

Ética Algorítmica: questões e desafios éticos do avanço tecnológico da sociedade da informação

Regina Rossetti¹

<https://orcid.org/0000-0002-3219-4570>

Alan Angeluci¹

<https://orcid.org/0000-0002-4093-0590>

¹ - Universidade de São Caetano do Sul.
São Caetano do Sul. (SP), Brasil.

Resumo: A Quarta Revolução Industrial inseriu na vida cotidiana a Inteligência Artificial (IA) e seus algoritmos. Essa nova realidade tecnológica requer uma profunda reflexão ética e enseja a necessidade de pensar sobre uma Ética Algorítmica. Partindo de uma pesquisa bibliográfica, este artigo propõe uma definição do termo e desenvolve uma reflexão sobre as principais questões suscitadas pela presença dos algoritmos nas mediações da vida humana contemporânea, referentes à falibilidade, opacidade, viés, discriminação, autonomia, privacidade e responsabilidade. Situando-a no domínio de uma ética aplicada, pondera-se sobre a conjunção das camadas humanas e computacionais na tomada de consciência diante dos dilemas.

Palavras-chave: ética algorítmica; comunicação; tecnologia; inteligência artificial.

Abstract: Algorithmic Ethic: ethical issues and challenges of the technological advancement of the information society - The Fourth Industrial Revolution has put Artificial Intelligence (AI) and its algorithms into everyday life. This new technological reality requires a deep ethical reflection and gives rise to the need of thinking about an Algorithmic Ethic. From a bibliographic research, this article proposes a definition of the ethics of the algorithms, presenting a reflection on the main ethical issues raised by the presence of the

algorithms in the mediations of contemporary human life regarding fallibility, opacity, bias, discrimination, autonomy, privacy and responsibility. Situating it in the domain of an applied ethic, one ponders about the conjunction of the human and computational layers in raising awareness in the face of dilemmas.

Keywords: algorithmic ethic; communication; technology; artificial intelligence.

Introdução

A Quarta Revolução Industrial levou a relação entre homem e máquina para um novo patamar de inovação em direção à Inteligência Artificial (IA) e seus algoritmos. Observa-se um acoplamento íntimo entre o *cyber* e o físico que se manifesta a partir da conexão entre o mundo cibernético da computação e as comunicações com o mundo físico - os propalados sistemas ciberfísicos (RAJKUMAR, 2010). Se, por um lado, essas aplicações têm enorme efeito social e benefícios econômicos a partir da otimização de recursos temporais e espaciais, por outro lado, suscitam questões éticas sobre as quais a humanidade ainda carece de reflexões, que emergem de maneira urgente na medida em que os actantes humanos e não-humanos entram em rotas, por ora, divergentes.

Um mundo ciberfísico pressupõe que os objetos de interesse humano e maquínico tornam-se disponíveis como dados digitais em redes, armazenados em tempo real. Desta maneira, tornam-se rastreáveis, exploráveis e analisáveis na rede (BRUNO, 2013). Isto leva a uma explosão de um *Big Data* potencialmente disponível e acessível a partir de qualquer lugar. Nesse deslocamento, instrumentos de controle precisam ser criados para proteger o *know-how* das empresas, em relação aos seus dados de produção e serviços, de forma a garantir a estabilidade do sistema e a privacidade dos dados. Adicionalmente, as indústrias também precisam prevenir o acesso não autorizado aos sistemas de produção para que não haja danos e prejuízos ambientais, econômicos e sociais aos humanos, assegurando a cibersegurança (DRATH; HORCH, 2014; HERMANN; PENTEK; OTTO, 2016). Nesse sentido, a real revolução da Indústria 4.0 não se centra somente nas implementações técnicas e talvez ainda esteja por ser feita: estariam os sistemas ciberfísicos plenamente habilitados para lidar com os desafios atuais e futuros da privacidade e da segurança dos dados? A história recente tem evidenciado que ainda há um caminho a percorrer no campo da ética.

O problema ético da privacidade dos dados quando aliada ao desafio da cibersegurança é um ponto nevrálgico da Quarta Revolução Industrial, que se

sustenta tão fundamentalmente no *Big Data*. O algoritmo se revela como um elemento central nessa problemática em tempos de extrativismo de dados (MOROZOV, 2018). Esses conjuntos de regras estão no cerne das programações dos sistemas digitais e de IA, pois aperfeiçoam técnicas, automatizam e agilizam ações e tornam os processos mais inteligentes.

Neste contexto dos avanços tecnológicos da atual sociedade da informação, o presente artigo debruça-se reflexivamente sobre as principais questões éticas suscitadas pela presença dos algoritmos nas mediações da vida humana. Primeiro, define-se algoritmo e ética algorítmica e, na sequência, são discutidas sete questões éticas fundamentais provocadas pela presença de algoritmos na rotina da humanidade contemporânea, ilustradas por casos recentes envolvendo algoritmos, pedindo uma reflexão mais acurada. Parte-se do mapeamento feito por Mittelstadt *et al.* (2016) que revela a existência de pelo menos sete tipos de problemas éticos considerados fundamentais quando se analisa a criação, o funcionamento e o uso dos algoritmos na sociedade atual fortemente midiaticizada: falibilidade, opacidade, viés, discriminação, autonomia, privacidade e responsabilidade. Os resultados indicam que essas questões éticas implicam em cuidados políticos, sociais e de governança dos algoritmos, que devem ser tratados efetivamente em *by design*, isto é, no cerne da própria criação dos algoritmos, para que princípios éticos já estejam presentes desde a concepção destes.

Ética algorítmica

O algoritmo é uma construção matemática antiga, usada para processar dados que produzem evidências e motivam ações específicas em direção à solução de determinado problema. Sua concepção ligada aos sistemas computacionais foi consolidada com estudos, como o da Máquina de Turing, fundantes da ciência da computação. Uma receita de um bolo, por exemplo, é um clássico exemplo metafórico de aplicação algorítmica: a correta execução de seus componentes em determinada sequência e medidas (uso dos ingredientes farinha, ovos e leite e seu modo de preparo passo-a-passo) conduzem ao objetivo final (o bolo pronto).

Como se pode supor, um algoritmo não é uma mera sequência de passos necessários para a realização de determinada tarefa. Sua implementação envolve diversos atores, autômatos ou humanos, tomadas de decisão por lógica e repetição de ações. Entendendo que atualmente o conceito

de algoritmo está intimamente ligado à construção da IA, alguns estudos (TADDEO; FLORIDI, 2018; MCCARTHY *et al.*, 2006; SAMUEL, 1960; YANG *et al.*, 2018) indicam que os algoritmos são algo além do que um pacote de modelos computacionais que dão base a uma arquitetura tecnológica. Os algoritmos aplicados à IA seriam, então,

um recurso crescente de agência de autoaprendizagem, interativa, autônoma, que permite que artefatos computacionais executem tarefas que, de outra forma, exigiriam que a inteligência humana fosse executada com sucesso (TADDEO; FLORIDI, 2018, p. 751. Tradução Nossa).

Usados em uma variedade de atividades que perpassam a vida humana, aos algoritmos são atribuídas uma série de demandas que beneficiam a sociedade da informação. Eles conseguem reduzir custos e riscos, além de trazer novas soluções aos problemas complexos; podem, no entanto, conduzir a consequências nocivas e inesperadas, sobretudo quando em seu processo de execução prescinde da nossa supervisão.

Dentro da evolução, surgem com a IA os algoritmos de aprendizagem. Quando se trata deste modelo, a definição e programação pode ter sido feita pelos próprios algoritmos. O aprendizado da máquina é definido como a capacidade de definir ou modificar regras de tomada de decisão de forma autônoma. As capacidades de aprendizado conferem aos algoritmos algum grau de autonomia. O impacto dessa autonomia deve permanecer incerto, mas até certo ponto. Como resultado, as tarefas executadas pelo aprendizado de máquina são difíceis de prever antecipadamente (como uma nova entrada será tratada) ou explicadas posteriormente (como uma determinada decisão foi tomada). A incerteza pode, assim, inibir a identificação e a correção de desafios éticos no projeto e operação de algoritmos de aprendizagem, que são capazes de tomar decisões. Então a questão fundamental que surge é: essa tomada de decisão algorítmica, por ser autônoma, é fundada em quais valores éticos?

No mundo contemporâneo, com o surgimento das máquinas e dos robôs surge a questão da possibilidade, e até da necessidade, de estender a ética aos seres semiautônomos. Uma das primeiras questões éticas a ser enfrentada diz respeito ao princípio da autonomia: até que ponto uma máquina é autônoma para tomar decisões? Outro problema, dentre os vários desafios éticos que o uso do algoritmo levanta, está na questão da identificação

da responsabilidade pelas ações induzidas pelos algoritmos, visto que se trata um ser autômato, isto é, uma máquina. Pode um robô ser responsabilizado pelos danos causados? Por um lado, Silveira (2020, p. 83) busca iluminar esta questão em seus estudos propondo a formulação da “pessoa eletrônica” — uma nova figura jurídica que pode colaborar com “a reivindicação da inescrutabilidade e da impossibilidade de interpretar as operações realizadas pelos algoritmos de aprendizado profundo”.

Em consonância com a tradição do pensamento filosófico que considera que a ética é exclusiva do humano, e que, conseqüentemente, as coisas em si não são nem boas nem más, Bill Gates afirmou que “a tecnologia é amoral” (IHU ONLINE, 2017), e nesse sentido, não seria possível falar de uma ética das coisas ou das máquinas. Essa é a tese da ética clássica que estabeleceu os limites deste conhecimento: ele é sempre humano. Nesse sentido, Lima Vaz (1999), define a ética como uma reflexão sobre a realidade histórico-social do conjunto dos costumes tradicionais de uma sociedade e sua presença no comportamento dos indivíduos considerados como valores e obrigações para sua conduta.

Todavia, os desafios da IA reabrem a questão da exclusiva humanidade da ética. Seria ela extensiva aos artefatos semiautônomos criados pelos homens? Ampliando a abrangência da ética e aplicando o conceito de algoritmo a este conceito, haveria uma ética algorítmica que pode ser definida como uma reflexão sobre a nova realidade histórico-social tecnológica, mediada por algoritmos, em que novos costumes e hábitos presentes na sociedade tecnológica, no comportamento dos indivíduos e na conduta das máquinas seriam reconsiderados em seus valores e obrigações. Nesse sentido, é possível falar de uma ética algorítmica como ética aplicada.

Existem três domínios da reflexão ética: primeiro, a ética geral ou ética fundamental, cujo objeto diz respeito aos fundamentos do saber ético; segundo, a ética aplicada aos vários ramos da atividade humana, é nesse sentido que se fala de ética ambiental, como uma reflexão sobre as relações éticas que o humano estabelece com o meio ambiente; e terceiro, a metaética que trata da linguagem ética e sua lógica. Este estudo aborda o domínio da ética aplicada ao ramo da atividade tecnológica e o objeto ao qual são aplicadas as investigações e reflexões sobre este conceito é o algoritmo. Nessa direção, fala-se aqui em uma ética algorítmica, não como uma criada pelos algoritmos, do mesmo modo que o meio ambiente não cria a ética ambiental, mas

como uma reflexão humana sobre algoritmos. Todavia, estar-se-ia diante de um novo fenômeno, porque os algoritmos, notadamente os de aprendizagem, por imitarem a inteligência humana, são capazes de tomadas de decisão, o que suscita uma questão ontológica sobre a natureza autônoma dos algoritmos, fato que o aproximaria cada vez mais do humano. Essa nova realidade também levanta inúmeras questões éticas relativas à transparência dos dados e do processo, a responsabilidade pelos danos causados, os efeitos discriminadores, a privacidade informacional, a autonomia das máquinas e dos usuários e a fabricação de resultados tendenciosos. Ainda, a ética possui papel central na garantia de que os algoritmos possam exercer seu potencial mitigando seus riscos.

Sete problemas éticos algorítmicos

No artigo intitulado *"The ethics of algorithms: mapping the debate"* (MITTELSTADT *et al.*, 2016) os autores constroem um mapa conceitual que contribui para o debate atual sobre uma possível ética dos algoritmos. Mittelstadt e seus colegas do *Oxford Internet Institute* apresentam as preocupações epistêmicas e normativas sobre o *modus operandi* dos algoritmos. Este mapeamento é concebido como uma estrutura de organização baseada em como algoritmos operam. Na figura 1, estão representados os seis tipos de preocupações éticas levantadas.

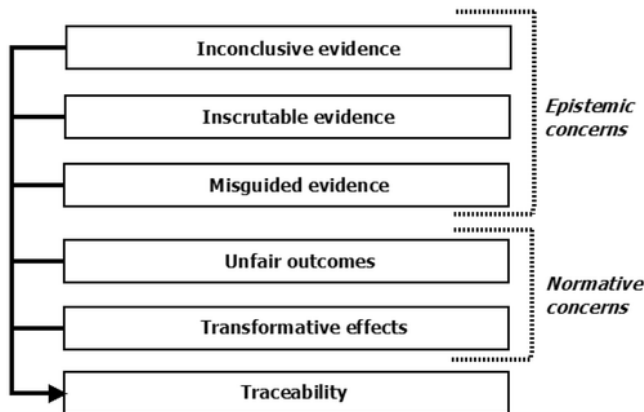


Figura 1. Mapa conceitual. Fonte: MITTELSATAD *et al.*, 2016, p. 4

No esquema dos autores é possível observar três conjuntos de preocupações éticas. O primeiro conjunto epistêmico se refere às características internas

do processo algorítmico como riscos presentes em seu desenho, composto por evidências inconclusivas, evidências inextricáveis e evidências errôneas mal direcionadas. O segundo conjunto normativo, composto por resultados injustos e efeitos transformativos, resulta das articulações entre os processos propriamente algorítmicos e seu acionamento social, ou seja, estes problemas normativos não estão embutidos nos algoritmos, mas sim na relação do usuário com o mesmo. O terceiro conjunto deriva dos outros dois conjuntos porque todos podem gerar problemas de rastreabilidade, tanto que as setas que partem das cinco preocupações vão em direção ao terceiro conjunto. Os autores partem do pressuposto de que: a) os algoritmos são usados para transformar dados em evidências para um determinado resultado; b) o resultado é usado para desencadear uma ação que pode não ser eticamente neutra; c) o processamento dos algoritmos é realizado de forma complexa e semiautônoma, o que torna difícil rastrear a responsabilidade pelos efeitos de ações impulsionadas pelo algoritmo. Portanto, a questão final e mais abrangente das implicações éticas dos algoritmos diz respeito à responsabilidade, ou seja, quem deve ser responsabilizado pelas falhas dos algoritmos? Deste modo, os autores indicam e organizam seis espécies de preocupações ou questões éticas acerca dos algoritmos.

Esse mapa conceitual em si não é diretamente direcionado como uma ferramenta para resolver dilemas éticos resultantes de ações problemáticas decorrentes de algoritmos, mas serve como estrutura para futuras discussões sobre as consequências éticas desse modo de funcionalidade dos algoritmos, como as que são desenvolvidas no presente artigo. Assim, é com base nesse mapa conceitual que passamos agora a apresentar sete questões éticas diretamente derivadas das seis preocupações apontadas no mapa conceitual dos autores do *Oxford Internet Institute*, conforme quadro 1.

Preocupações éticas levantadas por algoritmos	Questões Éticas discutidas neste artigo
Evidências inconclusivas	Falibilidade
Evidências inextricáveis	Opacidade
Evidências mal direcionadas	Viés
Resultados injustos	Discriminação
Efeitos transformativos	Autonomia
Efeitos transformativos	Privacidade de informações
Rastreabilidade	Responsabilidade

Quadro 1. Quadro de preocupações éticas. Fonte: o autor.

A relação estrita entre as preocupações do mapa conceitual e suas possíveis consequências éticas discutidas neste artigo (falibilidade, opacidade, viés, discriminação, autonomia, privacidade e responsabilidade) é que dá a dimensão e a estrutura exata dos sete riscos que passamos a apresentar.

A questão ética da falibilidade do algoritmo

O primeiro problema ético identificado, a falibilidade, diz respeito a evidências inconclusivas que levam às ações injustificadas. O algoritmo correlaciona os dados minerados de modo causal. Entretanto, nem todas as correlações efetuadas entre causa e efeito são corretas. As correlações estão sujeitas, então, a falibilidade de suas conclusões. A busca por ligações causais nem sempre é fácil por se tratar de uma quantidade enorme de dados processados. O problema se revela mais grave quando observados os dados do Instituto de Pesquisa Econômica Aplicada (IPEA) em 2016: “robôs investidores programados para reagir instantaneamente ante determinadas situações são responsáveis por mais de 40% das decisões de compra e venda no mercado de ações no país” (PIERRO, 2018, p.1). Foram atribuídos aos algoritmos tomadas de decisões que têm impacto direto em vidas humanas.

Empresas que levantam informações sobre o comportamento de consumidores, de opinião pública para venda de produtos ou campanhas políticas baseiam suas tomadas de decisão em análise de dados e precisam estar atentas à falibilidade. Muito embora algoritmos sejam programados para terem o apoio humano nas tomadas de decisão, é cada vez mais crescente o uso de técnicas de aprendizagem de máquina (*machine learning*) e a confiança irrestrita nesse método de previsão que dá certa autonomia ao algoritmo. Mittelstadt et al. (2016) destaca o uso de *insights* acionáveis, como realização complementar de *surveys* ou coleta de mensagens em redes sociais (*scraping*) associados à análise de *Big Data*. Eles podem ser mais eficientes e menos sujeitos a erros já que são métodos de obtenção direta de informações do usuário e não relações causais obtidas de dados brutos.

A falibilidade surge do processo de mineração de dados que ocorre de modo indutivo ao correlacionar os dados minerados. Os algoritmos são criados para relacioná-los de forma causal. Entretanto, nem sempre essa causalidade é legítima. A lógica programada pode associar um efeito a uma causa indevida, gerando conclusões injustas e inconclusivas. Em função da quantidade enorme de dados processados, as correlações funcionam bem para

populações, mas quando se trata de indivíduos, as transposições das relações causais encontradas podem gerar resultados espúrios e falaciosos.

Um exemplo de falibilidade do algoritmo pode ser ilustrado através de um uso compartilhado de mesmo perfil no *Netflix*: se duas pessoas passarem a assistir seus conteúdos prediletos usando o mesmo perfil em determinada conta, o algoritmo do sistema de recomendação passará a considerar todos os conteúdos assistidos como condizentes com o perfil de um único usuário imaginado. Certamente, não funcionará de forma adequada para cada pessoa individualmente, podendo gerar recomendações inverídicas ou, no mínimo, imprecisas.

A questão ética da opacidade do algoritmo

A questão ética da opacidade é aporética, pois surge da própria natureza complexa do algoritmo. O itinerário da criação de um algoritmo quase nunca está disponível ao usuário ou ao pesquisador. E mesmo que estivesse, a linguagem computacional é compreensível somente para os iniciados. Assim, compreender os processos que levaram um algoritmo a determinada conclusão não é acessível e transparente. Resta, portanto, ao usuário o que Bucher (2017) chama de imaginário algorítmico, ou seja, não só aquilo que ele sente e que interfere em seu estado de espírito, mas que também tensiona o processo interativo do aprendizado da máquina — tensão, aliás, já preconizada há duas décadas quando Manovich (2001) se referiu à transcodificação cultural que ocorre entre as camadas humanas e computacionais nos sistemas digitais.

O principal argumento dos programadores para a manutenção da opacidade dos dados e dos processos está fundado no princípio da privacidade. Eles consideram que os dados são privados e, portanto, não podem se tornar públicos. Eles alegam questões de segurança, garantindo que esses algoritmos não sejam usados de forma ilícita. E buscam também proteger seu trabalho da apropriação indevida pela concorrência. É sob essa ótica que operam as grandes empresas baseadas em serviços de Internet: os dados que a *Uber*, por exemplo, dispõe sobre trânsito e comportamento de usuários foram coletados e armazenados ao longo dos últimos anos e tratam de um rico ativo da empresa.

A transparência algorítmica é importante para que se tenha previsibilidade, controle, monitoramento e correção dos processos que envolvem o acesso

e compreensão das informações. Entretanto, os proprietários dos algoritmos mantêm segredo de seu funcionamento, seja por questões de competitividade, de segurança nacional e de privacidade informacional. Quando a relação entre os dados e seus resultados não é acessível ou inteligível para escrutínio, a opacidade conduz a uma espécie de algoritmo “caixa-preta”. Tem-se aqui um conflito ético entre o princípio da transparência e os princípios da privacidade e da autonomia.

A opacidade do algoritmo do aplicativo *Whatsapp* foi a principal razão de diversos bloqueios temporários ao seu serviço no território brasileiro em 2015. Na época, o Poder Judiciário não conseguia obter do *Facebook*, proprietário do aplicativo, dados e informações trocados entre criminosos e que poderiam auxiliar na resolução de crimes (ANGELUCI, 2019). O *Facebook* justificava que os dados eram criptografados e não eram armazenados pela empresa.

A questão ética do viés do algoritmo

O problema do viés está relacionado às decisões algorítmicas tendenciosas. Essas tendências são morais, pois envolvem a preferência por determinados valores. Porém, a questão que surge é se máquinas são capazes de valorar a realidade. Em certo sentido são, pois elas foram programadas por seres humanos, que são necessária e intrinsecamente avaliadores da realidade. Silva nos lembra em seus estudos que “problemas destes agentes artificiais que tomam decisões de visibilidade, acesso, classificação e processamento de materiais digitais também são frequentes, muitas vezes ligados a vieses de raça, gênero, classe, localidade, neuroatipicidade e outros” (SILVA, 2020, p. 431). É própria do ser humano a atribuição de valor a realidade e, nessa direção, não há como ser neutro. Assim, mesmo que supostamente um algoritmo não tenha preconceito, ele acaba refletindo os valores e escolhas de seus criadores.

Em 2015, um usuário descobriu que o *Google* possuía um algoritmo “racista” (SALAS, 2018): a IA do *Google Fotos* etiquetava todos os seus amigos negros como gorilas, simplesmente, por não conseguir distinguir a diferença da pele dos seres humanos e a de macacos. O erro foi corrigido, mas demonstrou como os algoritmos refletem o viés e o preconceito dos seres humanos que estão presentes nas bases de dados utilizadas pelos programadores — que também acabam espelhando seus preconceitos nos códigos que desenvolvem.

Não há neutralidade, pois os valores sociais são pré-existent e circundam os projetistas. Além disso, existem restrições técnicas e o contexto de uso do algoritmo que condicionam e circunscrevem o algoritmo. Assim, o algoritmo está sujeito a um viés social e cultural ao ser criado. Esse é o viés ético de seu criador, que acaba transferido para o algoritmo por meio das escolhas que o programador faz ao criar um algoritmo. Portanto, um algoritmo não é neutro, mas reflete os valores éticos de seus criadores humanos.

A questão ética da discriminação do algoritmo

A discriminação causada por resultados injustos é o quarto problema ético. O algoritmo minera dados, infere padrões e aplica os resultados às pessoas cujos perfis combinam com esse padrão. O problema surge quando os algoritmos discriminam perfis em função de atributos sensíveis como etnia, gênero e preferência sexual.

Algoritmos podem ser preconceituosos e discriminatórios ao aplicarem os resultados da mineração de dados para perfis de indivíduos que ele inferiu serem aderentes ao resultado. Um caso célebre é o do *software* COMPAS, utilizado pela Justiça dos EUA para análises de reincidência criminal: ele gerava resultados equivocados sobre o número de crimes cometidos por pessoas negras em comparação com brancas acusando o dobro da probabilidade (VICENTE, 2018).

Outro exemplo é o da *@TayandYou*, um perfil robô que emulava uma adolescente em redes sociais e que se tratava de um experimento de IA da *Microsoft* (MOREIRA, 2016). Ele teve curadoria de profissionais da empresa, mas se tornava cada vez mais “inteligente” à medida que interagiu com usuários reais nas redes sociais. O resultado: a robô passou a ser uma ferrenha seguidora do nazismo e a postar mensagens que reproduziam discursos racistas, homofóbicos e extremamente conservadores, que aprendeu com pessoas reais nas redes. O experimento foi encerrado.

A questão ética da autonomia nos algoritmos de tomada de decisão

O desafio ético da autonomia está ligado à questão da liberdade. Algoritmos são criados para influenciar o comportamento do usuário e, para tanto, filtram as informações que são oferecidas ao ser humano no intuito de adequá-las ao perfil do destinatário. O problema surge quando a escolha

reflete os interesses de terceiros e não os do indivíduo cuja autonomia é desrespeitada. A autonomia decisória do sujeito é corrompida quando ele é inserido em uma “câmara de eco”, que restringe as informações àquelas filtradas pelos algoritmos como sendo as preferidas pelo usuário.

Em entrevista concedida ao jornal *El País*, o filósofo sul-coreano *Byung-Chul Han*, um proeminente analista da sociedade da atual, pergunta se o algoritmo fará o homem e fala sobre o poder que o algoritmo possui para retirar a autonomia do homem contemporâneo.

Han – Os macrodados tornam supérfluo o pensamento porque se tudo é quantificável, tudo é igual... Estamos em pleno dataísmo: o homem não é mais soberano de si mesmo, mas resultado de uma operação algorítmica que o domina sem que ele perceba; vemos isso na China com a concessão de vistos segundo os dados geridos pelo Estado ou na técnica do reconhecimento facial. (EL PAÍS, 07/02/2018).

A pergunta que a reflexão suscita é: quais princípios éticos são desrespeitados em função da coerção exercida pela força invisível dos algoritmos na tomada de decisões dos sujeitos? Nesse sentido, trata-se de apontar, por exemplo, as implicações éticas dos algoritmos do *Facebook* na tomada de decisões de seus usuários. A hipótese é a de que o princípio ético da liberdade é infringido.

Os algoritmos mediam cada vez mais a vida digital e a tomada de decisões. Trata-se aqui de esclarecer a importância ética dessa mediação. Tomar uma decisão exige um sujeito autônomo e livre capaz de escolher sua ação. Contudo, na sociedade de informação, midiaticizada e informatizada, muitas decisões são tomadas por algoritmos. Os exemplos são inúmeros.

Algoritmos de tomada de decisão são usados em uma variedade de domínios, de modelos simplistas de tomada de decisão (LEVENSON; PETTREY, 1994) a algoritmos de perfil complexos (HILDEBRANDT, 2008). Exemplos contemporâneos notáveis incluem agentes de software online usados por provedores de serviços online para realizar operações em nome dos usuários (KIM *et al.*, 2014); algoritmos de resolução de disputas online que substituem os tomadores de decisão humanos na mediação de disputas (RAYMOND, 2014; SHACKELFORD; RAYMOND, 2014); sistemas de recomendação e filtragem que comparam e agrupam usuários para fornecer conteúdo personalizado (BARNET, 2009); sistemas de apoio

à decisão clínica (SADC) que recomendam diagnósticos e tratamentos aos médicos (DIAMOND *et al.*, 1987; MAZOUÉ, 1990); e sistemas de policiamento preditivo que prevêem pontos críticos de atividades criminosas. (MITTELSTADT *et al.*, 2016, p. 3).

Desse modo, existe uma classe particular de algoritmos que tomam decisões. Esses algoritmos atuam em situações que envolvem a análise de uma quantidade enorme de dados: é o caso dos algoritmos do *Google Analytics* e *Facebook Insights*, usados fortemente por diversas agências de *marketing* digital.

Entretanto, os algoritmos não são neutros, mas carregados de valores. Esses valores foram atribuídos pelos seus criadores e reforçados pelos seus usuários. As regras e os pesos utilizados pelos algoritmos foram definidos e programados manualmente pelos programadores.

A questão ética da privacidade no algoritmo

Um problema ético bem atual é o da privacidade informacional. Fuchs *et al.* (2012) destaca que nos últimos anos a Internet se transformou de um sistema primariamente orientado à provisão de informações para comunicação e formação de comunidades. A privacidade informacional diz respeito ao direito dos titulares de informação de protegerem seus dados pessoais de terceiros e de controlar o seu acesso e uso. Por outro lado, os algoritmos são construídos para garimpar dados no *Big Data*, assim, o esforço algorítmico é no sentido de romper as barreiras de proteção dos dados individuais para a construção do perfil do grupo no qual, supostamente, a identidade informacional do indivíduo é preservada. Com a emergência de noções como *web 2.0* e redes sociais,

está claro que o princípio que está na base dessas plataformas é o fornecimento e armazenamento massivo de dados pessoais que são sistematicamente avaliados, vendidos e usados para direcionar usuários com propaganda. A vigilância é, portanto, um tópico importante. (FUCHS *et al.*, 2012, p. 3. Tradução nossa).

O escândalo envolvendo o *Facebook* e a empresa de comunicação estratégica *Cambridge Analytica* é um notável exemplo da discussão sobre vigilância, privacidade e algoritmos. Com o vazamento de dados de usuários da rede social, obtidos por meio de algoritmos da *Cambridge Analytica*, foi possível obter dados sigilosos do comportamento de 50 milhões de usuários do *Facebook*,

utilizados para orientar campanhas nas redes sociais a favor do *Brexit* no Reino Unido e da candidatura de Donald Trump à presidência dos EUA.

Já durante a pandemia do novo coronavírus, o governo do Estado de São Paulo viu-se envolvido em uma polêmica sobre o monitoramento da geolocalização de celulares da população com o objetivo de aferir o cumprimento da quarentena. Por mais que governo e operadoras tenham alegado que os dados pessoais dos usuários não são passíveis de serem identificados nesse sistema, entidades civis alegam que esse tipo de monitoramento deveria ter anuência prévia dos usuários para o envio da informação. “Uma vigilância e um rastreamento de dados das pessoas nunca visto antes”, afirma Gabriela Zanfir-Fortuna, da *ONG Future of Privacy Forum*, sobre as consequências do combate à Covid-19 (PASSOS, 2020).

Há, portanto, um confronto entre a proteção dos dados privados e a necessidade essencial dos algoritmos de acessar esses dados. Trata-se de um dilema ético entre a preservação da privacidade do titular dos dados e a extração automatizada de conteúdo efetuada pelos algoritmos.

A questão ética da responsabilidade do algoritmo

E, por fim, a mais abrangente de todas as questões éticas aqui levantadas, a questão da responsabilidade, trazida pelo difícil rastreamento do responsável no processamento algorítmico. Em 2018, um carro autônomo da *Uber*, durante testes, atropelou e matou uma mulher nos EUA (G1, 2018). Nesse caso, de quem seria a responsabilidade pela morte da vítima?

Bostrom (2011, p. 203) elabora uma lista de critérios éticos que devem ser considerados em um algoritmo destinado a substituir o julgamento humano de funções sociais: responsabilidade, transparência, auditabilidade, incorruptibilidade, previsibilidade. Todos esses critérios são aplicados aos humanos que desempenham funções sociais. E o critério da responsabilidade é o primeiro a ser considerado.

Ao abordar especificamente a questão ética da responsabilidade, os autores (MITTELSTADT *et al.*, 2016, p. 10) consideram que, como artefatos de *software* usados no processamento de dados, os algoritmos herdam os desafios éticos associados ao design e à disponibilidade de novas tecnologias, bem como, desafios associados à manipulação de grandes quantidades de dados, o *Big Data*. Isso implica que os danos causados pela atividade algorítmica são

difíceis de serem depurados em vários níveis: detectar o dano, encontrar sua causa e, principalmente, identificar quem deve ser responsabilizado pelos danos causados. Os atores humanos envolvidos são vários: programadores, projetistas, fabricantes, designers e usuários. Contudo, a culpa somente pode ser atribuída se os atores envolvidos têm algum controle ou intencionalidade na execução da ação, como no caso dos algoritmos de não aprendizagem, como por exemplo, o algoritmo de personalização *EdgeRank* do *Facebook*.

A questão ética da responsabilidade é dependente da questão da transparência. Somente um algoritmo não opaco permite a identificação dos responsáveis pelo processo.

Considerações finais

Ao buscar entender as questões éticas relativas aos algoritmos, foram mapeados os seguintes aspectos: falibilidade, opacidade, viés, discriminação, autonomia, privacidade e responsabilidade. Os três primeiros (falibilidade da correlação, opacidade e viés) são de ordem epistêmica, pois abordam a qualidade da evidência produzida pelo algoritmo. Os três seguintes (discriminação, autonomia e privacidade) são de ordem normativa, pois dizem respeito à ação em si conduzida pelo algoritmo que pode ser justa ou injusta do ponto de vista ético. E todos desembocam no último problema, o da rastreabilidade do agente responsável.

Esses são os principais desafios éticos trazidos pelo avanço tecnológico da sociedade da informação com os quais a humanidade terá que lidar em um mundo cada vez mais conduzido por algoritmos. Nesse sentido, as propostas mais sensatas indicam que princípios éticos devem estar presentes *by design*. Por interagirem com a sociedade em geral, os algoritmos não podem simplesmente reproduzir os valores social vigentes. Faz-se necessário um mínimo de governança com a regulamentação ético-normativa que possa garantir que na própria concepção do algoritmo já estejam presentes os princípios da segurança, transparência, dignidade da pessoa humana, não discriminação, liberdade de escolha, proteção da privacidade e responsabilidade.

Para além do *by design* e da governança relacionada à concepção das tecnologias, existem os aspectos sociais que envolvem cuidados éticos, políticos e sociais, voltados para as consequências das interações da sociedade com os algoritmos. Portanto, pelos próprios aspectos tratados no artigo, derivados das preocupações normativas propostas por Mittelstadt e colegas,

observa-se que uma parte significativa das questões éticas não depende apenas do desenho das tecnologias e seus algoritmos; decorre, porém, da conjunção entre esse desenho e questões das relações de uso que a sociedade faz dos algoritmos. Consequentemente, a própria sociedade deve experimentar e desenvolver estratégias de ordem ética e política para evitar ou superar esse tipo de questões. A situação não pode ser resolvida apenas pelo aspecto tecnológico, mas envolvem também decisões e ações de ordem interacional, cultural e jurídica, diretamente relacionadas ao desenho algorítmico e ao seu uso. Esse tema da ética dos algoritmos é latente na sociedade: sob o aspecto interacional é discutida na UNESCO, ao tratar dos letramentos de mídia e informação (MIL); do ponto de vista cultural está presente em documentários, como “O dilema das redes” (NETFLIX, 2020) e em matérias jornalísticas sobre questões práticas que colocam em pauta essa discussão e do prisma jurídico a recente promulgação da Lei Geral de Proteção de Dados que busca, entre outros objetivos, proteger os dados pessoais do uso irregular dos algoritmos. Enfim, há de se empenhar um processo de tomada de consciência das decisões feitas tanto pelos planejadores e designers como pelos participantes sociais em geral que possam ser “capturados” pelos algoritmos, tornando-se alertas aos sentidos de seus acionamentos.

Regina Rossetti é doutora em Filosofia pela USP e professora do Programa de Pós-Graduação em Comunicação da USCS – Universidade Municipal de São Caetano do Sul.

regina.rossetti@online.uscs.edu.br

Alan Angeluci é Doutor em Ciências pela Escola Politécnica da Universidade de São Paulo – EPUSP e professor do Programa de Pós-Graduação em Comunicação da USCS – Universidade Municipal de São Caetano do Sul.

alan.angeluci@online.uscs.edu.br

Contribuições de cada autor: Regina Rossetti: supervisão e gestão do Projeto de Pesquisa, fundamentação teórica e conceituação, escrita – Primeira Redação, metodologia. Alan Angeluci: fundamentação teórica e conceituação, escrita – revisão e edição, metodologia, análise formal do *corpus*, construção de figuras e tabelas.

Referências

ANGELUCI, A. C. B.; DONATO, R.; SCOLARI, G. O WhatsApp como actante: o impacto do aplicativo interativo em redações jornalísticas. **Revista Mediação**, v. 19, n. 24, 2017.

BUCHER, T. The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms. **Information, Communication & Society**, 20:1, 30-44, 2017.

BOSTROM, N.; YUDKOWSKY E. A ética da inteligência artificial. **Fundamento – Rev. de Pesquisa em Filosofia**, v. 1, n. 3 – maio – ago. 2011.

BRUNO, F. **Máquinas de ver, modos de ser**: vigilância, tecnologia e subjetividade. Porto Alegre: Sulina, 2013.

DRATH, R.; HORCH, A. Industrie 4.0: Hit or hype? **IEEE industrial electronics magazine**, v. 8, n. 2, p. 56-58, 2014.

FLORIDI, L. **The Fourth Revolution**: How the Infosphere is Reshaping Human. Reality. Oxford: OUP, 2014.

FUCHS, C. *et al.* **Internet and Surveillance**: the challenges of Web 2.0 and social media. London: Routledge, 2011.

CARRO AUTÔNOMO DA UBER ATROPELA E MATA MULHER NOS EUA. **G1**. Disponível em: <https://g1.globo.com/carros/noticia/carro-autonomo-da-uber-atropela-e-mata-mulher-nos-eua.ghtml>.

GELI, C. Byung-Chul Han: Hoje o indivíduo se explora e acredita que isso é realização. **El País**. Disponível em: https://brasil.elpais.com/brasil/2018/02/07/cultura/1517989873_086219.html#:~:text=Big%20data.%E2%80%9DOs,vistos%20segundo%20os%20dados%20geridos.

HERMANN, M.; PENTEK, T.; OTTO, B. Design principles for industrie 4.0 escenarios. In: **System Sciences (HICSS)**, 2016 49th Hawaii International Conference, 2016. p. 3928-3937.

DILEMAS MORAIS DA QUARTA REVOLUÇÃO INDUSTRIAL. **Ihu Online**. 2017. Disponível em: <http://www.ihu.unisinos.br/78-noticias/571796-os-dilemas-morais-da-quarta-revolucao-industrial>.

MANOVICH, L. **The language of new media**. Cambridge, Massachusetts: MIT press, 2001.

MCCARTHY, J. *et al.* A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. **AI magazine**, v. 27, n. 4, p. 12, 2006.

MITTELSTADT, B. D.; ALLO, P. A.; TADDEO, M. R.; WACHTER, S.; FLORIDI, L. The ethics of algorithms: Mapping the debate. **Big Data & Society**. July–December 2016: 1–21.

MOREIRA, I. A Microsoft criou uma robô que interage nas redes sociais - e ela virou nazista. In: **Revista Galileu**, 2016. Disponível em: <https://revistagalileu.globo.com/blogs/buzz/>

noticia/2016/03/microsoft-criou-uma-robo-que-interage-nas-redes-sociais-e-ela-virou-nazista.html>.

MOROZOV, E. **Big Tech**: a ascensão dos dados e a morte da política. São Paulo: Ubu, 2018.

LIMA VAZ, H. C. **Escritos de Filosofia IV**: Introdução à Ética Filosófica. São Paulo: Loyola, 1999.

PASSOS, P. Combate à Covid-19 criará sociedade rastreada como nunca', diz pesquisadora. **Folha de São Paulo**. Disponível em: <<https://www1.folha.uol.com.br/mundo/2020/04/combate-a-covid-19-criara-sociedade-rastreada-como-nunca-diz-pesquisadora.shtml>>.

PIERRO, B. O mundo mediado por algoritmos. **Revista Pesquisa Fapesp**, abril de 2018. Disponível em: <<https://revistapesquisa.fapesp.br/o-mundo-mediado-por-algoritmos/>>.

RAJKUMAR, R. R. *et al.* Cyber-physical systems: the next computing revolution. In: **Proceedings of the 47th design automation conference**. ACM, 2010, p. 731-736.

SALAS, J. Google “conserta” seu algoritmo racista apagando Gorilas. In: **El País**. Disponível em: <https://brasil.elpais.com/brasil/2018/01/14/tecnologia/1515955554_803955.html>.

SAMUEL, A. L. Some moral and technical consequences of automation—a refutation. **Science**, v. 132, n. 3429, p. 741-742, 1960.

SILVA, T. Visão Computacional e Racismo Algorítmico: Branquitude e Opacidade no Aprendizado de Máquina. **Revista ABPN**, v. 12, p. 428-448, 2020.

SILVEIRA, S. A. Responsabilidade algorítmica, personalidade eletrônica e democracia. **Revista Eletrônica Internacional de Economia Política da Informação, da Comunicação e da Cultura**, v. 22, n. 2, p. 83-96, 2020.

TADDEO, M. R.; FLORIDI, L. How AI can be a force for good. **Science**, v. 361, n. 6404, p. 751-752, 2018.

VICENTE, J. P. Preconceito das máquinas: como algoritmos podem ser racistas e machistas. In: **Uol**, 2018. Disponível em: <<https://noticias.uol.com.br/tecnologia/noticias/redacao/2018/04/24/preconceito-das-maquinas-como-algoritmos-tomam-decisoes-discriminatorias.htm>>.

YANG, G.-Z. *et al.* The grand challenges of Science Robotics. **Science Robotics**, v. 3, n. 14, p. eaar7650, 2018.

Artigo recebido em 30/08/2020 e aprovado em 06/01/2021.