

A “WEAPONIZAÇÃO” DA IA: UMA CRÍTICA À MORAL CONTEMPORÂNEA

THE WEAPONIZATION OF AI: A CRITIQUE OF CONTEMPORARY MORALITY

LUIZ GUILHERME BAKKER DE PINHO E SOUZA¹

<https://orcid.org/0009-0007-3570-2984>

RESUMO: Recentemente, a inteligência artificial (IA) se tornou popular e facilmente acessível no mundo virtual. Assim, a IA também passou a ser utilizada para atacar outras pessoas, como em casos em que alguém se passa por um parente de uma vítima para aplicar golpes e roubar dinheiro, transformando-a em uma arma de ataque, ou seja, *weaponizando-a*. O objetivo do presente texto é analisar a inteligência artificial e seus riscos, esclarecer o significado do termo “weaponizar” e, por fim, realizar uma consideração sobre o que os usos espúrios dessa tecnologia avançada dizem sobre a moral no século 21.

PALAVRAS-CHAVE: IA; *Weaponização*; Moralidade.

ABSTRACT: Recently, Artificial Intelligence (AI) became popular and easily accessible in the virtual world. As such, AI also started to be used to attack other people, like in cases where someone impersonates their victim’s relatives in order to scam and steam money, turning this tool into an attack weapon, ergo, weaponizing it. The objective of the present text is to analyze the artificial intelligence and its risks, to clarify the meaning of the term “to weaponize” and, finally, to consider what the bad uses of this advanced technology are saying about morality on the 21st Century.

KEYWORDS: AI; *Weaponization*; Morality.

Introdução

O presente texto tem como objetivo tratar da relação entre ser humano e máquina, considerando, especificamente, a popularização de ferramentas baseadas em inteligência artificial. Com tantos trabalhos sobre a IA, é importante que se considere como as pessoas estão reagindo e o que estão fazendo com ela. O foco, portanto, será voltado mais para o usuário dessa tecnologia.

Hoje em dia, é possível encontrar diversos aplicativos e websites que funcionam com inteligência artificial. Certamente, um dos mais populares é o ChatGPT, que é uma IA generativa de textos. Seja como for, existe uma miríade de possibilidades de uso de uma IA: o usuário pode ter conversas fictícias com personagens virtuais graças aos LLMs (Large

¹ Doutorando em Filosofia na Pontifícia Universidade Católica do Rio de Janeiro. Bolsista da CAPES. Contato: luizbakker@protonmail.com. Link do currículo: <http://lattes.cnpq.br/7973245635877221>.

Language Models), criar representações visuais automatizadas utilizando, por exemplo, o DALL-E, configurar um amigo virtual e até mesmo solicitar textos sobre os mais diversos temas, em vários estilos possíveis, sendo possível, inclusive, emular o estilo de escrita de outras pessoas, novamente usando o ChatGPT. Por conta desse amplo leque de capacidades demonstrado pela inteligência artificial, é até previsível que exista um temor cada vez maior sobre certos riscos atrelados a esta tecnologia. Há diversas problematizações a respeito da inteligência artificial, existindo até mesmo discussões sobre a possibilidade que a IA assuma o controle da humanidade.²

Não se trata apenas do que a inteligência artificial pode fazer em sentido amplo. A questão também envolve o que exatamente pode ser feito com ela. Em casos mais simples, ela pode servir de instrumento para que golpistas enganem outras pessoas. Em casos mais complexos, ela pode controlar tecnologias, tanto no campo civil quanto no campo militar. Em casos como esses, encontrar um responsável pode se tornar uma tarefa bastante desafiadora.

Mas a inteligência artificial não possui, originalmente, fins ofensivos.³ Ela não foi criada para ser uma arma, e sim para realizar tarefas específicas conforme sua programação. Isso significa que, ao ser utilizada para montar uma mensagem de áudio falsa, ela está realizando a tarefa para o qual foi programada, mas com um objetivo ofensivo, diferente do original. Dessa forma, ela pode ser considerada uma arma. Essa ressignificação específica, no idioma inglês, é conhecida como *weaponization*, que, aqui, será traduzida como *weaponização*.

Este texto seguirá, assim, uma linha um pouco diferente do habitual: a problematização não será centrada na tecnologia em questão, mas sim na forma como ela está sendo utilizada pelo ser humano em geral. Então, resta indagar: como o indivíduo está reagindo a essas novas tecnologias? Será que a inteligência artificial representa uma ameaça à existência humana em si, ou ela também representa mais um risco à moralidade? Será realizada uma breve análise do comportamento das pessoas ao utilizar a inteligência artificial generativa, com o objetivo de convidar a refletir sobre como as novas tecnologias estão afetando as relações interpessoais em um sentido amplo.

² Há, inclusive, uma discussão sobre Inteligência Artificial Geral, ou Superinteligência artificial, que seria o futuro da IA, ou seja, sua evolução.

³ O conceito de inteligência artificial, no entanto, é bastante amplo e abrange diversas formas de IA. O ponto em comum entre elas seria a IA como um objeto autônomo capaz de cumprir tarefas voltadas para a resolução de problemas. Fonte: https://pt.wikipedia.org/wiki/Intelig%C3%A2ncia_artificial.

1. Sobre a palavra ‘*weaponização*’

A palavra aqui utilizada, *weaponização*, é baseada no vocábulo *weaponization*. Esse é um neologismo que significa “adaptar algo para utilizar como arma”.⁴ Basta que o objeto em questão seja utilizado de maneira agressiva, ou ao menos com intenção ofensiva, como, por exemplo, atirar um copo em outra pessoa.⁵ Ao fazer isso, o indivíduo está *weaponizando* o copo. Curiosamente, esse termo não é recente, tendo sido utilizado pela primeira vez em 1957, segundo o dicionário de Oxford.

É importante ressaltar que, contextualmente, o ato de recondicionar algo com objetivo ofensivo pode ocorrer de duas formas: no primeiro caso, o objeto pode ser diretamente utilizado para ataque, conforme o exemplo mencionado; no segundo caso, um veículo pode ter uma arma de fogo anexada a ele. Adequando o exemplo ao tema aqui proposto, uma inteligência artificial, originalmente programada para realizar uma função simples, pode ser usada para enganar outras pessoas, ou ser treinada para gerenciar sistemas bélicos.

Ainda em relação ao termo *weaponizar*, não existe uma tradução oficialmente estabelecida para esse vocábulo. Uma breve pesquisa em mecanismos de busca⁶ retorna, como possíveis traduções, os termos “militarização” e “armamentização”, sem que haja um consenso sobre qual termo é preferencial. Dado que essa falta de consistência prejudicaria a argumentação que se busca construir, optou-se, então, por abrigar o termo original, para fins de clareza contextual.

O tema da *weaponização* da inteligência artificial não é algo inexplorado. Em 2019, Joe Burton e Simona Soare publicaram um artigo intitulado *Understanding the Strategic Implications of the Weaponization of Artificial Intelligence*.⁷ O artigo trata precisamente do potencial de uma IA no ciberespaço, mas também como parte integrante de sistemas bélicos, nos dois contextos da palavra *weaponização*. O texto justifica a crescente preocupação com os perigos da inteligência artificial, por mais que Burton e Soare ressaltem que “a tecnologia pode claramente ser aproveitada para melhorar a segurança assim como para destruir”.⁸

⁴ Fonte: <https://www.oxfordlearnersdictionaries.com/definition/english/weaponization>. Último acesso em 05/08/2023.

⁵ Até emoções e comportamentos podem ser *weaponizados*.

⁶ Em outras palavras, uma busca no Google pela frase “*weaponizar* em português”.

⁷ Em português, *Implicações Estratégicas da Weaponização da Inteligência Artificial*.

⁸ Burton e Soare, 2019, p. 16.

2. Sobre a Inteligência Artificial e seus riscos

Normalmente, quando se fala de riscos relacionados à IA, é comum que se imagine um futuro distópico onde as máquinas tomaram controle da humanidade. Burton e Soare referenciam esse mesmo tipo de imaginação ao mencionar, na introdução de seu texto, o filme *Exterminador do Futuro*, que descreve esse exato cenário, em que uma inteligência artificial se manifesta primariamente, na forma de robôs assassinos decididos a eliminar a humanidade. Os próprios autores consideram esse tipo de debate como hiperbólico.

Entre os artigos que serão analisados a seguir, o já citado texto de Burton e Soare trata dos riscos da inteligência artificial no ciberespaço, em um sentido geral, embora falem sobre essa tecnologia também no campo bélico. Os autores afirmam:

Os avanços na IA também fazem com que os próprios malwares sejam mais danosos, mais sofisticados e mais capazes de mirar precisamente em seu destinatário desejado. Um exemplo recente é o software Deeplocker, desenvolvido pela IBM Research, que é altamente evasivo e capaz de ocultar suas intenções malignas antes de chegar ao seu alvo. O malware identifica alvos através de indicadores de mídias sociais, incluindo reconhecimento facial, geolocalização e reconhecimento vocal, e evita ser detectado até entregar sua ‘carga’ (...). Esse é apenas um exemplo do leque, em expansão, de capacidades ofensivas aprimoradas ou facilitadas pela IA.⁹

A IA, conforme é aprimorada, aumenta os danos dos malwares. Esses programas invasivos, outrora utilizados em ataques diretos, se tornaram mais furtivos e capazes de atingir seus alvos, pois a inteligência artificial sabe como evadir proteções, identificar possíveis alvos com base em indicadores de mídias sociais, e consegue se camuflar até atingir sua carga.

Em outro artigo relevante, intitulado *Malevolent soft power, AI, and the threat to democracy*, Kamarck aborda outro assunto relacionado, relatando que a IA pode ajudar a produzir edições de vídeos onde uma pessoa parece dizer algo que nunca falou:

O outro desenvolvimento da IA que oferece problemas reais para a democracia é a habilidade crescente de sistemas de IA de colocar palavras que nunca foram ditas nas bocas das pessoas (...).

Esse processo funciona utilizando algoritmos de reconhecimento facial e um clipe de áudio da pessoa falando, para manipular a boca e fazer parecer que ela está dizendo algo que não está. Até agora, a tecnologia se foca em mudar o formato da boca, mas mudar uma expressão facial ou uma postura completa certamente está no horizonte. Atualmente, é importante ter grandes quantidades de vídeo da pessoa falando para possibilitar este processo. Portanto, líderes mundiais como Barack Obama ou David Cameron estão

⁹ Burton e Soare, 2019, p. 10. Todos os trechos citados neste artigo foram traduzidos por mim.

particularmente vulneráveis a protagonizar vídeos com palavras falsas saindo de suas bocas.¹⁰

Na época em que Kamarck escreveu seu artigo, 2018, esta funcionalidade não era tão avançada e popular quanto é hoje, mas já possuía o mesmo potencial devastador, conforme exemplo explicitado pela autora. Hoje em dia, anos depois, não é necessário mais do que dois minutos de áudio para que a inteligência artificial possa produzir um vídeo ainda mais convincente, o que representa um risco ainda maior para a democracia.

Coeckelbergh, em seu artigo intitulado *Artificial Intelligence, Responsibility Attribution, and a Relational Justification of Explainability*, segue em uma direção diferente em seu texto. Ele reconhece que a IA pode ser usada de formas diferentes daquelas para as quais foi projetada, ou seja, mal-utilizada, mas insiste no valor moral e ético da ferramenta baseada em IA. Dado que o autor se baseia na ética aristotélica para argumentar em seu artigo, ele considera a inteligência artificial como um instrumento. Sendo assim, o problema está não no instrumento, mas nas consequências de seu uso:

Geralmente, programadores e usuários sabem o que eles querem fazer com a IA. Mais precisamente, eles sabem o que eles querem que a IA faça *para eles*. Eles sabem o objetivo, as consequências *desejadas*. Aristóteles diria *o fim*. Entretanto, usuários de IA não estão necessariamente a par das consequências *não desejadas* e da *significância moral* do que eles fazem. Por exemplo, eles podem não saber que há uma parcialidade no conjunto de dados que estão usando ou mesmo em seus algoritmos. Eles podem não estar nem a par de que o que eles fazem é moralmente significativo (...). E eles podem não saber as consequências exatas para aqueles que foram atingidos pelo algoritmo.¹¹

Os usuários da IA não pensam muito no funcionamento da tecnologia que utilizam, e acabam não levando em consideração que existem consequências indesejadas, visto que o código das IA possui uma parcialidade, ou mesmo um viés intrínseco. Eles não entendem completamente o funcionamento do algoritmo que estão utilizando, assim como não entendem completamente as consequências do uso indiscriminado da tecnologia que têm em mãos.

Mas o que mais chama atenção nessa afirmação de Coeckelbergh é a parte onde o autor se refere ao desconhecimento, da parte do usuário, da significância moral de suas ações com a inteligência artificial. Em outras palavras, alguém que *weaponiza* uma IA desconhece completamente, ou, pelo menos, ignora, o fato de que todo ato possui consequências, e isso inclui a maneira como uma tecnologia é utilizada. Nesse sentido, seria possível explorar mais

¹⁰ Kamarck, 2018, online.

¹¹ Coeckelbergh, 2019, p. 2059, grifos do autor.

profundamente como um indivíduo que *weaponiza* qualquer coisa ignora ou desconhece as consequências morais de seu ato, mas isso levaria a outro texto que seria ao menos duas vezes maior do que este artigo.

Embora os três textos relatem riscos da inteligência artificial, com Coeckelbergh indo em uma direção diferente dos outros autores, eles convergem no fato de que há alguém utilizando ou programando uma IA para realizar ações específicas. No caso da afirmação de Burton e Soare, onde essa tecnologia faz uso de um malware, ainda há um agente envolvido, pois alguém precisou programar a IA ou o malware em questão.

De fato, o exemplo fornecido pelos dois autores demonstra muito bem o que Coeckelbergh afirma sobre o algoritmo tendo efeitos não desejados que até mesmo os desenvolvedores desconhecem. Afinal, de certo modo, a inteligência artificial acabou se rebelando. Isso não significa, no entanto, que a humanidade esteja presenciando os primeiros estágios de uma “era das máquinas”, onde a tecnologia atacará seres humanos e assumirá controle do planeta, uma vez que a IA que utiliza o malware não atacou alvos indiscriminados ou realizou ataques a esmo. Na verdade, considerando o exemplo fornecido por Burton e Soare, a IA pode ser considerada uma arma cuja trava de segurança parou de funcionar e ela começou a atirar em seus alvos. Seguindo este raciocínio, ainda assim ela é curiosamente mais segura do que uma arma, pois os projéteis, quando a arma dispara, variam de direção indiscriminadamente, e o malware, que seria o equivalente à munição balística, não erra o alvo e nem atinge vítimas inocentes.

Essa *weaponização* da IA se mostra de maneira ainda mais direta ao considerar o texto de Kamarck, que trata especificamente deste tipo de mau uso da tecnologia com motivações espúrias. Nesse caso, inclusive, as pessoas que utilizam deste recurso para manipular as eleições não se importam com as consequências morais de seus atos.

Comparando os três artigos aqui mencionados, o que se pode perceber é que a inteligência artificial, por si só, não representa nenhuma ameaça, no sentido em que ela não vai além do que foi programada para fazer, embora, segundo Burton e Soare, existam riscos consideráveis:

O processo de weaponização – seja na área nuclear ou na área de informação – compreende riscos consideráveis. Esses estão associados com a instabilidade que a proliferação das tecnologias, dentro da arena internacional, cria, o prospecto das corridas de armas e dilemas de segurança, o risco de que os atores não estatais adquiram agentes weaponizados, o risco de que os estados não sejam capazes de controlar eficientemente a tecnologia

weaponizada, e que as tecnologias de IA serão incontíveis e resultarão em consequências indesejáveis quando foram utilizadas. Os riscos associados com a weaponização da IA não foram delineados sistematicamente, mas incluem o desenvolvimento de parcialidade dentro dos sistemas de IA. Essa dinâmica foi demonstrada recentemente quando um robô de conversa da Microsoft chamado ‘Tay’ ganhou sua própria conta do Twitter e foi deixado para interagir com o público e, como resultado de receber dados malignos, começou a exibir racismo, sexismo, e pontos de vista políticos extremistas (...). Outro risco significativo dos sistemas de IA é que eles podem ser manipulados e sua integridade pode ser alterada por atores malignos, e até mesmo programados para realizar funções indesejadas.¹²

O risco associado com a IA é decorrente do seu uso, mais especificamente de seu mau uso. Quando ela recebe dados malignos, ou dados parciais, a inteligência artificial acaba por assumir uma personalidade distorcida, baseada na informação recebida, como no caso do robô de conversação que se tornou racista e preconceituoso após navegar pelo Twitter. O outro risco associado à IA é precisamente o que está sendo abordado no presente texto: ela pode ser manipulada e alterada por atores malignos.

Com base no que foi apresentado nesses três artigos, pode-se concluir, portanto, que a inteligência artificial, por si só, é incapaz de agir fora de sua programação. No entanto, ela pode ser, de certo modo, distorcida, desenvolvendo uma parcialidade após ser exposta a informações malignas, e até reprogramada por indivíduos que planejam *weaponizá-la*. Sendo assim, o problema não está em seu potencial, mas sim na maneira como ela é utilizada.

3. IA e a Moral

Embora Burton e Soare foquem seu artigo na weaponização da IA para fins militares e estratégicos, eles admitem que o “risco significativo dos sistemas de AI é que eles podem ser manipulados e sua integridade alterada por atores maliciosos e até programados para desempenhar funções não-intencionais”, e que a “IA também causou preocupações sobre a manipulação social”.

Burton e Soare não estão sozinhos nessas enunciações. Kamarck denuncia uma possível manipulação das eleições norte-americanas de 2016, que teria ocorrido por meio do uso da inteligência artificial. Por conta disso, a autora chega a considerar a IA como uma potencial arma contra a democracia.

¹² Burton e Soare, 2019, p. 8.

Não há dúvidas de que a IA pode ser *weaponizada* e causar muitos estragos. No Brasil, por exemplo, golpistas têm utilizado essa tecnologia para ludibriar outras pessoas e roubar seu dinheiro. O golpe em questão consiste em utilizar a inteligência artificial para imitar a voz de outras vítimas, fazendo-as dizer qualquer frase que queiram.

Nesta matéria, publicada pelo site R7, intitulada *Golpe da Voz*, há o relato de um influenciador digital que teve sua voz copiada por uma inteligência artificial. O golpista, então, usou a cópia da voz para extorquir o pai deste influenciador digital e fazê-lo depositar dinheiro em uma conta falsa. O advogado especialista em crimes cibernéticos Luiz Augusto D’Urso relata que uma IA capaz de manipular a voz alheia não é recente, mas os golpes com esta tecnologia são. Ele ressalta que, quanto mais tecnologias surgem, mais opções de golpes surgirão com ela.

Enquanto Burton e Soare consideram a inteligência artificial como uma ferramenta *weaponizável* e Kamarck a vê como uma arma em potencial, Mark Coeckelbergh enxerga a questão sob uma ótica um pouco diferente: para ele, o ser humano “*pode* ser responsável e deve ser responsabilizado pelo que ele faz e decide quando desenvolve e utiliza IA”. Ele mesmo admite, no entanto, que esta questão é bastante complexa.

Os três textos se completam, de certo modo. O primeiro considera as possibilidades estratégicas e os riscos da weaponização da inteligência artificial; o segundo, por sua vez, trata da atribuição de responsabilidade pelo uso da IA, um ponto que os outros autores não abordam detalhadamente em seus respectivos textos. Burton e Soare falam da susceptibilidade da inteligência artificial para usos maliciosos, enquanto Kamarck dá maior ênfase à forma como ela foi utilizada para atacar a democracia norte-americana.

É verdade que o texto de Coeckelberg é voltado para um aspecto específico da explicabilidade. Ele afirma que seu objetivo é “a questão da atribuição de responsabilidade para tecnologias usadas na automação de ações e decisões normalmente tomadas por humanos”. Porém, o próprio autor afirma que o que ele trabalha no artigo é aplicável “além desses domínios para todos os usos da IA”. Ocorre que Coeckelbergh utiliza a ética aristotélica, que o faz considerar a IA como um mero instrumento, assim como está sendo feito neste texto. Enquanto o pesquisador admite que isso é limitante para o escopo de seu texto, essa limitação não se aplica no caso deste artigo, porque o tipo de situação considerada é de um indivíduo que utiliza a inteligência artificial para prejudicar outras pessoas. O caso dos golpes financeiros serve de exemplo: a IA foi completamente controlada, sendo apenas um instrumento em um

golpe perpetrado contra uma vítima. Nesse exemplo, a atribuição de responsabilidade é bastante óbvia. Comparando o que está sendo tratado aqui com o texto de Coeckelbergh, portanto, a diferença fundamental é que ele discute a responsabilização em casos de acidente, não em casos em que há dolo envolvido, como os que foram apresentados neste artigo. Aqui, o que está em questão é precisamente um tipo de caso em que a inteligência artificial é *weaponizado* por agentes humanos. O que se busca é entender por que as pessoas estão ressignificando essa tecnologia como uma arma, fora de contextos militares.

Aplicando rigorosamente a definição de *weaponização*, a simples divulgação de uma notícia falsa é, por si só, uma *weaponização*: o usuário está recondicionando o ambiente virtual como uma arma a ser utilizada contra o alvo da desinformação que está divulgando online. Então, mais do que a possibilidade de ser uma ameaça por conta de suas capacidades tecnológicas, a inteligência artificial acaba sendo utilizada como uma ferramenta conveniente para que o indivíduo ataque outras pessoas, sendo mais uma dentre as tecnologias que figuram na lista de opções de golpes virtuais conhecidos como *phishing*.¹³

A notícia, postada no site GaúchaZH sob o título *Inteligência Artificial Exige Atenção Redobrada de Internautas contra Golpes na Internet*, cita uma declaração do coordenador do curso de segurança da informação da Unisinos, Luciano Ignaczak, afirmando que o *phishing*, anos atrás, era mais artesanal, isto é, era redigido manualmente pelos golpistas, e a atenção à escrita era o bastante para perceber que uma mensagem virtual se tratava de um golpe. Ele nota que a IA não está apenas criando novas formas de enganar os outros, mas também sofisticando golpes mais antigos. A notícia infere, assim, que o *phishing* está evoluindo junto com a tecnologia.

Na mesma matéria, a professora da Escola Politécnica da PUC do Rio Grande do sul, Ana Benso, aponta a necessidade de criação de artifícios para saber se algo considerado verídico realmente é verdadeiro ou apenas produto de uma montagem realizada por inteligência artificial, e ressalta que é necessária a existência de um código de conduta para o ambiente virtual da internet.

Chama atenção como as duas matérias tratando do mesmo golpe evocam especialistas que oferecem conselhos diferentes, mas diagnósticos iguais: D'Urso sugere a criação de palavras-chave entre familiares e videochamadas para conferência da veracidade da mensagem

¹³ Fonte: <https://prodest.es.gov.br/entenda-o-que-e-phishing-e-adote-medidas-para-evita-lo>. Último acesso em 30/03/2024.

de áudio, enquanto Paulo Ricardo Muniz Barros, professor da Feevale, aconselha apenas uma simples ligação entre os familiares. Porém, todos eles são unânimes: mais tecnologias significam mais golpes virtuais.

Soare refere-se às implicações estratégicas de weaponizar a inteligência artificial, ou seja, ressignificar esta tecnologia com o objetivo de guerrear, em sentido amplo. Quanto mais tecnologias as pessoas têm à sua disposição, mais o ser humano estará sujeito a perigos, principalmente no ambiente virtual, onde a IA torna difícil de perceber a linha entre verdade e mentira.

Não é objetivo deste artigo, no entanto, generalizar ou partir do pressuposto de que o ser humano é necessariamente maligno e que obrigatoriamente buscará novas formas de atacar seus semelhantes. Mas é necessário refletir sobre para onde está indo a moral de um indivíduo capaz de aprender sobre e obter uma tecnologia avançada para deliberadamente *weaponizá-la* contra uma vítima.

Independentemente da ética considerada, a discussão enveredará por um caminho que se torna complexo demais para ser estendido aqui. O que se pode notar preliminarmente é que, se por um lado há extensa consideração a respeito dos riscos e potenciais ameaças relacionadas à inteligência artificial, é igualmente prioritário que se considere como as novas tecnologias estão afetando a moral, para o bem e para o mal, principalmente no aspecto das relações interpessoais. Sendo assim, em vez de fornecer uma argumentação a título de conclusão, posto que não há a intenção de se assumir uma posição específica em relação a isso, deixa-se, aqui, um convite para a reflexão: será que a moral está avançando junto com a tecnologia ou, ao contrário, está retrocedendo?

Conclusão

Coeckelbergh admite que utilizar a ética aristotélica é limitante, pois o obriga a considerar a inteligência artificial como mero instrumento. Porém, enquanto o autor trabalha com casos mais complexos, como veículos terrestres e aéreos autônomos, controlados por IA, os casos mostrados aqui consistem exatamente em situações nas quais ela se encaixa exatamente como mero instrumento, sendo *weaponizado* contra outrem. A abordagem aristotélica, no entanto, tem razão de ser quando é observada conforme a perspectiva do usuário que tem intenção de *weaponizar* a ferramenta.

Kamarck, ao tratar da inteligência artificial em seu texto, assume uma postura mais acusativa, até mesmo incisiva, considerando que, no campo político, essa tecnologia possui potencial para causar mais estragos do que trazer benefícios. Embora apresente relatos de maneira tão intensa, Kamarck não pode ter seus argumentos ignorados, uma vez que a IA pode ser manipulada e possui forte potencial *weaponizável*.

Burton e Soare também tratam brevemente sobre a capacidade de se manipular uma inteligência artificial. Porém, os dois autores não entram em muitos detalhes, preferindo apenas relatar o incidente das eleições norte-americanas superficialmente. No entanto, ao contrário de Kamarck, Burton e Soare não consideram a IA como uma ameaça em potencial. Eles não a problematizam diretamente.

Em realidade, os três artigos podem ser interpretados de maneira interligada, por causa de seus pontos comuns: Burton e Soare chamam atenção para o fato de que a inteligência artificial pode ser manipulada por atores maliciosos, Kamarck afirma que esse incidente já aconteceu em 2016 e pode acontecer novamente, e Coeckelbergh pergunta se os desenvolvedores da IA também deveriam ser responsabilizados por permitir, indiretamente, esse tipo de uso.

Por fim, reitera-se, aqui, o convite à reflexão. É bastante importante que se considere o quanto o avanço tecnológico está permitindo, ou estimulando, que as pessoas se deixem corromper. Talvez a questão seja ainda mais complexa, e seja necessário entender a origem deste comportamento. Por outro lado, talvez a questão seja mais simples do que aparenta, e o ser humano precisa apenas compreender que a ética e a moral também se estendem à tecnologia.

Referências

ARISTÓTELES. (1984). *Nicomachean ethics*. In BARNES, J. (Ed.), The complete works of Aristotle (Vol. 2, pp. 1729–1867). Princeton: Princeton University Press.

BURTON, J.; SOARE, S. R. *Understanding the Strategic Implications of the Weaponization of Artificial Intelligence*. In: h International Conference on Cyber Conflict: Silent Battle. T. Minárik, S. Alatalu, S. Biondi, M. Signoretti, I. Tolga, G. Visky. 2019: NATO CCD COE Publications, Tallinn. Disponível online em https://ccdcoe.org/uploads/2019/06/Art_14_Understanding-the-Strategic-Implications.pdf. Último acesso em 05/08/2023.

COECKELBERGH, M. *Artificial Intelligence, Responsibility Attribution, and a Relational Justification of Explainability*. Disponível online em <https://link.springer.com/article/10.1007/s11948-019-00146-8>. Último acesso em 05/08/2023.

GAÚCHAZH. <https://gauchazh.clicrbs.com.br/tecnologia/noticia/2023/05/inteligencia-artificial-exige-atencao-redobrada-de-internautas-contragolpes-na-internet-cli9987yo009j0165zwr5ecp0.html>. Último acesso em 05/08/2023.

KAMARCK, E. *Malevolent soft power, AI, and the threat to democracy*. Disponível online em <https://www.brookings.edu/articles/malevolent-soft-power-ai-and-the-threat-to-democracy/>. Último acesso em 05/08/2023.

OXFORD. <https://www.oxfordlearnersdictionaries.com/definition/english/weaponization>. Último acesso em 05/08/2023.

PORTAL R7. <https://noticias.r7.com/sao-paulo/golpe-da-voz-criminosos-usam-inteligencia-artificial-para-enganar-vitimas-e-roubar-dinheiro-20052023>. Último acesso em 05/08/2023.

PRODEST – Instituto de Tecnologia da Informação e Comunicação do Estado do Espírito Santo. <https://prodest.es.gov.br/entenda-o-que-e-phishing-e-adote-medidas-para-evita-lo>. Último acesso em 30/03/2024.

SCHOPENHAUER, A. *The World as Will and Representation*. Translation E. J. F. Payne. New York, Oxford University Press, 2012.

.