

# A ÉTICA COMO UM DOS DESAFIOS DA INTELIGÊNCIA ARTIFICIAL

Edgard Luiz Bernardes Valderramas <sup>1</sup>

Pontifícia Universidade Católica de São Paulo – PUC - SP, São Paulo, SP

**Abstract.** *This study deals with issues of ethics and challenges related to its application in the field of Artificial Intelligence. In an analysis of AI in its social context, the impacts that this emerging branch of science can bring and its consequences are examined. From a methodological point of view, the paper deals with ethical assumptions, the cybernetic challenges related to society for its implementation and its adherence with AI citing examples and carrying out some evaluations. In view of the rapid development of this area, the paper emphasizes the growing importance of the debate among the agents involved.*

**Resumo.** *Este trabalho trata as questões da ética e dos desafios relacionados à sua aplicação no campo da Inteligência Artificial. Por meio de uma análise da IA no contexto social em que se insere, serão enfatizados os impactos que este emergente ramo da ciência pode trazer e as suas consequências. Do ponto de vista metodológico, realizou-se uma discussão teórica a partir dos pressupostos da ética, dos desafios cibernéticos relacionados à sociedade para a sua implementação e por fim, a sua aderência com a IA citando-se exemplos e realizando algumas avaliações. Entre as conclusões, verificou-se a importância de se alargar o debate entre os agentes envolvidos, tendo em vista o rápido e importante desenvolvimento desta área.*

## Introdução

As mudanças provocadas pelos avanços da tecnologia e as suas ramificações, como por exemplo, a massificação das mídias na coleta e distribuição das informações, está fazendo cada vez mais com que as pessoas se atentem a usos e costumes antes ligadas apenas aos seres humanos e que agora precisam ser estendidos aos processos cibernéticos. Se antes as máquinas eram o meio, agora é a própria mensagem (McLUHAN, 1964).

A tecnologia alterou de tal forma a vida humana, que não se discute atualmente se as mudanças virão, mas como adaptar-se a elas, agregando-se valores às máquinas, plataformas móveis, *hardware*, *software* e tudo aquilo que se deriva desses elementos e nos mais diversos setores. Como toda nova área a ser explorada, esses novos processos carecem de parâmetros, principalmente os éticos, pois se trata de uma atividade que lida

---

<sup>1</sup> Doutorando em Tecnologias da Inteligência e Design Digital da Pontifícia Universidade Católica de São Paulo – PUC - SP

com informações, pessoas, formadores de opinião, grupos de interesse, sentimentos, destinos, entre outros e que merecem cuidado na sua forma de tratamento.

Este trabalho apresenta estudos sob o ponto de vista teórico em relação à ética a partir das contribuições de autores que já estudaram esses assuntos. O objetivo é contextualizar a ideia da ética, desenvolvida por Bucci (2000), Rachels (2004) e Abbagnano (2007), sobre cibernética e seus desafios Wiener (1950) e Ford (2015) e para tratar de Inteligência Artificial e sua conexão com as questões éticas Bostrom & Yudkowsky (2011), Di Blasi & Cantarino e Hammond (2017).

### **A Ética e os seus domínios**

A ética estuda os valores morais e princípios do comportamento humano, pois é a ciência da conduta que discute como o homem se relaciona com a sociedade. Etimologicamente, a palavra leva a duas ramificações de pesquisas da ética: a primeira, denominada filosófica, tem caráter normativo e busca estabelecer os princípios constantes e universalmente válidos para uma boa vida em sociedade, discutindo-se uma moral universal ideal e a segunda linha de estudos, chamada ética científica, tem caráter explicativo e busca compreender as morais históricas numa reflexão sobre os costumes.

As sociedades fornecem sistema de valores, instituições e modos de vida que moldam a existência dos indivíduos. As virtudes essenciais não são prescritas por convenções sociais, mas por fatos fundamentais sobre a nossa condição humana comum (RACHELS, 2004, p. 96).

Bucci (2000) nos traz a ideia de duas facetas da ética ao afirmar que uma se assenta sobre o indivíduo e sua conduta e outra na sociedade. Após fazer uma breve explicação sobre as bases teóricas das duas correntes, Bucci argumenta que a forma de agir de cada pessoa vai determinar o comportamento ético (2000, p. 16). Para tanto, argumenta ele, cada indivíduo recorre ao livre-arbítrio, foro íntimo no qual será definido o que é bom e ruim. Embora o próprio Bucci saliente o cuidado com a definição de termos tão amplos como bom e mau, o autor reforça a tese de que esses conceitos nortearão não só as atitudes individuais, mas igualmente as definidas em coletividade. Assim, é preciso ter senso de coletividade para tomar determinadas decisões, já que as atitudes individuais têm consequências sobre a comunidade (2000, p. 17).

Há uma sintonia entre os costumes e a boa conduta, pois a ética não está nem poderia estar fora dos costumes (BUCCI, 2000, p. 17). Bucci cita que dificilmente uma máquina com toda a racionalidade e precisão que a caracteriza teria condição de substituir um ser humano. O professor traz o exemplo de um suposto repórter de rádio que teria testemunhado um deslizamento de terra ao retornar para casa depois de um dia de trabalho. No momento chovia forte e havia pontos de alagamento. Ainda conforme o fato em questão, o jornalista tenta com o celular narrar ao vivo o fato dispondo de pouquíssimas impressões, a não ser o que está vendo e a partir da sua experiência. Neste exemplo, há a possibilidade de se transmitir partes da verdade, pois o profissional tem pouco controle sobre aquilo que está sendo dito ao cidadão. Bucci faz uma recomendação básica e elementar em uma situação como essa: não mentir, apenas agir de acordo com a sua consciência, esforçar-se para olhar o mundo com objetividade e não mascarar suas limitações com arrogância de quem tudo vê. (2000, p. 88-89). Em princípio, dá-se a impressão de que a sugestão seria facilmente aplicada pela máquina, visto que objetividade e precisão são características inerentes dos computadores. No entanto, é de se questionar até que o ponto um robô poderia exercer seguir as

recomendações a partir de termos como verdade parcial, consciência e arrogância. Se a ética está no modo de agir do jornalista, logo, é indispensável a visão de vida em sociedade para que ela possa existir.

O dicionário de filosofia de Nicola Abbagnano resume a ética como a ciência da conduta, a qual se pautaria por normas derivadas do próprio homem, fruto da sua racionalidade, ou de preceitos dos quais ele obedece. Nesta segunda possibilidade, “o bem não é definido com base na sua realidade ou perfeição, mas só como objeto da vontade humana ou das regras que a dirigem” (2007, p. 383). Outros exemplos de virtude são: honestidade, generosidade, equidade, paciência, compaixão, lealdade, benevolência, tolerância, cortesia, justiça.

### **Os desafios cibernéticos da Inteligência Artificial**

Em meados do século XX, muitos cientistas perceberam que os fenômenos não podem ser classificados em qualquer disciplina única, mas exigem uma compreensão interdisciplinar com base em fundamentos científicos que ainda não foram inventados. Várias tentativas de construir essas fundações incluem (entre outros) os campos da cibernética, sinergia, ciência de sistemas, e, mais recentemente, a ciência de sistemas complexos. É preciso outros olhares para compreender determinados fatos porque vê-los a partir das ferramentas das quais dispomos não são suficientes para entender esses sistemas. Não é possível separá-los para compreendê-los, como tradicionalmente a ciência o faz (VALDERRAMAS; ARAUJO apud MITCHELL, 2016).

Pantaleão (2010) descreve que a revolução digital trouxe em si uma profunda influência nas normas, princípios e valores que tangenciam a vida social e que problemas éticos de cunho social se desencadeiam uma vez que a relação homem-máquina é estabelecida de uma forma quase visceral. O mau uso e má manipulação da informação, acarretando um tipo de vandalismo informacional, é outra consequência ética da revolução tecnológica.

As máquinas robotizadas, que estão assumindo paulatinamente funções cada vez mais tipicamente humanas, podem ser definidas como um sistema mecânico integrado por conjuntos rígidos, articulados entre si e servido de garras ou dispositivo especializado que realiza atividades complexas ou de alto risco para o homem em processos (ORTEGA; SPONG, 1989, p. 878). Para eles esses objetos são construídos para manipular materiais, peças, ferramentas ou dispositivos especializados, através de movimentos variáveis programadas para a realização de tarefas diversas (1989, p. 877). Isto significa que, por exemplo, quando um software produz um texto a partir de estatísticas colhidas durante uma partida de futebol, está realizando uma rotina para a qual foi programado por meio da manipulação de materiais, no caso os números do jogo.

Ford (2015) apresenta em seu livro que a tarefa não é ficar de pé diante desse futuro, mas em vez disso se preparar para isso de tal maneira que mitigue os problemas enquanto acentua o potencial positivo. É um desafio, mas que só se tornará mais perigoso se for ignorado. Olhando para as tecnologias sensacionalistas e hipotéticas da nanotecnologia, para computadores superinteligentes, para “a singularidade”, todos retratam um mundo em que as mudanças nas condições de trabalho são quase insignificantes à luz das questões existenciais maiores que entram em jogo, pois os seres humanos criam máquinas mais inteligentes que eles próprios. É prematuro saber com certeza quais efeitos a impressão 3D, carros autônomos e computadores superinteligentes podem ter e, no entanto, todos eles apontam para um futuro em que

cada vez mais tarefas são tratadas por máquinas. Althoff (2007) afirma que existem duas ameaças concretas em relação aos nano-robôs que não podem ser ignoradas, quais sejam: a) eles podem ser controlados por entidades hostis de diversos tipos; b) eles podem autorreplicar-se como organismos autônomos e causar imenso dano à humanidade. A ressalva é extremamente pertinente porque elas tocam em questões ontológicas diversas, como o fato do homem poder replicar e amplificar sensações de outros, mas ainda deparar-se com situações éticas aparentemente sem respostas satisfatórias. Sob o ponto de vista da estratégia de gestão para as companhias, é bastante vantajoso substituir o trabalho humano pelo maquinário. Os mais visíveis repousam em atividades de alta periculosidade, impedindo o risco real para a saúde dos trabalhadores. No entanto, nas atividades intelectuais também existem diversos benefícios, entre eles reduzir custos e aperfeiçoar o processo produtivo. Se até pouco tempo atrás o emprego de maquinário estava restrito à execução de tarefas específicas do processo de produção, como pintura ou manuseio de objetos, mas recentemente adentrou também na realização de trabalhos intelectuais (TACHIZAWA et al, 2009, p. 2).

Além disto, a conjuntura global corrobora com a adoção dos robôs, em especial, pelo crescente mercado de consumo que exige a expansão da capacidade de produção, a reorganização da produção para maior eficiência energética e o avanço na colaboração homem-máquina, abrindo a possibilidade de novas aplicações das máquinas. Estas, por sua vez, tornaram-se onipresentes na sociedade contemporânea, como por exemplo fazendo-se presentes, amiúde, por meio da produção de reportagens textuais, notadamente nas editorias de esporte e de economia. Nos Estados Unidos o blog *Journalism in the Americas*, da Universidade do Texas em Austin, trouxe à baila diversos embates entre pesquisadores, jornalistas, políticos e juristas sobre a ética envolvendo o uso de dispositivos remotos para coleta de informações. O Laboratório de Jornalismo Robótico da Universidade de Nebraska-Lincoln's, por exemplo, construiu veículos aéreos não tripulados (VANT), conhecidos por drones, para realizar reportagens. A Escola de Jornalismo Missouri, do Programa de Tecnologia da Informação da Universidade de Missouri, e a estação da Rádio Pública Nacional (NPR) KBIA, lançaram um projeto em parceria que prevê o uso de drones para realização de reportagens científicas investigativas. Os defensores desses projetos argumentam que o uso dessas máquinas é algo inexorável, pois reduz custos das empresas de mídia, já bastante combatidas pela competição com a internet. Além disso, jornalistas e pesquisadores defendem o uso de máquinas em locais de difícil acesso para humanos, que seriam beneficiados com reportagens mais completas.

A maneira como uma máquina atua é diferente da forma como os seres humanos pensam. O eminente cientista norte-americano John Von Neumann publicou uma importante obra sobre esse aspecto. Em *O computador e o cérebro*, Neumann explicou que as máquinas necessitam de cálculos para realizar as tarefas solicitadas pelos humanos, pois as mesmas geralmente dizem respeito a respostas como sim e não, ou seja, processos automatizados, nos quais não são necessárias habilidades abstratas (2006, p. 108-109). Dessa forma, se um computador tivesse que dizer uma verdade parcial, ele simplesmente não diria, pois ele não foi programado para tanto. O problema seria ainda maior se considerar-se a palavra consciência. Segundo Abbagnano, consciência não é apenas estar ciente do que faz ou deixa de fazer, conforme compreende o senso comum. Nos termos da filosofia, consciência pressupõe:

Uma relação da alma consigo mesma, uma relação intrínseca ao homem, "interior" ou "espiritual", pela qual ele pode conhecer-se de modo imediato e privilegiado e por isso julgar-se de forma segura e infalível. Trata-se, portanto, de uma noção em que o aspecto moral— a possibilidade de autojulgar-se — tem conexões estreitas com o aspecto teórico, a possibilidade de conhecer-se de modo direto e infalível (2007, p. 185).

Alan Turing ficou conhecido pela criação de um dispositivo elétrico e mecânico que realizava cálculos matemáticos, denominada posteriormente de “A Máquina de Turing”, o autômato foi fundamental ao decifrar códigos nazistas para os países aliados (GLEICK, 2013, p. 222). Turing escreveu artigo seminal para os estudos de máquinas computadoras em 1950 quando descreve um jogo que teria como objetivo descobrir se uma máquina pode pensar. Chamado de O Jogo da Imitação, a ideia é identificar um ser humano entre duas pessoas e uma máquina a partir de perguntas e respostas, não chegando a conclusões taxativas em seu artigo, mas faz uma pergunta que pode ser compreendida como resposta: “Não podem as máquinas realizar algo que deveria ser descrito como pensar, mas que é muito diferente do que um homem faz? ” (1950, p. 435). Isto é, pode ser que as máquinas não façam exatamente aquilo que um ser humano no que diz respeito a pensar, porém, pode fazer algo com mecanismo semelhante que não seja necessariamente igual. Ao longo do artigo, busca desfazer mitos como o de que pensar era uma atribuição essencialmente humana, atribuída por Deus (1950, p. 443). Turing cita que no final do século, o uso de palavras e opinião educada geral vai ter alterado tanto que um será capaz de falar de máquinas pensantes sem esperar ser contrariada. (1950, p. 442). As conclusões de Turing reforçam a impressão de que uma máquina dificilmente conseguiria posicionar-se diante dos dilemas éticos, haja vista o fato de que o pensar da máquina é diferente do humano.

Norbert Wiener, em *Cibernética e sociedade*, afirma que a máquina digital funciona a partir do tudo ou nada e que “a nitidez da decisão entre ‘sim’ e ‘não’ permite-lhe acumular informação de maneira a facultar-nos discriminar diferenças muito pequenas em quantidades muito grandes” (1950, p. 64). Assim, como atestou Von Neumann, o cérebro humano não tem a precisão de uma máquina computadora no que diz respeito a discernir informações em meio a muitas outras, mas pode se sobressair em determinadas tarefas. “A deterioração da aritmética foi trocada por um aperfeiçoamento da lógica” (2006, p. 113), concluiu Neumann. Sendo a ética dependente de fatores externos, é demasiado improvável que uma máquina tenha informações exteriores suficientes para tomar uma decisão baseando-se em opiniões alheias. Ou ainda mais distante da realidade seria pensar que o robô teria condições de interagir com os humanos de tal forma que pudesse compreender seus comportamentos e tirar uma conclusão lógica disso.

Pode-se construí-la com uma preferência estatística por certa espécie de comportamento, que, não obstante, admite a possibilidade de outro comportamento; ou, então, certas características de seu comportamento podem ser rígida e inalteravelmente determinadas. Chamaremos a primeira espécie de determinação preferência e à segunda espécie coerção. Por exemplo, se as regras de xadrez oficial não forem introduzidas numa máquina de jogar como coerções, e se for dada à máquina a capacidade de aprender, ela poderá transformar-se, sem que se perceba, de uma máquina de jogar xadrez noutra que execute uma tarefa totalmente diversa (1950, p. 175).

Ainda que um robô possa aprender, esse aprendizado será limitado por dados estatísticos, isso é, conforme fatores de probabilidade, pois é inerente ao mecanismo autômato pautar-se por cálculos matemáticos como forma de aprendizado. Este seria o cerne para entender porque uma máquina dificilmente desempenhará funções estritamente humanas, como decisões éticas.

Sendo diferente a forma como o cérebro compreende números e ordens, crê-se que uma decisão ética provavelmente não fará parte do arcabouço matemático, pois ética diz respeito a comportamento individual inserido em um contexto social. Logo, uma máquina criada para executar comandos, ainda que possa aprender com o meio, não teria condições de elaborar um raciocínio nesse nível de complexidade (ARAÚJO; VALDERRAMAS apud WIENER, 2016).

Hofstadter formula algumas especulações e uma delas refere-se não diretamente à ética, mas às emoções de uma maneira geral:

As emoções serão explicitamente programadas em uma máquina?

Não. Isso é ridículo. Qualquer simulação direta de emoções - PARRY, por exemplo - não pode abordar a complexidade das emoções humanas, que surgem indiretamente da organização de nossas mentes. Programas ou máquinas irão adquirir emoções da mesma maneira: como subprodutos de sua estrutura, do modo como eles são organizados - não por programação direta. Assim, por exemplo, ninguém vai escrever uma sub-rotina de "apaixonar-se apaixonadamente", assim como eles não escreveriam um "erro" sub-rotina. "Apaixonar-se" é uma descrição que atribuímos a um complexo processo de um sistema complexo; não precisa haver nenhum módulo único dentro do sistema que é o único responsável por isso, no entanto! (1979, p. 674).

Vale ressaltar que primário e secundário não diz respeito à ordem de importância, mas a forma com a qual os processos se realizam no cérebro humano. Também é importante lembrar que a linguagem não representa necessariamente o alfabeto, mas a esquematização lógica pela qual nosso cérebro compreende o mundo. Dessa forma, além das relações humanas serem fortuitas e imprevisíveis a ponto de praticamente impossibilitar a precisão nos cálculos matemáticos de uma máquina, o cálculo não seria a forma mais apropriada de tomar decisões éticas haja vista que nosso cérebro tem melhor desempenho com a lógica se comparado à matemática.

### **A ética artificial**

De acordo com Di Blasi e Cantarino (2017) a Google introduziu em 2015 uma ferramenta inteligente que realiza o reconhecimento de formatos em fotos para agrupar em pastas distintas, o que ocasionou no triste incidente de reconhecer uma pessoa negra como gorila, gerando um grave incidente tecnológico. Em outro triste exemplo, mais recente, um palestino que trabalha em um assentamento judeu em Israel publicou em uma rede social uma foto com uma escavadeira, junto à frase “bom dia”, em árabe. Ocorre que o algoritmo deste website traduziu esta expressão para o hebraico e para o inglês como uma frase que incitaria violência, o que causou sua prisão, ao menos temporária, pela polícia local e, colaborando para o fato, escavadeiras eram usadas no passado em ataques terroristas.

Em páginas de busca como o Google e o Bing, a partir de inputs de usuários, passaram a exibir imagens de mulheres brancas quando se busca pelo termo “mulheres bonitas”, enquanto imagens de mulheres negras eram exibidas quando se buscava por imagens de

“mulheres feias”, de acordo com uma pesquisa feita em quarenta países por pesquisadores brasileiros da UFMG em 2016 (UFMG, 2016).

Em 2016, a Microsoft colocou uma Inteligência Artificial no Twitter para que ela aprendesse com os internautas. Em apenas um dia, ela se tornou preconceituosa, falando mal sobre judeus e mexicanos, e isso soou um alerta interno na empresa sobre a importância da ética para essa nova fronteira da tecnologia. Do ponto de vista de tecnologia, o experimento obteve sucesso, mas do ponto de vista social, foi uma experiência terrível. O algoritmo aprendeu do meio e que isso é um pouco do reflexo da sociedade na qual vivemos hoje. A questão da ética, do propósito e da visão do impacto do que fazer vai ser mais importante do que nunca. O poder também estará nas mãos das grandes organizações de tecnologia (EXAME, 2018).

Segundo Di Blasi e Cantarino (2017) apud Hammond (2017) a Inteligência Artificial somente deve ser lançada se seu criador puder explicar seu processo decisório e, em uma escala de requisitos mínimos para melhores padrões, as máquinas devem ser: auditáveis, terem capacidade de articulação dos dados para extrair as informações mais relevantes e, na melhor das práticas, explicar suas razões de decisão bem como considerar novos outputs caso seja fornecido um novo dado sobre a mesma questão.

De acordo com IERFH apud Bostrom e Yudkowsky (2011) uma reflexão sobre os parâmetros básicos que as máquinas que ocupam postos de dimensões sociais devem ter deve ser realizada: “responsabilidade, transparência, auditabilidade, incorruptibilidade e tendência para não fazer vítimas inocentes gritarem em desamparada frustração”. Ainda, citam dois critérios que são comumente propostos como importantemente relacionados com o estatuto moral, ou isoladamente ou em combinação: a senciência e a sapiência (ou personalidade). Estes podem ser caracterizados aproximadamente como segue: Senciência: a capacidade para a experiência fenomenal ou qualia, como a capacidade de sentir dor e sofrer; Sapiência: um conjunto de capacidades associadas com maior inteligência, como a autoconsciência e ser um agente racional responsável.

O dilema moral e ético passa a tomar forma quando, a partir do input de diversos dados em uma máquina programada com algoritmo de aprendizado em setores distintos e sem qualquer discernimento ético previamente programado, a máquina é capaz de realizar um cruzamento de dados, calcular índices estatísticos e fornecer um output que pode ser a fronteira entre aprovar ou não um empréstimo, criticar ou não um texto, prescrever ou não um medicamento ou, pior, atirar ou não em um cidadão (DI BLASI e CANTARINO, 2017).

Continuando com Di Blasi e Cantarino (2017) a *machine learning* em si faz com que as máquinas possuam, por um lado, uma capacidade de aprimoramento quase-humana, e, por outro, passíveis de tomar decisões que seriam minimamente questionáveis por humanos, mesmo consideradas antiéticas, imorais ou ofensivas de alguma forma.

A Inteligência Artificial é como uma criança que imita o comportamento de seus pais, em vez de fazer o que eles lhe falam. Ela absorve dados, encontra padrões e os copia. Embora empregadores possam dizer que querem diversidade, um programa de recrutamento pode, ao invés disso, seguir o padrão de comportamento dos empregadores.

O site UDACITY (2018) revela que o Fórum Econômico Mundial elencou nove preocupações em relação ao avanço da Inteligência Artificial ao redor do mundo: desemprego gerado pela substituição da mão de obra por máquinas autônomas, desigualdade causada pela má distribuição da riqueza gerada pelas máquinas, alterações do comportamento humano e das relações interpessoais, erros cometidos pelas máquinas autônomas, alguns podendo ser fatais aos seres humanos, preconceitos reproduzidos por máquinas e sistemas que utilizam bases de dados ruins ou pouco confiáveis, os riscos das armas autônomas, o receio de que a Inteligência Artificial venha a causar consequências ainda inimagináveis pelos humanos, o medo de que a Inteligência Artificial um dia supere as capacidades e a inteligência humana e a interação entre humanos e os robôs.

A boa notícia trazida por Di Blasi e Cantarino (2017) é que a IEEE – Institute of Electrical and Electronic Engineers criou uma iniciativa global de considerações éticas no ramo de Inteligência Artificial, em um documento chamado *Ethically Aligned Design*, para discutir o tema em questão, associando alguns princípios basilares para considerar a questão da ética na Inteligência Artificial, com os quais entendemos ser o início para o controle dos seus resultados: princípio do benefício humano, para que as inteligências artificiais não desrespeitem direitos humanos; princípio da responsabilidade, segundo o qual os poderes eleitos ou o próprio poder judiciário devem criar normas claras de responsabilidade jurídica envolvendo casos com Inteligência Artificial; princípio da transparência, e princípio da educação e consciência, de modo a minimizar os riscos de mau uso da Inteligência Artificial.

### **Considerações finais**

A ética não está em nenhum campo de atuação, pois trata-se de um exercício inerente de todas as profissões, portanto ela não pode e não deve em hipótese nenhuma ser desprezada, tampouco diante da enormidade de números, palavras, imagens e outros símbolos que remetem à intimidade das pessoas ou instituições e que estão presentes em proporções nunca antes vista na história da humanidade. Partindo-se da ideia de que sendo a ética um dos princípios norteadores do caráter do ser humano, as máquinas também deveriam ter o mesmo norte como base essencial de coleta, tratamento, classificação, distribuição e divulgação da informação com a singular diferença que este princípio estaria codificado por meio de símbolos.

Para Di Blasi e Cantarino (2017), embora a Inteligência Artificial atual nos ofereça algumas questões éticas que não estão presentes no design de automóveis ou de usinas de energia, a abordagem de algoritmos de Inteligência Artificial em relação a um pensamento mais humano prenuncia complicações desagradáveis e os papéis sociais podem ser preenchidos por meio de algoritmos de Inteligência Artificial, o que implica novas exigências de projeto, como transparência e previsibilidade.

O grande desafio apresentado é como a Inteligência Artificial atenderá aos preceitos éticos, não apenas do ponto de vista cibernético, mas muito mais do que isso, respeitando a moral, sendo responsável, transparente e incorruptível, ou seja, empregando não o que o programador inseriu nos seus algoritmos, mas fazendo aquilo que é certo, não reproduzindo comportamentos humanos errados. Este texto não tem a pretensão de esgotar o assunto, mas apenas contextualizá-lo por meio de uma fundamentação e evidenciá-lo pelos exemplos aqui citados, com o espírito vigilante de



correção e prevenção, entendendo que esse amadurecimento deverá se consolidar com o tempo, pois a única certeza que existe é que esse cenário veio para ficar.

## Referências

- ABBAGNANO, Nicola. **Dicionário de filosofia**. São Paulo: Martins Fontes, 2007.
- ALTHOFF, Fritz et al. **Nanoethics: The Ethical and Social Implications of Nanotechnology**. New Jersey: John Wiley & Sons, 2007.
- ARAÚJO, Lucas V.; VALDERRAMAS, Edgard L. B. Reflexões sobre a ética na produção de notícias por máquinas. Disponível em <http://docplayer.com.br/37230418-Reflexoes-sobre-a-etica-na-producao-de-noticias-por-maquinas-1.html>. 2016. Acesso em 18 de agosto de 2018.
- BOSTROM, Nick; YUDKOWSKY, Eliezer. The ethics of artificial intelligence. **Cambridge Handbook of Artificial Intelligence**. New York: Cambridge University Press, 2011.
- BUCCI, Eugênio. **Sobre ética e imprensa**. São Paulo: Companhia das Letras, 2000.
- DI BLASI, Gabriel e CANTARINO, Rodrigo. Limite da IA frente aos dilemas éticos e morais **Jota** 08/12/2017. Disponível em: <https://www.jota.info/opiniao-e-analise/artigos/limite-da-ia-frente-aos-dilemas-eticos-e-morais-08122017>. Acesso em 16 de maio de 2018.
- EXAME. **Robô preconceituoso soou alerta interno na microsoft**. Disponível em <https://exame.abril.com.br/tecnologia/robo-preconceituoso-soou-alerta-interno-na-microsoft/>. Acesso em 16 de maio de 2018.
- FORD, Martin. **Rise of robots**. New York: Basic Books, 2015.
- GLEICK, James. **A informação: uma história, uma teoria, uma enxurrada**. São Paulo: Companhia das Letras, 2013.
- HAMMOND, Kris. **Artificial intelligence: transparency isn't just a trend, Computerworld**. Jan 12, 2017. Disponível em <https://www.computerworld.com/article/3155534/artificial-intelligence/artificial-intelligence-transparency-isn-t-just-a-trend.html>. Acesso em 24 de maio de 2018.
- HOFSTADTER, Douglas. **Gödel, Escher, Bach**. New York: Basic Books, 1979.
- IERFH. **A Ética da Inteligência Artificial**. Disponível em <http://www.ierfh.org/br.txt/EticaDaIA2011.pdf/>. Acesso em 20 de maio de 2018.
- McLUHAN, Marshall. **Understanding media: The Extensions of Man**. Faltava o nome da cidade: McGraw-Hill Book, 1964.
- MITCHELL, Melanie. **Complexity: a guided tour**. New York: Oxford University Press, 2009.
- NEUMANN, John von. **O computador e o cérebro**. Lisboa: Relógio D'Água, 2006.
- ORTEGA, Romeo; SPONG, Mark. W. Adaptive motion control of rigid robots: A tutorial. **Automatica**, v. 25, n. 6, p. 877-888, 1989.
- PANTALEÃO, Cristina A. Máquinas e sociedade: uma abordagem ética acerca do artificial. Disponível em [http://www.marilia.unesp.br/Home/RevistasEletronicas/FILOGENESE/NathaliaPantaleao\(25-33\).pdf](http://www.marilia.unesp.br/Home/RevistasEletronicas/FILOGENESE/NathaliaPantaleao(25-33).pdf). 2010. Acesso em 18 de agosto de 2018.
- RACHELS, James. **Elementos da filosofia moral**. Lisboa: Gradiva, 2004.

- TACHIZAWA, Takeshi et al. O fator humano no contexto da cadeia produtiva sustentável: uma análise da qualidade de vida baseada em pesquisa empírica. In: **XII Simpósio de administração da produção, logística e operações internacionais da Fundação Getúlio Vargas**, 2009, São Paulo. Anais eletrônicos. Disponível em: [http://www.simpoi.fgvsp.br/arquivo/2009/artigos/E2009\\_T00025\\_PCN80703.pdf](http://www.simpoi.fgvsp.br/arquivo/2009/artigos/E2009_T00025_PCN80703.pdf). Acesso em 10 de junho de 2015.
- TURING, A. M. **Computing Machinery and Intelligence**. Mind, New Series, vol. 59, n.236, p. 433-460, 1950.
- UDACITY. **O avanço da Inteligência Artificial (e dos desafios éticos envolvidos)**. Disponível em <https://br.udacity.com/blog/post/desafios-eticos-inteligencia-artificial>. Acesso em 14 de maio de 2018.
- UFMG. **Pesquisadores do DCC identificam em mecanismos de busca na web estereótipos que relacionam raça, idade e beleza feminina**. Disponível em <https://www.ufmg.br/online/arquivos/044936.shtml>. 2016. Acesso em 16 de maio de 2018.
- VALDERRAMAS, Edgard L. B.; ARAÚJO, Lucas V. **Ética e o jornalismo de dados: discussões preliminares**. Disponível em <http://portalintercom.org.br/anais/nacional2016/resumos/R11-2248-1.pdf>. 2016. Acesso em 18 de agosto de 2018.
- WIENER, Norbert. **Cibernética e sociedade: o uso humano de seres humanos**. São Paulo: Cultrix, 1950.