JOURNAL OF
DEMOCRACY
EM PORTUGUÊS

Volume 8, Número 1, Maio de 2019

PLATAFORMA DEMOCRÁTICA
FUNDAÇÃO FHC
CENTRO EDELSTEIN

Apresentação

Os cinco artigos desta edição do *Journal of Democracy em Português* trazem *insights* sobre os rumos da democracia em um mundo impactado pela influência crescente das tecnologias digitais, entre elas as redes sociais e a Inteligência Artificial, e pela recente chegada ao poder de movimentos ou políticos de tendência antiliberal em países como Itália e Brasil.

Em "O que aconteceu com as democracias da terceira onda?", Scott Mainwaring e Fernando Bizzarro, ambos da Universidade Harvard, oferecem uma análise abrangente de 91 transições democráticas (em 79 países, pois alguns viveram mais de uma transição) ocorridas entre 1974 e 2012. A evolução de cada uma delas foi classificada em cinco categorias: avanço, estagnação, retrocesso ou colapso e alguns (poucos) casos que já partiram de um grau elevado de democracia, mas não avançaram significativamente no período.

Segundo os autores, que se basearam em dados da pesquisa *Variedades da Democracia (V-Dem)*, os resultados mais comuns foram colapso ou estagnação, que, somados, representam 62 dos 91 casos. "Os casos de aprofundamento democrático substancial resultando em democracias liberais robustas são exceções isoladas", afirmam. O Brasil é classificado entre os que tiveram avanço, com a ressalva de que o país vem enfrentando grandes desafios políticos.

Mas o que leva uma democracia a se consolidar, estagnar ou ruir?, perguntam. "Os fatores que podem reduzir as chances de colapso e aumentar as de avanços democráticos incluem um nível mais alto de desenvolvimento e desempenho econômico, maior capacidade estatal, a existência de países vizinhos democráticos e um passado democrático."

No segundo artigo, o cientista político Ronald J. Deibert (Universidade de Toronto) expõe o que ele chama de "três duras verdades sobre as redes sociais": (1) o modelo de negócio das redes está baseado na vigilância profunda e incansável dos dados pessoais dos consumidores; (2) voluntária e conscientemente toleramos esse nível desconcertante de vigilância; e (3) as redes sociais não apenas não são incompatíveis com o autoritarismo como, na prática, estão se mostrando uma de suas ferramentas mais efetivas.

"Somadas, essas três verdades pintam um quadro bastante sombrio da atual realidade social e política e pressagiam um futuro ainda mais sombrio", diz o canadense, que propõe uma reforma de longo prazo, estendendo-se do pessoal ao político, do local ao global, para restaurar a democracia liberal. "O mundo clama por inovações tecnológicas que ampliem as possibilidades para além das plataformas altamente centralizadas, intensamente vigiadas e facilmente instrumentalizadas pelos gigantes das redes sociais", escreve.

O terceiro texto, de Steven Feldstein, pesquisador do Fundo Carnegie para a Paz Internacional, revela o potencial da Inteligência Artificial (IA) como instrumento de repressão e controle por parte de regimes iliberais. "Essa é a simplicidade elegante da repressão por meio de IA: ela requer uma quantidade consideravelmente menor de atores humanos que a repressão convencional, menos perseguição física e menos recursos financeiros. E, no entanto, pode ter um impacto mais amplo e sistemático", afirma. Segundo o autor, esta nova tecnologia terá grande impacto na política global, pois "à medida que a IA provar seu valor repressivo para as autocracias, outros governos poderão começar a imitá-las".

Os dois últimos artigos da presente edição enfocam os processos políticos que desembocaram na eleição de governos de tendência politicamente antiliberal na Itália (março de 2018) e em seguida no Brasil (outubro de 2018).

Em “Como os populistas venceram na Itália”, o cientista político italiano Roberto D’Alimonte explica a trajetória de ascensão ao poder do Movimento Cinco Estrelas (M5S) e da Liga (antiga Liga Norte), que, apesar de consideráveis diferenças de origem e programáticas, hoje são sócios no governo em Roma. “Apesar de tentativas dos últimos governos de realizar reformas, houve um aprofundamento da sensação de que há uma lacuna separando as elites econômicas e políticas do resto da população. A Liga e o M5S souberam aproveitar esse descontentamento ao prometer mudanças radicais”, escreve o professor da Universidade LUISS Guido Carli, para quem “a volatilidade eleitoral tem sido a norma na Itália” e o país deve “permanecer no limbo por um bom tempo”.

Em artigo inédito para o *Journal of Democracy em Português*, a socióloga Angela Alonso (USP) analisa a gênese do que ela chama de “campo patriota”, que se esboça no governo Lula (2003-2010), emerge nas manifestações de junho de 2013 e ganha força em 2015 e 2016, durante o processo de impeachment da presidente Dilma Rousseff (2011-2016). “Ao longo desta mobilização intermitente, a rua passou da coexistência entre campos diferentes (o ‘autonomista’, o ‘socialista’ e o ‘patriota’) para o confronto entre frentes inimigas, com desfecho dramático na eleição de (Jair) Bolsonaro”, escreve a pesquisadora e ex-presidente do Cebrap.

“A insistência no simbolismo nacionalista, no combate à corrupção administrativa e de costumes e no antipetismo firmaram o campo político patriota. Este iceberg apareceu naquele junho de 2013. Foi festejado como primavera, mas era o começo do inverno”, afirma Alonso.

Boa leitura.

Bernardo Sorj e Sergio Fausto
Diretores de Plataforma Democrática

COMO A INTELIGÊNCIA ARTIFICIAL ESTÁ TRANSFORMANDO A REPRESSÃO

Steven Feldstein

***Steven Feldstein** é professor associado de gestão pública da Universidade Estadual de Boise (EUA) e pesquisador não residente do Programa de Democracia, Conflito e Governança do Carnegie Endowment for International Peace. De 2014 a 2017, foi subsecretário-adjunto do Bureau de Democracia, Direitos Humanos e Trabalho do Departamento de Estado dos EUA.*

No início de 2018, uma das principais forças de segurança da Malásia fez um anúncio surpreendente. A Polícia Auxiliar, uma divisão da Cooperativa da Polícia Real da Malásia, anunciou uma parceria com a empresa chinesa Yitu Technology para equipar os seus policiais com câmeras dotadas de reconhecimento facial. A força policial será capaz de comparar rapidamente imagens capturadas pelas câmeras acopladas ao corpo dos policiais com imagens de um banco de dados central. O chefe da Polícia Auxiliar explicou que esse uso de inteligência artificial (IA) era um "importante passo à frente" para melhorar a segurança pública. Também anunciou que a organização planejava

* Publicado originalmente como "The Road to Digital Unfreedom: How Artificial Intelligence is Reshaping Repression", *Journal of Democracy*, Volume 30, Número 1, Janeiro de 2019.

eventualmente aprimorar o sistema de câmeras para permitir "reconhecimento facial em tempo real e alertas instantâneos para a presença de pessoas procuradas pela polícia".[1]

A vizinha Singapura logo seguiu o exemplo, anunciando planos de um projeto-piloto de câmeras de vigilância dotadas de tecnologia de reconhecimento facial em todos os postes de iluminação pública. O projeto visa claramente a facilitar a "análise de multidões" e ajudar em operações antiterror. Grupos de defesa da privacidade, como a Electronic Frontier Foundation, alertaram que essa tecnologia permitiria que governos ampliassem o controle sobre opositores políticos e limitassem a liberdade de expressão, mas seus protestos foram em vão.[2]

Enquanto isso, em abril de 2018, a startup de IA CloudWalk Technology, com sede na cidade chinesa de Cantão, supostamente fechou um acordo com o governo de Zimbábue para prover tecnologia de reconhecimento facial que permitirá às forças de segurança do país construir um banco de dados de imagem nacional. A CloudWalk também é conhecida por fornecer tecnologia de reconhecimento facial e verificação de identidade a forças policiais da região autônoma de Xinjiang (China), uma das regiões do mundo com maior nível de repressão política. Sua nova parceira africana faz parte de uma política chinesa multicontinental de infraestrutura e investimento conhecida como "One Belt, One Road" (também apelidada de Rota da Seda do Século 21).[3] Os serviços da CloudWalk ameaçam agravar a repressão política no Zimbábue, onde o governo recentemente liderou uma violenta repressão pós-eleitoral.

Esses não são exemplos isolados. Ao redor do mundo, sistemas de IA vêm demonstrando seu potencial para incitar governos repressivos e afetar a relação entre cidadãos e o Estado, acelerando uma ressurgência global do autoritarismo. A República Popular da China (RPC) vem liderando a proliferação de tecnologia de IA entre regimes autori-

tários e iliberais, uma abordagem que se tornou componente-chave da estratégia geopolítica chinesa.

É difícil definir com precisão o conceito de IA. De maneira geral afirma-se que o objetivo da IA é "tornar as máquinas inteligentes", um conceito frequentemente explicado por meio de uma alusão à inteligência humana.[4] Outros, como o cientista da computação norte-americano Jerry Kaplan, questionam a utilidade de tais analogias. Kaplan afirma que o fato de as máquinas terem "autoconsciência como as pessoas" é irrelevante. Para ele, a essência da IA resume-se à capacidade de um computador de "fazer generalizações apropriadas e em tempo hábil com base em informações limitadas".[5]

Não é pretensão deste artigo superar essa controvérsia, mas focar nos efeitos práticos das novas tecnologias que vêm surgindo graças a três importantes fatores: 1) o aumento da disponibilidade de massivas bases de dados (Big Data) de fontes públicas e privadas; 2) aplicações avançadas de *machine learning* e algoritmos; e 3) o avanço da capacidade de processamento computacional. (*Machine learning*, que pode ser aplicada em tarefas diversas, desde vencer partidas de Go até identificar patógenos, é um processamento estatístico iterativo no qual um sistema de IA é alimentado com uma base de dados e "tenta derivar uma regra ou procedimento que explica esses dados ou é capaz de prever dados futuros".[6])

A importância dessa tecnologia tanto para regimes autoritários como para seus oponentes democráticos torna-se cada vez mais clara. Nos últimos anos, autocracias atingiram níveis inéditos de controle e manipulação, aplicando sistemas computacionais avançados a uma vasta quantidade de dados não estruturados atualmente disponíveis online e a imagens capturadas ao vivo por câmeras e outras fontes de monitoramento e vigilância. De tecnologias de reconhecimento facial que cruzam imagens em tempo real com grandes bases de dados a algoritmos que varrem as redes sociais em busca de sinais de atividades

de oposição, essas inovações são um divisor de águas para regimes autoritários em seus esforços para controlar o debate público e silenciar vozes da oposição.

A IA não é a única categoria de nova tecnologia sendo cada vez mais bem aproveitada por autocratas por razões políticas. Outras tecnologias de informação e comunicação, frequentemente usadas em conjunto com IA, estão produzindo efeitos igualmente alarmantes. Entre elas está a biometria avançada, hacking cibernético patrocinado por atores estatais e técnicas de distorção de informações.

O presente artigo destaca o impacto repressivo da tecnologia de IA por dois motivos. Em primeiro lugar, a IA representa uma capacidade avançada que integra e amplia as funções de outras tecnologias de novas e surpreendentes maneiras. Em segundo lugar, o entendimento comum do impacto político de tecnologias de IA ainda é limitado; formuladores de políticas públicas ainda não estão lidando de maneira suficientemente séria com as implicações repressivas da IA.

Por que IA é um trunfo para líderes autoritários

Embora a IA tenha um grande potencial de servir como ferramenta para qualquer governo, ela oferece uma série de vantagens em particular para regimes autoritários e iliberais. Apesar da ampla variedade de tipos de regimes não democráticos — de ditaduras de partido único a regimes híbridos ou semiautoritários, passando por ditaduras militares e autocracias personalistas —, a maioria desses governos se mantêm no poder por meio de uma mistura de coerção (ameaçando e intimidando potenciais rivais) e cooptação (subornando ou, de alguma maneira, induzindo atores políticos a fazer parte da coalizão governista).

Um líder que opte por reprimir precisa garantir que as forças de segurança do Estado apliquem as medidas coercitivas necessárias. Isso acarreta dois problemas. Primeiro, tal repressão é cara e intensiva em

mão-de-obra; com o tempo, exige cada vez mais recursos para se manter. Segundo, ela gera um problema de agente-principal: "os mesmos recursos que garantem que os agentes de repressão do regime oprimam a oposição também permitem que ajam contra o próprio regime".[7] Em outras palavras, à medida que um regime conta cada vez mais com policiais ou soldados para fazer seu trabalho sujo, também se torna mais vulnerável a pressões ou até mesmo insurgência dessas mesmas forças. Os líderes precisam avaliar se os benefícios de empregar forças de segurança para reprimir perigos externos ao sistema superam o potencial risco interno representado por essas mesmas forças dentro do sistema.

É aí que as vantagens da tecnologia de IA tornam-se aparentes. Em vez de depender de uma densa infraestrutura de forças de segurança para produzir vigilância, perseguição e intimidação em larga escala de oponentes por todo o território do Estado, líderes autoritários podem usar IA para desenvolver capacidade repressiva digital a um custo menor — e reduzir problemas de agente-principal.[8] Na verdade, as operações mais avançadas de vigilância dependem de poucos agentes humanos: muitas funções são automatizadas por meio de IA. Além disso, comparados com agentes humanos, que possuem um estoque limitado de tempo e atenção, sistemas de IA possuem um alcance muito maior. Por conta dessa onipresença, eles podem produzir mudanças no comportamento e criar um importante "efeito inibitório" [*chilling effect*] mesmo na ausência de violência física.

Se os cidadãos souberem que *bots* dotados de IA estão monitorando todas as comunicações e que algoritmos detectarão mensagens críticas ao regime e notificarão as autoridades, as pessoas se sentirão fortemente motivadas a se conformar. Essa é a simplicidade elegante da repressão por meio de IA: ela requer uma quantidade consideravelmente menor de atores humanos que a repressão convencional, menos perseguição física e menos recursos financeiros.[9] E, no entanto, pode ter um impacto mais amplo e sistemático.

Mesmo antes do início da repressão digital, o panorama do autoritarismo contemporâneo estava se transformando de maneiras dignas de nota. Em primeiro lugar, a erosão das instituições e normas democráticas acelerou em todo mundo. O relatório de 2018 do projeto Variedades de Democracia (V-Dem) estima que cerca de 2,5 bilhões de pessoas vivam atualmente em países afetados por essa "tendência global de autocratização".[10] De fato, um retrocesso democrático gradual tornou-se uma das rotas mais comuns em direção ao autoritarismo.

Em segundo lugar, a maneira pela qual os autocratas deixam o poder também está mudando. De 1946 a 1988, golpes de Estado eram a maneira mais comum pela qual um autocrata deixava o cargo e representavam 48,6% dos casos de fim de um regime autoritário. Mas, na era pós-Guerra Fria, casos de mudança por fatores externos ao regime foram mais frequentes do que os golpes de Estado. De 1989 a 2017, as causas mais comuns para a saída de ditadores foram revoltas populares e derrotas eleitorais. Fins de regime por vias de golpe de Estado despencaram, representando apenas 13% dos casos totais (até mesmo as saídas de lideranças devido a guerras civis superaram as saídas motivadas por golpes de Estado no período).[11]

Isso indica que as maiores ameaças à sobrevivência do autoritarismo atualmente vêm não de rebeliões internas, mas de grupos descontentes nas ruas ou nas urnas. A implicação para ditadores que desejam permanecer no poder é clara: redirecione recursos para manter os movimentos populares sob controle e seja cada vez melhor em fraudar as eleições. Nessas áreas, a tecnologia de IA oferece uma vantagem crucial. Em vez de depender de forças de segurança para reprimir sua população — com todos os custos e riscos políticos envolvidos —, líderes autocráticos estão abraçando táticas digitais de monitoramento, vigilância e perseguição de movimentos sociais e manipulação de eleições. Olhar para três possíveis cenários ajudará a tornar mais clara

a relevância da IA para alguns dos desafios mais urgentes que os autoritários contemporâneos enfrentam.

Cenário 1: monitorando o descontentamento popular e controlando protestos em massa. No primeiro cenário, um regime de partido único vê-se diante do crescimento da insatisfação com a estagnação econômica e a repressão política. No ano anterior houve protestos espontâneos que preocuparam as lideranças políticas. O regime deseja tomar medidas assertivas para impedir grandes mobilizações políticas, mas seus recursos limitados impedem o uso de detenções em massa. Ele também teme que uma repressão explícita desse tipo poderia desencadear uma revolta popular. Portanto, traça uma estratégia dividida em duas partes: 1) identificar, monitorar e, seletivamente, prender líderes da oposição e potenciais seguidores-chave; e 2) monitorar atentamente formações de multidões que possam se transformar em protestos em massa, mantendo forças de segurança em alerta para desmobilizar protestos antes que cresçam.

Para colocar essa estratégia em prática, o regime precisa, primeiro, identificar líderes dissidentes e seguidores-chave com grande probabilidade de mobilização. Inicia então uma extensa pesquisa em redes sociais e comunicações pessoais. Como certos grupos de conversa online usam configurações de privacidade ou criptografia para evitar espionagem do governo, as autoridades podem usar os serviços de empresas internacionais como FinFisher ou NSO Group, que comercializam softwares criados para penetrar esses grupos fechados. Alternativamente, o regime pode optar por uma opção mais barata e contratar os serviços de um "hacker de aluguel" internacional ou um outro desenvolvedor de *malware* para essa tarefa.[12]

Essa pesquisa de dados online ajuda o regime a reconhecer padrões, identificar indivíduos de interesse e focar em conversas relevantes. À medida que a operação de vigilância constrói perfis de ativistas políticos e mapeia redes de opositores, ela fornece essas informações ao

algoritmo de IA que, por sua vez, analisa múltiplos bancos de dados usando software de reconhecimento de padrões para identificar indivíduos com tendências à dissidência política. O algoritmo também ajuda o regime a monitorar questões que estão provocando a insatisfação popular e procura por comunicações que indiquem um protesto iminente. De posse dessas informações, o regime conduz detenções seletivas e preventivas para evitar que os protestos ocorram.

Se os protestos acontecerem apesar desses esforços, a IA pode ajudar o regime a contê-los. Uma tecnologia já disponibilizada pela popular plataforma chinesa de comunicação online WeChat produz "mapas de calor" que mostram densidade de multidões e mede tráfego de pedestres em locais específicos.[13] O regime pode embutir tecnologia de rastreamento em outras plataformas de comunicação online, permitindo-o saber instantaneamente quando multidões começam a se formar. Alternativamente, ele pode instalar sistemas de reconhecimento facial em espaços públicos urbanos (na linha da proposta de Singapura e seus postes de iluminação pública). Sistemas de IA com acesso a essas câmeras podem monitorar a densidade das multidões, procurar por indivíduos portando cartazes políticos e monitorar o paradeiro de pessoas procuradas.

Por fim, a IA aumenta a capacidade do Estado de empregar censura seletiva e campanhas online de desinformação para semear confusão e minar potenciais protestos. Isso pode assumir a forma de ataques de negação de serviço [*denial-of-service*] contra campanhas de protesto (minando a capacidade dos oponentes de se organizar e efetivamente censurando informações vitais) ou de campanhas de distorção de informações comandadas por *bots* (produzindo uma enxurrada de posts enganosos para ofuscar a mensagem dos oponentes e inundar de ruído os canais de informação).

Cenário 2: mantendo uma província rebelde sob controle. Neste exemplo, um regime autoritário lida com uma potencial instabilidade

em uma província distante onde grande parte da população é composta por uma minoria étnica. A legitimidade do regime nessa província é historicamente tênue, e a região frequentemente passa por crises de instabilidade civil. Recentemente, o governo central decidiu refrear a turbulência política por meio de uma repressão violenta que combina táticas tradicionais com nova tecnologia. Esse cenário remete à atual situação da região de Xinjiang, e os esforços da RPC de reprimir a dissidência política na região ilustram bem o vasto potencial repressivo da IA usada em conjunto com outras táticas coercitivas mais antigas.

Primeiro, o governo chinês faz uso generalizado de repressão física tradicional. As autoridades organizaram e vêm ampliando uma rede de campos de reeducação que, acredita-se, abrigam um milhão ou mais de prisioneiros de origem muçulmana, sobretudo membros da minoria étnica uigur. Isso representa uma parcela considerável dos cerca de 21 milhões de habitantes de Xinjiang. Esses campos de trabalho forçado envolvem doutrinação constante, incluindo autocrítica e repetição de frases ("nos oporemos ao extremismo, nos oporemos ao separatismo, nos oporemos ao terrorismo"). Prisioneiros são mantidos em quartos trancados em péssimas condições e são submetidos a uma disciplina draconiana.[14]

Segundo, as autoridades da RPC complementam a coerção física severa com uma abordagem mais ampla baseada em tecnologia avançada. Implementam uma "gestão social matricial", que envolve dividir as comunidades em "zonas geométricas para que as forças de segurança possam sistematicamente observar todas as atividades com o apoio das novas tecnologias".[15] Para atingir esse objetivo, o Estado colocou postos policiais a algumas dezenas de metros entre si em distritos selecionados, nos quais trabalham dezenas de milhares de agentes de segurança. Além disso, as autoridades chinesas estão equipando essa força com equipamentos avançados de vigilância e sistemas capazes de analisar uma grande quantidade de dados (*big data*).

Em particular, os chineses estão construindo um programa de policiamento preditivo que agrega e analisa múltiplas fontes de dados para identificar potenciais ameaças. A ONG Human Rights Watch expôs a criação, pelas autoridades de Xinjiang, de uma "Plataforma Integrada de Operações Conjuntas" (IJOP) que coleta informação de fontes incluindo câmeras de circuito fechado de tevê (muitas vezes equipadas com software de reconhecimento facial) e "farejadores de Wi-Fi" que coletam informações identificáveis de laptops e smartphones. A IJOP coleta informação adicional de placas de veículos e documentos de identidade solicitados durante inspeções policiais, bem como registros de saúde, bancários e outros.[16] Embora se desconheça o grau de integração entre a IJOP e outros esforços de coleta de dados da RPC, é importante ressaltar que as autoridades chinesas vêm crescentemente empregando dispositivos de varredura de eletrônicos para invadir smartphones e extrair contatos, comunicações em redes sociais, e-mails, fotos e vídeos. Além disso, a RPC criou recentemente um banco de dados de DNA oficial e compulsório com o objetivo de obter amostras de todos os residentes de Xinjiang com idades entre 12 e 65 anos.[17]

Uma vez que as informações relevantes sejam alimentadas nos computadores da IJOP, algoritmos vasculham por entre os dados à procura de padrões que possam indicar comportamento de risco. Não se sabe que limiares de confiança as autoridades chinesas estão usando para rodar esses testes, mas os algoritmos estão provavelmente gerando um número significativo de falsos positivos por limitações do sistema. Uma vez que a máquina identifica um indivíduo, aquela pessoa pode ser capturada pelas forças de segurança e detida por um período indefinido.

Desenvolver esse sistema não foi barato. A quantia alocada pelas autoridades de Xinjiang para "projetos de investimento relacionados à segurança" supostamente cresceu de apenas 27 milhões de dólares em

2015 para mais de 1 bilhão de dólares no primeiro trimestre de 2017.[18] No entanto, esse é um valor irrisório perto da quantia que o Estado teria que investir para construir um sistema comparável de vigilância e repressão sem o uso de tecnologia de IA.

Cenário 3: usando a desinformação para deslegitimar os oponentes. No terceiro cenário, um regime autocrático realiza eleições nacionais previstas na sua constituição. Ele planeja engajar-se em ações tradicionais de manipulação eleitoral, violação de urnas e supressão de votos, mas também busca novas estratégias que o ajudem a garantir a vitória sobre a oposição. A tecnologia de IA pode ajudar de diversas maneiras, em particular na frente de desinformação.

Em primeiro lugar, a IA pode manipular as informações disponíveis e disseminar mensagens-chave do regime. Por exemplo, as redes sociais usam algoritmos de curadoria de conteúdo para apresentar certos artigos aos usuários — e mantê-los fixados nos *feeds* de sua plataforma. Autoridades estatais podem explorar tais algoritmos para disseminar mensagens pró-regime usando exércitos de *bots* e *trolls* de aluguel. A IA pode ajudar a identificar importantes "influenciadores" das redes sociais, que as autoridades podem então cooptar para a disseminação de desinformações. Tecnologias emergentes de IA também podem facilitar o emprego, via redes sociais, de campanhas de desinformação automatizadas e hiperpersonalizadas — direcionadas a grupos ou indivíduos específicos —, na linha dos esforços russos de influenciar as eleições norte-americanas de 2016 ou os exércitos de *trolls* sauditas visando dissidentes como o jornalista assassinado Jamal Khashoggi. Nos últimos anos tem se observado uma tendência de disseminação de desinformação por atores políticos por esses e outros meios para estimular seus apoiadores e desorientar seus oponentes.

Em segundo lugar, a tecnologia de IA é cada vez mais capaz de produzir vídeos e áudios falsos com bastante realismo. Uma nova técnica, cujo potencial de desinformação preocupa em especial formula-

dores de políticas públicas, é o uso de redes generativas adversariais, que colocam sistemas de IA competindo uns contra os outros.[19] Em essência, a primeira máquina gera falsificações que a segunda máquina tenta detectar. O retorno do segundo sistema então ajuda o primeiro sistema a criar exemplos cada vez mais realistas. Ultimamente, isso pode resultar em falsificações sofisticadas que mesmo sistemas avançados de IA podem ser incapazes de detectar. Para líderes autoritários, a tecnologia de *deep-fake** oferece um meio de descreditar potenciais rivais, que podem se tornar vítimas de vídeos manipulados onde aparecem fazendo afirmações polêmicas ou praticando atos abjetos.

Principais desafios de política pública

A proliferação da tecnologias de IA e a ascensão da repressão digital representam importantes desafios de políticas públicas às democracias liberais. Uma questão-chave é se poderosas ferramentas de IA tornaram as próprias democracias mais repressivas. Em particular, será que a tentação gerada pelo potencial de vigilância oferecido pela IA acabará corroendo as salvaguardas democráticas?

A história sugere que os cidadãos têm motivo para se preocupar. Em 1975, alegações surpreendentes de conduta imprópria por parte da Comunidade de Inteligência dos Estados Unidos levaram o Senado americano a autorizar a criação do que veio a se tornar a Comissão Church. A CIA (Agência Central de Inteligência) havia sido acusada de conduzir tentativas de assassinato contra líderes estrangeiros, e outras agências haviam constituído redes de vigilância doméstica para monitorar e perseguir ativistas pelos direitos civis, manifestantes políticos e organizações indígenas.[20] Em seu último relatório, a Comissão Church alertou: "frequentemente, princípios constitucionais

* Junção das expressões "*deep learning*" (uma técnica de *machine learning*) e "*fake*" (falsificação), o *deep-fake* é uma técnica usada para produzir vídeos falsos de pessoas reais.

foram sujeitados a um pragmatismo segundo o qual os fins almejados ditavam e justificavam meios irregulares".[21] Apesar de uma longa tradição americana de proteção dos direitos individuais e contrapesos à autoridade governamental, o potencial para abusos do Estado permaneceu vasto.

O abuso intencional pelo aparato de segurança do Estado não é a única questão de liberdades civis levantada pela ascensão da IA. O viés implícito e a discriminação presentes em algoritmos também são motivos para preocupação. O aprendizado de IA usado no policiamento ou na atenção à saúde, por exemplo, podem reforçar a desigualdade e produzir ou perpetuar práticas discriminatórias. Um exemplo notório de viés implícito foi um incidente em 2015 no qual o sistema de indexação de imagens do Google descreveu fotos de afrodescendentes americanos como "gorilas". O maior culpado havia sido a "base de dados de treinamento" usada para "ensinar" o algoritmo a identificar rostos, que incluía predominantemente rostos caucasianos. Decerto não ajudou o fato de apenas 2% do quadro "profissional" do Google ser composto de afrodescendentes, algo que pode ter impedido a equipe de ter percebido antes esse problema.[22]

Pesquisas posteriores mostraram que o preconceito humano possui um efeito profundo no funcionamento de sistemas de IA. Um artigo de 2017 publicado pela *Science* documentou como programas de *machine learning* adquirem vieses a partir de dados textuais: o programa testado passou a associar descrições relacionadas à família, como "pais", com nomes femininos, enquanto relacionava nomes masculinos com termos como "profissional".[23]

A área da justiça criminal foi pioneira no uso de análise preditiva baseada em IA, mas estudos revelam que os programas usados frequentemente se baseavam em dados enviesados. Por exemplo, as estatísticas criminais indicam que afrodescendentes possuem mais chances de serem presos pela polícia do que seus pares caucasianos.

Mas algoritmos de máquina raramente consideram que o viés policial possa ter sido o motivo por trás da detenção desproporcional de afrodescendentes. Em vez disso, o pressuposto algorítmico padrão é que os afrodescendentes são mais propensos a cometer crimes. Essa conclusão dúbia forma a base para previsões subsequentes produzidas por esses algoritmos, ressaltando um princípio fundamental: máquinas com IA são tão boas quanto os dados usados para treiná-las.

Em 2018, não é difícil imaginar governos liberais-democráticos explorando tecnologia de IA de maneiras que violem os direitos dos cidadãos. Felizmente, cidadãos em democracias avançadas foram bem-sucedidos na luta contra abusos da vigilância estatal no passado, e há instrumentos robustos de freios e contrapesos capazes de frear abusos estatais. Embora não haja garantia de que a IA não irá enfraquecer os sistemas políticos democráticos, o risco nessas democracias é menor.

O perigo é muito maior no caso de democracias frágeis ou países com tendências autoritárias. Em regimes em retrocesso democrático como Polônia, Hungria ou as Filipinas, o potencial repressivo da IA pode levar a uma deterioração ainda maior. Governos iliberais prestes a enfrentar disputas populares possuem um interesse natural em tecnologia que possa ajudá-los a enfrentar o descontentamento em massa. Mesmo em sistemas políticos abertamente democráticos, governos possuem grande incentivo para armar as forças de segurança com tecnologia intrusiva, monitorar as atividades de oponentes políticos e da sociedade civil e agir preventivamente contra desafios potenciais à sua autoridade. Estados também monitoram de perto as ações uns dos outros. À medida que a IA prova seu valor repressivo para as autocracias que inauguram o uso de novas tecnologias, outros governos podem começar a imitá-las.

Os Estados Unidos e a China lideram o mundo em tecnologia de IA, mas oferecem visões bastante diferentes para seu uso. Para a China, a

IA é um componente essencial de um sistema mais amplo de controle que sustenta o governo do Partido Comunista. Além disso, oferecer novas possibilidades trazidas pela IA a outros regimes autoritários ajuda no avanço dos grandes objetivos estratégicos do regime, em particular o de "minar a ordem liberal ocidental e atingir a hegemonia da RPC na Ásia e a expansão da influência chinesa no mundo".[24]

Consequentemente, os chineses vêm trabalhando agressivamente para desenvolver novas capacidades de IA e oferecer seus novos produtos no exterior. Dos três principais componentes da IA — dados de treinamento para *machine learning*, poder computacional e algoritmos poderosos —, a China possui dados de treinamento em abundância e seus algoritmos estão melhorando, mas sua capacidade industrial de produção de chips de computador ainda está bem aquém da norte-americana.

Em contraste, os Estados Unidos possuem os microchips mais avançados do mundo, e seus algoritmos também são referência mundial em termos de sofisticação e complexidade. Mas os EUA estão cada vez mais atrás da China em termos de dados digitais disponíveis às empresas que trabalham com IA. Isso importa porque, cada vez mais, os dados "fazem toda a diferença" no que se refere a construir empresas baseadas em IA capazes de superar seus competidores.[25] Por meio da iniciativa "Made in China 2025", a RPC busca transformar sua capacidade de produção de chips por meio de investimentos e roubo de propriedade intelectual para dominar parte fundamental da indústria de alta tecnologia. Especialistas alertam que essa campanha sinaliza um desejo "nem tanto de se juntar a economias tecnologicamente desenvolvidas como Alemanha, EUA, Coreia do Sul e Japão, mas, sim, de substituí-las".[26] O rápido avanço da startup de IA Yitu é emblemático desse desejo chinês.

A Yitu foi fundada por dois especialistas chineses em IA em 2012 e, em apenas seis anos, alcançou diversos marcos impressionantes. Sua plataforma de imagens *Dragonfly Eye* [olho da libélula] já contém

mais de 1,8 bilhão de fotografias, e a empresa afirma que o sistema precisa de apenas três segundos para identificar um indivíduo em seu banco de dados, que inclui imagens da base nacional da RPC bem como cerca de 320 milhões de fotos de entrada e saída de pessoas das fronteiras do país. O valor de mercado da Yitu atingiu cerca de 2,4 bilhões de dólares em 2018, e a empresa emprega atualmente mais de 500 funcionários espalhados por Xangai, Singapura e Vale do Silício. Mais importante, seus algoritmos funcionam: as tecnologias de reconhecimento facial da Yitu venceram importantes prêmios do Instituto Nacional de Padrões e Tecnologia (NIST) e do programa Intelligence Advanced Research Projects Activity (IARPA) da Comunidade de Inteligência dos Estados Unidos.

À medida que a China desenvolve um forte setor de IA, utiliza a política "One Belt, One Road" para disseminar essa tecnologia sofisticada por governos do mundo inteiro. Entre os projetos ilustrativos estão a construção de uma rede de "cidades seguras" no Paquistão (com uma extensa tecnologia de monitoramento embutida diretamente na infraestrutura dessas cidades) e o fornecimento de software de IA e reconhecimento facial a autoridades da Argentina para o aperfeiçoamento da vigilância pública. A RPC considera, de maneira astuta, que, quanto mais ela puder alinhar os modelos de governança de outros países ao da China, menos esses países representarão uma ameaça à hegemonia chinesa. Ademais, à medida que os governos se tornam dependentes de tecnologia avançada chinesa para controlar suas populações, serão crescentemente pressionados a alinhar suas políticas aos interesses estratégicos da RPC.

De fato, a estratégia de IA da China é clara em relação aos benefícios percebidos da tecnologia: "A IA dará um novo impulso ao avanço das reformas estruturais pelo lado da oferta, uma nova oportunidade para revitalizar a economia real e será um novo motor para fazer da China uma superpotência industrial e cibernética".[27]

Respostas políticas

Nos próximos anos, a IA terá um grande impacto na política global. Embora não haja uma resposta política única capaz de resolver uma questão tão complexa e multifacetada, há diversas implicações importantes para os Estados democráticos.

Em geral, democracias avançadas deveriam reconhecer de maneira mais explícita o tamanho da ameaça representada pela tecnologia de IA a sistemas políticos abertos. Os esforços da China para construir sofisticadas capacidades de IA, bem como a proliferação dessa tecnologia para outros regimes autoritários, representam um grave risco no longo prazo. Formuladores de política pública do Ocidente deveriam ter como prioridade muito mais alta opor-se a esses esforços, tanto externa como domesticamente.

O abuso da tecnologia de IA não está limitado aos regimes autoritários. À medida que governos democráticos adquirem novas tecnologias que ampliam dramaticamente sua capacidade de monitoramento e vigilância, precisam estabelecer limites aceitáveis ao uso dessas tecnologias. Democracias devem olhar para dentro e liderar o desenvolvimento de um arcabouço regulatório. Esse processo será conturbado; a inovação tecnológica frequentemente está muitos anos à frente da capacidade de reguladores de desenvolver padrões e diretrizes razoáveis. Entretanto, as democracias avançadas estão mais bem posicionadas para pensar em como regular empresas privadas e evitar abusos.

Esforços domésticos deveriam vir acompanhados de uma ação internacional para criar arcabouços legais mais claros para o uso de IA. Iniciativas como os Princípios Orientadores sobre Empresas e Direitos Humanos da ONU são modelos úteis. O desenvolvimento de diretrizes internacionais sobre tecnologia de IA exigirá um processo *multistakeholder* que seja inclusivo por natureza; flexível o bastante para refletir novos avanços tecnológicos; e resistente à captura pela

China ou outros governos autoritários. Também é necessária uma discussão normativa muito mais profunda. A comunidade internacional ainda precisa enfrentar questões relacionadas a vieses algorítmicos, discriminação implícita e privacidade.

Por fim, as democracias precisam considerar maneiras de fortalecer a capacidade da sociedade civil de resistir à repressão facilitada pela IA e participar na construção de diretrizes para o uso de IA. Organizações da sociedade civil (OSCs) locais que operam em ambientes repressivos precisarão de mais recursos, treinamento e apoio tecnológico. Muitas dessas organizações migraram para o mundo online, mas não estão usando ferramentas de segurança digital amplamente disponíveis, como serviços de criptografia. Consequentemente, enfrentam riscos significativos de ataque cibernético, intrusão, monitoramento e vigilância. Para OSCs que operam em democracias, o grande desafio é conseguir monitorar propostas de regulamentação de forma abrangente, chamar a atenção pública para violações pelo uso indevido de IA e assumir um papel fiscalizador.

Investigações tais como a reportagem da ProPublica mostrando o viés implícito no sistema americano de justiça criminal estão fazendo uma diferença considerável em como governos usam tecnologias de IA. À medida que mais governos adotam plataformas de IA, haverá uma demanda crescente por esse tipo de trabalho. Internacionalmente, é vital que atores da sociedade civil possuam voz nas discussões sobre como regular adequadamente a IA.

As tecnologias de IA têm uma dupla função: podem ser empregadas para propósitos benéficos bem como usadas para fins militares e repressivos. Mas essa tecnologia não pode ser claramente classificada como "benéfica" ou "nociva". As funções que ganham valor a partir da automação também podem ser usadas tanto por regimes autoritários para fins maliciosos como por atores democráticos ou comerciais para fins benéficos. Para garantir que a IA seja usada de maneira res-

ponsável, ampliar as conexões entre o meio político e engenheiros e pesquisadores será essencial.

Em outras palavras, aqueles responsáveis por planejar, programar e implementar sistemas de IA também deveriam compartilhar a responsabilidade por aplicar e garantir padrões de defesa dos direitos humanos. Especialistas em políticas públicas deveriam manter um diálogo regular e aberto com engenheiros e tecnólogos para que todos os envolvidos tenham em mente os potenciais abusos da IA e possam desenvolver respostas apropriadas ainda na fase inicial do desenvolvimento de projetos de IA.

As autocracias do mundo, lideradas pela China, vêm crescentemente demonstrando os perigos representados pela interseção da tecnologia de IA de ponta, inovações mais amplas nas esferas da informação e da comunicação e projetos autoritários de coerção e controle. Para contrapor não apenas a disseminação da repressão de alta tecnologia no exterior, mas também abusos potenciais dentro de suas fronteiras, os formuladores de políticas públicas em Estados democráticos precisam considerar seriamente como mitigar o dano e desenvolver melhores práticas. Do Paquistão ao Zimbábue, uma perigosa visão autoritária do futuro da IA está sendo construída. Chegou a hora de os atores democráticos prepararem uma resposta à altura.

Notas

1. Li Tao, "Malaysian Police Wear Chinese Start-Up's AI Camera to Identify Suspected Criminals". *South China Morning Post*, 20 abr. 2018.

2. Aradhana Aravindan; John Geddie, "Singapore to Test Facial Recognition on Lampposts, Stoking Privacy Fears". *Reuters*, 13 abr. 2018.

3. Amy Hawkins, "Beijing's Big Brother Tech Needs African Faces". *Foreign Policy*, 24 jul. 2018.

4. Nils J. Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. Cambridge: Cambridge University Press, 2010.

5. Jerry Kaplan, *Artificial Intelligence: What Everyone Needs to Know*. Nova York: Oxford University Press, 2016, pp. 5-6.

6. Gabinete Executivo do Presidente dos Estados Unidos [EOP]. Conselho Nacional de Ciência e Tecnologia [NSTC]. Comitê de Tecnologia. "Preparing for the Future of Artificial Intelligence". Out. 2016. Disponível em: <https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf>. p. 8.

7. Milan W. Svolik, *The Politics of Authoritarian Rule*. Cambridge: Cambridge University Press, 2012, p. 124.

8. Lucan A. Way; Steven Levitsky, "The Dynamics of Autocratic Coercion After the Cold War". *Communist and Post-Communist Studies*, v. 39, n. 3 (set. 2006), pp. 387-410.

9. Em comparação, por exemplo, o serviço de segurança da Alemanha Oriental, a Stasi, contava com uma rede de informantes equivalente a 1% da população total do país, gerando custos econômicos vultosos e contínuos. Ver Andreas Lichter; Max Löffler; Sebastian Siegloch, "The Long-Term Costs of Government Surveillance: Insights from Stasi Spying in East Germany". SOEPpaper n. 865, Deutsches Institut für Wirtschaftsforschung, Berlin, 2016. Disponível em: <www.econstor.eu/bitstream/10419/146890/1/869045423.pdf>.

10. V-Dem Institute, *Democracy for All? V-Dem Annual Democracy Report 2018*, 2018. Disponível em: <www.v-dem.net/en/news/democracy-all-v-dem-annual-democracy-report-2018>. p. 19. Ver também Erica Frantz; Andrea Kendall-Taylor, "The Evolution of Autocracy: Why Authoritarianism Is Becoming More Formidable". *Survival*, v. 59, n. 5 (out.-nov. 2017), pp. 57-68.

11. As estatísticas e sistema de classificação usados para chegar a esse número baseiam-se em dados de 1946 a 2010 de Barbara Geddes; Joseph Wright; Erica Frantz, "Autocratic Breakdown and Regime Transitions: A New Data Set". *Perspectives on Politics*, v. 12, n. 2 (jun. 2014), pp. 313-31. Dados para o período 2010-17 foram atualizados pelo autor com o apoio de Erica Frantz.

12. Collin Anderson, "The Hollowing Middle of the Surveillance Malware Market". Motherboard, 14 dez. 2017. Disponível em: <https://motherboard.vice.com/en_us/article/595dkd/the-hollowing-middle-of-the-surveillance-malware-market>.

13. Josh Horwitz, "WeChat's New Heat Map Feature Lets Users—and Chinese Authorities—See Where Crowds Are Forming". *Quartz*, 7 out. 2015. Disponível em: <https://qz.com/518908/wechats-new-heat-map-feature-lets-users-and-chinese-authorities-see-where-crowdsare-forming>.

14. Gerry Shih, "China's Mass Indoctrination Camps Evoke Cultural Revolution". *Associated Press*, 18 maio 2018.

15. Adrian Zenz; James Leibold, "Chen Quanguo: The Strongman Behind Beijing's Securitization Strategy in Tibet and Xinjiang". *China Brief*, Jamestown Foundation, 21 set. 2017.

16. Human Rights Watch, "China: Big Data Fuels Crackdown in Minority Region". 26 fev. 2018. Disponível em: <www.hrw.org/news/2018/02/26/china-big-data-fuels-crackdown-minority-region>.

17. Cate Cadell, "From Laboratory in Far West, China's Surveillance State Spreads Quietly". *Reuters*, 14 ago. 2018; Human Rights Watch, "China: Minority Region Collects DNA from Millions". 13 dez. 2017. Disponível em: <www.hrw.org/news/2017/12/13/chinaminority-region-collects-dna-millions>.

18. Josh Chin; Clément Bürge, "Twelve Days in Xinjiang: How China's Surveillance State Overwhelms Daily Life". *Wall Street Journal*, 19 dez. 2017.

19. Will Knight, "The US Military Is Funding an Effort to Catch Deepfakes and Other AI Trickery". *MIT Technology Review*, 23 maio 2018.

20. Ver LeRoy Ashby; Rod Gramer, *Fighting the Odds: The Life of Senator Frank Church* (Pullman, WA, EUA: Washington State University Press, 1994), p. 478.

21. *Intelligence Activities and the Rights of Americans: Final Report of the Select Committee to Study Governmental Operations with Respect to Intelligence Activities*, Senado dos Estados Unidos, Livro II. Washington, D.C.: U.S. Government Printing Office, 1976. Disponível em: <www.intelligence.senate.gov/sites/default/files/94755_II.pdf>. p. III.

22. Wendy Lee, "How Tech's Lack of Diversity Leads to Racist Software". *San Francisco Chronicle*, 22 jul. 2015.

23. Universidade de Princeton, Faculdade de Engenharia, "Biased Bots: Human Prejudices Sneak into Artificial Intelligence Systems". *ScienceDaily*, 13 abr. 2017.

24. Minxin Pei, "China in Xi's 'New Era': A Play for Global Leadership". *Journal of Democracy*, v. 29, n. 2 (abr. 2018), p. 38.

25. Kai-Fu Lee, *AI Superpowers: China, Silicon Valley, and the New World Order*. Nova York: Houghton Mifflin Harcourt, 2018, p. 56.

26. Lorand Laskai, "Why Does Everyone Hate Made in China 2025?" *Net Politics*, Council on Foreign Relations, 28 mar. 2018. Disponível em: <www.cfr.org/blog/why-does-everyonehate-made-china-2025>.

27. Paul Triolo; Elsa Kania; Graham Webster, "Translation: Chinese Government Outlines AI Ambitions Through 2020". *DigiChina*, New America, 26 jan. 2018. Disponível em: <www.newamerica.org/cybersecurity-initiative/digichina/blog/translation-chinese-governmentoutlines-ai-ambitions-through-2020>.

Plataforma Democrática (www.plataformademocratica.org) é uma iniciativa da Fundação FHC e do Centro Edelstein de Pesquisas Sociais dedicada a fortalecer a cultura e as instituições democráticas na América Latina, por meio da produção de conhecimento e da promoção do debate pluralista de ideias sobre as transformações da sociedade e da política na região e no mundo. Realiza pesquisas e seminários para estimular o diálogo entre os produtores de conhecimentos e os diferentes atores sociais e políticos sobre temas da atualidade.

Plataforma Democrática oferece uma infraestrutura virtual com uma biblioteca de livre acesso que inclui milhares de textos sobre temas relacionados à democracia na América Latina e um banco de dados sobre instituições de pesquisa na região.

As principais áreas de trabalho da Plataforma Democrática são:

Transformações Geopolíticas Globais e instituições democráticas:
http://www.plataformademocratica.org/portugues/publicacoes#EstadoDemocracia
http://www.plataformademocratica.org/portugues/publicacoes#CambiosGeopoliticos

Meios de comunicação e Democracia:
http://www.plataformademocratica.org/portugues/publicacoes#MediosComunicacion
http://www.plataformademocratica.org/portugues/publicacoes#EnsaiosDemocracia

Sociedade civil e democracia:
http://www.plataformademocratica.org/portugues/publicacoes#CohesionSocial

Bibliotecas virtuais:
http://www.plataformademocratica.org/portugues/biblioteca
http://www.plataformademocratica.org/portugues/biblioteca-sociedade

Coleção Recursos de Pesquisa na Internet:
http://www.plataformademocratica.org/portugues/publicacoes#RecursosPesquisa