

# Interdependência Securitária: o Caso da Proliferação e Controlo de Armamentos no Domínio da Inteligência Artificial

Maria Francisca Saraiva

*Professora Auxiliar do Instituto Superior de Ciências Sociais e Políticas (ISCSP) da Universidade de Lisboa (UL) e Investigadora do Instituto da Defesa Nacional*

## Resumo

Este artigo analisa as dimensões mais obscuras do uso de sistemas de Inteligência Artificial a nível militar, com enfoque particular nos sistemas de armas letais autónomas. Com base na necessidade de regulação destas tecnologias disruptivas ao nível militar, o texto defende a proibição preventiva destes armamentos e faz propostas no sentido de uma regulação global da utilização da Inteligência Artificial em ambiente militar. O artigo argumenta que os sistemas autónomos agravam as dificuldades de gestão dos instrumentos da violência armada, podendo pôr em causa os fundamentos da Estratégia. Defende igualmente a necessidade de promover uma arquitetura de controlo de armamentos global, tendo em conta que hoje já é possível usar aplicações de Inteligência Artificial em todos os domínios operacionais e que estes se encontram cada vez mais inter-relacionados.

**Palavras-chave:** Controlo de Armamentos; Inteligência Artificial; Estratégia; Direitos Humanos; Direito Internacional Humanitário.

## Abstract

**Security Interdependence: the Case of Proliferation and Arms Control in the Field of Artificial Intelligence**

*This article examines the more obscure dimensions of the use of Artificial Intelligence systems in Defense, with a particular focus on lethal autonomous weapon systems. Based on the need to regulate these disruptive technologies in military applications, this paper defends the preventive prohibition of these armaments and makes proposals for a global regulation of the use of Artificial Intelligence in military strategy. The article argues that autonomous systems aggravate the difficulties in managing the instruments of armed violence, which may undermine the foundations of strategy. It also defends the need to promote a global arms control architecture, taking into account that today it is already possible to use Artificial Intelligence applications in all military operational domains and that these are increasingly interrelated.*

**Keywords:** Arms Control; Artificial Intelligence; Strategy; Human Rights; International Humanitarian Law.

Artigo recebido: 02.12.2019

Aprovado: 14.12.2019

<https://doi.org/10.47906/ND2020.155.04>

## **Introdução**

Temos assistido a uma crescente aplicação da Inteligência Artificial (IA) em ambiente civil e militar.

É um facto que os benefícios da IA já são perceptíveis em muitas áreas da Defesa. Mas também é verdade que estes desenvolvimentos, que têm sido muito rápidos, tornam necessária uma reflexão sobre as dimensões mais obscuras desta tecnologia aplicada à área militar.

Neste texto analisamos a questão da regulação destas tecnologias disruptivas ao nível militar, com especial enfoque nos sistemas de armas letais autónomas. Contudo, defende-se que o problema dos “robôs assassinos” não deve ser dissociado da questão mais abrangente que é cada vez mais necessário regular, isto é, os múltiplos usos militares da IA, devido às questões éticas, estratégicas e jurídicas que estes suscitam. São apresentadas propostas no sentido de regular globalmente a utilização da IA em ambiente militar, sob pena da atividade estratégica poder vir a perder o seu cunho antropocêntrico.

## **Controlo de Armamentos e Segurança**

No essencial, a necessidade de regular multilateralmente a proliferação de armamentos altamente destrutivos é uma preocupação que surge no início do século XX. A Sociedade das Nações, criada pelo Tratado de Versalhes no final da Primeira Guerra Mundial, entendeu a necessidade de desarmamento da Alemanha no quadro dos esforços de um desarmamento geral, que acabou por alcançar progressos pouco significativos, mas que deveria ser conjugado com a diplomacia e o desenvolvimento do Direito Internacional em prol do objetivo de manter a paz mundial. Deste modo, estabelece-se uma conexão entre desarmamento, sem dúvida a resposta mais direta ao problema das crescentes capacidades militares dos Estados (Buzan, 1987), e paz no quadro de uma diplomacia multilateral.

A noção de controlo de armamentos é um subproduto típico da teoria da dissuasão nuclear, desenvolvida em finais da década de 50 do século XX no quadro dos Estudos Estratégicos. Na década seguinte, o controlo de armamentos irá de facto substituir a tradicional lógica do desarmamento, desembocando na década de 70 nos primeiros tratados de controlo de armas nucleares estratégicas, o SALT I e II, acordados entre os Estados Unidos e a URSS. O desenho destes tratados difere profundamente dos esforços de desarmamento tradicionais por pretenderem somente limitar a competição no domínio dos armamentos nucleares controlando o aumento das existências e não a sua redução numérica, objetivo promovido pelo conceito de desarmamento através da destruição física das armas (Goldblat, 2003). De facto, a arma nuclear alterou profundamente a perspetiva estratégica sobre a melhor forma

de garantir a segurança internacional. Trata-se agora de assegurar a impossibilidade de uma guerra nuclear reduzindo os incentivos de uma escalada, objetivo que é atingido através do controlo de armamentos nucleares estratégicos, um dos temas centrais na doutrina da *Mutual Assured Destruction* (MAD), uma estabilidade estratégica que é salvaguardada pela estratégia de dissuasão nuclear credível que procura evitar o desenvolvimento de *first strike capabilities* associada à possibilidade de ataques preventivos (*idem*) e a uma gestão criteriosa das crises que envolvessem os superpoderes. Em resultado disto as disputas militares do período bipolar mantiveram uma natureza estritamente convencional (Buzan, 1987; Saraiva, 2010). Nesta mesma lógica, entendia-se que deveria ser combatida a proliferação horizontal de armamentos nucleares para outros Estados que não possuíam esta capacidade – assunto regulado pelo Tratado de Não Proliferação Nuclear de 1968 – e refreado o surgimento de novas armas e a produção de modelos mais evoluídos (proliferação vertical), dificultado por tratados como o Tratado que Proíbe Testes Nucleares na Atmosfera, Espaço e no Espaço Submarino (1963) e o Tratado que Limita os Testes Subterrâneos de Armas Nucleares, de 1974.

Importa notar que as evoluções tecnológicas no domínio militar ocorridas nas últimas décadas alteraram radicalmente esta perspetiva sobre as capacidades militares. Estas importantes transformações são a razão mais evidente para a ampliação da noção de controlo de armamentos, que nos nossos dias envolve um conjunto bastante abrangente de restrições que procuram atenuar as tensões internacionais incluindo, como os mais importantes, o congelamento, limitação, redução e abolição de certas categorias de armas, a proibição de testes de armas, bem como a proibição de certas atividades militares, a regulação do destacamento de Forças Armadas, a proibição de transferências de armamentos, medidas para a criação de confiança e a proibição de certos métodos de guerra, entre outras atividades (Goldblat, 2003). Na realidade, o controlo de armamentos envolve uma cooperação com potenciais inimigos com o objetivo de garantir a segurança mútua. Desta forma, o conceito insere-se numa lógica de resolução de conflitos e de prevenção de conflitos violentos no sentido “de evitar ou limitar os conflitos armados e minimizar os seus efeitos se tais conflitos eclodirem” (Pinto, 1988, p. 103)

Claro que, como reconhecido por muitos estudiosos, as funções do controlo de armamentos são muito diversas, sendo possível destacar as seguintes como as mais importantes (Goldblat, 2003; Morgan, 2012):

- Reduzir o risco de guerra por acidente ou relacionado com problemas na conceção das armas;
- Desacelerar corridas aos armamentos ao nível global ou regional;
- Aumentar a previsibilidade nas relações entre Estados hostis e reduzir medos em relação às intenções do potencial adversário;

- Proibir o desenvolvimento de novos tipos de armas e meios de guerra;
- Minimizar disparidades entre Estados altamente armados e Estados mal armados (redução das fontes de instabilidade);
- Libertar recursos para o desenvolvimento económico-social;
- Reduzir os custos políticos associados à manutenção de Forças Armadas e mitigar a destruição e o sofrimento nos conflitos armados;
- Diminuir os perigos para o ambiente;
- Promover um melhor entendimento entre as nações;
- Medo de um emprego não autorizado do sistema de armas;
- Medo das implicações na saúde humana.

Para Williams e Viotti (2012) é possível identificar três etapas no controlo internacional de armamentos: a primeira era do controlo de armamentos coincide, como referido, com os finais da Primeira Guerra Mundial; a segunda domina todo o período da Guerra Fria, e a terceira, que começou em 1989 e perdura até hoje, é dominada por uma certa desvalorização estratégica das vantagens da manutenção de uma *second strike capability* – capacidade de absorver um primeiro ataque nuclear através de uma retaliação devastadora –, ideia que dominou a estratégia de dissuasão nuclear no período do bipolarismo, e que é um reflexo direto da complexificação do ambiente estratégico.

Poucos contestam hoje que o complexo ambiente de segurança internacional exige uma nova visão do controlo de armamentos, mais ampla e integrada. Desse ponto de vista, os mecanismos internacionais atualmente existentes dão importância crescente ao problema dos conflitos internos e ao conceito de segurança humana: a Convenção de Otava, que proíbe as minas antipessoais, assinada em 1997, a Convenção contra as Bombas de Fragmentação, ou Convenção de Oslo, de 2008, o Protocolo à Convenção sobre Certas Armas Convencionais sobre Explosivos Remanescentes de Guerra (1980), bem como o Tratado de Comércio de Armas sobre Armas Ligeiras e de Pequeno Porte (2013), têm em comum a preocupação de aliviar o sofrimento humano e dificultar a transferência de armas para entidades não estatais. Mas se é verdade que a dimensão humanitária do controlo de armamentos é hoje uma preocupação internacional no quadro dos denominados regimes de segurança humanitários (Garcia, 2015), há que reconhecer a importância de tomar medidas para controlar sistemas tecnológicos mais avançados, emergentes e disruptivos, de que destacamos as aplicações no domínio militar da IA, havendo igualmente necessidade de regulação das atividades militares no espaço exterior e ciberespaço como novos domínios operacionais que não podemos ignorar.

O ponto que neste texto importa sublinhar é que o uso de sistemas de IA no campo militar provocou mudanças de tal modo profundas que não encontram precedente na história das tecnologias militares. Há que estar ciente que num futuro mais ou

menos próximo os algoritmos da IA poderão criar “campos de batalhas automatizados” onde armas totalmente autónomas, também conhecidas como sistemas de armas letais autónomas tenderão a dispensar seres humanos dos teatros de operações. Trata-se igualmente de refletir sobre o modo como estes progressos científicos criam dilemas éticos que afetam a sua utilidade estratégica. Como se verá mais adiante, pode-se ainda falar dos perigos da materialização de uma rutura estratégica provocada pelos sistemas autónomos. Neste sentido, atendendo a que esta nova realidade poderá de certo modo questionar os fundamentos da Estratégia enquanto atividade humana, é urgente uma reflexão sobre os desafios que se colocam à arquitetura internacional do controlo de armamentos.

### **A Militarização da Inteligência Artificial**

É sabido que as novas tecnologias influenciam os métodos de fazer a guerra e as doutrinas militares (Lele, 2017). Tanto a evolução como a difusão da tecnologia militar estão muito associadas ao processo de inovação e à procura da superioridade tecnológica e militar. Em particular, existe uma tendência para a militarização e armamentização (*weaponization*) das tecnologias disruptivas, apesar de se reconhecer que são potencialmente desestabilizadoras e ameaçadoras da estabilidade internacional e dos equilíbrios regionais.

Existem vários exemplos de que a IA tem um impacto considerável no nosso quotidiano. Na medicina, no trânsito, na aviação, na indústria, nos dispositivos móveis, por exemplo, já há incorporação desta tecnologia avançada, mas subsistem questões em aberto sobre os limites de interação dos seres humanos com esta tecnologia. As aplicações militares da IA são consideradas por muitos como a próxima grande vantagem militar, mas quando se trata do ambiente militar há que discutir limites e parâmetros de utilização com maior detalhe e cuidado.

A IA enquanto disciplina começou a ser estudada nos anos 40 do século XX e teve períodos em que os avanços foram pouco significativos, sem grandes progressos. Pelo contrário, noutras fases o interesse pelas oportunidades nesta área cresceu significativamente, como sucede atualmente. Desde 2010 que assistimos à criação de máquinas cada vez mais inteligentes (Sayler, 2019), avanços na denominada *narrow Artificial Intelligence*, em particular no quadro da abordagem do *machine learning*, que tem sido responsável pelas evoluções mais recentes dos sistemas autónomos (Boulanin, 2019). O *machine learning* define-se pela “*hability of a machine to execute a task or tasks without human input, using interactions of computer programming with the environment*” (Boulanin, 2019, p. 15). Trata-se de um procedimento matemático e estatístico através do qual os algoritmos procuram similitudes ou padrões em quantidades enormes de dados (Dahlmann e Dickow, 2019). Na realidade, esta técnica precisa de grandes bases de dados fornecidos previamente para estabelecer

relações estatísticas, o que significa que não aprendem de forma semelhante ao homem (Boulain, 2019). Para muitos, esta evolução tecnológica multiplicará o poder disruptivo da IA em múltiplas dimensões da vida, incluindo na esfera militar.

O *machine learning* e o *deep learning*<sup>1</sup>, por ser turno, estão muito associados à noção de general Artificial Intelligence, entendida como “systems capable of human-level intelligence across a broad range of tasks” (Sayler, 2019, p. 2) ou mesmo ao conceito de superinteligência. De qualquer modo, ainda que a ideia de um sistema superinteligente constitua uma possibilidade teórica, parece haver um forte ceticismo quanto à viabilidade e utilidade de tal desenvolvimento (UNIDIR, 2018; Livingsgton e Risse, 2019), mesmo que pudesse exceder largamente o ser humano em todos os domínios, atuando de forma genial.

De notar que os dois conceitos – *general Artificial Intelligence* e a superinteligência – reportam a realidades hipotéticas sobre as quais existem grandes dúvidas sobre a real possibilidade de alguma vez poderem vir a materializar-se. Para os mais otimistas demoraremos décadas a atingir este nível de evolução, na melhor das hipóteses. É, no entanto, significativo que um número significativo de especialistas oriundos de várias áreas envolvidas na conceção e no estudo destas tecnologias sob os mais variados prismas não acreditem que este objetivo seja exequível. Limitações técnicas de vária ordem podem inviabilizar que máquinas possam vir a agir de forma inteiramente autónoma (Boulain, 2019).

As ciências computacionais, em especial a IA, têm hoje muitas aplicações militares, por exemplo nos veículos não tripulados, mas é a possibilidade de uso de algoritmos para criar exércitos robóticos que mais perturba do ponto de vista estratégico mas também ético e na ótica do cumprimento dos princípios do Direito Internacional Humanitário (DIH) e do quadro normativo do Direito Internacional dos Direitos Humanos (DIDH). O método de programação da IA, a matemática e a estatística, permite hoje que as máquinas possam ser remotamente controladas ou, no futuro, podem vir a surgir máquinas autónomas capazes de desenvolver certas tarefas sem intervenção humana em ambientes dinâmicos aptas a desempenhar tarefas complexas (Dahlmann e Dickow, 2019). A definição não é consensual, mas tem vindo a ganhar adeptos a ideia de que os sistemas totalmente autónomos são aqueles que, uma vez ativados, podem selecionar e atacar objetivos sem intervenção de um operador humano, isto é, operam sem “meaningful human control” (WILPF, 2020; Dahlmann e Dickow, 2019; UNODA, 2019).

---

1 Segundo o United Nations Institute for Disarmament Research (UNIDIR) (2018, p. 3), “Deep learning is a type of machine learning approach that uses deep neural networks. Deep neural networks can be used for supervised, unsupervised, or reinforcement learning”.

---

Este texto foca-se primordialmente na possibilidade de regulação das armas letais autónomas, mas não esquece que os avanços da IA têm várias vantagens militares podendo ser aplicados a todos o tipo de sistemas militares – ciber, convencionais, ou nucleares –, sem que se conheçam bem os impactos estratégicos da sua utilização (Boulainin, 2019).

De notar que hoje em dia já existem sistemas totalmente autónomos de carácter defensivo que são programados para responder a ameaças (Zawieska, 2017). Falamos de sistemas de armas essencialmente utilizados para defesa aérea e defesa antimís-sil, o que significa que respondem a ameaças de mísseis, mas não são capazes de lançar um ataque de forma independente.

São muitas as vantagens militares apontadas às armas letais autónomas, tanto do ponto de vista tático como estratégico e a sua principal característica é o facto de aumentarem enormemente o poder de fogo (Payne, 2018). Adicionalmente, estes robôs não apresentam problemas comuns nos militares empenhados em missões, nomeadamente fadiga e emoção. Porém, os benefícios desta tecnologia são essencialmente políticos, envolvendo a possibilidade de uma redução substancial do número de militares empenhados no terreno (Dahlmann e Dickow, 2019) e uma diminuição do número de baixas que as características da tecnologia, obviamente, acabam por potenciar. Consequentemente, esta tecnologia permite atingir um maior grau de precisão e é apresentada como sendo uma opção economicamente mais barata.

Entende-se que as armas letais autónomas são militarmente mais ágeis e rápidas do ponto de vista do comando e controlo, por terem menos necessidade de comunicação constante entre o comandante militar e o robô (Sauer, 2016; Boulainin, 2019). Com efeito, a autonomia não necessita de uma ligação ao comando e controlo que é vulnerável à disrupção ou captura e pode revelar a localização do sistema. Assim, em tese, seria possível retirar importantes vantagens táticas da independência do robô face ao comando e controlo.

Outros benefícios frequentemente referidos são a menor dependência em relação a operadores e analistas, cujo número poderia ser reduzido, e a possibilidade de envolvimento em missões mais difíceis e eventualmente “suja”; o acesso a teatros operacionais anteriormente inacessíveis (Ártico, espaço exterior) e a possibilidade de operações colaborativas (*swams*) (Boulainin, 2019) para referir alguns dos empregos mais significativos.

Um dos principais argumentos a favor da robotização prende-se com a ideia de que esta tecnologia permite a condução de conflitos armados dentro dos parâmetros do DIH, contrariamente à situação atual vividas nos conflitos armados. Com efeito, a possibilidade de redução de baixas militares e de mortes de civis inocentes, a eliminação de sofrimentos desnecessários e a redução substancial do risco de serem cometidos crimes de guerra (Zawieska, 2017) são dimensões importantes no contexto dos conflitos armados.

Em suma, muitos consideram que nas operações de combate os sistemas autónomos permitiriam tomar uma decisão mais informada e o aumento da precisão, velocidade (*speed*) e escala da ação militar, oferecendo vantagens assimétricas (Sayler, 2019; Maas, 2019).

As inovações tecnológicas sempre influenciam os métodos de fazer a guerra e as doutrinas militares (Zawieska, 2017) mas as potencialidades do *machine learning* – bem como de outras técnicas de IA com aplicação militar igualmente disruptivas que estão fora do escopo deste texto – colocam desafios à Estratégia equivalentes ao advento das armas atômicas e nucleares, no final da Segunda Guerra Mundial. No caso da IA, os riscos que envolvem uma decisão de emprego da força armada, talvez o principal problema estratégico que os decisores têm que enfrentar, são aparentemente e em grande medida neutralizados pelas características da tecnologia. Ao contrário do que sucede com as armas nucleares que são, no essencial, armas ao serviço de uma estratégia de paz, tendo em conta as consequências devastadoras de uma guerra nuclear, a IA permite maior liberdade estratégica e de ação em relação aos condicionalismos políticos, estratégicos e jurídicos que envolvem o emprego da força armada. Os sistemas letais autónomos prometem revolucionar o *jus ad bellum* e o *jus in bello* e provocar profundas mudanças sociais, nas Forças Armadas e nas empresas (Boulanin, 2019). Os progressos nesta área são evidentes. Em 2017 a DeepMind anunciou que o seu novo algoritmo, o *AlphaGo Zero* aprendeu a jogar um jogo de estratégia chinês, o *Go*, sem treino com dados humanos, isto é, aprendeu exclusivamente a partir das regras do jogo. Pouco depois o algoritmo *Alpha Zero* aprendeu a jogar *Go*, *Shogi* e xadrez tendo-lhe sido dado a conhecer apenas as regras de cada um dos jogos (UNIDIR, 2018).

Tendo em conta estes desenvolvimentos, há desde logo uma questão a que importa responder: poderão as qualidades humanas ser literalmente reproduzidas numa máquina?

### **Implicações Estratégicas dos Sistemas de Armas Letais Autónomas**

É por demais evidente que os avanços no campo da IA e a possibilidade de construção de sistemas de armas letais autónomas têm favorecido as sociedades mais avançadas envolvidas na revolução da informação e na Revolução nos Assuntos Militares (Payne, 2018).

Estados Unidos, China, e em menor grau, a Rússia, disputam a supremacia tecnológica da IA e de outras tecnologias avançadas, à semelhança do que fizeram no passado com outras novidades tecnológicas. As razões são essencialmente políticas, mas também estratégicas. Como ocorre noutras dimensões da competição que envolve os Estados Unidos, que até aqui assumiu a liderança do mundo, e a China, potência em ascensão estratégica, alguns sectores nos Estados Unidos te-



mem que a China possa estar a caminho da supremacia na IA, na impressão a 3D (*3D printing*), no desenvolvimento de mísseis hipersónicos, na computação quântica (*quantum computing*) e na tecnologia 5G (*5G wireless connectivity*) (Krpooming e Gopalaswamy, 2018). A aposta chinesa em matéria de IA terá sido, portanto, a razão do anúncio da criação de um Joint Artificial Intelligence Center (JAIC) pelas autoridades dos Estados Unidos, em 2018 (Sharikov, 2018).

E também há que reconhecer as diferenças entre estes países, nomeadamente as vulnerabilidades americanas. A questão é que devido à complexidade dos sistemas robóticos as Forças Armadas americanas tenderão a estar cada vez mais dependentes das empresas privadas que os desenvolvem (Dahlmann e Dickow, 2019), o que não acontece necessariamente no caso chinês, enfrentando os problemas do duplo uso da IA amplamente disseminada em ambiente civil, numa escala que outras tecnologias – armas nucleares, armas químicas e armas biológicas, por exemplo – nunca conheceram.

Os “robôs assassinos” não têm emoções nem necessidade de se protegerem, por exemplo, embora os cientistas estejam a estudar a possibilidade de poderem vir a aderir a códigos morais. A IA toca nas áreas da cognição, neurociência e informática, tentando compreender o pensamento e o comportamento humano, a fim de reproduzi-lo artificialmente (Dahlmann e Dickow, 2019). No entanto, e de momento, não se prevê que o desenvolvimento tecnológico seja capaz de vir a reproduzir integralmente as qualidades humanas numa máquina (Zawieska, 2017). Discute-se se estes robôs poderão vir a desenvolver empatia e capacidade de decisão de improviso. Sabe-se que as máquinas não podem sentir, e muito menos compreender, mas executam pensamentos. Os valores têm que ser necessariamente programados.

Do ponto de vista dos desafios éticos, se algum dia lhes dermos autorização para decidir matar pessoas isso significará que não respeitámos a dignidade intrínseca do ser humano, a quem uma máquina nunca poderá tirar a vida por decisão sua. O problema ético resume-se ao facto da máquina não poder atuar moralmente, não tendo, por isso, a noção da mortalidade e do valor da vida humana (Dahlmann e Dickow, 2019). Apesar do imaginário popular ser muito fértil, nada indica, neste momento, que o desenvolvimento tecnológico possa vir a resolver este intrincado problema.

Na verdade, não são só os impactos políticos e militares desta decisão que importam mas sobretudo a conceção que temos da nossa condição de seres humanos que pode ser posta em causa (Zawieska, 2017). Como procuraremos mostrar, o corolário lógico deste dilema ético é que apenas a proibição deste tipo de tecnologias perturbadoras garante a preservação da dignidade da pessoa humana. Parece-nos, neste sentido, incontornável, a necessidade de controlo humano nas “funções críticas” de definição do alvo (*targeting*) e ordem de ataque (*engagement in attack*), garantindo deste modo a *accountability*, respeito pelos DH e DIH (Asaro, 2018) e ainda a observação das regras do *jus ad bellum*.

Importa sublinhar que a literatura tem refletido sobretudo sobre os desafios dos direitos humanos e do DIH associados à IA com aplicação militar. Embora na verdade estes condicionalismos sejam relevantes, não são, do nosso ponto de vista, o fulcro da questão. Ou seja, o argumento central deste texto é que a proibição deste tipo de tecnologias deve ser promovida por razões essencialmente estratégicas, sendo que todas as dificuldades jurídicas surgem a jusante, no sentido em que advêm das opções estratégicas que foram previamente tomadas. Isso significa que a primeira prioridade em termos de opções estratégicas a tomar deve ser a manutenção da estabilidade internacional. E são vários os desafios envolvidos.

Como afirmado por Payne (2018), a questão central é que a IA pode transformar a Estratégia comprometendo a essência psicológica da Estratégia e a sua natureza antropocêntrica, pondo em causa o princípio de que as decisões sobre a ameaça ou emprego da força armada são sempre tomadas por seres humanos. Colin Gray (2016) dedicou recentemente um livro, *Strategy and Politics* à relação umbilical, que no seu entender tem sido mal compreendida, entre Política e Estratégia. Merece ser realçado o facto de a Política ter a responsabilidade de clarificar o objetivo subjacente à atividade estratégica e que a Estratégia, por seu lado, oferece à Política os meios para a realização, na prática, dos fins definidos pela Política, o que sublinha, naturalmente, o carácter intrinsecamente humano desta atividade. Mas mais importante é concretizar as características centrais que tornam a IA particularmente desafiante em termos estratégicos.

Comparem-se os perigos da IA com os perigos das armas nucleares. Deve notar-se que se o fator nuclear introduziu a necessidade de se optar por uma estratégia de dissuasão, na essência uma estratégia de não guerra assente numa estratégia de forte cunho psicológico (Freedman, 2003), a IA pode ter consequências ainda mais disruptivas, podendo acabar por eliminar o elemento psicológico e contribuir para baixar o *threshold* da guerra em todo o tipo de ambientes operacionais. Ou seja, em termos de argumento, defende-se que a segurança do Estado deve estar no centro desta equação, tal como preconizam as correntes mais tradicionais das Relações Internacionais, o que não significa que a segurança do Estado não proteja os interesses dos indivíduos que se encontram sob a tutela desses Estados. Trata-se de uma posição que salvaguarda simultaneamente os interesses de segurança dos Estados e as preocupações humanitárias com a proteção dos indivíduos subjacentes aos regimes de segurança humanitários já referidos, um caso da política internacional em que claramente consideramos não existir oposição entre o objeto de referência para a segurança humana que é o ser humano enquanto indivíduo (Newman e Richmond, 2001; Ramcharan, 2002) e a segurança do Estado.

Imagine-se que o combate direto deixará de ser feito pelo homem, passando a ser tarefas de robôs. A velocidade (*speed*) da IA pode levar o Estado que procura defender-se dos ataques a sentir-se em situação de desvantagem, o que pode criar

incentivos para que desfira um primeiro ataque com capacidades semelhantes (Sayler, 2019). Sabendo-se que as tecnologias de *machine learning* podem ter comportamento imprevisível, podemos esperar uma escalada do conflito, razão pela qual não consideramos que se adequem a ambientes sociais complexos.

Caso permitíssemos a existência de robôs autónomos poderíamos perder o controlo sobre o *jus ad bellum*, isto é, sobre a capacidade de quando e em que circunstâncias deveríamos avançar para uma situação de conflito armado, iniciando as hostilidades. Isso porque a automatização da decisão poderia levar inadvertidamente ao início das hostilidades e, como reforça Payne (2018), a dinâmica do conflito poderia sair ainda mais do controlo caso houvesse capacidade automática de escalada em resposta a ameaças apercebidas.

Significa isto que as questões perenes da Estratégia não encontram nos robôs letais autónomos uma resolução satisfatória. Aliás, o argumento central deste texto é que os sistemas autónomos agravam as dificuldades de gestão dos instrumentos da violência armada, problema central na história da Estratégia e que os desenvolvimentos recentes da tecnologia e a visível proliferação de atores nos teatros de operações ajudou a agudizar.

O ambiente da guerra é por natureza confuso. Nas missões de combate há muitas variáveis que escapam ao controlo humano, entre as quais podemos destacar: a capacidade de adaptação do inimigo que nunca deve ser desprezada; é natural que ocorram riscos ambientais; existam civis no teatro de operações dificultando a execução das estratégias; há sempre a possibilidade de ocorrência de ciberataques e a fricção e nevoeiro da guerra são elementos sempre presentes, como referido por Clausewitz (1984). No caso de um cenário de robotização das missões de combate, sabemos que estas máquinas não desobedecem, mas se no contexto das suas ações ocorrem mudanças não têm flexibilidade para se adaptar por não dominarem o contexto específico. Assim, se ocorrerem situações não antecipadas os sistemas de *machine learning* poderão atuar de formas inesperadas, imprevisíveis e indesejáveis (Asaro, 2012). Sendo certo que nunca saberão lidar com o nevoeiro e a fricção da guerra. Também podem causar destruição por funcionarem mal ou por terem sido “hackeados”, continuando a matar, mesmo quando não era suposto que o fizessem. Finalmente, é fundamental perceber um último problema relacionado com a conveniência política de terminar uma guerra, mas tal não ser possível, por não haver possibilidade de comunicação com a máquina. Em suma concordamos com aqueles que defendem que a distinção binária homem/máquina é para manter (Payne, 2018; Leveringhaus, 2018). De outro modo, no limite, poderíamos correr o risco de perder o controlo humano sobre medidas letais – em relação ao início das hostilidades, no caso de escalada e quanto ao *terminus* do conflito armado.

Uma questão conexa com esta é o problema da observância do DIH e dos seus princípios de discriminação – distinção civil/combatente –, proporcionalidade dos

meios e fins e necessidade do uso da força. O facto de a IA não ser permeável à fadiga e emoção é visto como uma vantagem militar em situação de conflito. Esta posição não tem em conta o facto de estas características, na prática, poderem comprometer o desempenho estratégico. A vulnerabilidade estratégica da IA em ambiente militar reside precisamente na impossibilidade de categorizar e contextualizar informação, elementos essenciais na decisão humana. O problema é que a IA não consegue reproduzir estas dimensões de complexidade. O contexto e o dinamismo são desafios para os quais a engenharia ainda não encontrou soluções satisfatórias (Dahlmann e Dickow, 2019).

Do ponto de vista das exigências de DIH, estas armas não conseguem reconhecer que uma pessoa está fora de combate – juízo qualitativo contextualizado – pois não distinguem o lícito do ilícito. Também no critério da proporcionalidade estão envolvidas avaliações casuísticas segundo critérios subjetivos. O princípio da precaução, conceito jurídico abstrato que regula a proibição de vítimas civis inocentes, e a necessidade de equilibrar a necessidade militar com os princípios de humanidade não são conceitos que a máquina entenda (Jiménez-Segovia, 2019; Payne, 2018; Dahlmann e Dickow, 2019).

Conclui-se que é forçoso manter o controlo humano no *target cycle*, o ser humano deve ser o sujeito jurídico e deve tomar a decisão porque as máquinas não percebem que estão a matar pessoas (Dahlmann e Dickow, 2019; Sauer, 2016).

No geral, resulta óbvio que a incorporação de IA no processo de decisão estratégica e em sistemas de armamentos pode facilmente levar a uma corrida aos armamentos, bem como a problemas de *hacking*<sup>2</sup> e *spoofing*<sup>3</sup> dessas armas. A possibilidade de roubo dessas armas e o seu redireccionamento, por parte dos adversários, é também uma possibilidade.

É preciso perceber que todas as inovações tecnológicas criam novas vulnerabilidades e que estas necessitam, inevitavelmente, de estratégias de mitigação. Mas no caso da IA são muitas as questões em aberto pelo facto dos métodos de computação poderem ser facilmente incorporados em diferentes tipologias de armas e serem capazes de operar em diferentes ambientes operacionais.

O problema da atribuição é um deles. No caso dos ataques no ciberespaço o problema da atribuição, apesar de alguns avanços alcançados, é certo, ainda não se conseguiu mitigar completamente. Por isso resulta óbvio que a questão da atribuição é ainda mais problemática no caso de ataques cinéticos tradicionais que utilizam tecnologias de IA (Boulanin, 2019).

---

2 Entendido como intrusão não autorizada num computador ou rede. Ver <https://pt.malwarebytes.com/hacker/>. Acedido em 27 de dezembro de 2020.

3 Ou seja, envio de dados falsos (UNIDIR, 2018, p. 13).

São também de apontar os riscos associados ao ritmo do progresso técnico dos sistemas de IA, que podem ser potenciados no caso de os sistemas de IA serem combinados com outras tecnologias de informação inovadoras – *cloud computing*, *big data*, internet das coisas –, ameaçando a estabilidade das relações estratégicas (Sharikov, 2018).

Uma das questões que mais preocupa é a possibilidade de combinar tecnologias de IA com capacidades nucleares. Na verdade, hoje a estratégia de dissuasão nuclear está profundamente interligada com os sistemas de defesa antimíssil. Esta nova combinação, que na Guerra Fria era proibida por se entender que perturbava a estabilidade estratégica garantida pela estratégia de dissuasão nuclear, provoca o esbatimento da fronteira entre armamentos nucleares/não nucleares sendo igualmente certo que tanto o ciberespaço (Thränert, 2018) como o espaço exterior passam a fazer parte integrante da lógica nuclear. Como é sabido, a defesa antimíssil favorece a capacidade ofensiva e não a defensiva (Payne, 2018; Saraiva, 2010), o que significa que aumenta a probabilidade de um primeiro ataque de sucesso contra arsenais nucleares reduzidos e dissolve a distinção tradicional entre armamentos convencionais e nucleares (Hagström, 2019). Sendo assim, imaginem-se os riscos que a IA pode originar no campo do nuclear podendo por em causa décadas de consenso sobre as vantagens de uma estratégia de não guerra.

Desde este ponto de vista, parece evidente que os sistemas de IA podem ser facilmente explorados pelos adversários, podendo levar à proliferação de sistemas e a um aumento da possibilidade de serem “hackeáveis”. Na verdade, podem ser facilmente roubadas, estamos a falar de códigos de *software*.

É também de apontar, na linha de Hagström (*idem*), que a autonomia dos sistemas de IA pode ser bastante vulnerável a contramedidas dos oponentes o que inviabiliza a capacidade de serem previsíveis para quem os usa ao mesmo tempo que os torna altamente imprevisíveis para os oponentes. Para além do mais, o processo de verificação dos sistemas *black box* que é característico dos sistemas de *learning machine* não é fácil (*idem*), necessitando de uma grande quantidade de dados e é duvidoso que se possa aplicar aos sistemas de armas porque não existe um algoritmo aplicável a “multiple problem domains” (Livingston e Risse, 2019).

### **O Problema do Controlo das Tecnologias de Inteligência Artificial**

Tudo o que se disse até aqui sugere que o controlo das tecnologias de IA é urgente e necessário. Como referimos anteriormente, os robôs autónomos são a face mais visível do problema, é certo, embora não sejam necessariamente o problema mais complexo do ponto de vista dos objetivos da regulação internacional de armamentos. Tem-se procurado discutir a regulação das armas letais autónomas no quadro da Convenção das Nações Unidas sobre a Proibição ou Limitação de Certas Armas

Convencionais que Podem Ser Consideradas como Produzindo Efeitos Traumáticos Excessivos ou Ferindo Indiscriminadamente, de 1980. Desde 2014, foram promovidos três encontros informais de peritos com Estados, organizações não governamentais (ONG) e peritos independentes (Jiménez-Segovia, 2019) e há dois anos que se encontra a funcionar um Grupo de Peritos Governamentais (SIPRI, 2019), como originalmente sugerido pelo Relator Especial sobre Execuções Extrajudiciais ou Arbitrárias do Conselho de Direitos Humanos das Nações Unidas, Christof Heyns, no Relatório A/HRC/23/47, recomendando a sua proibição e a criação de um Grupo de Trabalho (*idem*). Tendo em conta o enquadramento da Convenção, pode desde logo observar-se que a discussão se tem centrado nos aspetos de DIH e não nas questões éticas e de estabilidade internacional que estas armas colocam. Também é verdade que transcorridos estes anos os progressos não têm sido particularmente significativos (UNIDIR, 2018).

Importa ter em conta a posição do Comité Internacional das Cruz Vermelha que tem defendido que a discussão deve focar-se na função das armas, isto é, deve centrar-se nas armas que selecionam e atacam objetivos sem intervenção humana (Jiménez-Segovia, 2019). Mas, como referido, estas armas ainda não existem, o debate é prospetivo.

Até ao momento há um conjunto de 30 Estados<sup>4</sup> favoráveis à ideia de banir estas armas através de um protocolo de proibição de desenvolvimento e uso. Contam com o importante apoio da coligação *Stop the Killer Robots*, uma coligação de ONG liderada pela Human Rights Watch.

Mas há quem se oponha a esta solução. Os Estados Unidos, Rússia, Coreia do Sul e Israel entendem que a automação tem grandes vantagens na redução de baixas civis e militares.

Reino Unido, Estados Unidos, Rússia, França, Itália, Japão, Israel, Coreia do Sul e China são os maiores produtores de armas do mundo. Estes países defendem que não devem definir-se sistemas de armas autónomas e querem excluir de uma eventual delimitação as já existentes e já em uso que utilizam IA (Jiménez-Segovia, 2019). Apesar da forte oposição destas potências, com exceção da China que mantém uma posição ambivalente, há países que procuram uma solução de compromisso sendo a ideia de uma declaração política comum para regular estas tecnologias a iniciativa mais relevante. Avançada em 2017 pela França e Alemanha<sup>5</sup>, procura salvaguardar

---

4 *Campaign Stop Killer Robots*, 2019. Ver <https://www.stopkillerrobots.org>. Acedido em 5 de janeiro de 2020.

5 *France and Germany, Joint Working Paper for Consideration by the Group of Governmental Experts on Lethal Autonomous Weapons Systems (LAWS) – CCW/GGE.1/2017/WP4*, 7 November 2017. Disponível em <http://www.reachingcriticalwill.org/images/documents/Disarmament-fora/ccw/2017/gge/documents/WP4.pdf>. Acedido em 5 de janeiro de 2020.

que a decisão última de usar a força deve ser humana. Também sugere uma série de medidas para fomentar a transparência e a confiança entre os Estados, a elaboração de um Código de Conduta e ainda a criação de um Grupo de Peritos Técnicos para monitorizar a evolução desta tecnologia (*idem*).

Em relação aos países em desenvolvimento e países emergentes a proibição preventiva é geralmente a posição adotada, mas como é preciso decidir por consenso a hipótese de criar um Protocolo à Convenção semelhante ao Protocolo sobre armas laser que causam cegueira, a elaboração de um instrumento jurídico autónomo fora do âmbito da Convenção, como sucedeu com as minas antipessoais e as bombas de fragmentação, começa a ser equacionada como uma opção viável para ultrapassar o relativo impasse que hoje se vive no quadro da Convenção.

Deve também ser referida a situação na União Europeia. Uma Resolução do Parlamento Europeu<sup>6</sup> de setembro de 2018 defende a proibição mas, como referido, o RU opõe-se à regulação, a França e Alemanha têm uma posição intermédia e a Áustria é o único país oficialmente favorável à sua proibição (Dahlmann e Dickow, 2019). Temos, portanto, um contexto político difícil. O problema do ponto de vista deste texto é que as características particulares da IA exigem uma nova abordagem ao controlo de armamentos não se resolvendo com medidas pontuais. Neste sentido, o problema já não passa por proibir uma determinada arma específica por violação do princípio da discriminação ou por produzir efeitos traumáticos excessivos. Em vez disso, a ênfase deve ser, como lembra Asaro (2012), no modo como se decide que se vão usar armas equipadas com IA, isto é, o que se procura salvaguardar é o tipo de decisão de emprego, que tem que ser uma decisão tomada por humanos, ficando vedada qualquer outra situação. Note-se ainda que os Estados Unidos deveriam assumir a liderança do processo de regulação destas tecnologias, à semelhança do que aconteceu com outras tecnologias disruptivas (Thränert, 2018). Contudo, isto não vai acontecer: o avanço tecnológico americano nesta e noutras tecnologias é brutal, pelo que o incentivo para a regulamentação é mínimo.

Por outro lado, estamos a falar de uma abordagem de controlo de armamentos particularmente intrusiva da soberania estando aqui em causa a proibição do desenvolvimento e colocação de armas em terreno operacional, e não a decisão do uso ou o condicionamento do momento em que a força é usada (Morgan, 2012). Neste sentido, seria importante criar incentivos estratégicos para o abandono dos programas de investigação que estão em curso. Embora não seja fácil alcançar este

---

6 European Parliament, *European Parliament Resolution of 12 September 2018 on Autonomous Weapon Systems* (2018/2752 RSP), 12 September 2018. Disponível em <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018IP0341&from=EN>. Acedido em 5 de janeiro de 2020.

objetivo, defendemos que, ainda assim, poderá ser eventualmente mais fácil banir um tipo de armamentos que ainda não se encontra ao serviço das Forças Armadas. Quanto à hipótese de um cenário de proibição que deixa de fora os países que estão na vanguarda das tecnologias, não é a solução ideal para o problema, mas é de realçar que esta solução pode ajudar a estancar a proliferação horizontal destas tecnologias para países que neste momento não dispõem de programas de investigação, mas que poderiam sentir-se tentados a seguir os passos dos países que lideram estas investigações. É ainda de assinalar que a proibição pode vir a criar um espaço normativo na arena humanitária capaz de constranger politicamente os países que defendem as armas autónomas. Faz, portanto, sentido estabelecer um paralelismo entre a regulação dos sistemas autónomos e o novo Tratado Sobre a Proibição de Armas Nucleares, que não conta com o apoio do Clube Nuclear, Estados no Limiar e países NATO, mas que poderá ter um efeito positivo na Conferência de Revisão do TNP que se realiza este ano, nomeadamente na possibilidade de evolução do posicionamento destes países em relação aos seus compromissos de desarmamento nuclear geral e completo (Saraiva, 2018).

Também é verdade que as extensas aplicações da IA às tecnologias dos militares já em uso ou em investigação sublinham a necessidade de um controlo de armamentos que não se cinja ao problema dos sistemas de armas letais autónomas.

Um ponto que deveria reunir consenso seria a inclusão do tema IA na agenda do controlo de armamentos de forma integrada, isto é, envolvendo armamentos convencionais, armamentos nucleares, o espaço exterior e o ciberespaço, tendo em conta a existência de novos teatros de operações e que estes se interrelacionam e se relacionam com o ambiente mais tradicionais, ambiente convencional e ambiente nuclear. O problema da regulação dos armamentos convencionais que incorporam IA mas não de forma autónoma ainda não equacionada, sendo importante começar a pensar-se num fórum onde se possa fazer a sua regulação. Em relação ao nuclear, o ponto principal parece-nos ser a necessidade urgente de inclusão da IA na agenda das Conferências de Revisão do TNP.

Por outro lado, o controlo de armamentos no ciberespaço, embora tema problemático, deve ser prosseguido no âmbito da 1.<sup>a</sup> Comissão da Assembleia-Geral das Nações Unidas, Comissão de Desarmamento e Segurança Internacional, e deve passar a incluir o tema da IA. Por iniciativa da URSS, foi adotada por consenso em 5 de novembro de 1998 a Resolução Desenvolvimentos no Campo da informação e telecomunicações no contexto da segurança internacional (A/RES/53/70) (Tikk-Ringas, 2012). A resolução enquadra todo o espectro das ciberameaças – ciberagressão, ciberterrorismo patrocinado ou não pelos Estados, passando pelo cibercrime, ameaças às infraestruturas críticas, e inclui a questão da ameaça de intervenção nos assuntos internos dos Estados. Desde então, várias resoluções têm sido apresentadas, nem todas apontando para o tipo de soluções que preco-



nizamos. Defendemos que os progressos que venham a ser atingidos no quadro destas resoluções devem incorporar a questão da IA no ciberespaço.

Por outro lado, a opção estratégica associada à construção de sistemas de defesa antimíssil no território americano e nos países NATO tem implicações importantes no domínio do ciberespaço mas também no espaço exterior. A implicação desta realidade é que também neste teatro de operações se deveria proibir a “armamentização”, isto é, a proibição de Armas Antissatélite (ASAT). É também na 1.<sup>a</sup> Comissão que têm sido aprovadas resoluções sobre o tema da corrida aos armamentos no espaço exterior e é aqui que devem prosseguir os esforços de manter o espaço exterior livres de todo o tipo de armas, uma vez que apenas se encontram proibidas as armas de destruição massiva<sup>7</sup>.

Também é urgente olhar para o nível mais global, havendo necessidade de envolver as empresas privadas que lideram este sector no sentido de se comprometerem a não fabricar sistemas totalmente autónomos e a seguir regras apertadas em relação a outras aplicações da IA. Outro aspeto fundamental são os bancos e fundos de pensões, que não deveriam investir no negócio dos sistemas autónomos por vontade própria (WILPF, 2019).

Em suma, a tendência cada vez mais generalizada de utilização da IA em ambiente militar sugere a necessidade urgente de adequar o sistema internacional de controlo de armamentos, o que na prática significa que é necessário gerar dinâmicas integradas em relação aos domínios operacionais que já não podem ser analisados, como no passado, como realidades separadas: terra, mar, ar, espaço e ciberespaço estão hoje interrelacionados e todo o tipo de armas – convencionais, nucleares e cibernéticas – utilizam aplicações de IA. Acresce que estas categorias de armas estão a ser empregues de forma crescentemente complementar e integrada, podendo assim falar-se dos perigos da materialização de uma rutura estratégica provocada pelos sistemas autónomos pelo que urge regular internacionalmente esta realidade em acelerada evolução.

### **Considerações Conclusivas**

Neste texto, refletimos sobre as consequências éticas, estratégicas e jurídicas das aplicações militares da IA, em particular no que respeita aos sistemas de armas letais autónomas. As armas letais autónomas são uma tecnologia disruptiva que ainda não se encontra à disposição das Forças Armadas mas que pelas suas características pode vir a colocar em causa a essência da Estratégia enquanto atividade humana.

---

7 *Tratado sobre os Princípios Que Regem as Atividades dos Estados na Exploração e Utilização do Espaço Exterior, Incluindo a Lua e Outros Corpos Celestes*, Nações Unidas, 1967.

O artigo defende a proibição preventiva deste tipo de armamentos com base na tese de que a manutenção da estabilidade estratégica do ambiente internacional deve ser assumida como a grande prioridade estratégica neste século. Nesse sentido, o texto procura demonstrar que os sistemas autônomos agravam as dificuldades de gestão dos instrumentos da violência armada contribuindo, deste modo, para a instabilidade estratégica mundial. Esta posição permite salvaguardar simultaneamente os interesses de segurança dos Estados e as preocupações com a proteção dos indivíduos subjacentes aos regimes de segurança humanitários.

Quanto à ampla utilização da IA em múltiplos usos militares, o artigo propõe um sistema integrado de regulação internacional da IA que deve incluir necessariamente as capacidades estratégicas que incorporam IA em terra, no mar, no ar, no ciberespaço e até no espaço, hoje domínios operacionais profundamente interrelacionados, uma vez que diversos tipos de armas – sejam convencionais, nucleares ou cibernéticas – utilizam atualmente aplicações de IA e tendem a ser empregues em simultâneo e de modo combinado.

Concluimos que é cada vez mais urgente redefinir a arquitetura internacional do controlo de armamentos nos termos sugeridos neste artigo.

### Referências Bibliográficas

- Asaro, P., 2018. Why the world needs to regulate autonomous weapons, and soon. *Bulletin of the Atomic Scientists*, April 27. Disponível em <https://thebulletin.org/2018/04/why-the-world-needs-to-regulate-autonomous-weapons-and-soon/>. Acedido em 20 de dezembro de 2019.
- Asaro, P., 2012. On banning autonomous weapon systems: human rights, automation, and the dehumanization of lethal decision-making. *International Review of the Red Cross*, 94(888), pp. 687-709.
- Boulanin, V., ed., 2019. *The impact of AI on strategic stability and nuclear risk. Euro-atlantic perspectives*. Volume I, Solna: SIPRI.
- Buzan, B., 1987. *An introduction to strategic studies: military technology and international relations*. Londres: Palgrave Macmillan.
- Clausewitz, C., 1984. *On war*. Rep. M. Howard and P. Paret, ed. and transl. Princeton: Princeton University Press.
- Dalhmman, A. e Dickow, M., 2019. Preventive regulation of autonomous weapon systems need for action by Germany at various levels. *SWP Research Paper 3*, Berlim, March. Disponível em [https://www.swpberlin.org/fileadmin/contents/products/research\\_papers/2019RP03\\_dnn\\_dkw.pdf](https://www.swpberlin.org/fileadmin/contents/products/research_papers/2019RP03_dnn_dkw.pdf). Acedido em 15 de dezembro de 2019.
- Freedman, L., 2003. *The evolution of nuclear strategy*, 3<sup>rd</sup> ed. Nova Iorque: Palgrave.

- Garcia, D., 2015. Humanitarian security regimes. *International Affairs*. 91(1), pp. 55-75.
- Goldblat, J., 2003. *Arms Control*. Londres: Sage.
- Gray, C. S., 2016. *Strategy and Politics*. Londres: Routledge.
- Hagström, M., 2019. Military applications of machine learning and autonomous weapons. Em V. Boulanin, ed., *The impact of AI on strategic stability and nuclear risk*. Euro-atlantic perspectives, Volume I. Solna: SIPRI, pp. 32-38.
- Jiménez-Segovia, R., 2019. Los sistemas de armas autónomos en la Convención sobre Ciertas Armas Convencionales: sombras legales e éticas de una autonomía bajo el control humano? *Revista Electrónica de Estudios Internacionales*, Vol. 37, pp. 36-69.
- Kroenig, M. e Gopalswamy, B., 2018. Will disruptive technology cause nuclear war? *Bulletin of the Atomic Scientists*, November 12. Disponível em <https://thebulletin.org/2018/11/will-disruptive-technology-cause-nuclear-war/>. Acesso em 15 de outubro de 2019.
- Lele, A., 2017. A Military perspective on lethal autonomous weapons systems. *UNODA Occasional Papers* 30. Nova Iorque: UNODA, pp. 57-61.
- Leveringhaus, A., 2018. What's so bad about killer robots? *Journal of Applied Philosophy*, 35(2), pp. 341-358.
- Livingson, S. e Risse, M., 2019. The future impact of Artificial Intelligence on humans and human rights. *Ethics & International Affairs*, 33(2), pp. 141-158.
- Maas, M., 2019. How viable is arms control for military artificial intelligence? Three lessons from nuclear weapons. *Contemporary Security Policy*, 40(3), pp. 285-311.
- Morgan, P., 2012. Elements of a general theory of arms control. Em Williams, R. and Viotti, P., eds., *Arms Control*. Santa Barbara: Praeger, pp. 15-40.
- Newman, E. e Richmond, O. P., 2001. *The United Nations and human security*. Londres: Palgrave Macmillan, pp. 168-189.
- Payne, K., 2018. Artificial Intelligence: A revolution in military affairs. *Survival*, October/November, pp.7-32.
- Pinto, S. C., 1988. As relações este-oeste e o controlo de armamentos. *Nação e Defesa*, n.º 48, pp.101-141.
- Ramcharan, B. G., 2002. *Human rights and human security*. Dordrecht: Martinus Nijhoff Publishers.
- Saraiva, F., 2018. O Tratado Sobre a Proibição de Armas Nucleares. *IDN Brief*, novembro. Disponível em [https://comum.rcaap.pt/bitstream/10400.26/25255/1/idnbrief\\_novembro2018.pdf](https://comum.rcaap.pt/bitstream/10400.26/25255/1/idnbrief_novembro2018.pdf). Acedido em 10 de dezembro de 2019.

- Saraiva, F., 2010. O novo conceito estratégico da NATO e a política nuclear da Aliança. *IDN Brief*, novembro, pp. 2-6. Disponível em [https://www.idn.gov.pt/publicacoes/newsletter/idnbrief\\_01.pdf](https://www.idn.gov.pt/publicacoes/newsletter/idnbrief_01.pdf). Acedido em 12 de dezembro de 2019.
- Sauer, F., 2016. Stopping “killer robots”. Why now is the time to ban autonomous weapons systems. *Arms Control Today*, October. Disponível em <https://www.armscontrol.org/act/2016-09/features/stopping-%E2%80%98killer-robôs%E2%80%99-why-now-time-ban-autonomous-weapons-systems>. Acedido em 12 de novembro de 2019.
- Sayler, K., 2019. Artificial Intelligence and national security. Washington: Congressional Research Service R45178. Disponível em <https://news.usni.org/2019/11/22/report-to-congress-on-artificial-intelligence-and-national-security>. Acedido em 20 de novembro de 2019.
- Sharikov, P., 2018. Artificial Intelligence, cyberattack, and nuclear weapons: a dangerous combination. *Bulletin of the Atomic Scientists*, 74(6), pp. 368-373.
- SIPRI Yearbook 2019*. Oxford: Oxford University Press/SIPRI.
- Thräner, O., 2018. *New challenges in nuclear arms control*. Center for Security Studies, ETH Zurich, 5 October. Disponível em <https://css.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/CSSAnalyse232-EN.pdf>. Acedido em 10 de outubro de 2019.
- Tikk-Ringas, E., 2012. Developments in the Field of Information and Telecommunication in the Context of International Security: Work of the UN First Committee 1998-2012. *Cyber Policy Process Brief*, CT4Peace Foundation. Disponível em <https://ict4peace.org/wp-content/uploads/2012/08/Eneken-GGE-2012-Brief.pdf>. Acedido em 12 de outubro de 2019.
- United Nations Institute for Disarmament Research (UNIDIR), 2018. The weaponization of increasingly autonomous technologies: Artificial Intelligence. A primer for CCW delegates. *UNIDIR Resources*, No. 8. Geneva: UNIDIR. Disponível em <https://unidir.org/sites/default/files/publication/pdfs/the-weaponization-of-increasingly-autonomous-technologies-artificial-intelligence-en-700.pdf>
- United Nations Office for Disarmament Affairs (UNODA), 2017. Perspectives on lethal autonomous weapons systems. *UNODA Occasional Papers*, No. 39. Nova Iorque: United Nations.
- Williams, R. e Viotti, P., eds., 2012. *Arms Control*. Santa Barbara: Praeger.
- Women’s International League for Peace and Freedom (WILPF), 2020. *A WILPF guide to killer robots*, 3<sup>rd</sup> ed. January, WILPF. Disponível em <http://www.reachingcriticalwill.org/images/documents/Publications/wilpf-guide-aws.pdf>. Acedido em 30 de janeiro de 2020.
- Zawieszka, K., 2017. An ethical perspective on autonomous weapons systems. *UNODA Occasional Papers*. No. 30. Nova Iorque: United Nations, pp. 49-56.