

EXPLICABILIDADE COMO CONDIÇÃO DE LEGITIMIDADE ALGORÍTMICA

EXPLAINABILITY AS A CONDITION FOR ALGORITHMIC LEGITIMACY

Article received on: 11/19/2025

Article accepted on: 2/19/2026

Sandro Marcos Godoy*

*Universidade de Marília (UNIMAR), Marília, São Paulo, Brasil

Orcid: [c](#)

sandromgodoy@uol.com.br

Luana Pereira Lacerda*

*Universidade de Marília (UNIMAR), Marília, São Paulo, Brasil

Orcid: <https://orcid.org/0009-0005-1520-9215>

luanaplacerda@gmail.com

Victoria Cássia Mozaner*

*Universidade de Marília (UNIMAR), Marília, São Paulo, Brasil

Orcid: <https://orcid.org/0009-0005-1023-4307>

victoriamozaner@gmail.com

Gabriel Guerra Miranda Muzeka dos Santos*

*Universidade de Marília (UNIMAR), Marília, São Paulo, Brasil

Orcid: <https://orcid.org/0009-0009-8278-6708>

gabrielmuzeka3@gmail.com

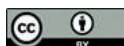
The authors declare that there is no conflict of interest

Resumo

Este artigo analisa a relação entre sistemas de Machine Learning e o direito à explicação nas decisões automatizadas, com enfoque na proteção dos direitos fundamentais no contexto da Lei Geral de Proteção de Dados. Parte-se do problema decorrente da crescente utilização de algoritmos na tomada de decisões que afetam diretamente os indivíduos, muitas vezes caracterizadas por elevado grau de opacidade e dificuldade de compreensão dos critérios utilizados. O objetivo da pesquisa consiste em examinar em que medida o direito à explicação pode atuar como mecanismo de transparência e controle jurídico sobre decisões automatizadas. Para alcançar esse objetivo, adotou-se abordagem qualitativa, com utilização do método dedutivo e da técnica de pesquisa bibliográfica, a partir da análise de literatura especializada em inteligência artificial, proteção de dados e direitos fundamentais. Os resultados indicam que, embora os sistemas de Machine Learning ampliem significativamente a capacidade de análise e processamento de dados, eles também introduzem desafios relevantes relacionados à opacidade algorítmica, à possibilidade de discriminação algorítmica e à limitação da autonomia informativa dos indivíduos. Conclui-

Abstract

This article analyzes the relationship between Machine Learning systems and the right to explanation in automated decision-making, with a focus on the protection of fundamental rights within the framework of the Brazilian General Data Protection Law (LGPD). The study addresses the problem arising from the increasing use of algorithms in decision-making processes that directly affect individuals, often characterized by a high degree of opacity and limited understanding of the criteria employed. The objective of this research is to examine to what extent the right to explanation can function as a mechanism of transparency and legal oversight over automated decisions. To achieve this objective, the study adopts a qualitative approach, employing the deductive method and bibliographic research techniques based on the analysis of specialized literature on artificial intelligence, data protection, and fundamental rights. The findings indicate that, although Machine Learning systems significantly expand the capacity for data processing and analysis, they also introduce important challenges related to algorithmic opacity, the potential for algorithmic discrimination, and limitations on individuals' informational autonomy. The study



se que o direito à explicação constitui instrumento essencial para a promoção da transparência, da responsabilização e da proteção dos direitos fundamentais no uso de sistemas de decisão automatizada.

Palavras-chave: Decisões Automatizadas. Dados Pessoais. Direito à Explicação. Direitos Fundamentais. Machine Learning.

concludes that the right to explanation constitutes an essential mechanism for promoting transparency, accountability, and the protection of fundamental rights in the context of automated decision-making systems.

Keywords: Automated Decisions. Fundamental Rights. Machine Learning. Personal Data. Right to Explanation.

1 INTRODUÇÃO

A expansão das tecnologias digitais e o avanço dos sistemas de Inteligência Artificial têm produzido transformações profundas nas formas de tomada de decisão na sociedade contemporânea. Nesse cenário, os sistemas baseados em *Machine Learning* passaram a ocupar posição central na análise de grandes volumes de dados e na formulação de decisões automatizadas capazes de influenciar diferentes dimensões da vida social. Tais sistemas são atualmente utilizados em áreas como concessão de crédito, seleção de candidatos para vagas de trabalho, definição de perfis de consumo e formulação de políticas públicas, demonstrando o crescente papel dos algoritmos na organização das relações sociais e econômicas.

Esse contexto suscita uma problemática relevante no campo jurídico, especialmente no que se refere à compatibilidade entre decisões automatizadas e a proteção dos direitos fundamentais. A utilização de algoritmos de aprendizado de máquina frequentemente envolve sistemas complexos e de difícil compreensão, caracterizados por elevado grau de opacidade quanto aos critérios utilizados na tomada de decisões. Diante desse cenário, emerge como questão central o direito à explicação, entendido como a possibilidade de o indivíduo compreender os fundamentos que sustentaram decisões automatizadas que afetem seus interesses.

Nesse sentido, coloca-se a seguinte questão de pesquisa: em que medida o direito à explicação pode atuar como instrumento de proteção dos direitos fundamentais diante da crescente utilização de sistemas de *Machine Learning* na tomada de decisões automatizadas? A formulação desse problema revela uma tensão estrutural entre, de um lado, a eficiência e a capacidade analítica proporcionadas pelas tecnologias baseadas em

dados e, de outro, as exigências de transparência, *accountability* e proteção da autonomia informativa dos indivíduos.

A relevância do tema decorre do fato de que decisões automatizadas têm impactado progressivamente o acesso a oportunidades sociais, econômicas e políticas, podendo produzir efeitos significativos sobre a vida das pessoas. A ausência de transparência nos sistemas algorítmicos, aliada à possibilidade de reprodução de vieses presentes nas bases de dados utilizadas para o treinamento dos modelos, levanta preocupações relacionadas à discriminação algorítmica, à limitação da autonomia informativa e à potencial violação de direitos fundamentais. Nesse cenário, o debate sobre o direito à explicação assume importância estratégica no âmbito da governança das tecnologias digitais e da proteção de dados pessoais.

Diante desse contexto, o presente artigo tem como objetivo geral analisar a relação entre os sistemas de *Machine Learning* e o direito à explicação nas decisões automatizadas, investigando seus impactos jurídicos e éticos na proteção dos direitos fundamentais. Como objetivos específicos, busca-se: (i) examinar os fundamentos conceituais e operacionais do Machine Learning e sua crescente utilização em processos decisórios automatizados; (ii) compreender o papel do direito à explicação no contexto da Lei Geral de Proteção de Dados; (iii) identificar os principais desafios relacionados à implementação desse direito, especialmente no que se refere à opacidade algorítmica, aos vieses discriminatórios e às limitações regulatórias existentes; e (iv) refletir sobre as perspectivas éticas e institucionais necessárias para o desenvolvimento responsável de sistemas de Inteligência Artificial.

A contribuição científica deste estudo consiste em articular a discussão tecnológica sobre sistemas de aprendizado de máquina com o debate jurídico relacionado à proteção de dados pessoais e aos direitos fundamentais, evidenciando como o direito à explicação pode funcionar como mecanismo de controle e transparência em ambientes decisórios automatizados. Ao explorar essa intersecção, o artigo busca contribuir para o avanço do debate jurídico acerca dos limites e das possibilidades da regulação das tecnologias baseadas em dados no contexto do Estado Democrático de Direito.

Para alcançar tais objetivos, o artigo foi estruturado em quatro seções principais, além desta introdução e da conclusão. Na primeira seção, apresentam-se os conceitos fundamentais relacionados ao *Machine Learning* e sua importância nos sistemas

contemporâneos, bem como sua intersecção com o direito à explicação. Na segunda seção, examinam-se os principais desafios relacionados à implementação desse direito, com destaque para a capacidade e o dimensionamento das explicações, a ocorrência de vieses e discriminações algorítmicas e as questões legais e regulatórias envolvidas. Na terceira seção, discutem-se as perspectivas futuras relacionadas à regulação ética da Inteligência Artificial, abordando especialmente a importância da participação pública, do engajamento institucional e da observância dos princípios fundamentais no desenvolvimento dessas tecnologias.

Do ponto de vista metodológico, a pesquisa foi desenvolvida por meio de abordagem qualitativa, utilizando-se do método dedutivo e da técnica de pesquisa bibliográfica. Foram analisadas obras doutrinárias, artigos científicos e documentos normativos relacionados à Inteligência Artificial, à proteção de dados pessoais e ao direito à explicação, com o objetivo de compreender os desafios jurídicos decorrentes da crescente utilização de sistemas de *Machine Learning*. A adoção desse percurso metodológico permite examinar criticamente os fundamentos conceituais, normativos e éticos que estruturam o debate contemporâneo sobre decisões automatizadas.

Assim, a análise proposta neste estudo parte da compreensão de que o avanço das tecnologias algorítmicas não representa apenas uma transformação tecnológica, mas também um desafio jurídico e institucional. É justamente nessa intersecção entre inovação tecnológica, proteção de dados pessoais e garantia dos direitos fundamentais que se insere a discussão desenvolvida nas seções seguintes.

2 CONCEITOS DE *MACHINE LEARNING*

Aborda-se, inicialmente, as interfaces do *Machine Learning*, amplamente conhecido como “aprendizado de máquina”, enquanto uma das vertentes compreendidas no âmbito da Inteligência Artificial, com reflexos significativos nos sistemas modernos e em uma diversidade de áreas essenciais para a sociedade contemporânea.

Conforme Paixão (2022), o aprendizado de máquina, também denominado *Machine Learning* (ML), configura-se como um ramo da inteligência artificial (IA) voltado ao estudo e à construção de algoritmos computacionais baseados em dados, cuja finalidade precípua consiste na criação de sistemas capazes de aprender a partir de um

banco de dados previamente definido e, ao final, gerar modelos de predição, classificação ou detecção.

Segundo Schleder e Fazzio (2020), o aprendizado de máquina (*machine learning*, ML) emerge nesse contexto como uma nova classe de ferramentas estatísticas voltadas à identificação automatizada de padrões em diferentes conjuntos de dados, contribuindo, assim, para a produção e a ampliação do conhecimento científico.

Tais apontamentos conduzem ao reconhecimento dos avanços inicialmente proporcionados pelo aprendizado de máquina, especialmente no que se refere à criação de sistemas concebidos para projetar decisões automatizadas, cujos contornos contemporâneos demandam análise e reflexão aprofundadas.

2.1 Definição e sua importância nos sistemas modernos

À medida que as tecnologias têm impulsionado uma ampla gama de transformações em diversos setores e atividades, à luz da Inteligência Artificial descortina-se um conjunto significativo de possibilidades, sobretudo quando se discute a existência de máquinas inteligentes capazes de gerar decisões independentes da vontade humana.

Certo é que o “*Machine Learning* (ML), também chamado de aprendizado de máquina, é considerado uma subárea da Inteligência Artificial capaz de aprender com os dados, a ponto de prever resultados sem interferência humana” (Gonçalves, 2022, p. 4).

O aprendizado por máquina opera, portanto, por meio do amplo volume de dados armazenados e transmitidos, permitindo que o próprio sistema produza uma variedade de resultados que implicam na construção de perfis, categorias e demais aspectos automaticamente inferidos pela máquina, gerando, assim, respostas cujos efeitos sociais, em um primeiro momento, não são plenamente previsíveis. Assim sendo:

Na base de toda essa revolução está a Inteligência Artificial e os múltiplos usos que ela é capaz de fazer do grande volume de dados disponíveis na internet ou fora dela, sobretudo por meio do chamado Aprendizado de Máquina (*Machine Learning*). O conjunto dos efeitos sociais desses usos ainda é um território desconhecido, pleno de possibilidades, de esperanças e de muitos receios também (Reis; Furtado, 2022, p. 5).

Diante desses apontamentos, observa-se que a Inteligência Artificial tem se configurado como uma verdadeira “revolução” no âmbito das transformações desencadeadas pelo surgimento de novas tecnologias, guardando, entretanto, certa indefinição quanto às diferentes formas que podem assumir os variados usos por parte de “máquinas” que simulam a inteligência humana.

Isso ocorre, inclusive, porque “na era das hipertecnologias, é imprescindível perquirir se o emprego da IA promove o direito à proteção de dados ou se, ao revés, serve a práticas discriminatórias, atingindo os direitos da pessoa ou dos grupos sociais que ela integra e representa” (Lima; Sá, 2020, p. 231).

Sustenta-se, portanto, que, por se tratar de autocomandos inteligentes, isto é, sistemas guiados por si mesmos, não há possibilidade, a priori, de assegurar de forma plena e efetiva que as “respostas”, em termos de categorização produzidas no âmbito do *Machine Learning*, representem necessariamente avanços em benefício da sociedade.

Por sua vez, os sistemas baseados em *Machine Learning* têm maior grau de complexidade, se comparados aos sistemas de algoritmo de análise, pois são capazes de prever ou generalizar padrões apreendidos com base num conjunto de dados utilizados para treinar o sistema. A construção algorítmica nesse sistema não depende de dados previamente escolhidos por seus operadores, isto é, o sistema aprende baseado na sua interação com um ambiente externo e dinâmico, por meio do qual realiza correlações com o objetivo de reconhecer padrões. Uma das diferenças entre o sistema de *Machine Learning* e o de *analytics* é que o primeiro é dotado de capacidade para analisar, fazer correlações e buscar padrões com base em dados não organizados, tais como fotos, vídeos ou textos coletados de *smartphones* (Alves Pinto, 2020, p. 460.)

Evidencia-se, desse modo, que o próprio sistema constrói seu repertório e suas respostas a partir de inferências derivadas da apropriação de uma ampla gama de informações e dados veiculados ou relacionados aos usuários, compilando tais dados com o objetivo de produzir proposições ou decisões automáticas que, especialmente no caso de sistemas não supervisionados, não se encontram necessariamente vinculadas à observância de comandos previamente estabelecidos.

Daí a necessidade de observar, com base em Alves Pinto (2020), que, de maneira geral, os sistemas de *Machine Learning* podem ser divididos em dois subgrupos: os supervisionados e os não supervisionados. No grupo supervisionado, o aprendizado da máquina ocorre por meio de correlações iniciais definidas por seres humanos; nesse caso,

mapeia-se um conjunto de informações para um determinado conjunto de resultados, incluindo métodos como árvores de classificação, redes neurais ou regressão linear. Em ambientes dinâmicos, são necessárias diversas interações iniciais com o objetivo de ajustar o sistema de IA por parte de alguém que detenha domínio específico da área de aplicação do sistema, até que este produza resultados mais precisos e minimamente satisfatórios.

No grupo não supervisionado, por sua vez, as informações que alimentam o sistema são rotuladas, mas os resultados não. Isso significa que o algoritmo precisa inferir a estrutura subjacente dos próprios dados, com o intuito de agrupar seus elementos em categorias semelhantes sem que se conheçam previamente a quantidade e a estrutura desses dados. Nesse modelo, dispensa-se o ajuste inicial realizado por um especialista da área na qual se pretende desenvolver a tecnologia de IA.

Não obstante a relevância da Inteligência Artificial diante do enorme volume de dados e informações que seria praticamente impossível de ser processado por um único ser humano, torna-se necessário considerar as implicações decorrentes da apropriação crescente de sistemas de *Machine Learning* em diferentes áreas essenciais para a coletividade.

Daí a razão pela qual Ludermir (2021) aponta que a inferência indutiva constitui um dos principais métodos utilizados para derivar novo conhecimento e prever eventos futuros no contexto do Aprendizado de Máquina. Nesse cenário, a generalização pode não ocorrer de maneira correta na inferência indutiva, razão pela qual as chances de acerto das generalizações aumentam proporcionalmente à qualidade dos dados utilizados.

Destaca-se ainda, com fundamento em Pinheiro e Coelho (2020), que, para além do armazenamento de dados e da segurança da informação, a inteligência artificial que anteriormente era utilizada predominantemente para decisões de baixo risco, como recomendações de propagandas, filmes ou músicas, passou a integrar diversas estratégias mercadológicas e soluções tecnológicas que atualmente auxiliam, por exemplo, na definição de scores de crédito, na seleção de candidatos para vagas de trabalho, na definição de prioridades em setores hospitalares de emergência ou mesmo na autorização de entrada de indivíduos em determinados países, entre outras situações consideradas de alto risco.

Reconhece-se, portanto, que os sistemas de *Machine Learning* não supervisionados têm ocupado posição de destaque em diversas instituições, uma vez que inúmeras empresas e organizações passaram a utilizar o aprendizado por máquina com o objetivo de agilizar processos e atividades. Todavia, tal dinâmica pode implicar na definição de categorias baseadas em critérios que podem induzir discriminações, exclusões e preferências, isto é, limitações ao exercício de direitos fundamentais, sem que as pessoas afetadas compreendam plenamente os mecanismos que conduzem a tais decisões.

No entendimento de Ludermir (2021), no aprendizado não supervisionado os exemplos são fornecidos ao algoritmo sem rótulos previamente definidos. O algoritmo, então, agrupa os exemplos de acordo com as similaridades de seus atributos, analisando os dados fornecidos e buscando identificar possíveis agrupamentos ou *clusters*. Após a determinação desses agrupamentos, torna-se necessária, em geral, uma análise adicional para compreender o significado de cada agrupamento no contexto do problema que está sendo analisado.

Em linhas gerais, portanto, respostas automatizadas podem gerar sérios entraves em razão da apropriação de dados e informações internas e externas de um determinado usuário, colocando-o à mercê de um sistema lógico e mercadologicamente estruturado para atender interesses predominantes da lógica capitalista vigente, ainda que exista um arcabouço jurídico voltado à proteção dos princípios e garantias fundamentais.

2.2 Direito à explicação e sua importância ética e jurídica

Em uma sociedade pautada pela informação, na qual, via de regra, cada vez mais as operações e transações realizadas pelos usuários ocorrem no ambiente on-line, seja na esfera privada, seja na esfera pública, o direito à explicação precisa ser compreendido sob os vieses ético e jurídico.

Do ponto de vista ético, reconhece-se que o indivíduo, enquanto membro da sociedade, precisa dispor do entendimento e da clareza necessários no que diz respeito à forma pela qual seus dados são tratados por organizações de diferentes naturezas, constituindo a obscuridade em relação a tais mecanismos uma postura antiética, que não deve se estabelecer entre as partes interessadas.

Em termos jurídicos, cabe atentar para o fato de que o ordenamento jurídico constitucional pátrio, mesmo antes da entrada em vigor da Lei Geral de Proteção de Dados (lei nº 13.709, de 14 de agosto de 2018), já assegurava ao indivíduo a possibilidade de obter informações sobre fatos que digam exclusivamente a seu respeito. Assim, a inviolabilidade no acesso a tais informações somente se justifica quando presentes interesses maiores do Estado e/ou da coletividade.

Sendo assim, sem descuidar da tutela constitucional que gravita em torno do direito à privacidade, cabe enfatizar que o direito à explicação se encontra contemplado expressamente no artigo 20 da Lei nº 13.709, de 14 de agosto de 2018, que assim dispõe:

Art. 20 (Lei nº 13.709/2018). O titular dos dados tem direito a solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem seus interesses, incluídas as decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade (Brasil, 2018, s/n).

É digno de nota, diante de tal dispositivo legal, que o titular de dados possui o direito expresso de solicitar a revisão de decisões tomadas exclusivamente com base em tratamento automatizado de dados pessoais quando estas venham a colidir com seus interesses. Evidencia-se, nesse contexto, a preocupação do legislador com a dinâmica que orienta tais decisões, especialmente no que se refere à criação de perfis pessoais e profissionais, bem como à influência na concessão de crédito ou no enquadramento do indivíduo em determinada categoria de consumo, ou ainda na atribuição de características relacionadas à sua personalidade.

É possível, portanto, sustentar que a gênese que fundamenta o direito à explicação, diante do uso crescente de sistemas de *Machine Learning* no âmbito da Inteligência Artificial, reside justamente na necessidade de esclarecer ao sujeito afetado quais critérios e elementos sustentaram decisões tomadas unilateralmente, isto é, decisões baseadas em processos de aprendizado de máquina.

Conforme Porto Júnior (2019), é nesse cenário que o direito à explicação pode ser compreendido tanto como (i) uma ferramenta jurídica de caráter preventivo, capaz de dissuadir o uso de sistemas de inteligência artificial discriminatórios ao permitir a fiscalização de seu modo de funcionamento; quanto como (ii) uma ferramenta de auxílio à atribuição de responsabilidade e à conseqüente compensação por eventuais danos,

podendo servir, por exemplo, como suporte para a demonstração do nexos causal no âmbito da responsabilidade civil. Tal perspectiva gera base para uma ampla gama de discussões tanto no campo técnico, acerca das possibilidades e limitações de se produzirem explicações sobre o funcionamento de sistemas de IA quanto no campo jurídico.

Reconhece-se, diante desses apontamentos, a necessidade de compreender que o direito à explicação tende a projetar-se a partir de uma dupla finalidade: de um lado, evitar o desenvolvimento de novas interfaces tecnológicas que possam implicar novas formas de discriminação e diferenciação; de outro, possibilitar a atribuição de responsabilidades quando decisões automatizadas, ao excluírem o indivíduo do exercício de seus direitos fundamentais, venham a ocasionar danos. Nesse contexto, discute-se, especialmente, o alcance e as limitações de natureza técnica e jurídica relacionadas à utilização de sistemas de aprendizado de máquina não supervisionados.

2.3 A intersecção entre a *Machine Learning* e o direito à explicação

Discutir a inter-relação entre o *Machine Learning* e o direito à explicação demanda reconhecer a extensa gama de aplicações que essas interfaces da Inteligência Artificial têm desenvolvido ao longo dos últimos anos.

Diante desse contexto:

O uso de programas de aprendizado por máquinas, conhecido pelo termo *Machine Learning*, permite que sejam criados sistemas de Inteligência Artificial (IA) que desenvolvem a capacidade de tomadas de decisão absolutamente autônomas em relação à interferência humana. Isto é, torna-se possível por meio de tratamento de dados em massa – *inputs* – o desenvolvimento de autoaprendizagem das máquinas – i.e. programas e sistemas – que permite o alcance de determinados resultados – *outputs* –, independentemente de qualquer mediação por um ser humano. Ou seja, o próprio sistema alcança resultados por meio de processos dedutivos e análises estatísticas que vão sendo determinados com base em correlações realizadas pela IA. Esses resultados, em não raras vezes, são obtidos sem que seja possível, a priori, reconhecer os padrões adotados pela IA para a análise de dados selecionados e o modo de trabalho que levaram a esses *outputs* (Mulholland; Frajoh, 2020, p. 1).

Reforça-se, portanto, o entendimento de que, se por um lado existe a possibilidade de utilizar sistemas capazes de alcançar resultados por si próprios, valendo-se de deduções

e análises estatísticas, por outro lado sabe-se que não há limites precisos ou padrões previamente estabelecidos para a apresentação dos resultados após a realização dessas correlações.

Nesse sentido, ainda que tais processos frequentemente passem despercebidos, inúmeras camadas da sociedade tendem a sofrer os efeitos decorrentes dessas decisões automatizadas.

Partindo desse pressuposto:

Por conta disso, as decisões têm sofrido um intenso processo de automatização baseada em critérios na maioria das vezes não conhecidos ou bem explicados por seus criadores, de modo que passam a ter grande influência no dia a dia das pessoas sem que elas necessariamente percebam (Alves Pinto, 2020, p. 50).

Emerge, desse modo, diante da ausência de critérios claros que ensejaram a categorização de um indivíduo, a necessidade de efetivação do direito à explicação, em um contexto no qual se busca esclarecer as proposições e os argumentos que conduziram a determinada decisão, bem como a eventual predisposição do indivíduo a determinada situação futura.

Diante do exposto:

Com efeito, métodos sofisticados de tratamento de dados, chamados genericamente de Inteligência Artificial, permitem a recopilação de dados dispersos para reconstituir ações humanas e para analisar, prever ou mesmo induzir comportamentos futuros. Enfim, permitem formar uma imagem completa do indivíduo, a partir dos vestígios digitais dos seus movimentos nas redes, prognosticando as suas ações, características, interesses e até pensamentos (Reis; Furtado, 2022, p. 4).

Desta forma, é indiscutível indagar até em que ponto se torna possível atestar por meio de certas deduções que tal indivíduo tende a ostentar tal modelo de comportamento no futuro.

É pertinente questionar até que ponto é possível afirmar, por meio de determinadas deduções, que um indivíduo tende a apresentar um determinado padrão de comportamento no futuro.

Torna-se, portanto, indispensável atentar, com fundamento em Reis e Furtado (2022), para o fato de que a decisão automatizada, para além do algoritmo, é fortemente

influenciada pelos dados, especialmente por três tipos distintos: (a) dados de treinamento; (b) dados de entrada; e (c) dados de *feedback*. Nesse contexto, os dados de treinamento constituem o *background* do modelo em fase anterior à sua implementação; os dados de entrada, já na fase de aplicação, indicam ao modelo as informações presentes no ambiente externo; e os dados de saída correspondem ao resultado do processo decisório artificial. Esses dados de saída podem, por sua vez, retornar ao sistema como *feedback* positivo ou negativo, permitindo que a máquina seja ajustada automaticamente ou pelo próprio desenvolvedor.

Em qualquer hipótese, não se pode admitir que os critérios adotados em cada uma dessas etapas não sejam passíveis de explicação e compreensão, especialmente por parte do usuário interessado ou afetado.

Como salientado por Pinheiro e Coelho (2020), esse aspecto assume grande relevância tanto sob o ponto de vista técnico, relacionado à programação e à construção dessas soluções, quanto sob o ponto de vista jurídico e ético, especialmente diante do aumento das pressões sociais e regulatórias por maior transparência, auditabilidade e explicação das decisões automatizadas. Tais demandas tornam-se ainda mais relevantes diante de modelos considerados opacos ou classificados como *black box*, caracterizados pela dificuldade de compreensão sobre como determinada decisão foi alcançada ou quais insumos foram determinantes para sua formação.

É indispensável, portanto, considerar que a complexidade inerente a esses sistemas pode impedir que os usuários afetados tenham a segurança de que seus dados foram legitimamente interpretados e analisados segundo critérios transparentes e juridicamente aceitáveis.

Nesse sentido, observa-se que:

Essa natureza adaptativa tem se tornado mais comum, graças a complexos sistemas de inteligência artificial e aprendizado de máquina capazes de influenciar as conclusões intermediárias – a ponto de não ser mais possível prever os resultados finais ou entender sua lógica subjacente. Essa opacidade impede que as pessoas entendam e verifiquem se seus dados pessoais são tratados de forma legítima, adequada e proporcional (MONTEIRO, 2018, p. 2).

Fala-se, assim, em opacidade ou, mais precisamente, na ausência de transparência em grau suficiente para comprometer a efetividade do direito à explicação previsto na Lei Geral de Proteção de Dados.

Mais ainda, conforme destaca Monteiro (2018), um dos setores da economia que mais se vale do uso e do tratamento de dados pessoais para a viabilização de decisões automatizadas é o setor de consumo. Nesse contexto, tais mecanismos são frequentemente utilizados para compreender o comportamento do consumidor e influenciar seus hábitos, situação que pode gerar práticas indesejadas, abusivas ou prejudiciais ao indivíduo.

Nesse sentido, a literatura internacional destaca que o chamado *right to explanation* emerge como uma resposta jurídica à crescente utilização de decisões automatizadas baseadas em sistemas complexos de aprendizado de máquina, cuja lógica frequentemente não é acessível aos indivíduos afetados, o que compromete a transparência e a possibilidade de contestação (Goodman; Flaxman, 2017, p. 6-7).

Em termos concretos, observa-se que o acesso a bens e serviços pode ser significativamente influenciado por decisões automatizadas baseadas em probabilidades, estatísticas e inferências, inclusive no que se refere à previsão de determinados eventos futuros.

3 DESAFIOS NA IMPLEMENTAÇÃO DO DIREITO À EXPLICAÇÃO

Coexistem inúmeros desafios no que diz respeito à implementação do direito à explicação, a começar pela cultura de muitas organizações que procuram manter suas políticas à margem da compreensão dos usuários.

Diante desses aspectos, reconhece-se a convergência entre opacidade, escala e dano:

A opacidade está relacionada à falta de transparência, como se tais mecanismos fossem uma verdadeira caixa-preta e de difícil compreensão para os sujeitos que são afetados por eles. A escala, relaciona-se com a capacidade de atingir milhares de pessoas e o dano refere-se aos impactos que esses mecanismos podem gerar aos indivíduos, como, por exemplo, dificuldades em ingressar na universidade, conseguir empréstimos, ser sentenciado a prisão, encontrar um emprego, em decorrência da utilização de algoritmos que propagam injustiças (GONÇALVES, 2022, p. 5).

Reitera-se, portanto, as implicações decorrentes das interfaces que permeiam o acesso aos critérios adotados, em um contexto no qual o uso crescente de sistemas informatizados acaba por retroalimentar tais mecanismos.

É nesse cenário que Pinheiro e Coelho (2020) enfatizam que um dos problemas mais complexos enfrentados pela humanidade nos últimos anos consiste no uso e na proteção de dados pessoais, especialmente pelo fato de que os principais insumos dessa revolução tecnológica são justamente as pessoas, seus hábitos e a produção diária de conteúdo em suas atividades profissionais e em sua vida cotidiana.

É possível, portanto, constatar que, na mesma proporção em que as organizações se valem do uso crescente do tratamento de dados pessoais, não se abrem aos usuários espaços suficientes para que estes compreendam como tais mecanismos efetivamente operam.

Nesse sentido, Pinheiro e Coelho (2020) destacam que, para além das questões envolvendo práticas ilegais, abusivas ou reputacionais, existem desafios relacionados à dificuldade de compreensão da diferença entre correlação e causalidade, bem como questões ligadas a vieses (bias), à codificação de preconceitos, à utilização de bases de dados não balanceadas e à presença de informações incompletas. Ademais, tais desafios também abrangem aspectos relacionados ao respeito à diversidade e à possibilidade de ocorrência de discriminação, sem descuidar da responsabilidade civil nesses casos, que frequentemente está vinculada a quem treinou a máquina ou selecionou os bancos de dados utilizados para determinada finalidade, considerando a necessidade de supervisão, balanceamento e seu impacto na modelagem dos sistemas.

Certo é que a indefinição de critérios e/ou a aplicação indiscriminada de decisões automatizadas devem ser analisadas à luz do princípio da transparência.

Como observa Monteiro (2018), decisões automatizadas controlam cada vez mais aspectos relevantes da vida cotidiana. Elas estão presentes na definição da melhor rota para evitar congestionamentos, na seleção de candidatos para vagas de trabalho e até mesmo na formulação de políticas públicas. Esses são exemplos de atividades quase inteiramente gerenciadas por algoritmos que podem exercer impactos significativos na vida dos cidadãos; todavia, a crescente presença desses sistemas no cotidiano é acompanhada de baixos níveis de transparência quanto ao seu funcionamento, o que torna

mais complexa a identificação de práticas abusivas, discriminatórias ou mesmo monopolísticas.

Reconhece-se, portanto, que a complexidade e a ausência de transparência na utilização de sistemas de *Machine Learning* constituem desafios relevantes a serem enfrentados, especialmente para possibilitar a aferição da ocorrência de práticas que possam favorecer determinados setores em detrimento dos interesses dos usuários.

3.1 Capacidade e dimensionamento do direito à explicação

Falar sobre a capacidade e o dimensionamento do direito à explicação implica compreender que a explicação consiste na oferta de elementos claros e precisos que permitam ao usuário identificar os critérios adotados nas decisões que afetam seus interesses.

Preceitua Monteiro (2018) que o direito à explicação deve ser compreendido como corolário do princípio da transparência, aplicável não apenas no território nacional, mas também no âmbito dos sistemas de proteção de dados pessoais existentes em todo o mundo, assegurando aos titulares de dados a possibilidade de conhecer, com precisão, o modo pelo qual se opera o tratamento das informações que lhes dizem respeito.

Tal condição deve compreender, necessariamente, a oferta ao usuário de informações claras e precisas, capazes de permitir a plena compreensão acerca da legitimidade do procedimento adotado.

3.2 Viés e discriminação algorítmica

Uma das maiores preocupações no campo das decisões automatizadas por algoritmos consiste na ocorrência de discriminação ou na adoção de critérios de diferenciação indevidos.

Diante desse cenário:

A capacidade de tratamento de dados pessoais das mais diversas ordens vem aumentando exponencialmente, principalmente devido ao advento de tecnologias avançadas de inteligência artificial, com o uso de algoritmos sofisticados e com a possibilidade de aprendizado por máquinas (*Machine*

Learning). Significa dizer que o tratamento de “*big data*” – literalmente, grandes bases de dados – por meio de técnicas computacionais cada vez mais desenvolvidas pode levar a análises probabilísticas e resultados que, ao mesmo tempo em que atingem os interesses de uma parcela específica da população, retiram a capacidade de autonomia do indivíduo e o seu direito de acesso ao consumo de bens e serviços e a determinadas políticas públicas, por exemplo (MULHOLLAND; FRAJHOF, 2020, p. 30.)

É possível, portanto, observar a dinâmica dos sistemas de Machine Learning, especialmente no que se refere à forma pela qual, a partir de decisões automatizadas, ocorre o enquadramento do usuário em determinadas categorias, o que pode resultar na limitação ou mesmo na exclusão do exercício de direitos fundamentais.

Nesse sentido, destaca-se que: “Ocorre que os sistemas de predição e outros sistemas de IA não estão livres de distorções no resultado. Afinal, os dados são inseridos por programadores humanos que, mesmo involuntariamente, podem contaminá-los com seus preconceitos” (Lima; Sá, 2020, p. 231).

No mesmo sentido, estudos internacionais demonstram que sistemas baseados em análise de grandes volumes de dados podem reproduzir e até amplificar desigualdades sociais preexistentes, fenômeno conhecido como *disparate impact*, no qual decisões aparentemente neutras resultam em efeitos discriminatórios sobre grupos específicos (Barocas; Selbst, 2016, p. 674).

Como se observa, embora os benefícios proporcionados pela Inteligência Artificial sejam amplamente reconhecidos, também se evidenciam riscos relevantes que podem afetar a autonomia do indivíduo no contexto do sistema social contemporâneo.

Sendo assim:

Com efeito, permitem-se às pessoas várias formas de conexões entre elas e de acesso a outros países, culturas e novas experiências; em contrapartida, existem também vários sinais de uma tendência à homogeneização nas formas de pensar, organizar a vida e planejar o futuro. De certo modo – e até contraditoriamente –, quanto mais variadas são as informações adquiridas nesse processo de alargamento dos horizontes, mais homogêneos se tornam os comportamentos experimentados pela sociedade (Alves Pinto, 2020, p. 44).

Ressalta-se, nesse contexto, a face perniciosa da discriminação algorítmica, que, conforme Gonçalves (2022), pode se manifestar por meio da reprodução e amplificação de preconceitos já existentes na sociedade, afetando principalmente grupos

historicamente vulnerabilizados, como mulheres, pessoas não brancas e indivíduos de baixa renda.

É nesse cenário que Mulholland e Frajhof (2020) apontam que uma prática com elevado potencial discriminatório é o *profiling*, ou perfilamento, que consiste na criação, por parte do controlador de dados, de perfis relacionados ao titular das informações. Tal prática tem como finalidade servir de parâmetro para a avaliação de determinados aspectos da personalidade do indivíduo, resultando em decisões automatizadas que delineiam o perfil do usuário a partir de inferências realizadas com base nos dados coletados.

3.3 Questões legais e regulatórias

No que tange às questões legais e regulatórias, conforme Lima e Sá (2020), é necessário ter em mente que a LGPD não especifica as hipóteses em que o processamento totalmente automatizado de dados pode ocorrer, disciplinando apenas que o direito à explicação emerge quando a decisão automatizada é tomada sem qualquer interferência humana.

Diante desse cenário, torna-se fundamental o desenvolvimento de mecanismos de controle e responsabilização dos sistemas automatizados, uma vez que algoritmos devem ser passíveis de auditoria e revisão, de modo a garantir sua conformidade com parâmetros jurídicos e éticos, conforme defendido pela literatura sobre *accountable algorithms* (Kroll *et al.*, 2017, p. 640).

Desse modo, com base nesses autores, observa-se que, embora se trate de um avanço legislativo relevante, os parâmetros legais para o exercício do direito à explicação podem não ser suficientes para assegurar a autonomia informativa do titular dos dados pessoais e para concretizar a principiologia estabelecida no art. 6º da referida lei, especialmente no que se refere aos princípios do livre acesso (inc. IV), da transparência (inc. VI) e da não discriminação (inc. IX).

Daí a razão pela qual Mulholland e Frajhof (2020) consignam que se torna indispensável a regulação da coleta, do uso, do tratamento e do compartilhamento de dados pela Lei Geral de Proteção de Dados, devendo tais atividades ser compatíveis com os princípios que a estruturam. Destaca-se, nesse contexto, que, no caso de dados

sensíveis, seu tratamento deve ocorrer de forma a respeitar o princípio da igualdade e evitar qualquer forma de discriminação. Assim, o princípio da não discriminação (artigo 6º, IX) deve ser observado em todas as circunstâncias em que o uso de dados, sejam eles sensíveis ou não, possa gerar algum tipo de desvalor ou conduzir a resultados inequitativos, servindo, conseqüentemente, como base de sustentação para a proteção dos dados sensíveis, especialmente quando estão em jogo aspectos relacionados ao exercício democrático e ao acesso a direitos sociais, tais como o direito ao trabalho, à saúde e à moradia.

Tem-se, desse modo, a relevância de se atentar para a diversidade de questões legais e para os marcos regulatórios capazes de servir de alicerce para a efetivação do direito ao esclarecimento no âmbito do tratamento de dados pessoais.

4 PERSPECTIVAS FUTURAS E REGULAÇÕES ÉTICAS

Lima e Sá (2020) levam à compreensão de que não há certeza de que as decisões automatizadas, isto é, decisões tomadas por máquinas dotadas de inteligência artificial, se revelem necessariamente mais sábias ou mais benéficas para a humanidade e para a biosfera.

Dessa forma, não se podem desconsiderar as implicações que tais tecnologias podem produzir para a própria humanidade, reconhecendo-se, ademais, que a tecnologia deve estar a serviço do ser humano, e não o contrário.

Diante do exposto:

O fato é que a tecnologia intelectual de hoje funciona como modalidade de cálculo racional da vida, o que faz com que essa extraordinária forma de pensar tenha um aumento exponencial de produtividade e eficiência. Isso não implica dizer que, paralelamente a todas essas relações entre pessoas e máquina, um de seus elementos mais caros, como a ética do comportamento humano, deva ser reduzido em sua complexidade – principalmente porque cada vez mais as relações humanas têm sido intermediadas por alguma modalidade de tecnologia da informação, seja ela de cunho inteligente ou não (Alves Pinto, 2020, p. 50).

É necessário, portanto, do ponto de vista das perspectivas futuras e das regulações éticas, refletir sobre a necessidade de conciliar o aumento esperado de produtividade e eficiência com a preservação dos princípios éticos que orientam as relações estabelecidas

em sociedade, cada vez mais influenciadas pelo avanço e pela incorporação de novas tecnologias.

4.1 Participação pública e engajamento

A participação pública e o engajamento social são fundamentais quando se está diante de temas que dizem respeito a toda a sociedade, incluindo-se, nesse contexto, a influência decorrente da utilização de sistemas de Inteligência Artificial.

É nesse cenário que Porto Júnior (2019) afirma que, para além de uma lógica excessivamente centrada no indivíduo e no direito à explicação, torna-se indispensável implementar mecanismos de transparência e *accountability* das inteligências artificiais por meio de instrumentos de natureza coletiva, como agências reguladoras, avaliações de impacto em proteção de dados (*Data Protection Impact Assessments – DPIAs*) e certificações de qualidade. Destaca-se, no contexto brasileiro, o papel estratégico da sociedade civil e do Ministério Público, inclusive por meio da propositura de ações civis públicas.

O referido autor também destaca, no âmbito nacional, a relevância da Autoridade Nacional de Proteção de Dados (ANPD), bem como de autoridades internacionais responsáveis pela proteção de dados, no que se refere à fiscalização de algoritmos de Inteligência Artificial. Ainda assim, permanece em debate se seria mais adequado adotar um modelo centralizado de agência reguladora específica para IA ou se a regulação deve ocorrer de forma descentralizada, por meio das instituições regulatórias já existentes, conforme suas respectivas áreas temáticas.

4.2 Reflexões sobre os princípios éticos e fundamentais

A reflexão sobre os princípios éticos e fundamentais deve levar em consideração o fato de que, não obstante os avanços proporcionados pela incorporação de novas tecnologias, sobretudo no campo da Inteligência Artificial, torna-se indispensável preservar a ética e o respeito intransigente aos direitos e garantias dos usuários.

É nesse cenário que Alves Pinto (2020) ressalta que tem se ampliado a percepção de que as transformações promovidas pelos avanços da Inteligência Artificial alterarão

decisivamente os rumos da sociedade moderna, tornando imperioso que a sociedade que se beneficia desses avanços tecnológicos saiba também lidar com as novas questões éticas que deles decorrem, com o objetivo de amenizar os efeitos colaterais produzidos pela era digital.

Em complemento a essa perspectiva, Reis e Furtado (2022) destacam que a crescente representação dos fatos da vida em termos de códigos e objetos digitais tem encurtado a distância entre o plano *off-line* e o plano *on-line*, gerando influências recíprocas entre esses dois ambientes e abalando modelos mentais anteriormente consolidados, especialmente aqueles relacionados à intimidade e à privacidade.

Cabe, portanto, atentar para o fato de que as inovações promovidas pela Inteligência Artificial e a capacidade de produção de decisões automatizadas não podem gerar situações que entrem em conflito com um dos principais fundamentos do Estado Democrático de Direito, qual seja, a dignidade da pessoa humana.

Nesse sentido, Monteiro (2018) ressalta que situações envolvendo áreas como saúde, educação, segurança, crédito, emprego, redes sociais, informação e até mesmo os rumos de um Estado Democrático de Direito dependem, cada vez mais, do uso massivo de dados pessoais e de processos automatizados de tomada de decisões. Tais processos podem produzir impactos diretos na vida das pessoas, inclusive com potencial para gerar práticas abusivas ou discriminatórias. Dessa forma, para prevenir violações de liberdades e direitos fundamentais, torna-se necessário compreender o funcionamento desses processos decisórios, o que possibilita sua contestação ou a revisão por pessoas naturais, de modo a evitar a reprodução de comportamentos enviesados decorrentes do processamento inadequado de dados ou da utilização de informações para finalidades ilícitas.

5 CONCLUSÃO

A presente investigação analisou a intersecção entre o desenvolvimento de sistemas baseados em *Machine Learning* e a proteção dos direitos fundamentais, com especial enfoque no direito à explicação previsto na Lei Geral de Proteção de Dados. Partiu-se da premissa de que a crescente utilização de decisões automatizadas em

diferentes esferas da vida social produz impactos diretos na autonomia informativa dos indivíduos e na própria estrutura de garantia dos direitos fundamentais.

A análise desenvolvida ao longo do estudo permitiu demonstrar que, embora os sistemas de aprendizado de máquina representem importante avanço tecnológico, especialmente em termos de eficiência e capacidade de processamento de grandes volumes de dados, tais sistemas introduzem novos desafios jurídicos e éticos relacionados à opacidade algorítmica, à possibilidade de reprodução de vieses e à dificuldade de compreensão dos critérios utilizados nas decisões automatizadas.

Nesse contexto, evidenciou-se que o direito à explicação assume papel central como mecanismo de proteção do titular de dados, na medida em que permite questionar, compreender e eventualmente revisar decisões tomadas exclusivamente por sistemas automatizados. Tal direito não se limita a uma dimensão meramente informativa, mas representa um instrumento de controle democrático sobre tecnologias capazes de influenciar o acesso a oportunidades sociais, econômicas e políticas.

Os resultados da pesquisa também indicam que os desafios para a efetivação do direito à explicação não são apenas técnicos, mas também institucionais e regulatórios. A ausência de parâmetros mais detalhados para a implementação desse direito, aliada à complexidade dos sistemas de inteligência artificial, evidencia a necessidade de aprimoramento dos mecanismos de transparência, auditoria e responsabilização no tratamento de dados pessoais.

Além disso, verificou-se que a discussão sobre o direito à explicação não pode ser dissociada da problemática da discriminação algorítmica, especialmente quando decisões automatizadas passam a influenciar processos como concessão de crédito, seleção profissional, acesso a serviços ou formulação de políticas públicas. Nesses contextos, a transparência e a possibilidade de contestação das decisões automatizadas tornam-se condições essenciais para a preservação da igualdade e da dignidade da pessoa humana.

Diante desse cenário, sustenta-se que a consolidação de um modelo de governança de sistemas baseados em *Machine Learning* deve necessariamente incorporar princípios de transparência, responsabilidade e supervisão institucional. Tal perspectiva exige não apenas o fortalecimento de instrumentos jurídicos já existentes, como os previstos na LGPD, mas também a ampliação da participação social, da atuação de órgãos reguladores e da fiscalização institucional sobre o uso de tecnologias de inteligência artificial.

Conclui-se, portanto, que o direito à explicação constitui elemento estruturante para a construção de um ambiente tecnológico compatível com os fundamentos do Estado Democrático de Direito. Mais do que um direito individual, trata-se de um mecanismo de equilíbrio entre inovação tecnológica e proteção dos direitos fundamentais, capaz de assegurar que o desenvolvimento e a aplicação de sistemas de inteligência artificial ocorram de forma ética, transparente e socialmente responsável.

REFERÊNCIAS

ALVES PINTO, Henrique. A utilização da inteligência artificial no processo de tomada de decisões Por uma necessária *accountability*. **RIL Brasília** a. 57 n. 225 p. 43-60 jan./mar. 2020. Disponível em: https://www12.senado.leg.br/ril/edicoes/57/225/ril_v57_n225_p43. Acesso em: 28 fev. 2024.

BAROCAS, Solon; SELBST, Andrew D. Big data's disparate impact. **California Law Review**, v. 104, n. 3, p. 671-732, 2016. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899. Acesso em: 29 fev. 2024.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. **Lei Geral de Proteção de Dados Pessoais (LGPD)**. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/113709.htm. Acesso em: 25 fev. 2024.

FACELI, Katti; LORENA, Ana Carolina; GAMA, João; CARVALHO, André Carlos Ponce de Leon Ferreira de. **Inteligência Artificial: Uma Abordagem de Aprendizagem de Máquina**. Rio de Janeiro: LTC, 2011.

GONÇALVES, Gabriella Almeida de Oliveira. **Discriminação algorítmica e decisões automatizadas: análise do art. 20 da LGPD como instrumento de proteção dos interesses dos titulares**. Artigo apresentado como requisito parcial para a conclusão do curso de bacharelado em Direito da Universidade São Judas Tadeu. 2022. Disponível em: <https://repositorio.animaeducacao.com.br/items/10919c89-78d6-4025-80c2-dead924579a4>. Acesso em: 29 fev. 2024.

GOODMAN, Bryce; FLAXMAN, Seth. European Union regulations on algorithmic decision-making and a “right to explanation”. **AI Magazine**, v. 38, n. 3, p. 50-57, 2017. Disponível em: <https://doi.org/10.48550/arXiv.1606.08813>. Acesso em: 26 fev. 2026.

KROLL, Joshua A. *et al.* Accountable algorithms. **University of Pennsylvania Law Review**, v. 165, n. 3, p. 633-705, 2017. Disponível em: <https://repository.law.upenn.edu/Documents/Detail/accountable-algorithms/155735>. Acesso em: 26 fev. 2024.

LIMA, Taisa Maria Macena; SÁ, Maria de Fátima Freire. Inteligência artificial e lei geral de proteção de dados pessoais: o direito à explicação nas decisões automatizadas. **Revista Brasileira de Direito Civil – RBDCivil**, Belo Horizonte, v. 26, p. 227-246,

out./dez. 2020. Disponível em: <https://rbdcivil.ibdcivil.org.br/rbdc/article/view/584>. Acesso em: 26 fev. 2024.

LUDERMIR, Teresa Bernarda. Inteligência Artificial e Aprendizado de Máquina: estado atual e tendências. **Estudos Avançados**, vol. 35 (101), 2021. Disponível em: <https://www.scielo.br/j/ea/a/wXBdv8yHBV9xHz8qG5RCgZd/?format=pdf&lang=pt>. Acesso em: 29 fev. 2024.

MONTEIRO, Renato Leite. Existe um direito à explicação na Lei Geral de Proteção de Dados do Brasil? **Instituto Igarapé**, Artigo Estratégico 39, Dezembro de 2018. Disponível em: <https://igarape.org.br/wp-content/uploads/2018/12/Existe-um-direito-a-explicacao-na-Lei-Geral-de-Protecao-de-Dados-no-Brasil.pdf>. Acesso em: 27 fev. 2024.

MULHOLLAND, Caitlin; FRAJHOF, Isabella Z. **Inteligência artificial e a Lei Geral de Proteção de Dados pessoais**: breves anotações sobre o direito à explicação perante a tomada de decisões por meio de machine learning. São Paulo: Revista dos Tribunais, 2020. Disponível em: <https://www.jusbrasil.com.br/doutrina/secao/inteligencia-artificial-e-a-lei-geral-de-protecao-de-dados-pessoais-breves-anotacoes-sobre-o-direito-a-explicacao-perante-a-tomada-de-decisoes-por-meio-de-machine-learning/1196969650>. Acesso em: 26 fev. 2024.

PAIXÃO, Gabriela Miana de Mattos *et al.* Machine Learning na Medicina: Revisão e Aplicabilidade. **Arq Bras Cardiol**. 2022; 118(1):95-102. Disponível em: <https://www.scielo.br/j/abc/a/WMgVngCLbYfJrkmC65VFCkp/?format=pdf&lang=pt>. Acesso em: 28 fev. 2024.

PINHEIRO, Patrícia Deck. **Direito Digital**. 7.ed. São Paulo: Saraiva Educação, 2021.

PINHEIRO, Patricia Deck; COELHO, Alexandre Zavaglia. Transparência do algoritmo: decisões automatizadas e direito à explicação. **FebrabanTech**. Artigo publicado em 14 de maio de 2020. Disponível em: <https://febrabantech.febraban.org.br/especialista/patricia-peck-pinheiro/transparencia-do-algoritmo-decisoes-automatizadas-e-direito-a-explicacao>. Acesso em: 27 fev. 2024.

PINHEIRO, Patricia Peck. **Direito Digital**. 5. ed. rev., atual. e ampl. de acordo com as Leis n. 12.735 e 12.737, de 2012 — São Paulo : Saraiva, 2013.

PORTO JÚNIOR, Odélio. **Existe um Direito à Explicação para Decisões Automatizadas realizadas por Algoritmos de Inteligência Artificial? - Uma análise da GDPR e LGPD**. Monografia apresentada à banca examinadora, como requisito para conclusão do curso de Direito e obtenção do título de bacharel em Direito na Universidade Federal de Minas Gerais. Belo Horizonte/MG Faculdade de Direito da UFMG 2019. Disponível em: https://www.academia.edu/43798533/Existe_um_Direito_%C3%A0_Explic%C3%A7%C3%A3o_para_Decis%C3%B5es_Automatizadas_realizadas_por_Algoritmos_de_Intelig%C3%Aancia_Artificial_Uma_an%C3%A1lise_da_GDPR_e_LGPD. Acesso em: 29 fev. 2024.

REIS, Nazareno César Moreira; FURTADO, Gabriel Rocha. Decisões automatizadas: definição, benefícios e riscos. **Civilistica.com**, [S. l.], v. 11, n. 2, p. 1–44, 2022.

Disponível em: <https://civilistica.emnuvens.com.br/redc/article/view/763>. Acesso em: 5 mar. 2024.

SCHLEDER, Gabriel R; FAZZIO, Adalberto. Machine Learning na Física, Química, e Ciência de Materiais: Descoberta e Design de Materiais. **Revista Brasileira de Ensino de Física**, vol. 43, suppl. 1, e20200407, 2021. Disponível em: <https://www.scielo.br/j/rbef/a/qzcfSKw4nzBK5Mddr8ZXy4v/?format=pdf&lang=pt>. Acesso em: 28 fev. 2024.

Contribuição dos autores

Todos os autores contribuíram igualmente para o desenvolvimento deste artigo.

Disponibilidade dos dados

Todos os conjuntos de dados relevantes para as conclusões deste estudo estão totalmente disponíveis no artigo.

Como citar este artigo (APA)

Godoy, S. M., Lacerda, L. P., Mozaner, V. C., & Santos, G. G. M. M. dos. (2026). EXPLICABILIDADE COMO CONDIÇÃO DE LEGITIMIDADE ALGORÍTMICA. *Veredas Do Direito*, 23(5), e235463. <https://doi.org/10.18623/rvd.v23.5463>