico – são interdependentes e formam uma ecologia sociotécnica na qual decisões e transformações em um subsistema impactam significativamente os demais. A compreensão dessa estrutura sistêmica é essencial para propor intervenções viáveis, identificar conflitos éticos e delinear caminhos para uma governança global robusta e responsável das armas autônomas baseadas em IA.

4.2. Estágio 3 e 4

Tabela 1 - CATWOE

CATWOE para transformação 1	
C	Forças Armadas e população
A	Governos
T	Não há regulamentações internacionais unificadas quanto ao desenvolvimento tecnológico da IA e suas aplicações bélicas → Regularizar internacionalmente o desenvolvimento tecnológico da IA e suas aplicações bélicas
W	Balizamento jurídico e ético no uso da tecnologia de IA em fins bélicos
O	Organizações Mundiais
E	Interesses individuais das nações

Definição raiz T1 → Um sistema que atende Forças Armadas e população geral, operado por Governos Nacionais, que têm a capacidade de aplicar a regulamentação internacional sobre o desenvolvimento tecnológico da IA e suas aplicações bélicas em função de balizamento jurídico e ético no uso da tecnologia de IA em fins bélicos, que está sob domínio de Organizações Mundiais e limitada por interesses individuais das nações.

CATWOE para transformação 2	
C	População
A	Forças Armadas
T	A decisão sobre a execução de um alvo pode ser feita automatizado pela IA → A decisão sobre a execução de um alvo deve ser feita com supervisão humana
W	Responsabilização individualizável pela decisão sobre a execução de um alvo
O	Governos
E	Possibilidade de ficar em desvantagem em combate caso a Força oponente não siga o mesmo preceito

Definição raiz T2 → Um sistema que atende população, operado por Forças Armadas, que têm a capacidade de realizar a decisão sobre a execução de um alvo com supervisão humana, para que exista responsabilização individualizável pela decisão sobre a execução de um alvo, que está sob domínio de Governos e limitada pela possibilidade de ficar em desvantagem em combate caso a Força oponente não siga o mesmo preceito.

CATWOE para transformação 3	
C	População
A	Empresa de tecnologia e Acadêmicos
T	A tecnologia de IA que controla máquinas automatizadas armadas amplamente acessível → Acesso restrito à tecnologia de IA que controla máquinas automatizadas armadas
W	Riscos de a tecnologia cair em mão de organizações terroristas e ter o seu uso desvirtuado
O	Governos
E	Interesses econômicos e interesses políticos

Definição raiz T3 → Um sistema que atende a população, operado por empresa de tecnologia e acadêmicos, que têm a capacidade de realizarem acesso restrito à tecnologia de IA que controla máquinas automatizadas armadas, evitando os riscos de a tecnologia cair em mão de organizações terroristas e ter o seu uso desvirtuado, que está sob domínio de Governos e limitada por Interesses econômicos e interesses políticos.

4.2.1. Tarefas necessárias

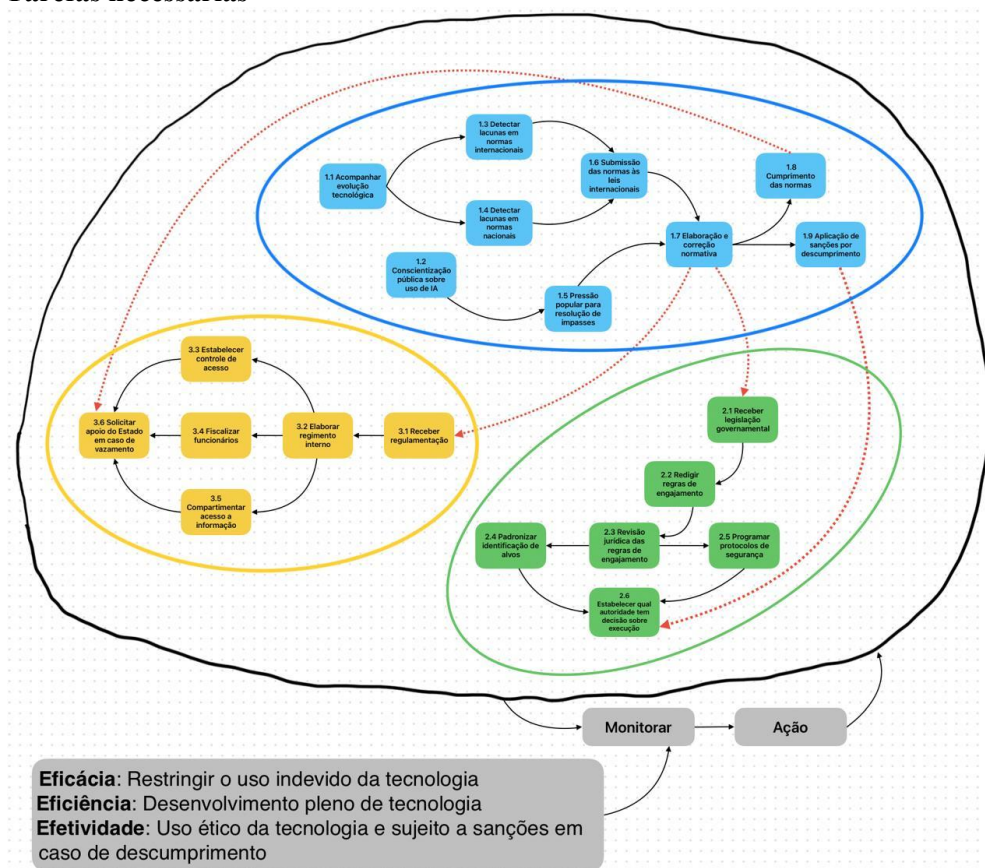


Figura 3 - Supersistema

A construção de um modelo ético, jurídico e operacional para o uso da IA em aplicações bélicas exige a coordenação de um conjunto de tarefas estruturadas, interligadas e distribuídas entre diferentes sistemas relevantes, reguladores, de defesa e de desenvolvimento tecnológico. Tais tarefas devem ser guiadas por três dimensões avaliativas fundamentais: eficácia (alcançar os objetivos propostos), eficiência (atingi-los com o uso racional de recursos e governança) e efetividade (obter impacto concreto e duradouro no sistema internacional).

O primeiro eixo de ação (grupo azul na Figura 3) diz respeito à regulamentação internacional do desenvolvimento tecnológico da IA e suas aplicações militares. Essa tarefa requer inicialmente a identificação de lacunas nos atuais acordos e normas internacionais, muitas das quais não contemplam adequadamente os riscos e especificidades da inteligência artificial em armamentos autônomos. A partir dessa identificação, torna-se necessário corrigir e atualizar continuamente essas normas, incorporando os princípios do DIH e dos Direitos Humanos, de modo que acompanhem o ritmo acelerado das inovações tecnológicas. Para garantir sua eficácia, devem ser previstas sanções concretas para o descumprimento das normas e mecanismos multilaterais para sua fiscalização.

Além disso, é essencial provocar um apelo social, sensibilizando a opinião pública global quanto aos riscos éticos e existenciais do uso indiscriminado da IA em contextos bélicos. A pressão popular pode, assim, servir como força catalisadora para que os governos assumam compromissos vinculantes. A eficácia desta tarefa será verificada quando houver a criação de normas robustas e abrangentes; a eficiência será medida pela adesão dos países desenvolvedores; e a efetividade será alcançada quando todos os membros da ONU se tornarem signatários e cumpridores desses regulamentos.

O segundo conjunto de tarefas fundamentais (grupo verde na Figura 3) refere-se à garantia de supervisão humana nas decisões sobre a execução de alvos. Este ponto é central para evitar a perda de controle ético no uso de armamentos autônomos letais. A primeira etapa é o estabelecimento de padrões rigorosos de identificação de alvos e a redação de regras claras de engajamento, submetidas à revisão jurídica especializada. Tais regras devem definir inequivocamente quem, dentro da cadeia de comando, possui autoridade decisória para a execução de um alvo, assegurando que nenhuma máquina opere de forma letal sem validação humana. Adicionalmente, protocolos de segurança devem ser incorporados aos sistemas de IA embarcados em armamentos para que não haja disparo sem supervisão.

Essa estrutura deve ser respaldada por legislações nacionais, que regulamentem detalhadamente o uso dessa tecnologia. Sua eficácia será aferida pela ausência de vítimas civis; sua eficiência, pela aplicação exclusivamente contra alvos militares legítimos; e sua efetividade, quando os altos comandos das Forças Armadas adotarem tais práticas de forma sistematizada e preventiva.

O terceiro eixo (grupo amarelo na Figura 3) diz respeito ao controle de acesso à tecnologia de IA aplicada a armamentos automatizados. Dada sua natureza sensível e seu potencial destrutivo, é imprescindível regular rigorosamente quem pode desenvolver, manipular e acessar essa tecnologia. Para tanto, é necessário regulamentar a cadeia de acesso, criando regimentos internos baseados na legislação vigente, estabelecendo critérios rigorosos de segurança para o pessoal autorizado e aplicando fiscalização constante. Caso haja suspeita de corrupção ou vazamento de informações, devem ser acionados os órgãos de Estado competentes.

Além disso, recomenda-se a compartimentação dos processos de desenvolvimento tecnológico, impedindo que um único indivíduo ou grupo tenha pleno conhecimento de todos os componentes críticos do sistema. A eficácia será observada no grau de controle sobre o fluxo de informações; a eficiência, na continuidade segura da pesquisa e desenvolvimento; e a efetividade, na limitação do acesso apenas a agentes submetidos às normativas legais.

Esses três eixos operam sob a lógica de um supersistema normativo e operacional (Figura 3), que articula o desenvolvimento tecnológico à governança ética e jurídica. A finalidade do supersistema é, de forma ampla, restringir o uso indevido da tecnologia, sem inibir seu avanço em segurança e capacidade de resposta. Para isso, o sistema deve garantir eficácia ao evitar violações éticas e humanitárias, eficiência ao promover inovação responsável e efetividade ao assegurar que o uso da IA esteja permanentemente submetido ao império das leis, com sanções claras e dissuasivas em caso de descumprimento.

Adicionalmente, tarefas executadas de forma independente, como a redação de regras de engajamento ou a compartimentação técnica, não surtem o impacto desejado, além disso, não possuem legitimidade, uma vez que dependem de uma estrutura integrada de acompanhamento e avaliação. Assim, é fundamental estabelecer critérios de controle para monitorar a execução de cada tarefa em tempo real, garantindo coerência entre os objetivos estratégicos, os padrões éticos e os resultados práticos. Esse modelo integrado de regulação, supervisão e controle se configura como uma das únicas formas viáveis de assegurar que a inteligência artificial, ao adentrar o domínio bélico, não se torne uma ameaça à própria condição humana.

4.3. Estágio 5, 6 e 7

A comparação entre o modelo conceitual do supersistema e a realidade empírica revela uma implementação parcial e fragmentada das transformações propostas. Algumas ações, como a redação e revisão jurídica das regras de engajamento (2.2 e 2.3), bem como a definição da autoridade decisória na cadeia de comando (2.6), já são praticadas de forma consistente. No entanto, grande parte das atividades essenciais para a governança responsável da IA militar permanece subdesenvolvida ou inexistente. Ações como a identificação de lacunas normativas (1.4), a submissão das normas às leis internacionais (1.6), a correção normativa (1.7), a implementação de protocolos de segurança que impeçam disparos autônomos (2.5) e o controle de acesso à tecnologia sensível (3.1 e 3.5) ainda não são realizadas. Outras medidas, como campanhas

de conscientização pública (1.2), a detecção de lacunas em tratados internacionais (1.3), a aplicação de sanções por descumprimento normativo (1.8), e a fiscalização interna de pessoal (3.4), ocorrem apenas de forma esporádica ou reativa. Essa assimetria entre o ideal proposto e a prática vigente evidencia a necessidade urgente de fortalecimento institucional, revisão normativa e engajamento multissetorial para garantir que os princípios éticos e legais acompanhem a evolução tecnológica no campo da inteligência artificial aplicada à defesa.

A partir da modelagem realizada com base na SSM, foram identificadas mudanças sistêmicas que, além de desejáveis do ponto de vista ético, operacional e estratégico, também se mostraram culturalmente viáveis dentro dos contextos institucionais e sociais analisados. A avaliação seguiu o princípio de que qualquer transformação proposta deve atender simultaneamente à lógica sistêmica, ou seja, contribuir para a funcionalidade e coerência do sistema como um todo, e ser compatível com os valores, estruturas e práticas existentes nos diferentes atores envolvidos, como governos, forças armadas, organizações internacionais, empresas de tecnologia e a população.

Portanto, as transformações 1, 2 e 3 formam um conjunto integrado de mudanças que não apenas atendem às exigências sistêmicas da problemática analisada, ao reforçar a segurança, a previsibilidade e a responsabilização, como também se ajustam, sem rupturas, às culturas políticas, jurídicas e institucionais já existentes nos principais sistemas envolvidos. O sucesso na identificação dessas transformações marca um avanço significativo na proposição de soluções estruturadas para os dilemas éticos e operacionais relacionados ao uso da inteligência artificial em armamentos autônomos, pavimentando o caminho para ações concretas, coordenadas e legitimadas globalmente.

Com base nos aprendizados obtidos ao longo da análise, foram delineadas ações estratégicas voltadas à implementação de transformações sistêmicas, com foco na regulação, controle, uso ético e governança responsável da IA aplicada ao domínio militar. No eixo da governança normativa, propõe-se a criação de um comitê permanente de atualização legislativa tecnológica, responsável por revisar periodicamente as normas à luz dos avanços em IA. Essa medida deve ser acompanhada por análises comparativas entre tratados internacionais existentes e a legislação nacional, a fim de identificar lacunas e propor ajustes normativos. A compatibilização das novas regras com os princípios do DIH exige a elaboração de pareceres técnicos-jurídicos, que deverão ser submetidos a instâncias competentes, como o Comitê Internacional da Cruz Vermelha. Em complemento, sugere-se a inclusão de cláusulas específicas sobre o uso ético da IA nas emendas legislativas e a implementação de sistemas de auditoria e sanção para garantir o cumprimento das normas. Para assegurar a aplicação efetiva dessas diretrizes, recomenda-se ainda a realização de treinamentos obrigatórios para operadores militares e desenvolvedores de tecnologia sensível.

No campo da mobilização social e conscientização, torna-se fundamental envolver a população civil na compreensão dos riscos do uso bélico da IA. Campanhas educativas em mídias digitais, documentários e eventos públicos podem auxiliar na construção de uma cultura crítica sobre o tema. Simultaneamente, a articulação com ONGs, por meio de petições e manifestações, pode gerar pressão social para a assinatura de tratados internacionais de controle sobre armamentos autônomos.

Quanto à regulamentação do uso tático, é essencial acolher as legislações que venham a ser sancionadas pelos poderes executivos nacionais, bem como estabelecer regras claras de engajamento operacional. Tais regras devem ser submetidas à revisão jurídica, garantindo sua adequação legal. Para reduzir os riscos de erros críticos, recomenda-se o estabelecimento de padrões rigorosos para a identificação de alvos, com treinamento da IA por meio de aprendizado supervisionado e testes em ambientes controlados. Além disso, devem ser programados protocolos de segurança que impeçam disparos sem supervisão humana e estabelecida uma cadeia de comando clara, restringindo a decisão sobre execuções letais a autoridades militares devidamente capacitadas.

Por fim, o controle de acesso à tecnologia representa um eixo fundamental para a prevenção de abusos e vazamentos. É necessário regulamentar e estruturar o acesso aos ambientes de pesquisa e desenvolvimento de IA aplicada a armamentos, com mecanismos robustos de identificação, monitoramento e registro de entrada. A fiscalização interna deve ser intensificada, com o monitoramento de atividades suspeitas, e a compartimentação do conhecimento técnico deve ser adotada como estratégia preventiva. Em casos de corrupção ou vazamento de informações, recomenda-se o acionamento imediato de órgãos estatais responsáveis pela segurança institucional e aplicação da justiça.

A modelagem da problemática por meio da SSM possibilitou estruturar, de forma coerente e compreensiva, os desafios multifacetados relacionados ao uso da inteligência artificial em sistemas bélicos. A abordagem destacou a presença de múltiplos atores, como governos, organizações internacionais, empresas de tecnologia, forças armadas, acadêmicos, população e até a própria IA como agente autônomo, e revelou as complexas dinâmicas de poder, interesse e responsabilidade entre eles. Através da construção das definições CATWOE e dos modelos conceituais, foi possível compreender a interação entre os sistemas de regulamentação, defesa e desenvolvimento tecnológico, identificando pontos críticos e relações de interdependência.

Esse processo analítico permitiu, de forma fundamentada, identificar transformações sistêmicas desejáveis e culturalmente viáveis, como a regulamentação internacional do uso bélico da IA, a garantia de supervisão humana nas decisões letais e o controle rígido de acesso à tecnologia sensível. A partir dessas transformações, foram propostas tarefas específicas, organizadas em três eixos estratégicos, e definidos critérios claros de eficácia, eficiência e efetividade, tanto para a implementação quanto para a avaliação contínua dessas ações.

O contraste entre os modelos ideais elaborados e a realidade atual evidenciou lacunas preocupantes na regulamentação, fiscalização e responsabilização pelo uso de tecnologias emergentes em contextos militares. Tais lacunas foram agravadas pela velocidade do avanço tecnológico, pela ausência de consensos globais e pela resistência cultural à mudança em setores estratégicos. Nesse cenário, a abordagem sistêmica da SSM mostrou-se particularmente eficaz ao lidar com a ambiguidade, a incerteza e os conflitos de interesse que permeiam o tema. Ela forneceu uma estrutura lógica e dialógica que não apenas organiza o debate, mas também aponta caminhos viáveis para intervenções éticas, seguras e colaborativas, orientadas por princípios de governança internacional, respeito aos direitos humanos e responsabilidade compartilhada. Assim, o estudo avança na direção de uma modelagem propositiva, capaz de contribuir efetivamente para o desenvolvimento de políticas públicas e mecanismos de controle sobre o uso da inteligência artificial em sistemas de armamento autônomo, dentro de uma perspectiva sistêmica e multidisciplinar.

5. Conclusões

Este estudo estruturou, por meio da SSM, alguns dos desafios éticos, jurídicos e geopolíticos associados ao uso da inteligência artificial em armamentos autônomos letais, revelando a complexidade multifacetada da problemática. A modelagem permitiu identificar os principais *stakeholders*, suas dinâmicas de influência e os conflitos de interesse que dificultam a implementação de uma governança global robusta.

Ao construir modelos conceituais e definições CATWOE, o estudo apontou três transformações sistêmicas centrais: regulamentação internacional, supervisão humana obrigatória e controle de acesso à tecnologia, que se mostraram simultaneamente desejáveis e culturalmente viáveis. A análise evidenciou lacunas normativas, riscos de responsabilização difusa e ausência de mecanismos de supervisão consistentes, exigindo intervenções que articulem direito, segurança e ética em um ambiente de rápida inovação tecnológica e frágil regulação internacional.

Embora a SSM tenha se mostrado adequada para compreender e estruturar o problema, reconhece-se que sua eficácia analítica pode ser ampliada por uma abordagem multimetodológica [Mingers 1997]. A integração com outros PSMs, como o SODA ou o *Drama Theory*, pode intensificar a participação dos atores sociais, ampliar a análise de conflitos e tornar as

recomendações mais sensíveis ao contexto político-institucional. Conforme defendem [Mingers e Rosenhead 2004], a combinação de métodos permite lidar com incertezas e pluralidade de valores de forma mais integrada.

Assim, recomenda-se que pesquisas futuras aprofundem o diálogo entre abordagens qualitativas e quantitativas, ampliando o escopo das análises e contribuindo para a construção de um marco regulatório internacional que promova o uso responsável da IA militar, ancorado na dignidade humana, no direito humanitário e na paz duradoura.

Agradecimento

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Referências

Afina, Y; Paoli, P (2024). *Governance of Artificial Intelligence in the Military Domain: A Multi-stakeholder Perspective on Priority Areas*. Disponível em: https://undir.org/wp-content/uploads/2024/09/UNIDIR_Governance_of_Artificial_Intelligence_in_the_Military_Domain_A_Multi-stakeholder_Perspective_on_Priority_Areas.pdf. Acessado: 2025-05-07.

Apps, P (2024). *Ukraine, Taiwan face-offs help drive drone, AI revolution*. Reuters World News. Disponível em: <https://www.reuters.com/world/ukraine-taiwan-face-offs-help-drive-drone-ai-revolution-2024-12-13/>. Acessado: 2025-05-06.

Blanchard, A (2024). Ethics in the international regulatory debate on autonomous weapon systems. Disponível em: <https://blogs.icrc.org/law-and-policy/2024/04/25/the-road-less-travelled-ethics-in-the-international-regulatory-debate-on-autonomous-weapon-systems/>. Acessado: 2025-05-05.

Blanchard, A.; Bruun, L (2024). Bias in military artificial intelligence. *SIPRI Background Paper*, dez. 2024. Stockholm International Peace Research Institute.

Checkland, P (1981). *Systems thinking, systems practice*. Chichester. John Wiley & Sons.

Checkland, P.; Poulter, J (2020). *Learning for Action: A Short Definitive Account of Soft Systems Methodology and its Use for Practitioner, Teachers and Students*. Wiley.

ICRC and Geneva Academy (2024). Artificial Intelligence and Related Technologies in Military Decision-Making on the Use of Force in Armed Conflicts. Geneva Academy of International Humanitarian Law and Human Rights. Disponível em: <https://www.geneva-academy.ch/joomlatools-files/docman-files/Artificial%20Intelligence%20And%20Related%20Technologies%20In%20Military%20Decision-Making.pdf>. Acessado: 2025-05-05.

Mingers, J.; Gill, A. (1997) *Multimethodology: The theory and practice of combining management science methodologies*. Chichester. Wiley.

Mingers, J.; Rosenhead, J. (2004). Problem structuring methods in action. *European Journal of Operational Research*, v. 152, n. 3, p. 530–554, 2004., J.; ROSENHEAD, J. Problem structuring methods in action. *European Journal of Operational Research*, v. 152, n. 3, p. 530–554.

United Nations (2023). *Governing AI for Humanity: Final Report of the High-Level Advisory Body on Artificial Intelligence*. United Nations Office of the Secretary-General. Disponível em: https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf. Acessado: 2025-05-05.

UNODA - United Nations Office For Disarmament Affairs (2023). Lethal Autonomous Weapons Systems (LAWS). Disponível em: <https://disarmament.unoda.org/the-convention-on-certain-conventional-weapons/background-on-laws-in-the-ccw/>. Acessado: 2025-05-05.

West, L (2024). Moving toward best practices in accountability and military use of AI. *CIGI Papers*, n. 301, set. 2024. Disponível em: https://www.cigionline.org/static/documents/Leah_West-no.301_EhDzqMx.pdf. Acessado: 2025-05-05.