

## Da black box à transparência: proteção de dados e ética na era da inteligência artificial

### From Black Box to Transparency: Data Protection and Ethics in the Age of Artificial Intelligence

Renata Petreli Piae\*  
Adriano Fernandes Ferreira\*\*  
Tiago Esashika\*\*\*

#### RESUMO

O presente trabalho tem como pergunta principal: Como convergir para a aliança entre o princípio da transparência e a proteção de dados no desenho industrial das novas tecnologias de Inteligência Artificial e Aprendizado de Máquina? Na tentativa de dirimir o dilema, parte-se de uma pesquisa bibliográfica e do método hipotético-dedutivo, dividindo-se em três partes, além de introdução e considerações finais. Em um primeiro momento, relata-se a história da Inteligência Artificial e da técnica de Aprendizado de Máquina para então apresentar o conceito de Black Box. Após, narram-se determinados eventos atinentes à falta de transparência nos processos decisórios algorítmicos, organizando-os entre os que envolvam empresas privadas e entes públicos. Ainda, analisa-se a legislação pátria, com ênfase na Lei Geral de Proteção de Dados, para aferir a existência, ou não, de mecanismos jurídicos capazes de proteger o usuário das práticas obscuras das corporações de tecnologia. Ao final, propõem-se expedientes para o emprego da Inteligência Artificial de forma ética e responsável.

**Palavras-Chave:** inteligência artificial; black box; transparência; ética.

#### ABSTRACT

The main question of this paper is: How to converge towards the alliance between the principle of transparency and data protection in the industrial design of new technologies of Artificial Intelligence and Machine Learning? To solve the dilemma, it is based on bibliographic research and the hypothetical-deductive method, divided into three parts, in addition to an introduction and final considerations. At first, the history of Artificial Intelligence and the Machine Learning technique is reported, and then the concept of Black Box is presented. Afterwards, certain events related to the lack of transparency in algorithmic decision-making processes are narrated,

---

Artigo submetido em 16 de setembro de 2024 a aprovado em 9 de julho de 2025.

\* Mestranda em Constitucionalismo e Direitos na Amazônia no Programa de Pós-Graduação em Direito - PPGDir da Universidade Federal do Amazonas - UFAM. Especialista em Direito Administrativo pela Universidade Anhanguera-Uniderp, Direito Constitucional por Damásio de Jesus e Direito Processual Civil pela Universidade Anhanguera-Uniderp. Graduada em Direito pela Pontifícia Universidade Católica do Paraná - PUC/PR. Advogada. E-mail: renatapetrelip@gmail.com

\*\* Professor da Faculdade de Direito da Universidade Federal do Amazonas – UFAM. Mestre em Direito pela Universidade Gama Filho. Doutor em Ciências Jurídicas pela Universidade Castilha la Mancha. Pós-doutor em Direito pela Universidade de Santiago de Compostela. E-mail: adrianofernandes3@hotmail.com

\*\*\* Mestrando em Constitucionalismo e Direitos na Amazônia no Programa de Pós-Graduação em Direito - PPGDir da Universidade Federal do Amazonas - UFAM. Especialista em Direito Direito Civil e Processo Civil pelo Instituto de Estudos Jurídicos (Iprojude, 2021) e em Direito Público pela Universidade do Estado do Amazonas (UEA, 2022). Residente Jurídico na Procuradoria Geral do Município de Manaus. E-mail: ti.esashika@gmail.com

organizing them between those involving private companies and public entities. In addition, the national legislation is analyzed, with emphasis on the General Data Protection Law, to assess the existence, or not, of legal mechanisms capable of protecting the user from the shady practices of technology corporations. In the end, expedients are proposed for the use of Artificial Intelligence in an ethical and responsible manner.

**Keywords:** artificial intelligence; black box; transparency; ethics.

## 1 INTRODUÇÃO

Escolher um bom filme conforme sugestões da plataforma de streaming, percorrer o trajeto mais rápido aconselhado pelo aplicativo de trânsito, das tarefas mais triviais à intrincadas cirurgias robóticas, o cotidiano da humanidade foi integralmente permeado pela Inteligência Artificial.

Tamanho o avanço tecnológico sucedido nos últimos anos e tão rapidamente a sociedade pós-moderna rodeou-se de inovações, que subtraía o estabelecimento de premissas éticas e jurídicas para as plataformas digitais e suas relações com os usuários.

Muito embora fortemente dominado por grandes empresas de tecnologia, o ambiente cibernético ainda é um território selvagem desbravado mediante práticas indômitas, sem qualquer ponderação acerca dos custos e benefícios sociais ou, raras vezes, responsabilização por eventuais danos.

Atualmente, viver em sociedade implica, em maior ou menor medida, expor-se a tecnologias cujo funcionamento não é amplamente conhecido pelo público, razão pela qual se intenta trazer luz aos expedientes obscuros das Big Techs ao propor acautelada reflexão para promoção de soluções idôneas à efetiva segurança dos usuários.

Partindo da busca pela criação de seres autômatos e inteligentes, estabelecidas as delimitações históricas da Inteligência Artificial e Aprendizado de Máquina, o presente artigo apresenta algumas das problemáticas advindas da carência regulatória no campo cibernético, em especial quanto ao dilema da Black Box ou caixa preta algorítmica e da falta de transparência das decisões emanadas de ferramentas de tecnologia, apresentando ao final as primeiras tratativas para normatização do ramo no âmbito internacional e nacional.

Assim, os procedimentos metodológicos empregados neste ensaio compreendem pesquisas bibliográficas, utilizando-se de livros e artigos publicados sobre a ampla temática Inteligência Artificial, Proteção de Dados e Ética, além de pesquisa documental e normativa para analisar o estado da arte da disciplina jurídica, ou sua completa ausência, no ambiente virtual.

## 2 APRENDIZADO DE MÁQUINA

Desde o mito grego de Prometeu, ao mito hebraico de Golem, a humanidade concentra esforços na criação de seres autômatos que simulam a inteligência humana. Todavia, foi na Conferência de Dartmouth College, em New Hampshire, EUA, que o termo Inteligência Artificial foi utilizado pela primeira vez, dando origem à nova seara de estudos científicos. (De Castro Barbosa, 2020, p. 93)

Tocante às produções científicas, em 1943, Warren McCulloch e Walter Pitts redigiram um ensaio sobre estruturas artificiais de raciocínio sob a forma de modelo matemático e, abstraída qualquer ressalva ética, ao cabo da Segunda Guerra Mundial, concretizaram-se relevantes inovações tecnológicas, além de estudos nas áreas médicas e psicológicas concentradas pelo Simpósio de Hixon, em 1948. (De Castro Barbosa, 2020, p. 93-94)

Na década seguinte, em 1950, o matemático Alan Turing, pai da computação, desenvolveu um molde abstrato de computador capaz de simular a comunicação escrita. Para certificar se o maquinário conseguiria persuadir-se por ser humano ao comunicar-se por meio de textos datilografados, considerado, assim, inteligente, criou o Teste de Turing. (Onody, 2021)

Posteriormente, em 1951, Marvin Minsky concebeu uma espécie de calculadora artificial baseada nas redes neurais humanas e, em 1957, Frank Rosenblatt engendrou o Perceptron, algoritmo que classificava resultados para, em 1958 advir a linguagem Lisp de programação, a qual se tornaria um protótipo da Inteligência Artificial. (De Castro Barbosa, 2020, p. 94)

Ainda nos anos cinquenta, Frank Rosenblatt, ao criar o Perceptron, também inspirado nas estruturas neurais do cérebro humano, possibilitou o primeiro arquétipo de rede neural artificial cuja forma de aprendizado assemelhava-se à dos animais. Contudo, foi o teorema de Novikoff, ao converter as transcrições presentes no algoritmo do Perceptron, que propiciou o avanço da tecnologia que hoje conhecemos como Inteligência Artificial. (Fradkov, 2020, p. 01)

Ultrapassadas as informações históricas, cumpre conceituar a cadeia de ações correspondente ao algoritmo.

Cuida-se de uma sequência delimitada para a resolução dos mais variados problemas, desde uma receita culinária a um projeto de edificação. Nesses casos, há uma solução exata, procedidos os passos na perfeita ordem algorítmica, o resultado será aquele antecipado pela fórmula matemática. Ressalve-se que nem todos os problemas comportam uma solução exata, sendo melhor executados pelo homem, momento no qual o algoritmo sugerirá possíveis soluções, como sucede na medicina diagnóstica. (Sichman, 2021, p. 02-03)

Passando a temática deste tópico, a principal ferramenta utilizada pela Inteligência Artificial corresponde à técnica de Aprendizado de Máquina, a qual permite a acumulação de novas aptidões por sistemas computacionais, melhorando seu desempenho com o aumento do volume de dados aos quais são submetidos. Em apertada síntese, o sistema algorítmico aprenderá por meio das informações que lhes são alimentadas, gerando hipóteses para eventos futuros capazes de prover soluções para a problemas cada vez mais complexos. (Ludermir, 2021)

Quanto às espécies de inteligência artificial, acham-se a IA Focada ou IA Fraca, a qual, muito embora responsável pelo armazenamento de grande volume de dados, somente será apta a executar a tarefa para a qual foi concebida. Noutra banda, a IA Generalizada ou IA Forte, atual estágio da Inteligência Artificial, a partir de técnicas de aprendizado de máquina, adquire novas habilidades distintas da sua programação inicial resolvendo problemas substanciais tal qual o cérebro humano. Por fim, futuro estágio da Inteligência Artificial, por meio da IA Superinteligente, tamanha a habilidade dos sistemas operacionais, que estarão aptos a superar o homem na execução de qualquer encargo. (Ludermir, 2021)

A Inteligência Artificial, por meio do Aprendizado de Máquina, trouxe infundáveis ganhos à sociedade, como a possibilidade de comunicação em tempo real entre pessoas situadas nas mais longínquas partes do globo, a ampliação da eficiência e produtividade das indústrias, estudos cartográficos cada vez mais precisos, a substituição humana para realização de tarefas perigosas, entre outras.

No âmbito da saúde, sistemas de Inteligência Artificial já são capazes de analisar imagens radiográficas, auxiliando profissionais da medicina diagnóstica (Lemes et al., 2019, p. 82). Ainda, o sistema *Watson for Oncology* utiliza Inteligência Artificial para assessorar os processos de tomada de decisão em tratamentos oncológicos (Aikemu, 2021) e, noutra banda, o robô concebido pela chinesa iFlytec foi aprovado no exame admissional de médicos do país. (Saracco, 2017)

Não se olvidam os proveitos que a tecnologia logrou à humanidade, contudo, tendo em vista que o Aprendizado de Máquina somente será possível por meio dos dados fornecidos pelo homem e por conseguinte, sua percepção de mundo, o estabelecimento dos padrões para treinamento de aprendizagem do algoritmo presente nos Sistemas de inteligência Artificial pode se dar de forma enviesada, não constituindo fiel representação da realidade ou reproduzindo concepções preconceituosas.

Explica-se.

Um algoritmo pode ser desenvolvido para distinguir gatos de cachorros, porém, em relação à primeira espécie somente lhes são fornecidos animais pretos e, da segunda, brancos, de modo que o sistema passa a compreender que todos os cachorros são pretos e todos os gatos são brancos ou, noutro tanto, somente distingue as cores que lhes foram fornecidas. (Ludemir, 2021)

Dessa forma, muito embora as técnicas de Aprendizado de Máquina representem o que há de mais moderno na ciência computacional, partindo das premissas inseridas no seu código, somado ao aspecto imprevisível de sua evolução, dentre as possíveis falhas dessa ciência cita-se a superveniência de racismo algoritmo.

Em sede de racismo algoritmo, uma simples busca por cabelos bonitos utilizando-se de ferramentas online trará como resultado, majoritariamente, pessoas brancas e com cabelos que se aproximam ao liso (Nyland, 2023, p. 01), além de casos emblemáticos, como quando o sistema de etiquetamento de imagens do Google identificou duas mulheres negras como gorilas (Da Silva, 2020, p. 8) e a ferramenta Faceapp, do Facebook, foi acusada de embranquecer os usuários. (Sulaimon et al., 2020, p. 01)

Além do caráter imponderável do Aprendizado de Máquina, pois não há completa compreensão pelos desenvolvedores acerca de todas as etapas de evolução e desenvolvimento dos algoritmos iniciais, inexistem, outrossim, transparência quanto aos processos decisórios tomados pelas tecnologias de Inteligência Artificial. Ademais, falta clareza acerca da utilização de dados pessoais tanto pelas empresas provedoras de tecnologia, como pelas entidades públicas e privadas.

Concluído o introito, passa-se à análise do dilema da Black Box e a necessidade de se assegurar o controle sobre sistemas inteligentes, resguardada a transparência das decisões.

### **3 BLACK BOX**

O termo Black Box ou caixa preta possui duplo sentido. O primeiro remete à aviação, correspondente ao dispositivo de gravação instalado nas cabines de comando para fornecer informações em caso de acidente, ao passo que o segundo corresponde aos processos de entrada e saída de componentes em sistemas computacionais nos quais não será possível aferir como uma informação transformou-se em outra, ou seja, desconhece-se o funcionamento interno dos aparatos de inteligência artificial, inexistindo qualquer transparência para o usuário ou observador externo. (Pasquale, 2015, p. 03)

Profundamente transformada pelas tecnologias informação, a sociedade contemporânea passa por uma reconfiguração de princípios organizacionais e, após a disseminação da internet, nesse cenário de constante evolução tecnológica, os dados pessoais se tornaram matéria-prima altamente lucrativa, favorecendo a supremacia de grandes corporações do ramo sobre a nova ordem mundial, sintetizando-se na máxima “*o Facebook define quem somos, a Amazon define o que queremos e o Google define o que pensamos*”. (Pasquale, 2015)

A partir das informações pessoais coletadas pelas plataformas digitais, a vida contemporânea passou a ser regida por algoritmos, substituindo o homem nos mais variados processos decisórios e, uma vez observado que corporações privadas e governos estatais omitem a finalidade do emprego desses dados ou se tais informações poderão ser transferidas

a terceiros, volta-se a atenção à segunda acepção de Black Box, em que pese a ausência de transparência ou clara compreensão pelos indivíduos acerca da utilização de seus dados pessoais.

Exsurge então um modelo de negócio denominado Capitalismo de Vigilância, termo cunhado pela economista alemã Shoshana Zuboff (2015) no artigo “*Big Other: capitalismo de vigilância e perspectivas para uma civilização da informação*”, consoante o qual há um novo impulso de acumulação capitalista realizado através da exploração, tratamento e comercialização de dados. Nessa nova forma de capitalismo baseada em informações, busca-se antecipar e alterar o comportamento humano como estratégia para gerar lucros e controlar o mercado.

Sob a égide do mais recente modelo capitalista, desenrolado através da supervisão meticulosa de cada aspecto da vida das pessoas, as plataformas digitais têm sido maliciosamente utilizadas por governos de estado e corporações privadas para colher informações pessoais dos cidadãos, conferindo-lhes profundo conhecimento acerca dos indivíduos desprovido de qualquer prestação de contas, resultando na formação da sociedade de vigilância. (Frazão, 2019)

A partir da vigilância e da falta de transparência na manipulação dos dados, proliferam-se imbróglios éticos e jurídicos e, para fins de análise neste artigo, concentra-se em dois episódios específicos: O primeiro envolvendo o setor público e o segundo relacionado à iniciativa privada.

Dentre as controvérsias associadas aos algoritmos de caixa preta, cita-se o episódio envolvendo o Sistema de Indicação de Riscos – SyRI implementado pelo governo holandês para prevenir e combater fraude previdenciária. No caso, o algoritmo de processamento de dados analisa, conecta e anonimiza as informações em um ambiente seguro, produzindo relatórios de risco que indicam a probabilidade de fraude à seguridade social.

Não obstante a eficácia do SyRI para prevenção de fraudes previdenciárias, o sistema levantou preocupações importantes, notadamente em relação à privacidade dos dados, à transparência algorítmica e à possibilidade de discriminação, tendo em vista que o relatório produzido é capaz de identificar uma pessoa física ou jurídica suscetível ao cometimento de infrações, sugestionando, inclusive, que sejam investigadas pelas autoridades holandesas. (Lazcoz Moratinos, 2020)

Uma das principais críticas perpetradas em face do SyRI foi a falta de transparência sobre como o sistema realizava previsões e, a partir delas, apresentava um veredito. Cuidava-se, portanto, de nítido exemplo de caixa preta algorítmica, uma vez que muito embora aferível que os dados pessoais estivessem sendo utilizados para identificar potenciais fraudadores, desconhecia-se os processos decisórios do sistema, embaraçando quaisquer questionamentos pelos cidadãos e, em última análise, obstando eventual contestação ou apelação contra as deliberações do SyRI.

A controvérsia foi levada ao Tribunal Distrital de Haia, quando se reconheceu que o SyRI só havia sido empregado pelo governo holandês nos bairros mais humildes, medida que amplificou padrões discriminatórios ao estigmatizar determinados grupos.

Assim, em 5 de fevereiro de 2020, decidiu-se que o uso do SyRI estaria em conflito com o artigo 8º da Convenção Europeia dos Direitos do Homem - CEDH, consoante a qual todos têm direito ao respeito à vida privada e familiar, ao domicílio e correspondência, não se permitindo qualquer interferência da autoridade pública, salvo diante de expressa previsão legal, ou, em se tratando de medida essencial à segurança pública, à economia ou bem-estar do país, à defesa da ordem e prevenção de crimes, à proteção da saúde ou da moral ou, não menos relevante, à proteção dos direitos e liberdades.

Ao final, declarou o Tribunal Distrital de Haia que o sistema SyRI seria incompatível com a CEDH ante a interferência no direito à privacidade dos cidadãos, em que pese o descumprimento pelo governo holandês das garantias de necessidade e proporcionalidade previstas na legislação. (Lazcoz Moratinos, 2020)

Noutra banda, como exemplo de episódio emblemático de caixa preta algorítmica sucedido na iniciativa privada, cita-se o sistema computacional desenvolvido pela Amazon para avaliação de candidatos a vagas de emprego.

No caso, o sistema se baseava em padrões identificados nos currículos encaminhados ao longo de uma década e, em razão da predominância de candidatos homens, perpetuou-se o domínio masculino na medida que os algoritmos adquiriram a predisposição de favorecer candidatos desse gênero. Logo, o sistema penalizava currículos que incluíam a palavra "mulheres", como por exemplo na expressão "capitã do clube de xadrez das mulheres", bem como depreciava candidatos formados em universidades exclusivamente femininas. (Mendoza, 2020, p. 01-03)

À época, por mais que a Amazon tentasse neutralizar termos específicos do funcionamento da inteligência artificial, não impediu que os algoritmos encontrassem outras maneiras discriminatórias para categorizar candidatos, razão pela qual a utilização do sistema foi abolida.

Percebe-se que o sistema de inteligência da Amazon foi abastecido com informações viciadas ao longo de dez anos ao reproduzir vieses - domínio masculino - presentes na estrutura da sociedade. Logo, decisões tomadas por algoritmos, quando treinados com uma base de dado viciado ou não representativa reproduzirão os mesmos vieses, fenômeno denominado pela ciência dos dados como "*entra lixo, sai lixo*". (Da Silva e César, 2022, p.47)

Não obstante a inquestionável relevância da inteligência artificial para a sociedade pós-moderna, os exemplos narrados demonstram a complexidade inerente ao uso de ferramentas de aprendizado de máquina e das práticas de exploração de informações pessoais associadas ao capitalismo de vigilância, cujo obscurantismo alerta para dilemas éticos e jurídicos e impõe acautelada reflexão com vistas a práticas mais transparentes.

#### **4 IMPLICAÇÕES JURÍDICAS E DESAFIOS**

Com a crescente utilização de algoritmos de caixa preta e da coleta de dados pessoais para finalidades desconhecidas, emerge a necessidade de regulação jurídica das plataformas digitais.

Caminhando em direção à normatização das plataformas digitais, o Parlamento Europeu aprovou, em abril de 2022, a Lei dos Serviços Digitais - DAS (Parlamento Europeu, 2022), que prevê a responsabilização das empresas de tecnologia pela propagação de conteúdos ilícitos e propensos a geração de riscos aos usuários ou à sociedade. Para conferir transparência às operações digitais, o marco regulatório determina o fornecimento de informações às autoridades reguladoras relativas à moderação do conteúdo ofertado aos cidadãos e do algoritmo de recomendação. (Ruediger, 2022, p. 09-10)

A DAS, em vigor desde fevereiro de 2024, impõe às plataformas a obrigação de fornecer aos usuários ferramentas para identificação de conteúdos ilegais, como apologia ao terrorismo, discurso de ódio, abuso infantil ou fraudes comerciais, regulando-se, ademais, conteúdos publicitários baseados em dados pessoais sensíveis como religião, gênero, etnia e orientação sexual.

Sancionou-se ademais, em março de 2022, com o fim de controlar a hegemonia das empresas de tecnologia, o Ato de Mercados Digitais - DMA, estabelecendo-se regras mais rígidas para inovação, crescimento e competitividade em mercados digitais, especialmente

atinentes ao comércio eletrônico, governança de algoritmos, ferramentas de busca, anúncios on-line, serviços de mensagens e outras ferramentas digitais. (Ruediger, 2022, p. 09-10)

No Brasil, desde 2014 com o Marco Civil da Internet, Lei nº 12.965/2014 (Brasil, 2014), que disciplina uso da internet no país, escorada nos princípios da garantia da liberdade de expressão, comunicação e manifestação de pensamento, nos termos da Constituição Federal, bem como, dentre outros, nos princípios da proteção da privacidade e dos dados pessoais, intenta-se promover maior segurança aos cidadãos na utilização de ferramentas tecnológicas.

Caminhando para a plena normatização do universo virtual, a promulgação da Lei nº 13.709/2018 (Brasil, 2018), Lei Geral de Proteção de Dados - LGPD, avança na proteção do direito à privacidade ao regulamentar a coleta e tratamento de dados pessoais sensíveis, especialmente atinentes à raça, etnia, religião e orientação política.

Ainda, com amparo na boa-fé, a LGPD determina que as atividades de tratamento de dados pessoais observarão, dentre outros, os seguintes princípios: a) finalidade, garantindo-se o tratamento com propósitos legítimos, específicos, explícitos e informados ao titular, defeso qualquer tratamento posterior cuja finalidade seja incompatível com aquela previamente informada; b) adequação, estabelecendo-se a compatibilidade do tratamento com as finalidades informadas ao titular; c) necessidade, limitando-se o tratamento ao mínimo necessário e; d) livre acesso, garantindo aos titulares, a consulta facilitada e gratuita sobre a forma e a duração do tratamento, bem como sobre a integralidade de seus dados pessoais.

Tal como a garantia de livre acesso, a LGPD determina a observância do princípio da transparência, assegurando aos usuários das plataformas digitais, o acesso a informações claras, precisas e amplamente disponíveis sobre a realização do tratamento e seus respectivos agentes.

Quanto ao princípio da transparência, a nova LGPD confere aos titulares das informações pessoais a necessidade de divulgação sobre a metodologia de tratamento de dados empregada pelas plataformas digitais, permitindo-lhes avaliar a legalidade, legitimidade e segurança em relação aos propósitos de finalidade, adequação e necessidade, capacitando-os a consentir – ou não – com o tratamento de modo informado. (Maldonado e Blum, 2020, p. 116)

Em razão da inobservância do princípio da transparência, somada aos conceitos de caixa preta algorítmica e capitalismo de vigilância, impõe-se atenção à vulnerabilidade dos titulares de dados pessoais em face do comportamento silencioso das empresas de tecnologia, tendo em vista as infinitas possibilidades de processamento das informações. Ademais, considerando-se a rápida evolução das técnicas de tratamento de dados, sobressai-se ainda mais a vulnerabilidade ante a falta de transparência das plataformas digitais, agravando-se a necessidade de fornecimento de informações claras aos usuários do ambiente cibernético.

Ressalte-se, outrossim, que a LGPD determina a obediência aos princípios da a) segurança mediante a adoção de medidas técnicas e administrativas aptas a proteger os dados pessoais de acessos não autorizados; b) prevenção de danos em virtude do tratamento de dados pessoais e, não menos relevante; c) responsabilização e prestação de contas pelo agente acerca da adoção de medidas eficazes ao cumprimento das normas de proteção de dados pessoais e; d) vedação à discriminação algorítmica ao impossibilitar a realização do tratamento para fins segregatícios ilícitos ou abusivos.

Outra inovação trazida pela LGPD que merece destaque é o direito à revisão da decisão tomada unicamente com base no tratamento automatizado de dados. Em sua redação original<sup>1</sup>, a normativa permitia ao usuário solicitar que a revisão do tratamento fosse procedida por pessoa natural ou física. Contudo, após a edição da Medida Provisória nº 869/2018 (Brasil, 2018), a

---

<sup>1</sup> Art. 20. O titular dos dados tem direito a solicitar revisão, por pessoa natural, de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem seus interesses, inclusive de decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade.

norma foi alterada para excluir a necessidade de revisão por pessoa natural e, ao cabo da redação final assentada pela Lei nº 13.853/2019 (Brasil, 2019), muito embora mantido o direito de revisão, dispensa-se a obrigatoriedade de análise humana, a qual será mera faculdade do controlador.

Mister assinalar que não obstante a modificação da redação original do dispositivo atinente ao direito de revisão, subsiste a unicidade do sistema jurídico pátrio, a qual impõe a interpretação constitucional da LGPD associada, outrossim, aos princípios advindos do Código de Defesa do Consumidor, Lei nº 8078/90 (Brasil, 1990). Cuida-se da aplicação do princípio da transparência, conferindo-se ao atingido pelas decisões automatizadas, o direito a informações claras a respeito dos critérios e procedimentos empregados pelas ferramentas tecnológicas, o que somente pode ser garantido por meio da explicação. (Dourado, 2022, p. 04)

Ressalte-se, ademais, que em 2022 o Congresso Nacional aprovou a Emenda Constitucional nº 115 (BRASIL, 2015), acrescentando a redação do artigo 5º, inciso LXXIV da Constituição Federal, o direito fundamental à proteção dos dados pessoais. Contudo, apesar do avanço rumo a transparência algorítmica, sua implementação efetiva enfrenta constantes desafios, a exemplo do Projeto de Lei nº 2.630/2020. (Senado, 2020)

Ainda em debate no Congresso Nacional, o Projeto de Lei nº 2.630/2020 (Senado, 2020), que institui a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet para desestimular o abuso ou manipulação com potencial de dar causa a danos individuais ou coletivos, estabelecendo novas regras para transparência de redes sociais e de serviços de mensagens privada da internet, ficou popularmente conhecido como “PL das Fake News”.

A despeito do esforço regulatório para proteger o cidadão em face da obscuridade dos algoritmos de caixa preta adotados pelas plataformas digitais, o que aproximaria o Brasil do asilo conferido pela Lei dos Serviços Digitais - DAS em vigência nos países da União Europeia, o Projeto de Lei nº 2.630/2020 (Senado, 2020), tendo em vista o escopo de combater a desinformação e prover um ambiente digital mais íntegro, tem sofrido duras críticas, especialmente nos anos eleitorais, subsistindo um espinhoso trajeto até a plena normatização do espaço virtual e da transparência das plataformas digitais.

## 5 CONSIDERAÇÕES FINAIS

Muito embora inegáveis esforços para regulamentar a matéria, como o Projeto de Lei nº 2.630/2020 (Senado, 2020) que, no Brasil, propõe um marco legal para a Inteligência Artificial ou, no âmbito internacional, a Lei dos Serviços Digitais – DAS (Parlamento Europeu), em vigor nos países da União Europeia desde fevereiro de 2024, consoante atual estado da arte, inexistente uma moldura ética presente desde a criação dos sistemas inteligentes.

Isso porque, para além da utilização ética dos algoritmos, a criação e o desenvolvimento dos sistemas computacionais devem amparar-se, desde o nascedouro, sobre rigorosos parâmetros éticos e normativos. Ou seja, a ética deve constituir o próprio código daquele software em construção.

Para a consolidação de uma Inteligência Artificial confiável, bem como prevenir eventuais danos derivados do mau funcionamento das plataformas digitais, a sua gênese deve observar parâmetros éticos, como o princípio da transparência, de modo que as decisões tomadas pelos algorítmicos sejam passíveis de tradução ao público por meio de técnicas explanatórias. Igualmente, o princípio da justiça deve nortear a formação dos algoritmos ao estipular parâmetros não discriminatórios e imparciais e, por fim, o princípio da não maleficência, determinando-se que a manipulação da Inteligência Artificial não se preste a causar mal aos demais indivíduos da sociedade.

Tendo em vista o caráter transnacional das tecnologias em exame, conquanto louvável as iniciativas legislativas mencionadas, sem uma governança cooperativa entre entidades

internacionais públicas e privadas, com a colaboração dos envolvidos no desenho de determinado programa computacional, vinculante a, senão todos, grande parte dos países do globo, a sociedade civil não estará protegida do uso malicioso da Inteligência Artificial.

Sugestiona-se aos legisladores e gestores públicos o compromisso com o uso ético da Inteligência Artificial, promovida a governança descentralizada entre atores da administração pública, respeitados os direitos individuais, o acesso à informação, a proteção de dados, observadas as características culturais de cada país, além de despertar o interesse dos cidadãos para o uso da tecnologia, ressaltando-se a necessidade de educação da população para compreensão das novas ferramentas informáticas e, destarte, dos intrincados sistemas computacionais.

Com vistas à Inteligência Artificial ética, deve-se recorrer aos pilares da governança cooperativa, quais sejam, prestação de contas, que exige a elucidação dos processos decisórios realizados pelas plataformas; responsabilidade, a qual impõe a capacidade de identificação de erros e resultados inesperados pelo sistema, além da proba utilização dos dados pessoais dos usuários e, por fim; transparência, tema abordado neste ensaio quanto ao dilema da Black Box, que impõe a capacidade de tradução dos algoritmos.

Em desfecho, assinala-se que a despeito dos benefícios trazidos pelos sistemas inteligentes, qualquer exercício profissional ou atividade mercadológica cerca-se não somente dos aludidos preceitos éticos, mas, sobretudo, de parâmetros jurídicos, restando urgente que as tecnologias de Aprendizado de Máquina abandonem o esboço de velho-oeste a ser desbravado e assumam contornos civilizatórios por meio de mandatária regulamentação.

## REFERÊNCIAS

AIKEMU, Batuer et al. Artificial intelligence in decision-making for colorectal cancer treatment strategy: an observational study of implementing Watson for Oncology in a 250-case cohort. **Frontiers in Oncology**, v. 10, p. 594182, 2021.

BASTOS, Elísio Augusto Velloso; SOUZA, Jozeh Julyo Queiroz Jatene; DA SILVA, Natalia Gemaque. O direito ao esquecimento enquanto ferramenta jurídica de autodeterminação informativa em face das violações à privacidade decorrentes do capitalismo de vigilância. **1ª Edição, Santa Catarina-2021**, p. 523.

BRASIL. **Emenda Constitucional nº 115, de 10 de fevereiro de 2022**. Diário Oficial da União, 11 fev. 2022. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/constituicao/Emendas/Emc/emc115.htm#:~:text=EMENDA%20CONSTITUCIONAL%20N%C2%BA%20115%2C%20DE,e%20tratamento%20de%20dados%20pessoais](https://www.planalto.gov.br/ccivil_03/constituicao/Emendas/Emc/emc115.htm#:~:text=EMENDA%20CONSTITUCIONAL%20N%C2%BA%20115%2C%20DE,e%20tratamento%20de%20dados%20pessoais). Acesso em: 12 set. 2024.

BRASIL. **Lei nº 13.853, de 08 de julho de 2019**. Diário Oficial da União, 09 jul. 2019. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/\\_Ato2019-2022/2019/Lei/L13853.htm#art1](https://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2019/Lei/L13853.htm#art1). Acesso em: 12 set. 2024.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Diário Oficial da União, 15 ago. 1018. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/\\_ato2015-2018/2018/lei/l13709.htm](https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm). Acesso em: 12 set. 2024.

BRASIL. **Lei nº 12.964, de 23 de abril de 2014**. Diário Oficial da União, 24 abr. 2014. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/ato2011-2014/2014/lei/112965.htm](https://www.planalto.gov.br/ccivil_03/ato2011-2014/2014/lei/112965.htm). Acesso em: 12 set. 2024.

BRASIL. **Lei nº 8.078, de 11 de setembro de 1990**. Diário Oficial da União, 12 set. 1990. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/leis/18078compilado.htm](https://www.planalto.gov.br/ccivil_03/leis/18078compilado.htm). Acesso em: 12 set. 2024.

BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil**. Brasília: Senado, 1988.

DA SILVA, Dandara Ramos Silvestre; CÉSAR, Daniel. 3. Discriminação algorítmica: estudo sobre hipóteses causadoras e soluções regulatórias. **Título: Direitos humanos, ética e democracia na sociedade da informação. III–Discurso de Ódio na Sociedade em Rede Formato: Livro Digital, p. 47**. 2022.

DA SILVA, Tarcízio. Visão computacional e racismo algorítmico: branquitude e opacidade no aprendizado de máquina. **Revista da Associação Brasileira de Pesquisadores/as Negros/as (ABPN)**, v. 12, n. 31, 2020.

DE CASTRO BARBOSA, Xênia. Breve introdução à história da Inteligência Artificial. **Jamaxi**, v. 4, n. 1, 2020.

DOURADO, Daniel de Araujo; AITH, Fernando Mussa Abujamra. A regulação da inteligência artificial na saúde no Brasil começa com a Lei Geral de Proteção de Dados Pessoais. **Revista de Saúde Pública**, v. 56, 2022.

LAZCOZ MORATINOS, Guillermo; CASTILLO PARRILLA, José Antonio. Valoración algorítmica ante los derechos humanos y el Reglamento General de Protección de Datos: el caso SyRI. **Revista chilena de derecho y tecnología**, v. 9, n. 1, p. 207-225, 2020.

LEMES, Marcellle Martins et al. O uso da inteligência artificial na saúde pela Administração Pública Brasileira. **Cadernos Ibero-Americanos de Direito Sanitário**, v. 9, n. 3, p. 88, 2020.

LUDERMIR, Teresa Bernarda. Inteligência Artificial e Aprendizado de Máquina: estado atual e tendências. **Estudos Avançados**, v. 35, 2021.

FRADKOV, Alexander L. Early history of machine learning. **IFAC-PapersOnLine**, v. 53, n. 2, 2020.

FRAZÃO, Ana. Capitalismo de vigilância e Black Box Society. **JOTA**, 2019. Disponível em: <https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/capitalismo-de-vigilancia-e-black-box-society-28022019>. Acesso em: 24 jun. 2023.

MALDONADO, Viviane Nóbrega; BLUM, Renato Ópice. LGPD: Lei geral de proteção de dados comentada. **São Paulo: Revista dos Tribunais**, 2019.

MENDES, Laura Schertel; MATTIUZZO, Marcela. Discriminação algorítmica: conceito, fundamento legal e tipologia. **Direito Público**, v. 16, n. 90, 2019.

MENDOZA, Daniela. Racismo y roles de género, conductas perpetuadas en algoritmos de inteligencia artificial. **Coloquio**, n. 65, p. 131-135, 2020.

NYLAND, Joana Josiane Andriotte Oliveira Lima. Racismo algorítmico: uma revisão de literatura. **Research, Society and Development**, v. 12, n. 2, 2023.

ONODY, Roberto Nicolau. **Teste de Turing e inteligência artificial**. 2021.

PARLAMENTO EUROPEU. **A Lei dos Mercados Digitais e da Lei dos Serviços Digitais da UE em detalhe**. 2022. Disponível em: <https://www.europarl.europa.eu/topics/pt/article/20211209STO19124/a-lei-dos-mercados-digitais-e-da-lei-dos-servicos-digitais-da-ue-explicadas>. Acesso em: 12 set. 2024.

PASQUALE, Frank. The black box society: The secret algorithms that control money and information. **Harvard University Press**, 2015.

RUEDIGER, Marco Aurélio et al. Regulação de plataformas digitais: uma contribuição para a análise do debate nacional frente a um desafio global. 2022.

SARACCO, R. Congrats Xiaoyi. You are now a medical doctor. **IEEE Future Directions**. 2017. Disponível em: <https://cmte.ieee.org/futuredirections/2017/12/02/congrats-xiaoyi-you-are-now-a-medical-doctor/>. Acesso em: 7 fev. 2021.

SENADO FEDERAL. **Projeto de Lei nº 2.630/2020**. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento?dm=8110630&ts=1712173889269&disposition=inline>. Acesso em: 12 set. 2024.

SICHMAN, Jaime Simão. Inteligência Artificial e sociedade: avanços e riscos. **Estudos Avançados**, v. 35, p. 37-50, 2021.

SULAIMON, I. A.; GHONEIM, A.; ALRASHOUD, M. Adaptation of machine learning based fairness algorithm for real time decision in autonomous systems. **Num. Com. Meth. Sci. Eng.** 2, No. 2, 73-77, 2020.

ZUBOFF, Shoshana. Big other: surveillance capitalism and the prospects of an information civilization. **Journal of information technology**, v. 30, n. 1, p. 75-89, 2015.