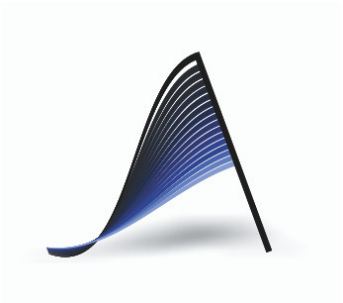
Cadernos do Ateliê

Incertezas da Inteligência Artificial (2/4):
a guerra inteligente potencializada

ISSN: 2596-2566

Cadernos do Ateliê.Vol.1, n.1, fascículo 2, fevereiro, 2018

https://ateliedehumanidades.com/cadernos-do-atelie/



Cadernos do Ateliê

Fascículo 2. Incertezas da Inteligência Artificial (2/4): a guerra inteligente potencializada

André Magnelli - IESP/UERJ
Renato Magnelli - UFBA

ISSN: 2596-2566

Direção:
André Magnelli
e-mail: direcao.ateliedehumanidades@gmail.com
Contato:
e-mail: ateliedehumanidades@gmail.com
Telefone: (021) 9 7979-3743
Site: www.ateliedehumanidades.com
Redes sociais: @ateliedehumanidades

INCERTEZAS DA INTELIGÊNCIA ARTIFICIAL (2/4):

A GUERRA INTELIGENTE POTENCIALIZADA

No dia 20 de fevereiro, foi lançado o relatório The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation por meio do qual vinte e seis especialistas em Inteligência Artificial oriundos de centros universitários (Yale, Stanford, Cambridge e Oxford) e de organizações não-governamentais (como Electronic Frontier Foundation e OpenAI), assumiram uma posição sobre os potenciais usos maliciosos das IAs com ameaças à segurança digital, física e política.

Este ensaio é o segundo de uma série que analisa e reflete sobre o relatório em questão, tendo sido produzida em parceria do Ateliê de Humanidades com o Blog do Sociofilo, escrita pela equipe do Plano de Convergência do Ateliê "Tecnociências & Sociedades: Interflúvios e Porvires da Máquina, da Vida e do (Pós-)Humano". Esse ensaio apresenta e reflete criticamente sobre os riscos gerados à segurança física pelo uso de IAs em agressões físicas, guerras e atos terroristas. A ele se seguirão dois outros, dedicados, sucessivamente, à segurança política e às propostas de regulamentação e intervenção sobre as IAs

PARCERIA

ATELIÊ DE HUMANIDADES - Espaço de livre estudo, pesquisa, escrita e formação

Plano de Convergência - Tecnociências & Sociedades: Interflúvios e Porvires da Máquina, da Vida e do (Pós-)Humano / (Ateliê de Humanidades - RJ)
contato: ateliedehumanidades@gmail.com
site: www.ateliedehumanidades.com (em construção)

BLOG DO SOCIOFILO - (CO)Laboratório de Teoria Social
site: www;blogdosociofilo.wordpress.com

Introdução

O relógio do apocalipse (*doomsday clock*), mantido desde 1947 através do Quadro de Ciência e Segurança do Boletim de Cientistas Atômicos (*Bulletin of the Atomic Scientists*), representa uma analogia à iminência de um holocausto termonuclear ao badalar da meia noite. Seu horário vem sendo ajustado dinamicamente através das análises dos riscos básicos contínuos ao qual a humanidade está submetida em uma era nuclear. Atingiu seu nadir, em 1953, ao marcar 2 minutos para a meia-noite após o sucesso em 1952 da detonação das bombas de hidrogênio na operação Ivy. Desde o evento, perseverou acima desse patamar, mesmo durante a crise dos mísseis de Cuba em 1962. Hoje, com a banalização da doutrina de destruição mútua assegurada e as escatologias trazidas pelo contínuo avanço das capacidades humanas por meio da inovação tecnológica, quase se passou despercebido que revivemos aquela mais crítica situação: de acordo com o Relatório, publicado em 25 de janeiro de 2018, são novamente 2 minutos para a meia noite.

Desde 2007, o Relatório foi modificado para incluir ameaças oriundas de tecnologias e impactos ainda não previstos ou subestimados, passando a incluir mudanças climáticas, bioterrorismo, e, como pertinente ao tema tratado, Inteligência Artificial. Mas o avanço do relógio foi motivado principalmente pela retórica hiperbólica e pelo circo de encenação de virilidade por parte dos governantes americano e norte-coreano, que veio se arrastando ao longo do ano passado e só recentemente tem se apaziguado. Os argumentos do conflito diplomático

foram protagonizados pelos *mísseis* balísticos intercontinentais (*ICBMs*) termonucleares americanos e, mais recentemente - e principalmente - norte-coreanos.

> a passagem para uma era de ciberguerras e, possivelmente, de guerras inteligentes automatizadas abre a possibilidade de que os chamados "atores semi-clandestinos" adquiram um poder de violência desproporcional a seus recursos, números e territórios

Este contexto nos traz à memória os tempos passados da Guerra Fria. Uma breve leitura do clássico *Paz e Guerra entre as Nações* (1962), de Raymond Aron, permite-nos mensurar as mudanças nos desafios de nosso tempo. Ao analisar a dialética do antagonismo na década de 1960, Aron diagnosticava um contexto de convergência entre duas séries históricas: de um lado, a inovação constante das tecnologias militares (bombas termonucleares e engenhos balísticos) conduzia a um aumento da capacidade de destruição; de outro lado, a acentuação dos elementos psicológicos dos conflitos levava à diminuição da violência física. Esta situação convergente lhe autorizava o argumento de uma dissuasão recíproca:

> O caráter desproporcional da tecnologia bélica leva a guerra à sua essência - uma prova de vontade -, seja pela substituição da ação pela ameaça, seja pela impotência recíproca das grandes potências, que impede os conflitos diretos e, ao mesmo tempo, expande o espaço onde se manifesta a *violência clandestina ou dispersa*, que não acarreta um risco excessivo para a humanidade (ARON, Raymond, p. 245, grifo nosso).

A Guerra Fria era, segundo ele, uma Era de “paz do terror” porque constituída pela iminente possibilidade de destruição mútua. Com ela, as populações encontram-se suspensas sobre uma “ameaça global e monstruosa”, uma guerra psicológica, que “reduz os homens a uma forma de passividade coletiva” (ibid., p.246).

De fato, há semelhanças com nosso tempo: continuam a nos caracterizar a passividade das populações diante do terror de uma sempre iminente ameaça de destruição global, com uma dissuasão de guerra entre potências nucleares somada a uma expansão de uma violência clandestina e dispersa. Na espiral de agressões verbais de 2017, ao se presumir uma racionalidade de americanos e norte-coreanos, supõe-se que o que está em jogo é uma performance de engajamento própria de uma situação em que o lançamento de uma bomba atômica deve ser uma terrível possibilidade na defesa de seu interesse, e nada mais que isso. Contudo, para além das permanências, houve uma mudança considerável no tocante à presença dos “atores semi-clandestinos” e suas violências dispersas: agora, elas poderão acarretar um “risco excessivo para a humanidade”, pois a passagem para uma era de ciberguerras e, possivelmente, de guerras inteligentes automatizadas abre a possibilidade de que seja adquirido um poder de violência desproporcional a seus recursos, números e territórios.

Isso nos leva a um ponto atrator das discussões do Relatório, que por ora analisamos, e dos debates mais amplos sobre armamento civil, qual seja: a existência de empoderamento de indivíduos, grupos e instituições tidos por ilegítimos, incompetentes ou, simplesmente, mal-

intencionados em suas capacidades de ameaça ou violência física; possibilidade dada pela existência de um mercado de suprimentos amplamente disponível, pela potencialização do poder destrutivo de armas automatizadas e pela possibilidade de conversão para uso militar (*weaponizing*) de ferramentas de uso originalmente civil.

Neste segundo ensaio de nossa série, começaremos por analisar as ameaças à segurança física, apresentadas pelo Relatório, que são oriundas do desenvolvimento de IAs e da robótica; em seguida, dissertaremos sobre as propostas de regulamentação e intervenção feitas pelos autores, a fim de, por último, realizar algumas reflexões e considerações críticas.

A potencialização das guerras "inteligentes"

No Relatório *The Malicious Use* é construído um amplo cenário de ameaças. Como dissemos em nosso primeiro ensaio, as IAs possuem uma capacidade de exceder as habilidades humanas na realização de um mesmo objetivo por causa de sua eficiência, escalabilidade, pronta difusão e possibilidade de cumprimento dos fins com anonimato e distância espaço-temporal e psicológica. Cada uma destas características prepara o terreno para potenciais aplicações disruptivas das IAs e, consequentemente, a curto prazo, para seu uso malicioso. Com isso, são expandidas as ameaças físicas existentes, alterados os caracteres típicos delas e introduzidas novas.

Devemos analisar a potencialização das ameaças físicas distinguindo dois aspectos das IAs: o dos *hardwares* e dos *softwares*. Embora estejam inteiramente articulados entre si, é frutífero distinguir os

dois aspectos do processo para que não prendamos nossa atenção apenas em um fato mais ostensivo - por exemplo, o surgimento de robôs utilizados para a agressão física - esquecendo as ameaças mais sutis - por exemplo, o uso de fragilidades de softwares para gerar danos.

Evidentemente, o cenário deve ser considerado, antes de tudo, no tocante ao uso militar das tecnologias de IAs, ou seja, sobre a possibilidade de surgimento de uma nova era da tecnologia bélica, onde se pergunta, dentre outras coisas: quais são as ameaças decorrentes de tecnologia militar automatizada por IAs e robôs assassinos? Mas o que preocupa os autores do Relatório não se restringe à automatização da tecnologia bélica, pois o seu desenvolvimento civil também pode ser apropriado de forma maliciosa.

Comecemos a ver então os cenários possíveis de tal uso dual (militar/civil) tomando como referência as capacidades específicas das IAs.

1. No que diz respeito aos ***hardwares***, ressalta-se, em primeiro lugar, uma das capacidades, a sua *ponta difusão*. Quando tratamos da possibilidade de seu uso em tecnologias bélicas, vêm à mente, de imediato, os drones, que são uma realidade já existente e em pleno desenvolvimento (por exemplo, os *Predators* produzidos pela *General Atomics*). Embora eles não sejam ainda inteiramente automatizados e não possam ser qualificados de “armas autônomas”, eles são tecnologias de IAs que prenunciam o desenvolvimento de armas letais autônomas (*Lethal autonomous weapons - LAWs*), sobretudo quando se vislumbra a

possibilidade de construção de “robôs assassinos” (*killer robots*).[1] Diante disso, existe um debate sobre a necessidade de banimento de tais tecnologias (ver carta aberta em caixa de texto abaixo), ou, no mínimo, de regulação. O fato que importa, por enquanto, é perceber que, uma vez construída tal tecnologia bélica, ela será de pronta difusão – muito maior do que a das armas termonucleares do século XX – com um quase certo desencadeamento de corrida dos atores estatais por sua posse e sofisticação crescente.

Mas a questão da pronta difusão não se coloca apenas no âmbito militar, mas também no civil, por causa de dois fatos: a existência de um mercado global de robótica e o uso de robôs para as mais diversas aplicações. Em primeiro lugar, o mercado de robótica já é globalizado e

[1] Em Relatório recente, publicado em novembro de 2017, Boulanin e Verbruggen mapearam o desenvolvimento da autonomia em sistema de armas letais, concluindo que a autonomia já é usada em uma ampla gama de de sistema de armas, inclusive relacionadas ao uso de força. BOULANIN, Vincent; VERBRUGGEN, Maaike. *Mapping the Development of Autonomy in Weapon Systems*. SIPRI, novembro de 2017. Como diz Chaperon (2017), mencionando o próprio Relatório, as empresas americana Northrop Grumman, a britânica BAE Systems, a chinesa LeiShen Intelligent System e a israelita IAI possuem drones aéreos, tanques e outras armas lasers crescentemente automatizadas, mas ainda não existe arma letal inteiramente autônomas. A respeito disso, Grégoire Chamayou (2015) contribui teoricamente ao refletir sobre como os drones expressam uma mutação do próprio Estado moderno enquanto tentativa de resolução do problema da soberania: “Ao inventar o drone armado, descobriu-se também, quase por acaso, outra coisa: uma solução para a contradição principal que afetava em seu centro havia vários séculos a teoria moderna da soberania política em sua dimensão guerreira. A generalização dessa arma implica a tendência a uma mutação das condições de exercício do poder de guerra, e isso na relação do Estado com seus próprios sujeitos. Seria um erro reduzir a questão das armas à esfera da violência externa. O que implicaria, para uma população, tornar-se o sujeito de um Estado-drone?”. Refletindo sobre o militarismo democrático em voga, ele vê perfilar-se no horizonte o desenho dos “autômatos políticos” e “o pesadelo de que as armas se tornem elas próprias os únicos agentes discerníveis da violência que conduzem”. Não acompanharemos o autor aqui, mas sua teoria é uma útil contribuição para estender a discussão aqui feita para o âmbito da teoria política.

possui uma cadeia de produção e distribuição de suprimentos difundida por quase todos os continentes.[2] Além disso, existem múltiplos usos da robótica e das inteligências artificiais não apenas policiais e militares, mas também humanitários, recreativos, comerciais, etc. A este respeito, sabe-se que os drones possuem diversas funções, sendo utilizados em corridas, fotografias, tomadas cinematográficas, mapeamento topográfico, etc.; e adquirem a cada dia novas funções (como, por exemplo, no seu uso em serviços de entregas).[3]

Tais fatos possuem consequências significativas. Ainda que se cumpra a hipótese de um *banimento* dos sistemas letais autônomos, é possível, ainda assim, que haja uma apropriação maliciosa de tecnologias civis por atores civis, paraestatais ou terroristas a fim de ameaçar ou causar danos físicos. Essa é uma questão importante no tocante à relação geral entre tecnociências e sociedade. Muito embora as tecnologias sejam construídas para fins muito específicos, por meio de sistemas especializados, o fato é que a maioria daquelas empregadas em IAs são suficientemente genéricas para serem customizáveis para uma variedade de propósitos muito distintos daqueles originalmente concebidos. Neste sentido, agentes mal-intencionados podem adquirir suprimentos de

---

[2] Segundo dados do próprio Relatório (p. 38), há uma crescente produção e venda de robôs de aplicação industrial (mais que dobrando de 2010 a 2015, de 121 mil para 254 mil), de robôs de serviços profissionais e de robôs domésticos (41 mil e 5.4 milhões em 2015, respectivamente). Além disso, nos anos de 2016 e 2017, houve mais de 670 mil drones registrados na FAA (*Federal Aviation Administration*) dos EUA.

[3] Este foi o projeto lançado pelo diretor executivo da Amazon, Jeff Bezos em 2013. Em 7 de dezembro de 2016 ocorreu a 1ª entrega de produtos da Amazon por meio de um drone. Mas não é somente a Amazon que está testando tal possibilidade, pois a chinesa Jd.com, a Domino's da Nova Zelândia, a 7-Eleven e o Google também estão fazendo seus testes, assim como a padaria Pão to Go do interior de São Paulo...

sistemas de IA feitos para usos comerciais ou civis e adaptá-los para realizar um atentado.

Mas não é somente a pronta difusão das tecnologias de IAs que preocupa o Relatório, mas também suas características intrínsecas. A possibilidade de realização de um fim instrumental com distância espaço-temporal e psicológica, em pleno anonimato, possui consequências realmente aterradoras. De fato, as guerras e ações militares contemporâneas ainda dependem, em grande parte, da participação humana. É o caso dos drones usados para fins militares, que dependem de olhar, identificar, mirar e decidir alvejar o alvo. Neste caso, os comportamentos dos sistemas de IAs passam por um processo decisório que, por mais impessoal que seja por causa do uso de máquinas, probabilidades e procedimentos, são imputáveis, ao fim ao cabo, a um julgamento e a uma decisão feitos por mente humana.[4] Isso permite uma imputabilidade moral e jurídica menos problemática, assim como pode acarretar sofrimentos ou distúrbios psicológicos (depressão e stress pós-

[4] Isso não garante, de forma nenhuma, uma decisão razoável ou, mesmo, juridicamente respaldada. Para certificar-se disso, não é preciso chegar a analisar os documentos vazados pelo Wikileaks, bastando, para tanto, refletir sobre os próprios dados disponíveis pelo governo americano desde a implementação, no Governo Obama, do programa de ataques secretos por drones. Desde então, já foram mortos em território estrangeiro milhares de humanos tidos por terroristas e muitos civis como "efeitos colaterais" por meio de decisões que, como parte da chamada "guerra ao terror", ficam ao encargo cotidiano da CIA e do Pentágono, que realiza execuções extrajudiciais de "inimigos" de guerra. O frágil controle de tais procedimentos, elaborado pela *Presidential Policy Guidance, or P.P.G* por Obama em 2013, está em vias de flexibilização pelo Governo Trump, que busca limpar o caminho para alvejar terroristas islâmicos de baixo escalão, mesmo sem a presença de um líder de alto nível e sem estar em vias de realizar ação de ataque, ao passo que, na política vigente até agora, somente se pode atacar terroristas que não sejam líderes caso eles estejam em vias de causar danos a forças americanas.

traumático) naqueles que atuam em uma tecnologia com poder de vida e morte sobre indivíduos monitorados remotamente.[5]

Contudo, o desenvolvimento crescente das armas letais autônomas pode estar a mudar tal lógica, envolvendo questões éticas e jurídicas inteiramente novas às quais retornaremos mais à frente.[6] A possibilidade de cumprimento dos fins de forma anônima e à distância se articula com uma eficácia (super-humana) e a escalabilidade próprias das IAs. Isso aumenta, evidentemente, o potencial destrutivo do poder militar: a potência de violência e destruição aumenta instrumentalmente, ao mesmo tempo em que se perde grande parte dos freios ético-existenciais implicados pela mediação da ação humana.

Mas isso não é tudo, pois, ao estar dedicado a analisar os cenários de usos maliciosos – deixando de lado, portanto, o de uso militar oficial –, todo o Relatório tem como preocupação subjacente a possibilidade das IAs conduzirem a uma perda relativa de potência militar dos Estados-

---

[5] É o que é dito de modo expressivo por um ex-operador de drone, Michael Hass, em reportagem para o *The Guardian*: "Sempre pisar as formigas e nunca pensar nisso de outra forma? Isso é o que você está fazendo para pensar nos alvos - como tão somente bolhas pretas em uma tela. Você começa a fazer essas ginásticas psicológicas para tornar mais fácil fazer o que você precisa fazer - eles mereceram isso, eles escolheram o lado deles. Você teve que matar parte de sua consciência para continuar fazendo o seu trabalho todos os dias - e ignorar aquelas vozes que dizem que isso não era certo" (THE GUARDIAN. Life as a drone operator: 'Ever step on ants and never give it another thought?'. 19 November 2015).

[6] As questões éticas e jurídicas são muito complexas. Como diz Chaperon (2017), o problema começa pela própria definição do que é uma "arma completamente autônoma" e sobre o que significaria conservar um "controle humano significativo": "mas qual? Programar uma máquina é suficiente? É preciso que um operador autorize o robô a utilizar a força ou basta prever uma possibilidade de abortar a missão?". O fato é que, com diz a autora citando Boulanin, não existe robôs que saberiam matar seres humanos respeitando regras do direito humanitário e que soubessem distinguir entre combatentes e civis, nem que possam analisar os danos colaterais proporcionais à vantagem militar buscada, análise que é imposta por convenções internacionais.

nacionais. Isso porque existe um potencial já presente de modificação de sistemas robóticos comerciais com IA que transformem uma baixa capacidade técnica em ataques sofisticados em grande escala. A existência de robôs customizados e equipados com bombas (*payloads*) permite que se consiga realizar vários ataques físicos com precisão e a longa distância, o que antes eram capacidade exclusiva de países que detinham mísseis de cruzeiro.

Os ataques conduzidos por inteligências autônomas são, portanto, liberados da necessidade de intervenção humana espaço-temporal e possibilitam uma grande distância e duração da ação. Há de se ressaltar que a autonomia de tais tecnologias bélicas não depende apenas da capacidade de operação e decisão, mas também, mais prosaicamente, da capacidade de operação por longas durações, ou seja, de sua fonte e duração de energia. Também nisso, o processo em curso é crescente, pois os produtos (comerciais ou militares) buscarão sempre o mesmo: mais e mais autonomia dos robôs, de modo a permitir operações na maior duração e distância com o menor custo possível. Ora, a maior autonomia dos robôs habilita-os a realizar ataques ou manter metas em risco por um longo período de tempo. Além disso, como são capazes de transitar em terrenos diferentes, com amplas e diversificadas capacidades físicas e de percepção, eles transpassam barreiras e proteções projetadas para impedir acesso de humanos e tornam vulnerável uma variedade de novos espaços.

As IAs possibilitam, desta forma, construir um sistema plenamente autônomo, dando a países menores, pequenos grupos ou mesmo simples pessoas (os chamados “lobos solitários”) uma capacidade de realizar

ataques, à grande distância, com uma letalidade antes exclusiva a altas tecnologias militares em mãos de poucas potências. Essa é uma realidade já em curso, pois há relato de robôs pilotados conduzidos por grupos como ISIS e Hamas (SOLOMON, 2017; COHEN, 2017; FRANKE, 2016). Soma-se a este temor o fato de que a ameaça é relativamente imprevisível, pois, não sendo algo típico ou com precedentes, os ataques são difíceis de serem prevenidos ou antecipados pelas autoridades de segurança.

2. Passemos agora da análise do ponto de vista do hardware para outra da perspectiva dos *softwares*. Ambos estão estreitamente vinculados e são impossíveis de serem dissociados. Podemos distinguir as ameaças deste quesito em duas frentes: a de desenvolvimento de softwares com IAs para uso militar e a de exploração por parte de terceiros de vulnerabilidade de tais softwares.

De um lado, ressalta-se que qualquer cenário de desenvolvimento de hardwares – como os robôs assassinos – depende do desenvolvimento paralelo de uma série de softwares, tais como algoritmos de detecção de face e algoritmos de planejamento e navegação. Tais softwares estão em desenvolvimento e se tornando cada vez mais maduros e – ponto central para o Relatório, como veremos adiante – boa parte deles é de código aberto e está disponível para todos. Alguns destes softwares possuem uma possibilidade não apenas de aumentar a escala e a eficácia de ataques, como também modificar inteiramente seu *modus operandi*. É o caso do uso de sistemas multiagentes (SMA) para uma *inteligência de enxame* (*Swarm intelligence – SI*) que, operando com Inteligência Artificial distribuída, permitem desenvolver agentes robóticos autônomos e coordenados entre si. Em novembro de 2017, um movimento para o

banimento de armas letais autônomas (*Ban Lethal Autonomous Weapons*) publicou o vídeo Slaughterbots com um terrível cenário distópico em que, dentre outras coisas, aparece tal possível tecnologia, onde enxames de drones com tamanhos de insetos atacam civis impotentes.

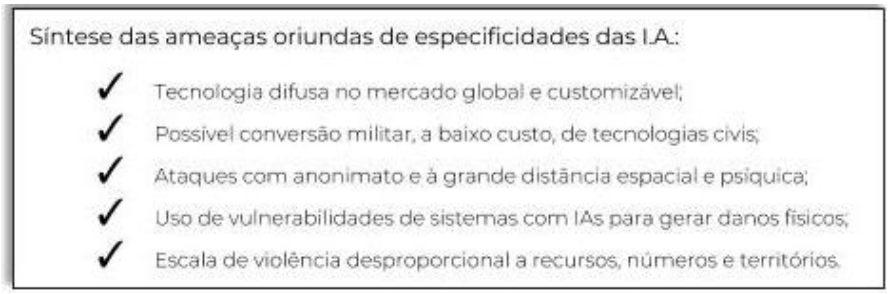
De outro lado, quando focamos no uso de IAs em softwares para tecnologias de uso social e institucional cotidianos, deparamo-nos com uma interseção entre as ameaças advindas dos *sistemas ciberfísicos crescentemente autônomos* com aquelas oriundas dos *sistemas de cibersegurança* (analisadas no nosso primeiro ensaio). Como dizem os autores do Relatório, somam-se às vulnerabilidades tradicionais de segurança cibernética, as vulnerabilidades próprias das IAs, usadas em sistemas IoT e em sistemas robóticos, tornando-os alvos de invasões e ataques por envenenamento, como é o caso dos *exemplos adversários* (adversarial examples) usados em sistemas supervisionados por algoritmos de aprendizado de máquina.

Tais novos riscos podem ocorrer em todos aqueles sistemas quase ou totalmente automatizados que conduzem atividades de alto risco e tendem a ser crescentemente operados por IAs, como, por exemplo, os sistemas de transporte (terrestre, aéreo ou marítimo), de lançamento de mísseis, de abastecimento[7] (energia, água, aquecimento, etc.) ou de segurança de instituições as mais diversas (como represas, usinas nucleares, plataformas de petróleo, etc.). Surgem, com isso, diversos cenários de subversão de sistemas habilitados para IA com o intuito de

[7] Recentemente, os Estados Unidos estabeleceram novas sanções contra a Rússia, acusando-a de conduzir ciberataques contra o seu setor energético.

ocasionar danos físicos; ou seja, agentes maliciosos podem encontrar *cibervulnerabilidades para realizar, a partir de qualquer lugar e de forma anônima, ciberataques destrutivos fisicamente*.

O supramencionado *Relatório do Boletim de Cientistas Atômicos* (NECKLIN, 2018) já havia mencionado que sérios danos físicos podem vir de ferramentas originalmente voltadas para danos digitais e políticos. Pela própria feição do Relatório, é feito um alerta acerca do aumento do risco de armas nucleares serem utilizadas intencionalmente ou por erro de cálculo (p. 2); contudo, riscos atômicos podem ser gerados também pela exploração de falhas em sistemas digitais diretamente vinculados à tecnologia bélica, ou, o que é mais provável, de subversão de dispositivos de segurança de usinas nucleares - ou seja, pode-se tanto detonar uma bomba nuclear, quanto transformar uma usina nuclear em uma "bomba".



No último dia 15 de março, por exemplo, foi noticiado, no NY Times, que, em agosto de 2017, uma companhia petroquímica na Arábia Saudita foi atingida por um ciberataque voltado não para destruir dados ou roubar segredos industriais, mas sim para sabotar as operações da companhia e gerar uma explosão. A única razão pela qual não houve uma explosão foi um *bug* no código de computador dos ofensores que inadvertidamente desligou o sistema de produção da planta ao invés de sabotá-lo secretamente. A notícia sinaliza bem a mudança em jogo:

> o ataque foi uma perigosa escalada no hacking internacional, na medida em que inimigos sem rosto demonstraram a intenção e a habilidade de infligir um sério dano físico. E funcionários do governo dos Estados Unidos, seus aliados e pesquisadores de cibersegurança receiam que os culpados podem replicar o ataque em outros países, uma vez que milhares de plantas industriais [cerca de 18.000] por todo o mundo fazem uso do mesmo sistema de computador de fabricação americana [controlador *Schneider Triconex*] que foi comprometido (PERLROTH, Nicole; KRAUSS, Clifford, 15/03/2018).

Mas há novas ameaças, também, com o surgimento da introdução crescente e confiante de robôs móveis autônomos no cotidiano das instituições e das cidades. A Internet das Coisas (sistemas IoT) pode ser usada como *vetor de ataque* através dos quais sistemas de IA que controlam sistemas chaves seriam subvertidos para causar muito provavelmente mais dano do que seria possível caso estivessem sob controle humano (p. 40). O desenvolvimento tecnológico atual, quase irresistível, de carros autônomos (auto-pilotados) – que já são testados em ambientes não-controlados e ganham as primeiras manchetes sobre pequenos acidentes e, nos últimos dias, de um atropelamento mortal – é um significativo exemplo trazido pelo Relatório. Segundo os autores, a liberação em grande escala de tais carros requererá a consideração de diversas questões técnicas, posto que, controlados por IAs acessíveis por ciberataques, podem ser vulneráveis a uma manipulação que transforme carros em armas (com pessoas dentro ou não) sob controle remoto de hackers. Outros robôs mais rotineiros introduzidos na vida cotidiana das cidades, como de entrega de encomendas, transporte ou limpeza, podem também virar bombas disfarçadas.

Armas Autônomas: uma carta aberta de pesquisadores de IA e Robótica

As armas autônomas selecionam e atiram em alvos sem intervenção humana. Elas poderiam incluir, por exemplo, um quadricóptero armado que pode buscar por ou eliminar pessoas ao encontrá-las por meio de certos critérios pré-definidos; mas elas não incluem os mísseis de cruzeiro ou os drones controlados de modo remoto, pelo fato de que os humanos podem decidir todos os alvos. A tecnologia de Inteligência Artificial (IA) alcançou um ponto em que o desenvolvimento de tais sistemas é – prática, se não legalmente – viável dentro de alguns anos, e não décadas, e as apostas são altas: as armas autônomas têm sido descritas como a terceira revolução na guerra, após a pólvora e as armas nucleares.

Muitos argumentos têm sido feitos contra e a favor das armas autônomas; por exemplo, que substituir soldados humanos por máquinas é bom para reduzir baixas para o proprietário destas; mas más por reduzir, da mesma forma, a motivação necessária para se ir ao campo de batalha. A questão chave para a humanidade hoje é se devemos começar uma corrida global em busca de armas de IA ou se devemos prevenir o seu início. Se qualquer potência militar maior der partida no desenvolvimento de arma de IA, uma corrida global pelas armas será virtualmente inevitável, e o ponto final desta trajetória tecnológica é óbvio: as armas autônomas se tornarão os Kalashnikovs de amanhã. Diferentemente das armas nucleares, elas não requerem matérias-primas custosas ou difíceis de se obter, logo se tornarão ubíquas e baratas para produção em massa por todas as potências militares significativas. Será somente uma questão de tempo até elas aparecerem no mercado negro e nas mãos de terroristas, ditadores desejando melhorar o controle sobre a população, senhores da guerra desejosos em perpetrar limpeza étnica, etc. As armas autônomas são o ideal para tarefas tais como assassinatos, desestabilização de nações, subjugamento de populações e assassínio seletivo de grupos étnicos. Nós acreditamos, portanto, que a corrida militar por armas de IA não seria benéfica para a humanidade. Há muitas maneiras pelos quais a IA pode tornar os campos de batalhas seguros para humanos, especialmente meios civis, sem criar novas ferramentas para matar pessoas.

Assim como muitos químicos e biólogos têm nenhum interesse em construir armas químicas e biológicas, muitos pesquisadores de IA têm nenhum interesse em construir armas de IA – e não querem que outros manchem o seu campo fazendo isso, criando potencialmente uma reação pública contra a IA que restrinja seus benefícios sociais futuros. De fato, químicos e biólogos têm defendido amplamente acordos internacionais que proibiram de forma bem sucedida as armas químicas e biológicas, tal como muitos físicos defenderam os tratados para banir armas nucleares espaciais e armas a laser para cegar pessoas.

Em suma, nós acreditamos que a IA tem um grande potencial para beneficiar a humanidade em muitas maneiras, e que o objetivo do campo deveria ser o de fazer isso. Começar uma corrida militar por armas de IA é uma má ideia, e deveria ser prevenida por um banimento de armas agressoras autônomas independentes de controle humano significativo.

Carta na Abertura da International Joint Conference on Artificial Intelligence (IJCAI). Institute of Future Life, 28 de julho de 2015.

Signatários (até a data de 18-03-2018): 3.724 pesquisadores de IA e Robótica e 20.487 outros.

A lista de signatários inclui: Stephen Hawking, Elon Musk, Stuart Russell, Nils J. Nilsson, Barbara J. Grosz, Eric Horvitz et al.

Propostas de Regulamentações para um Ecossistema Robótico Livre de Violência Maliciosa

As recomendações do Relatório podem ser apresentadas igualmente, de partida, distinguindo os dois aspectos: hardware e software. Os autores se preocupam com o fato de uma rápida difusão de tecnologia ocasionar, atualmente, um fosso crescente entre capacidades de ataque e defesa. Se, de fato, está em curso o desenvolvimento de sistemas de defesas contra robôs, especialmente os drones aéreos, as regulações e pesquisas têm sido lentas, segundo eles, em controlar a proliferação global de robôs que podem ser usados para ataques, até porque há um descompasso entre o investimento intensivo de capital necessário para os sistemas de defesa em relação à disponibilidade já existente de hardwares e softwares. Isso porque muitas iniciativas de defesa são muito caras (exigindo instalações físicas de controle e defesa) e/ou exigem trabalho humano altamente capacitado e intensivo, sendo viáveis por isso somente para defender grandes alvos (como aeroportos), deixando desprotegidos "pequenos alvos" (a própria aglomeração urbana). O maior desafio é, segundo os autores, encontrar métodos efetivos de defesa que tem uma razão custo/benefício aceitável, de forma a proteger não apenas nações com capacidade de investimento militar, como os EUA (que já lançou um programa de defesa contra drones, testando lasers e redes contra ataques terroristas), mas também nações mais vulneráveis econômica e politicamente.

Quanto à indústria e ao mercado de *hardware*, por causa deste descompasso existente no desenvolvimento de sistemas de defesa em grande escala, eles propõem que sejam empreendidas intervenções e

medidas políticas nos *ecossistemas de fabricação e distribuição de suprimentos*.

De um lado, assim como ele já havia proposto, na questão da cibersegurança (tratada em nosso primeiro ensaio) uma centralização do mercado digital, do qual o papel do Google é o exemplo maior, o Relatório indica a necessidade de uma concentração do mercado, citando elogiosamente companhias como a DJI (Dà-Jiāng), responsável por 70% do mercado global de venda de drones, pois a concentração/centralização tornaria "o ecossistema de *hardware* mais compreensivo e governável do que o ecossistema análogo de desenvolvimento de *software* IA" (ibid., 41). Mas, prevendo uma descentralização e difusão do mercado, eles não se prendem no papel dos fabricantes como ponto focal de governança e sugerem medidas indiretas de segurança com regulamentações de mercado, tais como a exigência de padrões mínimos de segurança contra ataques e adulterações, bem como o aumento do nível de habilidade requerido para conduzir ataques por estes meios e, também, dos custos de aquisição de dispositivos não-controlados. Por outro lado, o ecossistema de distribuição é mais difuso e complexo que ao nível de fabricação. Eles propõem uma mitigação de riscos por meio de ações nos pontos de distribuição, o que é mais factível para *hardwares*, com restrições a vendas de drones ou sistemas robóticos potencialmente letais, tal como se faz para controle de venda de armas ou drogas.

A respeito do mercado e da distribuição de *softwares*, eles se preocupam com o seu fácil acesso por agentes maliciosos, seja porque existem bibliotecas abertas para navegação e visão robótica, seja porque os próprios produtos comerciais, como os drones, são distribuídos com

softwares, seja, por fim, porque os softwares desenvolvidos por grandes empresas, cujo treinamento e aplicação requerem grandes conjuntos de dados e poder computacional, estão disponíveis na computação em nuvem. Para eles, tudo isso sugeriria um *ponto de controle adicional*, fortalecendo a centralização e evitando um acesso difuso.

Mas não basta, para eles, controlar o acesso a hardware e software, pois importa em igual medida regulamentar o uso daqueles que estão disponíveis comercialmente no mercado. Neste sentido, há um movimento mundial de regulamentação de tais tecnologias, como, por exemplo, no caso dos drones[8], com a exigência de registros para uso, requisitos de treinamento de pilotagem e o estabelecimento de zonas de voo proibidas (como proximidade a aeroportos ou regiões de alto risco). Os autores do Relatório chegam a sugerir que tais regulamentações sejam impostas, pelos fabricantes e governos, no próprio software dos robôs.

Contudo, a regulamentação não se aplica somente no acesso ou uso de softwares, pois, para que haja, por exemplo, um ataque aéreo por meio de um drone é necessário que exista, também, acesso a suprimentos perigosos usados como munição, carga ou ogiva (os "*payloads*"), que podem ser químicos, biológicos ou balísticos (pólvora, ácidos, dinamites, ogivas nucleares, agentes químicos, etc.). Aqui se impõe o mesmo problema que existe na discussão sobre regulamentação de armamentos civis, em que o controle do mercado de munições é visto

[8] Diversos órgãos de aviação civil, como a *Federal Aviation Administration* (FAA) nos Estados Unidos, a *Civil Aviation Safety Authority* (CASA) na Austrália, a *European Aviation Safety Agency* (EASA) na União Europeia e a Agência Nacional de Aviação Civil (ANAC) no Brasil, já adotam limites e regulações para aeromodelos e drones particulares.

como uma forma estratégica de minimizar o risco de seu uso para fins de atentado (sabe-se, contudo, o quanto é difícil que isso se efetive na prática). Para este controle, o Relatório sugere o uso das próprias IAs, que permitem rastrear possíveis agressores e fazer, por meio da inteligência de sinais (intelligence-gathering by signals intelligence - SIGINT), uma análise, pelo cruzamento de dados feitos em comunicações sociais, em busca de comportamentos anômalos de risco, o que já é feito, inclusive, com resultados nem sempre animadores (e sempre perigosos) na identificação de "potencial terrorista".

Por fim, como já vimos na caixa de texto acima, existem discussões em favor do banimento ou da *regulamentação* de sistemas de armas autônomas letais.[9] Aqui, retornarmos sobre as questões éticas e jurídicas em torno de tais tecnologias, que poderiam gerar um banimento prévio de seu uso militar, antecipando-se a seu desenvolvimento. Contudo, existem países centrais, como os EUA, que se opõem ao banimento forte (*strong ban*), o que torna mais provável, neste caso, a curto prazo, tão somente o estabelecimento de normas que as regulamentem. De todo modo, há bases jurídicas e éticas que estreitam a possibilidade de sua liberação. Como menciona o próprio Relatório, o Departamento de Defesa dos EUA tem uma diretiva que define políticas para o desenvolvimento e uso de autonomia em armas que a enquadra juridicamente. Além disso, existem documentos que, a princípio, limitam a autonomia de armas letais, como é o caso do manual da Lei de Guerra americana, que define os humanos como os responsáveis primários por ataques em conflitos

[9] cf. também a carta aberta para a Convenção das Nações Unidas sobre Certas Armas Convencionais de 21 de agosto de 2017.

armados, bem como o Comitê Internacional da Cruz Vermelha, que estabelece que "a decisão de matar e de destruir vem da responsabilidade humana e não pode ser delegada a uma máquina ou a uma arma". Mas, além de toda esta discussão, que abordaremos com mais detalhes em nosso último ensaio, o ponto principal, para o Relatório, é que “enquanto tais discussões de controle de armas e processos de desenvolvimento de normas são críticos, é improvável que eles interrompam atores não-estatais motivados em conduzir ataques” (p.42), ou seja, que impeçam seu desenvolvimento por usos maliciosos a partir da tecnologia civil.[10]

> Não pensemos apenas no potencial uso de IAs por terroristas, pois devemos temer, em igual medida, a centralização do mercado em oligopólios, o desenvolvimento militar de robôs assassinos e a automatização pelo Estado do seu reivindicado monopólio de violência e de vigilância sobre a população

[10] Para o debate sobre questões éticas e legais envolvidas nos sistemas letais autônomos, cf. CROOTOF, Rebecca (2015), LEVERINGHAUS, Alex (2016), NASU, Hitoshi; McLAUGHLIN, Robert (2014). Uma perspectiva diferente é encontrada em CHAMAYOU (2015).

Figura 2. *Raining in desert*. Imagem gerada por deep learning, março de 2018.

Considerações Críticas

Quando olhamos sob a ótica das questões de guerra e paz, é um truísmo afirmar que é comum o uso de alta tecnologia com fins militares. Uma breve história das guerras no século XX mostraria farta documentação em torno da via de mão dupla entre inovação tecnológica e tecnologia militar. Já há algum tempo que a eletrônica e os computadores sem IAs são utilizados militarmente potencializando os sistemas de ataque e defesa. Desnecessário lembrar, neste sentido, as histórias da IBM e da General Eletrics (GE). Todavia, com a introdução das IAs, as tecnologias militares (ou civis voltadas a danos físicos) entram em uma espécie de nova geração quando passam a incorporar as capacidades próprias de tais tecnologias.

Poderíamos analisá-las a partir das características específicas de um cenário de guerra e paz. Em situações de guerra iminente ou em curso, as atividades de ataque e defesa são ações instrumentais ou estratégicas, em que existe um cálculo utilitário das relações meios/fins/consequências, estando dado o fim de vencer o inimigo – o que pode ser feito, muito bem, provocando-lhe danos físicos, mas também por meio de sua incapacitação, dissuasão, fuga, rendição, etc. Neste caso, a reflexão sobre o uso de tecnologias em situações de guerra é feita tanto do ponto de vista de sua eficácia e dos riscos nela envolvidos (como o de destruição mútua, por exemplo), quanto também das regulamentações ético-jurídicas (direitos de guerra, mas também direitos humanos) a ele aplicadas. Vimos, nas páginas anteriores, quais são os novos desafios e riscos da possível introdução de IA e robótica na tecnologia militar.

Mas os autores do Relatório não pensam predominantemente neste registro, pois sua maior preocupação é com dispositivos de controle e manutenção de um estado de paz civil contra seus usos maliciosos. Contudo, quando se pensa em tais "malícias", percebemos que tudo se joga em uma garantia de estado de paz interna em uma situação de novas formas de guerra: aquela deslocada para a esfera do virtual (as ciberguerras analisadas no nosso primeiro ensaio) e aquela, que está no centro aqui, da guerra contra um terrorismo difuso.

Na análise do cenário e nas proposições de regulamentação, o Relatório possui uma premissa maior não refletida suficientemente: com Max Weber (cf. 1974 [1919]), chamamo-la de uma reivindicação de monopólio de uso da violência legítima por Estados-nações, a partir do qual são identificadas as potenciais ameaças de Estados adversários ou inimigos, ou de indivíduos isolados e grupos paraestatais, notadamente terroristas; monopólio que se soma a outro, teorizado por Anthony Giddens(2001[1985]) sob a influência de Michel Foucault: a saber, além do processo de industrialização da guerra e da centralidade da violência militar, o Estado-Nacional reivindica o monopólio de sistemas de vigilância e um poder administrativo responsável pela pacificação interna. Aceitemos, momentaneamente, tais premissas e analisemos suas propostas a partir delas.[11]

[11] Estamos cientes de que poderíamos seguir a análise por um rumo teórico mais politizado, presente de modo exemplar no pós-estruturalista, por exemplo, com a teoria das governamentalidade de Michel Foucault e discípulos, da sociedade de controle de Gilles Deleuze e do Estado de Exceção de Giorgio Agamben. Alguns de nós, do *Ateliê de Humanidades*, o farão futuramente de um ponto de vista mais antropológico. Mas optamos por sustentar aqui, no máximo que conseguimos, uma postura interpretativa de "liberdade em relação aos valores" (*Wertfreiheit*), defendida por Max Weber, pois alguns de nós consideram que tal atitude teórica permite-nos ingressar no campo de investigações que por

No tocante aos riscos físicos, vemos que o Relatório traz sugestões semelhantes àquelas de segurança digital, ao propor uma restrição das tecnologias, pela regulamentação e monitoramento da cadeia de suprimentos e pela limitação dos possíveis usuários. Entendemos que são soluções muito fáceis para um problema complexo.

A primeira razão disso se deve a uma questão cara às ciências sociais, a saber, a das consequências sociais e éticas das tecnologias decorrentes de seus usos. Infelizmente, temos que reconhecer que a letalidade em potencial de uma tecnologia é um conceito deveras impreciso, isso porque muitas aplicações pacíficas - como veículos de transporte autônomos (carros, caminhões e até navios) ou drones de grande porte - são deveras letais pela simples combinação da massa com grandes velocidades. Além disso, várias tecnologias são de usos tão múltiplos, difusos e indeterminados. O temor de seu uso malicioso por alguns pode acarretar em medidas danosas para todos. Um exemplo disso pode ser visto em um problema análogo, aquele da difusão doméstica da manufatura aditiva (as chamadas impressoras 3D), pois ela facilitou a fabricação de armas de fogo, a exemplo da pistola Liberator impressa em plástico ABS e capaz de disparar projéteis de diferentes calibres. Na verdade, a fabricação de armas artesanais sempre foi possível, apenas potencializada por uma tecnologia emergente. Por si, sem cartuchos de munição, detentores da energia letal de impelir o projétil a penetrar a vítima, a arma é inútil. Em termos instrumentais, controlar a munição já bastaria, pois embora também possa ser fabricada, é mais complexa e necessita de

---

ora desbravamos com uma visão ampla dos cenários possíveis e das questões em jogo, mantendo em suspenso nossas eventuais aversões particulares à violência e predileções pacificistas.

substâncias químicas controladas. Questiona-se se com os robôs não seria o mesmo, já que os *payloads* mencionados invariavelmente envolveriam substâncias controladas, como explosivos e agentes biológicos ou químicos.

Mas a situação é mais grave ainda quando levamos em consideração outro fator: os terroristas são na maioria engenheiros. Oito dos membros do grupo que conduziu o ataque ao World Trade Center eram estudantes de engenharia. A proficiência de grupos terroristas como o Estado Islâmico no ambiente digital, se utilizando da produção de conteúdo multimídia e de redes sociais para divulgação e recrutamento nos mostram que os agressores provavelmente não são iletrados nas técnicas de interesse, pelo contrário, a dominam quando necessário. Seriam de uso restrito tão somente os produtos comerciais inteligentes, ou também os componentes utilizados para construí-los? Motores, sensores, microchips, softwares? O fato de uma tecnologia (hardware ou software) ser de possível uso malicioso justifica sua centralização sobre controle de uma grande companhia ou do Estado, mesmo que ela tenha, em igual medida, possibilidades de usos benéficos individual e socialmente? E mesmo que esta fosse a opção, é forçoso reconhecer: os mestres na arte de dispositivos explosivos improvisados (IEDs) poderiam tornar-se também mestres na arte dos dispositivos autônomos improvisados, se assim necessitarem ou quiserem. O problema do terrorismo passa, evidentemente, pela questão tecnológica - a saber, os meios instrumentais disponíveis para uma ação terrorista -, mas eles não são estritamente técnicos nem passam por soluções meramente tecnológicas de controle, vigilância e defesa: na verdade, combatê-los apenas por estes meios é, no fim das contas, ineficiente e ineficaz.

Neste ponto, retornamos sobre o perigo instrumental e político das propostas de centralização ou concentração tecnológica sobre poucos atores supostamente responsáveis pela segurança de todos. Em termos funcionais, se a difusão de fabricantes de *hardware* pode resultar em produtos de baixa qualidade, mal manutenidos, desatualizados e inseguros, a centralização do hardware sob um fabricante, como reconhecido pelo Relatório no tocante à segurança de sistemas digitais, pode aumentar a vulnerabilidade. Como exemplo podemos citar o caso *Schneider Triconex* mencionado acima, em que um ataque bem sucedido de violação de um sistema pode significar vulnerabilidade de 18.000 plantas industriais em todo mundo.[12]

Isso é agravado pelo grande interesse que governos têm demonstrado em empresas de tecnologia à luz dos recentes vazamentos de informações sobre programas de espionagem em massa e cooperação entre empresas e agências de inteligências, cuja principal justificativa foi o antiterrorismo. A utilização de IA no monitoramento e espionagem em massa ameaçam a privacidade e geram preocupações sobre a criação de pontos de escuta e acesso remotos em robôs, principalmente por serem conduzidos por

[12] Outro caso é o da Intel, fabricante de 90% dos processadores para computadores pessoais e 99% dos processadores dos servidores que efetivamente rodam a internet, que recentemente anunciou a descoberta de vulnerabilidade em seu módulo de gestão remota (*Intel ME*), incluso em seu *hardware*: um sistema operacional completo, invisível para o usuário, capaz de acesso irrestrito ao processador e às interfaces de rede. Pesquisadores em cibersegurança já haviam descoberto *bugs* através de engenharia reversa meses antes do anúncio. No início do ano, foram reportadas mais duas vulnerabilidades nos processadores da Intel, ainda não totalmente sanadas, causando prejuízos de difícil contabilização relacionados à severa redução de performance e custos associados às suas correções e a custos de oportunidade, mesmo sem nenhum ataque reportado explorando essas falhas. O que tal vulnerabilidade poderia ter por resultado se fosse descoberta e manipulada por agentes maliciosos em articulando ciberataques com ataques ciberfísicos?

agressores poderosos[13]. A DJI, cuja concentração do mercado de drones aéreos foi elogiada no relatório, recentemente foi acusada também por oficiais americanos de transmitir dados à China, onde é sediada.

É evidente que o acesso de agentes maliciosos a recursos que permitem potencializar as consequências de suas ações deve ser limitado, apesar de dificilmente extinto. A elevação dos custos de aquisição e a criação de dificuldades ajudarão até certo ponto: afinal, grupos que contam com financiamento de nações adversárias, regimes autoritários ou que são financiados por atividades altamente lucrativas conduzidas em paralelo (como produção e transporte de petróleo e drogas ilícitas), possuirão os recursos suficientes para adquirir a maioria das ferramentas e pessoas que necessitam. Se hoje já possuem acesso a veículos de combate e armamento de guerra refinado, e os utilizam contra civis, a exemplo dos equipamentos utilizados pelo ISIS e na interceptação do vôo da Malasian Airlines por rebeldes russos, eles não hesitarão em fazer o mesmo com equipamento letal autônomo disponível para as forças militares, tão logo se torne barato e difundido o suficiente para ser contrabandeado.

Isso reforça a importância da discussão de armamento letal autônomo como um todo. O mesmo boletim de cientistas atômicos da nossa introdução acima já havia alertado sobre o fato em 2015, em apoio à *Campaign to Stop Killer Robots*. Mais do que ameaças físicas futuras causadas por adaptações de produtos comerciais, que são o foco do Relatório, os produtos militares aumentados pelo uso de tecnologias de IA tão recentes quanto o *deep learning* estão em uso, hoje, para identificação e rastreamento de alvos,

[13] Os vazamentos de informações sobre as operações da *National Security Agency* (NSA) também denunciou a adulteração de *hardware* pela agência.

trilhando lentamente seu rumo à aplicação em armas letais através de ferramentas d'o estado da arte.

Por essa razão, um eventual monopólio estatal do uso policial e militar de tecnologias de IA e robóticas, ocorrido em nome da segurança e da mitigação de riscos físicos, poderá acarretar em uma concentração da possibilidade de violência, de controle e de gestão das populações que, enfim, tornar-se-á perigosa para as democracias e as próprias seguranças físicas a serem preservadas; afinal, é frágil e tênue a fronteira entre uma segurança desejada pelas populações e um estado *orwelliano* efetivo que prive o público geral das reais potencialidades das tecnologias inteligentes. Não pensemos, portanto, apenas no potencial uso de IAs por terroristas, pois devemos temer, em igual medida, a centralização do mercado de IA sobre controle de oligopólios, o desenvolvimento militar de robôs assassinos e a automatização, por parte do Estado, do seu reivindicado monopólio de violência e de vigilância sobre a população, com uma estreita cooperação entre oligopólios de IAs e agências de inteligência.

De fato, o cenário futuro de um desenvolvimento militar de robôs matadores é típico de uma ficção distópica. Contudo, vale lembrar, a morte de humanos causada por robôs transcende, há muito, o mundo ficcional e está integrada à nossa realidade desde 25 de janeiro de 1979. Mais do que trágicas mortes causadas por dispositivos autônomos programados por indivíduos extremistas, devemos evitar, de modo efetivo, que haja uma competição entre as nações rumo à expansão dos limites da inteligência artificial, levando-nos a uma terceira guerra mundial. Se fomos levados a temer, na Guerra Fria, que uma destruição total poderia advir de um apocalipse nuclear, não

esqueçamos que, já há quase uma centena de anos, um outro imaginário ficcional nos alerta sobre os perigos de brincar de aprendizes de feiticeiro: o da total aniquilação dos criadores de robôs através das suas próprias e infatigáveis criações. Entre pesadelos, temores e desejos, sonhos sempre podem se tornar realidades enquanto os humanos vivem entre o encantamento e o sonambulismo.

Referências bibliográficas

ALAOUI, Youness. Reverse-engineering the Intel Management Engine's ROMP module. 10 de maio de 2017.

ARON, Raymond. *Paz e Guerra entre as Nações*. Brasília: Imprensa Oficial do Estado / Editora Universidade de Brasília: Instituto de Pesquisa de Relações Internacionais, 2002 [1962].

BAN LETHAL AUTONOMOUS WEAPONS. Site.

BBC. North Korea launches 'highest ever' ballistic missile. 29 de novembro de 2017.

BBC. Trump to Kim: My nuclear button is 'bigger and more powerful'. 3 de janeiro de 2018

BERREBY, David. Engineering Terror. 10 de setembro de 2010.

BOULANIN, Vincent; VERBRUGGEN, Maaike. *Mapping the Development of Autonomy in Weapon Systems*. SIPRI, novembro de 2017.

BUMP, Philip. No, Really, Regulate the Bullets. 16 de dezembro de 2012.

CANO, Rosa Jiménez. Primer atropello mortal de un coche sin conductor. 20 de março de 2018.

CHAFKIN, Max; KING, Ian. Intel Has a Big Problem. It Needs to Act Like It. 18 de janeiro de 2018.

CHAMAYOU, Grégoire. *Teoria do Drone*. São Paulo: Cosac Naify (Coleção Exit), 2015.

CHAPERON, Isabelle. Les robots tueurs menacent-ils notre sécurité ?. Le Monde, 21 de novembro de 2017.

COHEN, G. Israel Shoots Down Hamas Drone Over Gaza Strip. *Haaretz*, 23 de fevereiro de 2017.

CROOTOF, Rebecca. The Killer Robots Are Here: Legal and Policy Implications (December 5, 2014). 36 *Cardozo L. Rev.* 1837 (2015). Available at SSRN.

GABBATT, Adam. New York woman visited by police after researching pressure cookers online. 01 de agosto de 2013.

GIBBS, Samuel. Google's AI is being used by US military drone programme. 07 de março de 2018.

GIDDENS, Anthony. *O Estado-Nação e a Violência*. São Paulo: Editora da Universidade de São Paulo, 2001 [1985].

GREENBERG, Andy. Feds tighten restrictions on 3-D printed gun files online. 11 de junho de 2015.

INSTITUTE OF FUTURE LIFE. Autonomous Weapons: An Open Letter from AI & Robotics Researchers (Open letter at the opening of the IJCAI). Institute of Future Life, 28 de julho de 2015.

KOERNER, Brendan I. Why ISIS is winning the social media war. Abril de 2016.

KRAVETS, David. JAN. 25, 1979: ROBOT KILLS HUMAN. 25 de janeiro de 2010.

LANDLER, Mark. North Korea Asks for Direct Nuclear Talks, and Trump Agrees. 8 de março de 2018.

LEVERINGHAUS, Alex. *Ethics and Autonomous Weapons*. London: Palgrave Macmillan, 2016,

LAUMOND, Jean-Paul. Equiper un drone d'un armement change totalement le paradigme. Le Monde, 06 de setembro de 2017.

MAGNELLI, Renato. *Raining in the desert*. Aplicação de estilo impressionista sobre colagem de fotos de drone *Predator* lançando míssil *hellfire* sobre paisagem afegã, com *DeepStyle*. Março de 2018.

MAGNELLI, Renato. *Virtual Militant*. Aplicação de estilo digital sobre foto de militante em funeral do Hamas, com *DeepStyle*. Março de 2018.

MECKLIN, John (ed.) *It is 2 minutes to midnight. Doomsday Clock Statement*. Science and Security Board, Bulletin of the Atomic Scientists. 25 de janeiro de 2018.

NASU, Hitoshi; McLAUGHLIN, Robert (ed.). New Technologies and the Law of Armed Conflict. Asser Press/Springer, 2014.

MOZUR, Paul. Drone Maker D.J.I. May Be Sending Data to China, U.S. Officials Say. 29 de novembro de 2017.

PERLROTH, Nicole; KRAUSS, Clifford. A Cyberattack in Saudi Arabia Had a Deadly Goal. Experts Fear Another Try. NY Times, 15 de março de 2018.

PHYS.ORG. Norway to build first self-sailing electric cargo ship. 10 de maio de 2017.

PILKINGTON, Ed. Life as a drone operator: 'Ever step on ants and never give it another thought?'. *The Guardian*, 19 November 2015.

Slaughterbots (legendado em espanhol disponível no youtube), publicado em 12 de novembro de 2017.

SAVAGE, Charlie; SCHMITT, Eric. Trump Poised to Drop Some Limits on Drone Strikes and Commando Raids. NY TYMES, 21 de setembro de 2017.

SCHMITT, Eric. Pentagon Tests Lasers and Nets to Combat a Vexing Foe: ISIS Drones. 27 de setembro de 2017.

SOLOMON, B. Witnessing an ISIS Drone Attack. New York Times. 2017.

STEPHEN, Jonhson. A.I. "predator" drones can now spot and track illegal poachers. 6 de janeiro de 2018.

STEPHENS, Rachel. Initial Cost Analysis of Meltdown and Spectre. 09 de fevereiro de 2018.

SULLIVAN, Ben. America Is Going to Fight ISIS With Algorithms. 16 de maio de 2017.

ULANOFF, Lance. Now there are bullets that won't break your 3D-printed gun. 06 de novembro de 2014.

UNTERSINGER, Martin. Comment l'armée américaine prévoit la guerre du futur. Le Monde, 14 de março de 2017.

WEBER, Max. Política como vocação (1919). In: Ensaios de sociologia. Rio de Janeiro: Zahar, 1974.

WINGFIELD, Nick; SCOTT, Mark. In Major Step for Drone Delivery, Amazon Flies Package to Customer in England. 14 de dezembro de 2016.