Armas inteligentes no ciberespaço: oportunidades inovadoras e desafios prementes

Intelligent weapons in cyberspace: innovative opportunities and pressing challenges

Rev. Bras. Est. Def. v. 7, nº 2, jul./dez. 2020, p. 133-157
DOI: 10.26792/RBED.v7n2.2020.75211
ISSN 2358-3932

ANA CAROLINA DE OLIVEIRA ASSIS
NATHALIA VIVIANI BITTENCOURT
SANDRA MARIA BECKER TAVARES

INTRODUÇÃO

As definições de espaço cibernético e de inteligência artificial (doravante, IA) são controversas e difusas, cada qual à sua maneira. O ciberespaço é, via de regra, um domínio de exploração, criação e tráfego de informações através do uso de eletrônicos em redes interconectadas de sistemas de computadores,[1] enquanto o conceito de IA abrange uma diversidade de técnicas de computação capazes de executar tarefas e resolver problemas antes limitados à cognição humana[2] (Boulanin *et al.* 2019). No contexto contemporâneo, as tecnologias de informação e comunicação estão massivamente presentes na sociedade, desde os computadores pessoais a infraestruturas críticas que dependem do funcionamento remoto de redes (por exemplo, eletricidade, redes de esgoto, sistema financeiro). Da mesma forma, podemos observar a aplicação de IA em diversos setores da vida

Ana Carolina de Oliveira Assis — Doutoranda no Programa de Pós-Graduação em Ciência Política pela Universidade Federal de Pernambuco. Graduada no curso de Relações Internacionais da Universidade Federal da Paraíba e Mestrado em Ciência Política pela Universidade Federal de Pernambuco. Membro do Grupo de Pesquisa O Brasil e as Américas (UFPE) e Grupo de Pesquisa sobre Estratégia e Segurança Internacional (UFPB).
Nathalia Viviani Bittencourt — Doutoranda em Ciência Política com ênfase em Relações Internacionais pela Universidade Federal de Pernambuco. Graduada em Direito pela mesma Universidade e advogada. Membro do Grupo de Pesquisa O Brasil e as Américas (UFPE) e pesquisadora da Rede de Ciência e Tecnologia e Inovação em Defesa: Cibernética e Defesa Nacional (Pró-Defesa IV).
Sandra Maria Becker Tavares — Doutora em Bioética, Ética Aplicada e Saúde Coletiva pela Fundação Oswaldo Cruz (2014). Professora Adjunta 3 da Universidade Federal do Rio de Janeiro. Coordenadora Adjunta (vice-diretora) do Instituto de Relações Internacionais e Defesa (IRID/UFRJ) e Coordenadora de Extensão da Decania do Centro de Ciências Jurídicas e Econômicas (CCJE/UFRJ). É membro titular do Conselho de Extensão Universitária da UFRJ.

cotidiana, como aplicativos que otimizam nossas escolhas e a produção de artefatos de reconhecimento facial e de voz. No que concerne às suas aplicações na esfera Estatal, ambos os termos podem ser considerados em ampla ascensão estratégica à segurança e à defesa nacional.

Nessa perspectiva, governos estão dedicando mais possibilidades para questões referentes ao espaço cibernético — tanto na criação de instituições, subdivisões e meios para defesa/ataque aos inimigos, como também ao emprego da IA para, *inter alia*, construção de armas crescentemente autônomas, sistemas robóticos, sensores responsáveis para identificar alvos (Morgan *et al.* 2020). A título ilustrativo, o Departamento de Defesa dos Estados Unidos (***DOD —Department of Defense***), em 2018, teve o seu Comando Cibernético (CyberCom) unificado para um Comando de Combate, o que significa mais autonomia na condução de operações, considerando que o domínio cibernético está mudando muitos aspectos da guerra (Ferdinando 2018). Por seu turno, as capacidades da IA têm influenciado muitos Estados a elaborarem documentos estratégicos que incentivem seus investimentos em setores de interesse nacional, a exemplo da China, Rússia, Índia, Canadá, União Europeia, Reino Unido e Estados Unidos (EUA) (Future of Life Institute 2020).

No que tange ao governo brasileiro, algumas iniciativas coordenadas têm recentemente atribuído grande relevância ao ciberespaço e à IA. O Ministério da Tecnologia, Ciência e Inovações (MCTIC) está elaborando um documento que identifica as áreas prioritárias que podem ter seus processos aperfeiçoados pelos benefícios dessa tecnologia (Brasil 2019). Ademais, em fevereiro de 2020 foi aprovada a Estratégia Nacional de Segurança Cibernética (E-Ciber) pelo Decreto nº 10.222. Dentre os seus objetivos de tornar o espaço cibernético mais resiliente e seguro, o documento apresenta a necessidade do desenvolvimento de mecanismos avançados que maximizem o combate a ameaças cibernéticas, a saber:

> As ameaças cibernéticas [...] têm o escopo de alcançar grande número de organizações, inclusive as representantes das infraestruturas críticas, que, por prestarem serviços essenciais à sociedade, possuem elevado nível de criticidade. Por isso, essas organizações necessitam de meios para identificar, proteger, detectar, avaliar, responder, recuperar e assim gerenciar o risco das ameaças cibernéticas, e também de ferramentas de automação de segurança que usam inteligência artificial e aprendizado de máquina, que permitam analisar, identificar e conter os ataques cibernéticos. (Brasil 2020, 6)

Diante disso, a exploração da IA seria um desses meios que poderiam contribuir na segurança e resiliência cibernética. Segundo Taddeo *et al.*

(2019), estima-se que os investimentos em IA com a finalidade de cibersegurança pelos países vão ter um salto de US$1 bilhão para US$ 34 bilhões no intervalo de nove anos (entre 2016 e 2025). Ainda sobre esse aparato, os autores demonstram que o fenômeno é de tamanha relevância que as Estratégias de Segurança e de Defesa Cibernética de alguns países como a Austrália, Japão, Singapura, Estados Unidos, China e Reino Unido já mencionam a utilização de IA para aprimorar a proteção de suas infraestruturas críticas. De fato, o pensamento de Morgan *et al.* (2020) assenta que essa tecnologia tem o potencial de auxiliar na identificação de *malwares* e das próprias vulnerabilidades das redes — o que permite, assim, o aprimoramento da defesa — e na localização de brechas dos sistemas de inimigos a fim de explorá-las para, eventualmente, promover um ataque.

Isso posto, o objetivo do presente artigo é descrever como o emprego de IA tem influenciado no ataque e ao combate de ameaças cibernéticas pela Defesa, assim como inserir o leitor no debate acerca dos benefícios e dos perigos do uso dessa tecnologia no ciberespaço. A questão investigativa que norteia este estudo é: quais são as oportunidades e desafios que a expansão do uso das tecnologias de IA proporciona ao armamento e à segurança cibernética no ramo militar? Acredita-se que o presente estudo seja relevante devido a fatores de duas ordens. No âmbito teórico, a consulta bibliográfica permitiu inferir ser ainda incipiente o debate nas academias de Relações Internacionais (RIs) e de Defesa Nacional sobre a interseção da Inteligência Artificial no espaço cibernético, de modo que se busca contribuir à literatura, especialmente aos estudos de segurança das RIs e de Defesa. De ordem prática, é significativo apontar a importância estratégica que Estados têm conferido a esse assunto.

Para atender ao objetivo do estudo, optou-se pela abordagem qualitativa. No primeiro momento, foi utilizada uma revisão bibliográfica da literatura dos Estudos de Segurança e de Defesa sobre o espaço cibernético e a exploração da IA. Em sequência, foi realizada, com o auxílio da Teoria da Balança Ofensiva-Defensiva Cibernética, uma análise das relações internacionais em relação ao emprego da IA neste domínio, introduzindo conceitos como poder cibernético (*cyberpower*), dissuasão cibernética (*cyberdeterrence*) e dilema da segurança cibernética (*cyber security dilemma*). Com efeito, entendemos que os pressupostos trazidos pela Teoria à análise dos custos das operações de ataque e de defesa revelou-se pertinente ao objetivo do artigo, sendo este a descrição das oportunidades e desafios que as aplicações de IA introduzem às operações realizadas no domínio cibernético. Por fim, foram levantadas, com apoio também de uma revisão bibliográfica e de documentos oficiais, as limitações e oportunidades da IA para Segurança e Defesa Cibernética.

Como resultados, entendemos que o desenvolvimento de IA no espaço cibernético é uma faca de dois gumes: ao mesmo tempo em que o processamento rápido de grande volume de dados e o reconhecimento de padrões pela programação computacional podem ser grandes aliados na antecipação e combate de ameaças a redes, por outro, novas vulnerabilidades exigem cautela em sua operacionalização. Além disto, acreditamos que, no que tange à Teoria da Balança Ofensiva-Defensiva Cibernética, as capacidades da referida tecnologia podem promover o favorecimento da defesa em detrimento da primazia do ataque.

A REVOLUÇÃO DA INFORMAÇÃO E A MILITARIZAÇÃO DO ESPAÇO CIBERNÉTICO

A partir da ideia de que a informação sempre foi um recurso de poder para os Estados, Nye (2010) assume que o poder cibernético (*cyberpower*) possui idiossincrasias que tornam o seu conceito mais difuso e volátil. Ao contrário dos domínios tradicionais, o ciberespaço é uma criação artificial do homem que permite a redução de custos no emprego de ataques, dificulta a atribuição de responsabilidade e facilita a proliferação de ofensivas por atores não-estatais e por Estados considerados menos influentes na política internacional. Nessa acepção, o autor define o poder cibernético como a capacidade dos Estados em manipular, criar e explorar informações eletrônicas e baseadas em sistemas de computador que podem promover vantagens neste domínio ou influenciar eventos em outros ambientes operacionais (Nye 2010)

Nessa perspectiva, a capacidade das operações cibernéticas afetarem outros âmbitos da esfera social e militar, bem como as suas singularidades de ataques e do anonimato de atores promovem, inevitavelmente, a sua securitização na agenda política dos Estados (Waever *et al.* 1993; Waever 1995; Buzan *et al.* 1998; Cavelty 2008). Sob esse prisma, destaca-se o trabalho pioneiro de Hansen e Nissenbaum (2009), cuja análise dos ataques cibernéticos de negação de serviços (DDoS — Distributed Denial of Service[3]) na Estônia, em 2007, foi realizada à luz da Escola de Copenhague. Dessa forma, as autoras situam a questão cibernética como um setor específico de segurança — não apenas subordinado a um problema econômico, técnico ou criminal — mas que se apresenta de forma independente e que provoca consequências de ordens políticas e normativas próprias. Ato contínuo, concluem que existe a necessidade premente de que os Estudos de Segurança adotem uma perspectiva interdisciplinar na abordagem teórica do fenômeno cibernético. Com efeito, a avaliação das operações neste domínio exige que a abordagem seja

multifacetada, inovadora, e que inclua outros ramos do conhecimento para sua compreensão.

Outrossim, a securitização e a consequente politização do espaço cibernético tornam as suas operações matérias de grande interesse e mudanças no setor militar. Em verdade, na literatura (Schneider 2019, Cavelty 2012) argumenta-se que a revolução da informação exerce influência de modo decisivo ao seu *modus operandi*, tendo em vista o crescente desenvolvimento de tecnologias que tornam a sociedade e governos cada vez mais dependentes de serviços que envolvem processos digitais em infraestruturas críticas. Krepinevich (1994), ao fazer uma análise histórica sobre as Revoluções Militares e o fator decisivo de tecnologias emergentes como recurso de poder, argumenta que a Guerra do Golfo pode ser considerada precursora de uma nova Era no campo de batalha, haja vista a coordenação e integração das operações proporcionadas pela informação em redes.

Ao debruçar-se sobre o estudo cronológico e argumento desse autor, Schneider (2019) aponta que a exploração das atuais tecnologias de ponta pela Defesa, a exemplo da IA, computação quântica e armazenamento em nuvem torna o recurso de dados como uma recente etapa da Revolução Militar[4], ao lado do marco da Infantaria, Artilharia, dentre outros. Ademais, a autora conclui que existe um paradoxo de vulnerabilidade\vantagem ao espaço cibernético militarizado, tendo em vista que o aumento da dependência de redes ao funcionamento de serviços pode provocar brechas de segurança que permitem ataques ocultos e a escalada, por consequência, das tensões em conflitos. Esse argumento também é explorado por Liff (2012), ao afirmar que a rapidez nos processos de tecnologias de informação e comunicação e a expansão da digitalização nas operações podem representar um verdadeiro "calcanhar de Aquiles" em campanhas militares dos EUA (Liff 2012).

ARMAS CIBERNÉTICAS, AUTONOMIA E IA

De modo geral, pode-se afirmar que as armas cibernéticas (*cyber weapons*) são aplicações da tecnologia de informações que buscam criar efeitos negativos na disponibilidade, integridade e\ou confidencialidade nos dados de um computador individual ou de sistemas complexos de comunicação (Lin 2016). Além disso, o autor salienta que enquanto os ataques cibernéticos (*cyberattacks*) têm como objetivo afetar a disponibilidade e integridade dos dados e do funcionamento de sistemas, a exploração ou espionagem cibernética (*cyber exploitation*) se refere à busca por informações sigilosas e confidenciais. Por último, torna-se relevante destacar dois mecanismos comumente utilizados no emprego das armas cibernéticas: para Lin (2016, 115), intrusão (*penetration*) e carga útil (*payload*), os quais significam, respectivamente, a

obtenção do acesso a sistemas de computadores a partir das vulnerabilidades encontradas e a execução programada para agir após a intrusão, o que pode, a depender do canal de comunicação, ser remotamente atualizada.

Além disso, Kello (2013) revela alguns termos técnicos comuns a respeito do alcance das armas cibernéticas e suas complicações em estratégias de defesa cibernética. Ao indicar argumentos de pesquisadores sobre a vantagem da ofensiva ou da defesa em operações neste meio, o autor assinala perigos que desafiam o equilíbrio da balança. Assim, a imprevisibilidade de ataques que atingem vulnerabilidades dia-zero (*zero-day vulnerabilities*), por exemplo, torna o ataque muito difícil de ser rastreado em função da capacidade do *malware* de mascarar os seus efeitos danosos ao computador (Kello 2013). A título ilustrativo, o amplamente conhecido *worm Stuxnet* é um exemplo desse tipo de ataque, o qual foi capaz de explorar uma brecha de segurança dos sistemas de enriquecimento de urânio do Irã que não tinha cobertura (*patch*) para corrigir a sua vulnerabilidade.

Ademais, o autor identifica outros fatores de desequilíbrio, como o ataque de distribuição de negação de serviços, a complexidade da superfície de defesa (associada ao atual desenvolvimento intrincado dos sistemas de software e hardware, os quais exigem mais *expertise* na implementação de ataques e na defesa) e a presença maciça da indústria privada na cadeia de suprimentos e na infraestrutura críticas de computadores, o que dificulta a elaboração de políticas coesas no espaço cibernético pelos governos. Diante dessas assimetrias, tem-se que a atribuição de valor e custo a essas problemáticas podem conferir vantagens ao mecanismo de ataque. Entretanto, conforme poderemos observar nas próximas seções, avanços em IA na segurança cibernética podem oferecer razões que desafiam a primazia da ofensiva neste meio.

No que concerne à implementação de IA no ramo militar, Johnson (2019) aponta que não há dúvidas de que o uso de processos habilitados para essa tecnologia pode promover vantagens decisivas, haja vista a sua capacidade de fomentar, por exemplo, o sensoriamento remoto e a compressão do ciclo de tomada de decisão, sobretudo em casos de situações hostis que exijam respostas rápidas. De modo amplo, o autor categoriza em diferentes áreas de atuação as possibilidades do aumento de performance de sistemas automáticos e autônomos pela IA, sendo as mais relevantes para o nosso escopo do artigo: área de raciocínio e de tomada de decisão, as quais se relacionam com resolução de problemas e planejamento; a de representação de aprendizagem e conhecimento (aprendizagem de máquina e aprendizagem profunda[5]) e a de autonomia.

Há importante debate em setores acadêmicos e em Organizações Internacionais acerca da definição de autonomia e de suas diferenças em

relação a sistemas automáticos. Os sistemas automáticos são considerados processos que respondem a regras pré-programadas, além de apresentarem resultados previsíveis, com pouca ou nenhuma capacidade para lidar com variâncias. Autonomia, por seu turno, possui um grau mais complexo e inteligente de autocontrole, tendo em vista que pode selecionar diferentes caminhos para alcançar um objetivo (Boulanin *et al.* 2019; Morgan *et al.* 2020). Nesse contexto, tal distinção possui relevância às operações desempenhadas no ciberespaço em razão da crescente autonomia das armas desenvolvidas nesse meio, seja por intermédio de IA ou não, conforme afirmam Liivoja *et al.* (2019). A partir de alguns exemplos de capacidades cibernéticas autônomas, como o *Stuxnet*, os autores trazem à tona questões do *jus ad bellum* e *jus in bello* ao debate do *compliance* dessas operações às normas e princípios do Direito Internacional, bem como argumentam que as matérias de sistemas de armas autônomas (*autonomous weapons systems*) e das intervenções cibernéticas não podem ser tratadas de forma isolada pela comunidade internacional. Diante dessas discussões que a IA suscita nas relações internacionais, assim como as suas possibilidades de otimizar processos e decisões em diversas tecnologias no meio militar, alguns Estados, especialmente a China, Rússia e EUA têm aplicado e aprimorado algumas de suas capacidades cibernéticas por intermédio da IA com finalidade de alcançar as vantagens do pioneiro (*first-mover*) (Johnson 2019).

Nessa perspectiva, a respeito da China, sua indústria de defesa tem desenvolvido armas com considerável capacidade de inteligência e autonomia, as quais permitem alto grau de precisão nas operações, o que soma à sua capacidade de dissuasão, conforme pontua Kania (2019). Utiliza-se como exemplo dessas habilidades o uso de algoritmos inteligentes ao processamento de imagens de satélites e a criação de armas hipersônicas. Além disso, a autora afirma que o Exército de Libertação Popular, por intermédio de sua Força de Apoio Estratégico, tem integrado suas práticas relativas à guerra cibernética, eletrônica e espacial, bem como explorado pesquisas de tecnologias de IA para aprimorar a sua segurança e defesa cibernética.

No que tange à indústria de defesa do governo russo, Thornton e Miron (2020) destacam a relevância estratégica da IA ao fortalecimento de duas facetas da guerra cibernética que são levantadas pela perspectiva russa: a psicológica (*cyber-psychological*) e a técnica (*cyber-technical*). Quanto à primeira, armas cibernéticas impulsionadas pela IA seriam capazes de disseminar informações a um escopo e velocidade maiores, a exemplo das *fake news* e da produção de *deepfakes*,[6] com o intuito de exercer maior influência na consecução de seus interesses. Quanto à perspectiva técnica, aludem os autores que essa faceta configura o uso de espionagem e desenvolvimento

de malwares inteligentes que buscam vulnerabilidades em sistemas de tecnologia da informação.

Os EUA, por seu turno, têm demonstrado preocupação com a aplicação generalizada de IA no espaço cibernético. Em 2018, o evento CyCon US (2018) promoveu um debate acerca do uso da IA na segurança cibernética, oportunidade na qual dois oficiais das Forças Armadas discutiram sobre os mais recentes avanços dessa tecnologia e seus limites. Na ocasião, o Major Nathaniel D. Bastian e o Brigadeiro-General Matthew Easley defenderam que a IA tem sido muito útil na descoberta e correção automática de *malwares* e anomalias de forma rápida e eficaz, mas acreditam que essa tecnologia ainda não é capaz de proteger sistemas de ataques mais complexos, como os de dia-zero. Além disso, afirmaram que a aprendizagem de máquina tem demonstrado avanços na análise e classificação de padrões, porém também permite que novos ataques com essa estrutura sejam mais sofisticados (*adversarial-learning attacks*[7]) (Bastian and Easley 2018).

Diante do exposto, percebe-se que as tecnologias de IA têm sido amplamente utilizadas no domínio cibernético. Entretanto, as suas capacidades ofensivas e defensivas tornam o debate acerca dos custos de suas operações complexo à medida que novos tipos de ataque e de defesa baseados nessa tecnologia são desenvolvidos. Na próxima seção, vamos abordar com mais profundidade a Teoria da Balança Ofensiva-Defensiva no Ciberespaço para, em seguida, analisar algumas oportunidades e desafios que a IA possibilita à indústria de Defesa Cibernética.

TEORIA DA BALANÇA OFENSIVA-DEFENSIVA NO ESPAÇO CIBERNÉTICO

Uma vez que as pesquisas sobre as dinâmicas no ciberespaço nas Relações Internacionais são recentes, parte dos autores tenta adaptar diversos conceitos e teorias tradicionais utilizadas nos Estudos Estratégicos, nas RIs e em outras disciplinas para o novo contexto em que se inserem. Alguns estudiosos, por exemplo, utilizam a teoria da guerra proposta por Clausewitz em estudos sobre cibernética. Outros dedicam-se à própria concepção da guerra à análise de seus aspectos fundamentais e da sua natureza para travar o debate se a guerra cibernética é um fenômeno bélico de fato (Rid 2012; Liff 2012; Stone 2013), ou até para entender como a tecnologia cibernética pode transformar as estratégias e táticas no campo de batalha (Teixeira Júnior *et al.* 2017).

Além dessa perspectiva sobre o fenômeno cibernético, alguns estudos passaram a abordar a questão da securitização e politização do espaço cibernético de modo semelhante ao trabalho de Hansen e Nissenbaum

(2009). Dentre eles, citamos o de Lacy e Price (2018) e pesquisas até mais específicas que se destinam a tratar do caso brasileiro (Muggah *et al.* 2014; Medeiros *et al.* 2019).

Outro conceito afeito aos estudos das RIs refere-se à questão do Dilema de Segurança. Inicialmente, essa concepção foi elaborada por Jervis (1978) em *Cooperation under the Security Dilemma*, no qual o autor vai abordar as dinâmicas que levam à percepção de ameaça entre os atores no nível internacional e como isso impacta a relação entre eles. No que tange ao Dilema de Segurança Cibernética, é argumentado de forma inovadora por Buchanan (2017) que os princípios levantados por Jervis aplicam-se a vários momentos da história mundial, inclusive no contexto atual da era da informação. De acordo com o referido autor, no que concerne à segurança cibernética, assim como em outras esferas, as características estruturantes do sistema internacional, assim como as particularidades de cada tipo de operação, geram apreensão nos atores; por conseguinte, esse pavor pode escalar conflitos.

Por sua vez, Jervis (1978), em seu estudo seminal, complementa a sua discussão trazendo a Teoria Ofensiva-Defensiva, na qual avalia o efeito de cada uma das abordagens para escalada do conflito ou sobre os incentivos à Corrida Armamentista — por exemplo, no momento em que a balança pende para a ofensiva, o dilema de segurança tende a se intensificar, o que pode, por consequência, levar a conflitos (Garfinkel and Dafoe 2019). Sobre a Teoria da Balança Ofensiva-defensiva, uma definição mais tradicional, também atrelada aos estudos de RI pode ser apontada como:

> a razão dos custos das forças que o atacante requer para tomar o território ao custo das forças do defensor emprega. Isto é, se o defensor investir X em ativos militares, quão maior deve ser o investimento Y do atacante para adquirir as forças necessárias para tomar o território: A balança ofensiva-defensiva é a razão Y/X. Razões maiores indicam uma balança mais favorável à defesa (Glaser and Kaufmann 1998, 3).[8]

Embora a definição acima descreva como o objetivo final a tomada de território, as finalidades definidas pelos atores envolvidos podem ser distintas. De forma mais informal, a balança ofensiva-defensiva relaciona-se com os custos ou facilidades em realizar um ataque ou adotar uma postura defensiva. Ou seja, os tomadores de decisão levam em consideração nos cálculos racionais a influência da tecnologia militar nos dispêndios da defensiva e da ofensiva. (Garfinkel and Dafoe 2019, Slayton 2017).

No que tange à aplicação ao setor cibernético, Rebecca Slayton (2017) aponta para três principais perspectivas implícitas do que seria a Balança

Ofensiva-Defensiva Cibernética — implícitas porque muitos dos autores não apresentam diretamente uma definição, mas a avaliam empiricamente. Nesse sentido, a primeira perspectiva se refere aos custos relativos da ofensiva e defensiva. Dominante na academia, essa abordagem abrange a atribuição de valores (principalmente materiais) a cada operação para medir a sua equação. Já a perspectiva da Ofensiva Cibernética concentra-se mais nos resultados (*payoff*) em detrimento dos custos, na medida em que a balança está relacionada com a eficácia das operações. Por fim, a perspectiva da vantagem do pioneiro identifica as vantagens de ser o primeiro a tomar iniciativas de conflito.

Em termos gerais, a Balança Ofensiva-Defensiva Cibernética está relacionada com a análise dos custos para empreender ataques cibernéticos em relação aos custos para edificar uma estratégia de resiliência. Cada opção estratégica terá investimentos e resultados diferenciados e cabe ao tomador de decisão optar pelo caminho a seguir. Algo recorrente e que transpassa a literatura sobre a Balança Ofensiva-Defensiva é a questão do *culto à ofensiva* ou *primazia do ataque*.[9] Para adeptos dessa argumentação, a postura ofensiva normalmente é tida como a melhor opção, uma vez que os custos são menores em relação à defesa e até a própria configuração das tecnologias atuais facilitam um ataque em detrimento da defesa — um exemplo disso é a questão da vulnerabilidade dos Sistemas e a facilidade dos 'atacantes' para encontrar apenas uma fragilidade e partir para ação, em contrapartida da defesa, a qual tem que identificar todas as vulnerabilidades possíveis e mesmo assim pode ser que não seja a que o atacante vai explorar (Slayton 2017).

Entretanto, alguns acadêmicos não concordam que a ofensiva permaneça como vantagem. Slayton (2017) lança uma proposta com quatro argumentos para explicar por que as reivindicações abrangentes sobre as vantagens da ofensiva no ciberespaço são equivocadas. Em primeiro lugar, nos estudos que defendem a primazia do ataque, a Balança Ofensiva-Defensiva é analisada apenas em relação aos custos das operações, entretanto, há a necessidade de avaliar também o valor atribuído pelos tomadores de decisão aos objetivos (caso uma postura agressiva tenha consequências mais dispendiosas para os tomadores de decisão, eles devem optar por uma postura defensiva). Em segundo lugar, a tecnologia não é o fator determinante único das vantagens da defensiva e da ofensiva, sendo imprescindível avaliar o processo organizacional que rege as relações entre tecnologia e operadores habilidosos. Em outras palavras, a Balança é determinada tanto pelas habilidades possuídas pelos atores em comparação com seus adversários, como pelas complexidades dos objetivos (não adianta ter um objetivo complexo sem habilidade e vice-versa, ou sem a organização do processo da tradução dos objetivos em ações).

Em terceiro lugar, a autora demonstra que a ofensiva é colocada como superior porque em grande parte das análises, o sucesso dessa estratégia é resultado de objetivos mais limitados e da má gestão dos adversários, isto é, em muitos casos a vantagem da ofensiva recai sobre fragilidades do inimigo ao invés da superioridade tecnológica nos ataques. Por fim, a autora realiza um estudo de caso sobre o *Stuxnet*, calcula em termos materiais o custo-benefício tanto para os atacantes como para os defensores e obtém que a defensiva seria menos custosa, financeiramente, do que a ofensiva (Slayton 2017).

Garfinkel and Dafoe (2019), por seu turno, argumentam que a defensiva apresenta vantagens superiores à ofensiva conforme a quantidade de investimentos aplicados às tecnologias. Inicialmente, quando os investimentos ainda são discretos, a ofensiva é mais vantajosa, pois necessita de menos recursos, uma vez que os defensores não conseguem cobrir todos os pontos de vulnerabilidade (o que o autor denomina como *gap exploitation*). Entretanto, para os autores, à medida que as aplicações aumentam, a defesa é mais vantajosa porque os pontos de ingressão (vulnerabilidades) vão sendo cerceados gerando o desgaste ou uma *deffensive saturation.*

Ademais, Schneier (2018) pontua que a IA pode favorecer a defensiva na medida em que suas tecnologias realizam a segurança em sistemas de informação de forma mais rápida e em maior escala do que os humanos. Sendo assim, ainda que suas capacidades no espaço cibernético também favoreçam o escopo de ataques, o autor acredita que a IA tornará a defesa mais eficaz ao combate de *malwares* e anomalias tendo em vista o seu potencial de analisar um grande volume de dados a uma velocidade superior à humana.

Por último, outro conceito também relacionado à Teoria da Balança Ofensiva-Defensiva é o de dissuasão (*deterrence)*. Quanto à sua influência no espaço cibernético, autores como Taddeo (2018) e Nye (2017) argumentam que a sua concepção tradicional não é adequada para abordar as questões singulares de alcance global, anonimidade e interconectividade atribuídas a esse domínio. Nye (2017), por exemplo, entende que existem quatro meios de emprego da dissuasão no ciberespaço: a ameaça de punição (*threat punishment)*; a defesa por negação (*defense denial*), emaranhamento (*entanglement)* e taboos normativos (*normative taboos)*. Quanto ao primeiro, o autor baseia-se na obra pioneira de Libicki (2009) para defender a possibilidade do emprego de punição como forma de dissuasão no espaço cibernético, ainda que o problema de atribuição persista. O segundo se refere à redução de incentivos a ataques em razão de a defesa ser capaz de torná-la ineficaz. Acerca deste ponto, Nye (2017) entende que as novas tecnologias terão papel fundamental na mudança da primazia da ofensiva no espaço

cibernético pelo emprego da dissuasão por meio da defesa por negação. Quanto aos aspectos do emaranhamento e os *taboos* normativos, o autor entende que o fenômeno da interdependência poderá tornar os custos de um ataque muito superiores do que seus benefícios devido a retaliações por outros meios, como respostas econômicas e diplomáticas.

Taddeo (2018), por sua vez, fornece argumentos para uma Teoria da dissuasão cibernética. Assim, a dissuasão *cyber* é composta por três elementos centrais: Identificação do alvo; Retaliação; e Demonstração. Primeiramente, a identificação dos sistemas que realizaram o ataque é essencial para que ocorra a dissuasão — independentemente do conhecimento ou não de quem são os agressores — pois, dessa forma, o defensor pode isolar a ameaça, contra-atacar e obter a retaliação almejada. Acerca da Retaliação, a autora afirma que apenas a intimidação do oponente não é suficiente. Para deter novos ataques a seus sistemas, o ator envolvido deve infringir danos nos sistemas do inimigo para que ele perca a intenção de atacar novamente. Nesse sentido, a imprevisibilidade e inevitabilidade de ataques podem tornar a retaliação aconselhada como forma de defesa. Com efeito, se tradicionalmente os Estados realizam demonstrações públicas nas relações internacionais com armamentos ou ataques para intimidação e ameaça ao oponente, isso não teria efeito dissuasório no domínio cibernético em razão de que Estados não divulgam muitas informações quando são vítimas ou quando atacam. O argumento geral de sua teoria pode ser resumido através da Figura 1:

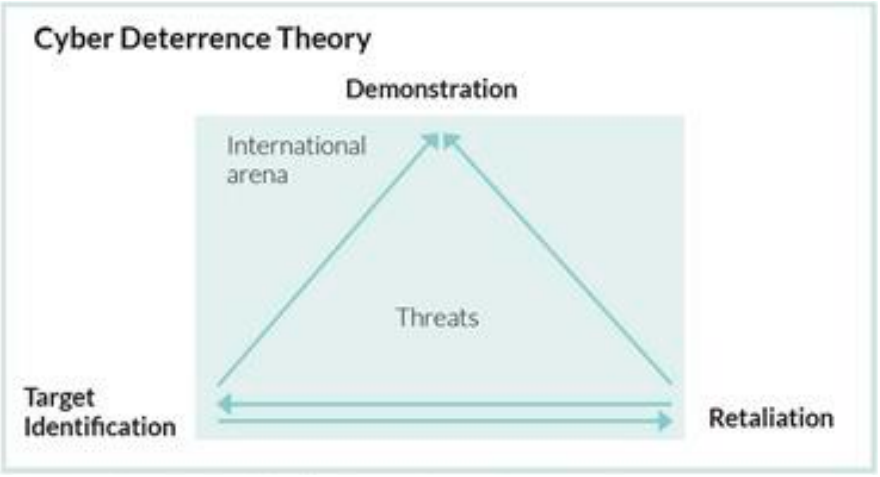

Figura 1 — Teoria da Dissuasão Cibernética
Fonte:Taddeo 2018, 5.

OPORTUNIDADES E DESAFIOS PROPORCIONADOS PELO NEXO CYBER-IA

Nesta seção, vamos apresentar algumas das oportunidades que as tecnologias de IA proporcionam à segurança cibernética, bem como seus riscos inerentes às operações e às possibilidades de ataques coordenados a esses sistemas. Sendo assim, torna-se relevante destacar, neste primeiro momento, recente White Paper sobre Segurança e Inteligência Artificial publicado por um think thank do Ministério da Indústria e Tecnologia da Informação da China, em 2018, no qual uma estrutura de segurança em IA foi analisada sob a perspectiva de três áreas estratégicas, como explicitado na Figura 2:

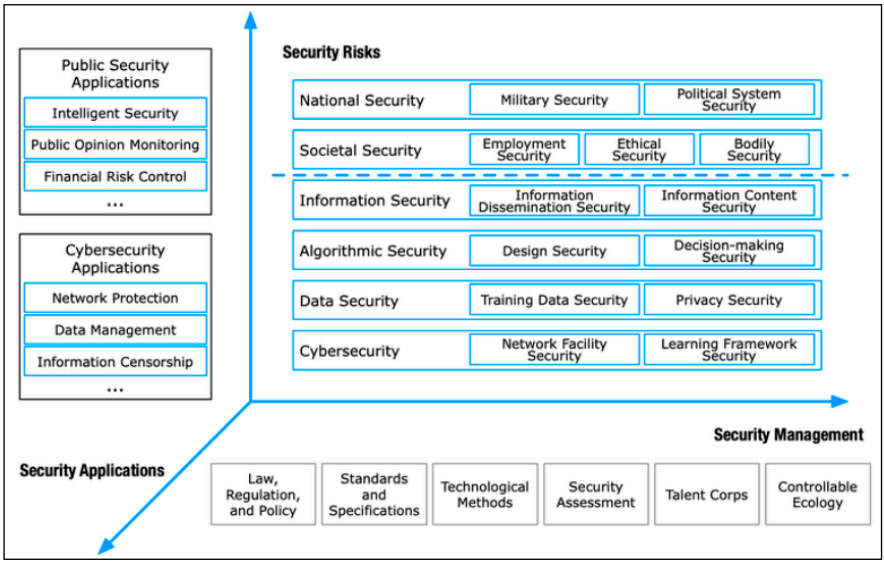

Figura 2 — Quadro do Sistema de Segurança em Inteligência Cibernética
Fonte: China Academy of Information and Communications Technology (CAICT), 2018, traduzido por *Kania et al.* (2019).

No âmbito da relação entre cibernética e IA, dentro do quesito de riscos de segurança (*Security Risks*), podemos observar que o grupo de pesquisa chinês verifica quatro ameaças que a IA pode influenciar no espaço cibernético. Em relação à segurança na informação (*Information Security*), considera-se fator de atenção a forma pela qual a IA potencialmente afeta

a velocidade na disseminação e manipulação de seu conteúdo. Além disso, a segurança algorítmica (*Algorithmic Security*) requer mais estudos acerca de seu *design* (maturidade técnica do algoritmo) e da sua tomada de decisão segura (*decision-making security*) no que tange a questões de explicação dos processos dos algoritmos de caixa-preta.[10] Quanto à segurança dos dados (*Data Security*), levantam-se problemas em matéria de privacidade dos indivíduos e do seu tratamento pela programação. Por último, no que se refere a questões de cibersegurança, revela-se a preocupação sobre as vulnerabilidades na infraestrutura das redes e riscos sistêmicos causados por aplicações maliciosas da IA.

Por outro lado, a Figura 2 compartilhada acima também apresenta algumas aplicações de IA (*Security Applications*) em segurança cibernética que têm demonstrado avanços na detecção e neutralização de *malwares* de forma automática. Sendo assim, a IA pode envolver formas dinâmicas de defesa de elementos maliciosos nas redes a partir do seu potencial de aprendizagem. Algumas empresas já utilizam essas aplicações de IA para segurança cibernética, a exemplo da *Darktrace*, proveniente do Reino Unido, que propõe um método de antivírus inovador que reconhece aspectos de softwares maliciosos sem a necessidade de confiar em uma lista predefinida, já que os métodos tradicionais dependem de ameaças históricas que se baseiam em "assinaturas" de vírus (Babuta *et al.* 2020, 10).

Ademais, torna-se importante apresentar os resultados de entrevistas a especialistas em IA sobre benefícios e malefícios gerais que suas aplicações promovem ao setor militar (Morgan *et al.* 2020). A Figura 3 reflete as principais concordâncias acerca das suas oportunidades, a saber:

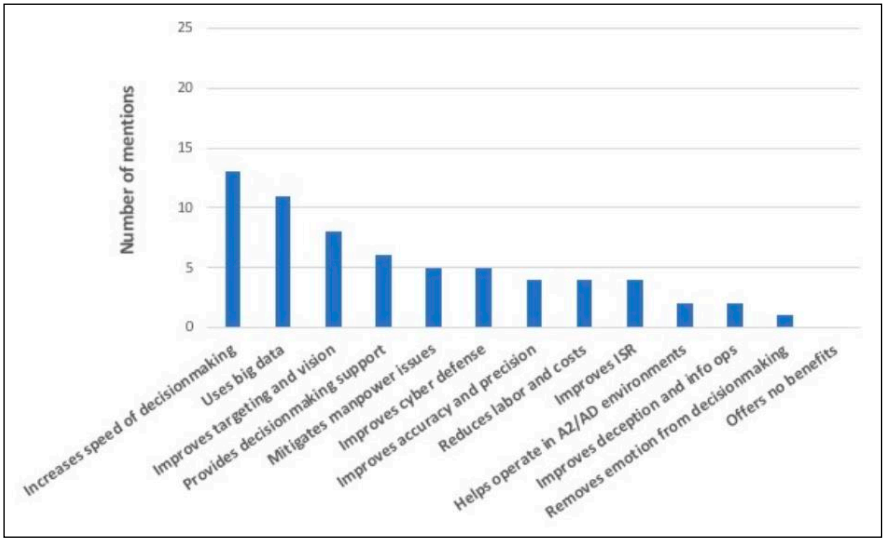

Figura 3 — Benefícios Potenciais de Aplicações Militares de Inteligência Artificial Identificadas nas Entrevistas Estruturadas
Fonte: Morgan et al. 2020, 16.

Dentre elas, destacam-se o aumento da velocidade da tomada de decisão e seu suporte, o uso de *big data*, o aprimoramento da precisão e acurácia de equipamentos, bem como o aperfeiçoamento da segurança cibernética. De fato, a IA é capaz de promover mais eficiência ao fortalecimento da robustez, resposta e resiliência dos sistemas de computadores, conforme assentam Taddeo *et al.* (2019). Na visão dos autores, o primeiro potencial torna o sistema capaz de continuar a se comportar conforme previsto mesmo com estímulos anômalos em seus *inputs*, o que pode reduzir o impacto de ataques de dia-zero, enquanto o segundo significa o aumento da capacidade de autonomia na defesa e em contra-ataques, a exemplo dos *honeypots*. Quanto à resiliência, a IA enriquece a capacidade dos sistemas a suportarem ataques, o que facilita a análise de detecção de ameaças. Tendo em vista essas vantagens táticas e estratégicas, o artigo argumenta que há muita expectativa acerca de suas aplicações à segurança cibernética, porém isso não garante que o sistema se torne totalmente imune a ameaças.

Nessa perspectiva, apresentamos as respostas acerca dos riscos que a IA provoca no ramo militar respondidas pelos especialistas, as quais podem ser conferidas na Figura 4:

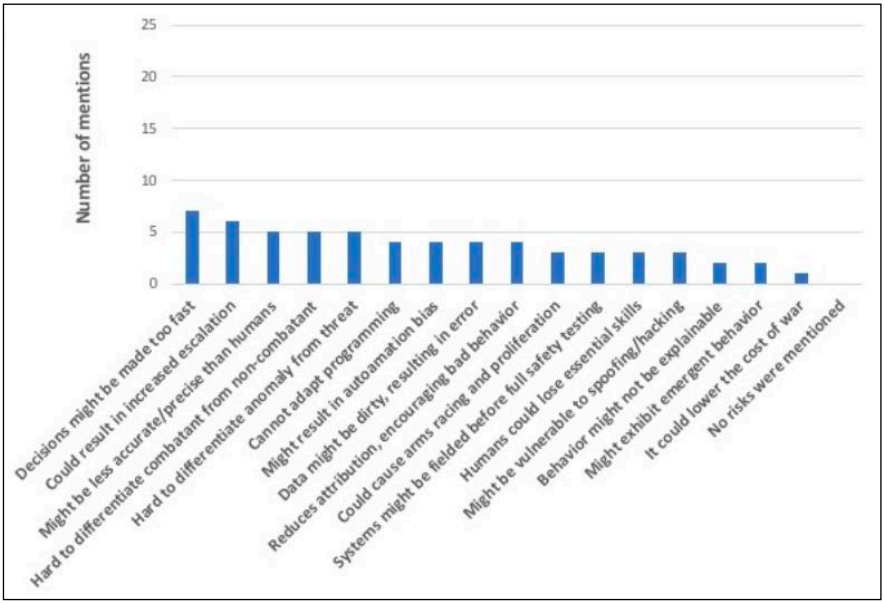

Figura 4 — Riscos das Aplicações Militares da Inteligência Artificial Identificadas nas Entrevistas Estruturadas
Fonte: Morgan et al. 2020, 21

Dentre os riscos mencionados, nota-se o incremento da escalada de conflitos, da corrida armamentista, dos erros surgidos em decorrência da manipulação errada de dados e da vulnerabilidade a *hacking* e a ataques de *spoofing*. Brundage *et al.* (2018) fazem uma análise acerca das ameaças de segurança de IA em três domínios: o digital, físico e político. O relatório menciona a necessidade de tornar os modelos de IA mais maduros a fim de mitigar a sua exploração maliciosa. Além disso, no que concerne aos riscos identificados no âmbito digital, os autores enfatizam o envenenamento de dados (*data poisoning*) e o aprendizado de máquina adversário como possíveis ataques que tentarão acessar os sistemas de forma sorrateira.

Diante dessas oportunidades e desafios, Johnson (2019) argumenta que as características disruptivas da IA podem provocar uma corrida armamentista, especialmente no atual cenário de competição entre a China e os EUA por influência na política internacional. Além disso, as questões de dilema de segurança assinaladas também estimulam Estados a se protegerem por intermédio de um aumento de suas capacidades cibernéticas, sendo a IA uma dessas opções. Com efeito, as suas ferramentas que fortalecem os sistemas de segurança da informação são aliadas importantes no aspec-

to da dissuasão por negação (*defense denial*), o que poderá desestimular o emprego de ataques no ciberespaço. Contudo, para que as tecnologias de IA tenham impacto relevante na segurança cibernética, especialmente em relação à proteção de redes de infraestruturas críticas dos Estados, torna-se fundamental que sejam desenvolvidas políticas coordenadas acerca da exploração da IA na defesa de sistemas de computadores.

CONSIDERAÇÕES FINAIS

Nosso artigo buscou ilustrar algumas das oportunidades que as tecnologias de IA têm proporcionado ao espaço cibernético, sobretudo em relação às expectativas de avanços no combate a *malwares* em uma velocidade e escala notáveis. Entretanto, também foram demonstrados alguns desafios quanto ao desenvolvimento de ataques mais sofisticados às redes de computadores, o que aumenta a complexidade na aferição da relação dos aspectos ofensivos e defensivos nos conflitos cibernéticos.

Quanto às oportunidades promovidas pela IA, mencionou-se a sua capacidade de fortalecer mais robustez, tempo de resposta e resiliência aos sistemas de computadores sobre ameaças do espaço cibernético, na medida em que os seus modelos podem auxiliar na manutenção e resistência dos programas, bem como no combate a anomalias. Quanto aos desafios, por sua vez, identificamos o envenenamento de dados (*data poisoning)* e os ataques de aprendizado adversário (*adversarial-learning attacks)* como riscos que essa tecnologia pode provocar no meio digital, especialmente em relação ao funcionamento de sistemas.

Outrossim, no que concerne à concepção da primazia do ataque ou culto à ofensiva no mundo cibernético, entendemos que esta nem sempre deve ser considerada opção estratégica superior à defensiva, uma vez que os danos causados podem ser muito inferiores ao seu custo material e político. Além disso, sugerimos que os avanços dos modelos e técnicas de IA podem favorecer a segurança dos sistemas em razão de sua eficiência em termos de escala e velocidade no combate a ataques. Diante dessas razões, programas computacionais impulsionados pela IA podem fazer a balança pender para a defesa. Sobre a possibilidade de se investir em políticas cibernéticas retaliatórias como forma de dissuasão, salientamos que essa alternativa pode levar à escalada do dilema de segurança e de conflitos entre os atores.

Nessa esteira, acreditamos que seja necessário tornar os sistemas de IA mais confiáveis (*reliability*) e explicáveis (*explainability*) para garantir que suas tarefas sejam executadas em consonância com o planejado e para permitir que seus processos autônomos de segurança cibernética sejam mais robustos (Taddeo *et al.* 2019) na mitigação de vulnerabilidades e neu-

tralização de ataques. Contudo, isso torna fundamental a implementação de regulamentações e estandardizações profissionais para aqueles que produzem e administram ferramentas com IA, de modo a criar uma cultura padronizada e, por consequência, mais segura acerca do desenvolvimento de suas aplicações.

Em arremate, entendemos que as políticas de ataque e defesa no espaço cibernético tornam indispensáveis a colaboração multissetorial de técnicos, pesquisadores e militares para que as suas formulações sejam coordenadas e estratégicas ao interesse nacional. Além disso, recomendamos que as questões de *compliance* às normas e princípios do Direito Internacional sejam contempladas por intermédio de medidas didático-pedagógicas em conteúdos teórico-práticos, baseadas em conhecimentos de Psicologia, Ética e Sociologia ao longo da formação profissional do combatente virtual.

REFERÊNCIAS

Babuta, Alexander, Marion Oswald, and Ardi Janjeva. 2020. "Artificial Intelligence and UK National Security". *RUSI Occasional Paper*. Royal United Services Institute.

Bastian, Nathaniel D., and Matthew Easley. 2018. "Artificial Intelligence in Cyber Security". Filmed 2018 at CyCON US, Army Cyber Institute. Vídeo, 1:14:22. https://www.youtube.com/watch?v=YEejuT2s5QQ&t=2582s

Battaglia, Rafael. 2020. "Afinal, o que são *deepfakes*?". Revista *Superinteressante*. https://super.abril.com.br/tecnologia/afinal-o-que-sao-deepfakes/.

Boulanin, Vincent, Lora Saalman, Petr Topychkanov, Fei Su, and Moa Carlsson. 2019. *Artificial Intelligence, Strategic Stability and Nuclear Risk*. Solna: SIPRI Publications.

Boulanin, Vincent, and Maaike Verbruggen. 2017. *Mapping the Development of Autonomy in Weapon Systems*. Solna: SIPRI Publications.

Boulanin, Vincent, Neil Davison, Netta Goussac, and Moa Carlsson. 2020. *Limits on Autonomy in Weapon Systems: Identifying Practical Elements of Human Control*. Solna: SIPRI Publications.

Brasil. Ministério da Ciência, Tecnologia e Inovações. 2019. "Estratégia Brasileira de Inteligência Artificial". www.mctic.gov.br/mctic/opencms/inovacao/paginas/politicasDigitais/Inteligencia/Artificial.html.

_____. 2020. "Decreto nº 10.222, de 5 de fevereiro de 2020". Diário Oficial da União (DOU). http://www.in.gov.br/web/dou.

Brundage, Miles, *et al.* 2018. *The Malicious Use February 2018 of Artificial Intelligence: Forecasting, Prevention, and Mitigation.* https://maliciousaireport.com/.

Buchanan, Ben. 2017. *The Cybersecurity Dilemma: Hacking, Trust and Fear Between Nations.* New York: Oxford University Press.

Buzan, Barry, Ole Wæver, and Jaap de Wilde. 1998. *Security: A New Framework for Analysis.* Boulder: Lynne Rienner.

Cavelty, Myriam Dunn. 2008. *Cyber-Security and Threat Politics: US Efforts to Secure the Information Age.* London: Routledge.

_____. 2012. "The Militarization of Cyber Security as a Source of Global Tension". *Cyber Security* 22: 103–24.

Coady, C.A.J. 2012. The Jus Post Bellum. In *New wars and new soldiers: military ethics in the contemporary world*, edited by P. Tripodi, and J. Wolfendale. Farnham: Ashgate

Ferdinando, Lisa. 2018. "Cybercom to Elevate to Combatant Command". U.S. Department of Defense. www.defense.gov/Explore/News/Article/Article/1511959/cybercom-to-elevate-to-combatant-command/.

Future of Life Institute. 2020. "National and International AI Strategies". *Future of Life Institute Organization.* https://futureoflife.org/national-international-ai-strategies/.

Garfinkel, Ben, and Allan Dafoe. 2019. "How does the offense-defense balance scale?". *Journal Of Strategic Studies* 42, no. 6: 736–63. http://dx.doi.org/10.1080/01402390.2019.1631810.

Glaser, Charles L., and Chaim Kaufmann. 1998. "What is the Offense-Defense Balance, and Can We Measure it?". *International Security* 22, no. 4: 44–82.

Hansen, Lene, and Helen Nissenbaum. 2009. "Digital Disaster, Cyber Security, and the Copenhagen School". *International Studies Quarterly* 53, no. 4: 1155–75.

Jervis, Robert. 1978. "Cooperation under the Security Dilemma". *World Politics* 30, no. 2: 167–214.

Johnson, James. 2019. "The AI-Cyber Nexus: Implications for Military Escalation, Deterrence and Strategic Stability". *Journal of Cyber Policy* 4, no. 3: 442–60.

Kania, Elsa B. 2019. "Chinese Military Innovation in the AI Revolution". *The RUSI Journal* 164, no. 5–6: 26–34.

Kania, Elza, Dahlia Peterson, Lorand Laskai, and Graham Webster. 2019. "Translation: Key Chinese Think Tank's 'AI Security White Paper' (Excerpts)". *New America.* http://newamerica.org/cybersecurity-initiative/digichina/blog/translation-key-chinese-think-tanks-ai-security-white-paper-excerpts/.

Kello, Lucas. 2013. "The Meaning of Cyber Revolution. Perils to Theory and StateCraft". *International Security* 38, no. 2: 7–40.

Krepinevich, Andrew. 1994. "Cavalry to Computer: The Pattern of Military Revolutions". *National Interest* 37: 29–40. https://nationalinterest.org/article/cavalry-to-computer-the-pattern-of-military-revolutions-848?page=0%2C2.

Kuehl, Dan. 2009. "From Cyberspace to Cyberpower: Defining the Problem". In: *Cyberpower and National Security*, edited by Franklin D. Kramer, Stuart H. Starr, and Larry Wentz: 24–42. Washington D.C.: National Defense University Press.

Lacy, Mark, and Daniel Prince. 2018. "Securitization and the global politics of cybersecurity". *Global Discurse* 8, no. 1: 101–15.

Libicki, Martin C. 2009. *Cyberdeterrence and Cyberwar*. Santa Monica, CA: RAND

Liff, Adam. 2012. "Cyberwar: A New 'Absolute Weapon'? The Proliferation of Cyberwarfare Capabilities and Interstate War". *Journal of Strategic Studies* 35, no. 3: 401–28.

Liivoja, Rain, Maarja Naagel, and Ann Väljataga. 2019. "Autonomous Cyber Capabilities Under International Law". *NATO CCDCOE Publications.* https://ccdcoe.org/library/publications/autonomous-cyber-capabilities-under-international-law/

Lin, Herbert. 2016. "Governance of Information Technology and Cyber Weapons". *Governance of Dual-Use Technologies: Theory and Practice*, edited by Elisa D. Harris: 112–58. American Academy of Arts & Sciences.

Lynn, William J. 2010. "Defending a New Domain: The Pentagon's Cyberstrategy". *Foreign Affairs* 89, no. 5: 97–108.

Maness, Ryan C., and Brandon Valeriano. 2016. "The Impact of Cyber Conflict on International Interactions". *Armed Forces & Society* 42, no. 2: 301–23.

Medeiros, Breno P., Alessandra C. Carvalho, and Luiz R. F. Goldoni. 2019. "Uma análise sobre o processo de securitização do ciberespaço". *Coleção Meira Mattos* 13, no. 46: 45–66.

Morgan, Forrest E., Benjamin Boudreaux, Andrew J. Lohn, Mark Ashby, Christian Curriden, Kelly Klima, and Derek Grossman. 2020. *Military Applications of Artificial Intelligence: Ethical Concerns in an Uncertain World*. Santa Monica: Rand Corporation.

Muggah, Robert, Misha Glenn, and Gustavo Diniz. 2014. "Securitização da cibersegurança no Brasil". *Cadernos Adenauer* 15, no. 4: 69–109.

Nye, Joseph S. 2010. *Cyber Power*. Cambrige: Belfer Center for Science and International Affairs.

_____. 2017. "Deterrence and Dissuasion in Cyberspace". *International Security* 41, no. 3: 44–71.

Rid, Thomas. 2012. "Cyber War Will Not Take Place". *Journal of Strategic Studies* 35, no. 1: 5–31.

Schneider, Jacquelyn. 2019. "The Capability/Vulnerability Paradox and Military Revolutions: Implications for Computing, Cyber, and the Onset of War". *Journal of Strategic Studies* 42, no. 6: 841–63.

Schneier, Bruce. 2018. "Artificial Intelligence and the Attack/Defense Balance". *IEEE Security & Privacy* 16, no. 2: 96-96. Institute of Electrical and Electronics Engineers (IEEE).

Shires, James, and Max Smeets. 2017. *Contesting "Cyber"*. Washington: New America.

Slayton, Rebecca. 2017. "What is the Cyber Offense-Defense Balance? Conceptions, Causes, and Assessment". *International Security* 41, no. 3: 72–109.

Stone, John. 2013. "Cyber War Will Take Place!". *Journal of Strategic Studies* 36, no. 1: 101–8.

Taddeo, Mariarosaria. 2018, "How to Deter in Cyberspace". *The European Centre of Excellence for Countering Hybrid Threats* 6: 1–10.

Taddeo, Mariarosaria, Tom Mcctcheon, and Luciano Floridi. 2019. "Trusting artificial intelligence in cybersecurity is a double-edged sword". *Nature Machine Intelligence* 1, no. 12: 557–60.

Teixeira Júnior, Augusto, Gills Villar-Lopes, and Marco T. Freitas. 2017. "As três tendências da guerra cibernética: novo domínio, arma combinada e arma estratégica". *Revista Carta Internacional* 12, no. 3: 30–53.

Thornton, Rod, and Marina Miron. 2020. "Towards the 'Third Revolution in Military Affairs': The Russian Military's Use of AI-Enabled Cyber Warfare". *The RUSI Journal*: 1–10 (Maio).

Turing, A. 1950. "Computing Machinery and Intelligence". *Mind, New Series* 59, no. 236: 433–60.

Wæver, Ole. 1995. "Securitization and Desecuritization". In *On Security*, edited by Ronnie Lipschutz. New York: Columbia University Press.

Wæver, Ole, Barry Buzan, Morten Kelstrup, and Pierre Lemaitre. 1993. *Identity, Migration and the New Security Agenda in Europe*. London: Pinter.

Whyte, Christopher. 2020. "Problems of Poison: New Paradigms and 'Agreed' Competition in the Era of AI-Enabled". 12th International Conference on Cyber Conflict. *NATO CCDCOE Publications*.

NOTAS

1. Neste artigo, consideramos a proposta de Kuehl (2009) para atribuir significado ao espaço cibernético. Utilizamos também a concepção de Smeets e Shires (2017), que trazem uma perspectiva acerca de como o entorno dos avanços da interconectividade de sistemas de computadores passou de um entusiasmo inicial de mundo sem fronteiras para um ambiente de constantes ameaças, crimes e espionagem.
2. Alan Turing é amplamente considerado como o pioneiro na criação de sistemas inteligentes, os quais teriam capacidade de substituir os seres humanos em algumas atividades. A partir de um jogo de imitação, Turing (1950) investiga a capacidade de uma máquina de responder a certas perguntas da mesma forma que um indivíduo. Devido aos rápidos avanços da computação neste aspecto, sobretudo em relação aos programas de aprendizagem de máquina (*machine learning)* e aprendizagem profunda (*deep learning)*, capazes de processar e identificar padrões em um grande volume de dados, adotamos o conceito geral de que IA constitui uma série de aplicações que reproduzem tarefas que geralmente requerem inteligência humana (Morgan *et al.* 2020).
3. Maness e Valeriano (2016) argumentam que o ataque de negação de serviços é um dos únicos métodos sorrateiros que afetam as dinâmicas de relações internacionais, na medida em que tem um efeito psicológico e público na sociedade conectada que exige uma resposta Estatal coordenada. No caso da Estônia (2007), sites oficiais do governo, da mídia e de bancos ficaram fora do ar, bem como diversas difamações foram direcionadas ao Primeiro Ministro. Apesar de a origem dessas ofensivas ser atribuída à Rússia, ela permanece desconhecida.
4. A autora não utiliza o termo Revolução dos Assuntos Militares (RAM) em razão de não querer se comprometer com as políticas de defesa específicas que o conceito, muito em voga nos anos 90 e 2000, exigem (Schneider 2019, 844). Dessa forma, prefere utilizar a expressão genérica de Revolução Militar para categorizar a potencialidade dessas tecnologias em promover transformações no setor.
5. De acordo com Boulanin e Verbruggen (2017, 16), aprendizagem de máquina é um método de IA capaz de executar mudanças em sua estrutura, programa ou dados de acordo com os insumos recebidos, motivo pelo qual é esperado que a sua performance tenha um melhor desempenho à medida que seu funcionamento avance. A aprendizagem profunda, por seu turno, é um tipo de aprendizagem de máquina que transforma os dados brutos em representações. Por exemplo, esta se refere a técnicas de reconhecimento facial, enquanto a primeira se relaciona com artefatos que fazem reconhecimento de voz.
6. *Deepfake* é uma aplicação da tecnologia de IA de *deep learning* capaz de manipular o conteúdo de vídeos de tal modo que torna muito difícil de serem desmascarados (Battaglia 2020).

7. Whyte (2020) afirma que o termo aprendizado adversário se refere a algoritmos de modelos de aprendizado de máquina (*machine learning*) capazes de se adaptar em diferentes ambientes hostis de sistemas de computadores. Um *malware* desenvolvido com essa aplicação poderia perturbar os dados de entrada nos sistemas, o que causaria problemas de falsos positivos ou negativos, por exemplo.
8. Do original: "the ratio of the cost of the forces the attacker requires to take territory to the cost of the forces the defender has deployed. That is, if the defender invests X in military assets, how large an investment Y must the attacker make to acquire the forces necessary for taking territory? The offense-defense balance is the ratio Y/X. Larger ratios indicate a balance more in favor of defense" (Glaser and Kaufmann 1998, 3).
9. A ideia de que a ofensiva é mais vantajosa em relação à defesa é compartilhada não apenas entre acadêmicos, mas também por militares e políticos, a exemplo do Ex Secretário de Defesa William Lynn, em 2010, o qual demonstra que os EUA não devem utilizar uma postura defensiva (o autor faz alusão à Linha Maginot, Estratégia defensiva francesa que fracassou na Segunda Guerra mundial) e que a ofensiva é superior (Lynn 2010).
10. O problema relacionado à caixa preta (*black box problem*) dos algoritmos se insere no contexto de que a complexidade de seus sistemas não oferece explicações plausíveis acerca dos caminhos percorridos ao encontro dos seus resultados. À medida que se ampliam as aplicações desse tipo de IA na vida cotidiana, há uma crescente preocupação sobre o direito dos indivíduos a uma explicação do processo de tomada de decisão algorítmica, e a respeito da confiabilidade do modelo pelos operadores.

ARMAS INTELIGENTES NO CIBERESPAÇO: OPORTUNIDADES INOVADORAS E DESAFIOS PREMENTES

RESUMO

A característica disruptiva e inovadora da Inteligência Artificial (IA) tem provocado uma miríade de benefícios e especulações em múltiplos setores da sociedade, em especial no militar. O presente artigo busca trazer à tona o debate das oportunidades e desafios que a IA pode promover em relação a ataques e à segurança no espaço cibernético. Com efeito, argumenta-se que essa nova modalidade de capacidade cibernética possui a característica de faca de dois gumes: se, por um lado, o processamento rápido de grande volume de dados e o reconhecimento de padrões pela programação computacional podem ser grandes aliados na antecipação e combate de ameaças a redes, por outro, novas vulnerabilidades exigem cautela em sua operacionalização. Em termos metodológicos, foi realizada revisão bibliográfica e documental de abordagens sobre a interseção da inteligência artificial no ramo da segurança e defesa cibernética. Além disso, nosso arcabouço teórico debruça-se sobre alguns conceitos das Relações Internacionais, em especial o da Teoria da Balança Ofensiva-Defensiva Cibernética para elucidar como as capacidades dessa tecnologia podem influenciar na dissuasão e favorecimento da defesa em detrimento da primazia do ataque.

Palavras-chave: Ciberataque; Defesa Cibernética; Vulnerabilidade Cibernética; Inteligência Artificial.

ABSTRACT

The disruptive and innovative feature of Artificial Intelligence (AI) has provoked a myriad of benefits and speculation in multiple sectors of society, especially in the military. This article seeks to bring up the debate about the opportunities and challenges that AI can promote in relation to cyberattacks and cybersecurity. Indeed, it is argued that this new type of cyber capability has a double-edged sword feature: if, on the one hand, the rapid processing of large volumes of data and the recognition of patterns by computer programming can be great allies in anticipation and combating network threats, on the other hand, new vulnerabilities require caution in their operation. In methodological terms, a bibliographic and documentary review of approaches on the intersection of artificial intelligence in the field of cyber security and defense were carried out. In addition, our theoretical framework focuses on some concepts of International Relations, especially that of the Theory of Cyber-Offensive-Defensive Balance to elucidate how the capabilities of this technology can influence deterrence and favoring defense over the primacy of attack.

Keywords: Cyberattack; Cyberdefense; Cyber Vulnerability; Artificial Intelligence.

Recebido em 01/07/2020. Aceito para publicação em 20/04/2021.