XIII ENCONTRO INTERNACIONAL DO CONPEDI URUGUAI – MONTEVIDÉU

GOVERNO DIGITAL, DIREITO E NOVAS TECNOLOGIAS I

DANIELLE JACON AYRES PINTO

YURI NATHAN DA COSTA LANNES

LAURA INÉS NAHABETIÁN BRUNET

Copyright © 2024 Conselho Nacional de Pesquisa e Pós-Graduação em Direito

Todos os direitos reservados e protegidos. Nenhuma parte destes anais poderá ser reproduzida ou transmitida sejam quais forem os meios empregados sem prévia autorização dos editores.

Diretoria – CONPEDI
Presidente - Profa. Dra. Samyra Haydêe Dal Farra Naspolini - FMU – São Paulo
Diretor Executivo - Prof. Dr. Orides Mezzaroba - UFSC – Santa Catarina
Vice-presidente Norte - Prof. Dr. Jean Carlos Dias - Cesupa – Pará
Vice-presidente Centro-Oeste - Prof. Dr. José Querino Tavares Neto - UFG – Goiás
Vice-presidente Sul - Prof. Dr. Leonel Severo Rocha - Unisinos – Rio Grande do Sul
Vice-presidente Sudeste - Profa. Dra. Rosângela Lunardelli Cavallazzi - UFRJ/PUCRio – Rio de Janeiro
Vice-presidente Nordeste - Prof. Dr. Raymundo Juliano Feitosa - UNICAP - Pernambuco

Representante Discente: Prof. Dr. Abner da Silva Jaques - UPM/UNIGRAN - Mato Grosso do Sul

Conselho Fiscal:
Prof. Dr. José Filomeno de Moraes Filho - UFMA - Maranhão
Prof. Dr. Caio Augusto Souza Lara - SKEMA/ESDHC/UFMG - Minas Gerais
Prof. Dr. Valter Moura do Carmo - UFERSA - Rio Grande do Norte
Prof. Dr. Fernando Passos – UNIARA - São Paulo
Prof. Dr. Edinilson Donisete Machado - UNIVEM/UENP - São Paulo
Secretarias
Relações Institucionais:
Prof. Dra. Claudia Maria Barbosa - PUCPR - Paraná
Prof. Dr. Heron José de Santana Gordilho - UFBA - Bahia
Profa. Dra. Daniela Marques de Moraes - UNB - Distrito Federal
Comunicação:
Prof. Dr. Robison Tramontina - UNOESC - Santa Catarina
Prof. Dr. Liton Lanes Pilau Sobrinho - UPF/Univali - Rio Grande do Sul
Prof. Dr. Lucas Gonçalves da Silva - UFS - Sergipe
Relações Internacionais para o Continente Americano:
Prof. Dr. Jerônimo Siqueira Tybusch - UFSM - Rio Grande do Sul
Prof. Dr. Paulo Roberto Barbosa Ramos - UFMA - Maranhão
Prof. Dr. Felipe Chiarello de Souza Pinto - UPM – São Paulo
Relações Internacionais para os demais Continentes:
Profa. Dra. Gina Vidal Marcilio Pompeu - UNIFOR - Ceará
Profa. Dra. Sandra Regina Martini - UNIRITTER / UFRGS - Rio Grande do Sul
Profa. Dra. Maria Claudia da Silva Antunes de Souza - UNIVALI – Santa Catarina
Eventos:
Prof. Dr. Yuri Nathan da Costa Lannes - FDF - São Paulo
Profa. Dra. Norma Sueli Padilha - UFSC - Santa Catarina
Prof. Dr. Juraci Mourão Lopes Filho - UNICHRISTUS - Ceará

Membro Nato – Presidência anterior Prof. Dr. Raymundo Juliano Feitosa – UNICAP – Pernambuco

D597

GOVERNO DIGITAL, DIREITO E NOVAS TECNOLOGIAS I

[Recurso eletrônico on-line] organização CONPEDI

Coordenadores: Danielle Jacon Ayres Pinto, Yuri Nathan da Costa Lannes, Laura Inés Nahabetián Brunet – Florianópolis: CONPEDI, 2024.

Inclui bibliografia
ISBN: 978-85-5505-986-5
Modo de acesso: www.conpedi.org.br em publicações
Tema: ESTADO DE DERECHO, INVESTIGACIÓN JURÍDICA E INNOVACIÓN

1. Direito – Estudo e ensino (Pós-graduação) – 2. Governo digital. 3. Novas tecnologias. XIII ENCONTRO INTERNACIONAL DO CONPEDI URUGUAI – MONTEVIDÉU (2: 2024 : Florianópolis, Brasil).

CDU: 34



Conselho Nacional de Pesquisa
e Pós-Graduação em Direito Florianópolis
Santa Catarina – Brasil
www.conpedi.org.br

XIII ENCONTRO INTERNACIONAL DO CONPEDI URUGUAI – MONTEVIDÉU

GOVERNO DIGITAL, DIREITO E NOVAS TECNOLOGIAS I

Apresentação

O XIII ENCONTRO INTERNACIONAL DO CONPEDI URUGUAI – MONTEVIDÉU, realizado na Universidad de La República Uruguay, entre os dias 18 a 20 de setembro de 2024, apresentou como temática central "Estado de Derecho, Investigación Jurídica e Innovación". Esta questão suscitou intensos debates desde o início e, no decorrer do evento, com a apresentação dos trabalhos previamente selecionados, fóruns e painéis que ocorreram na cidade de Montevidéu-Uruguai.

Os trabalhos contidos nesta publicação foram apresentados como artigos no Grupo de Trabalho "DIREITO, GOVERNANÇA E NOVAS TECNOLOGIAS I", realizado no dia 20 de setembro de 2024, que passaram previamente por no mínimo dupla avaliação cega por pares. Encontram-se os resultados de pesquisas desenvolvidas em diversos Programas de Pós-Graduação em Direito, que retratam parcela relevante dos estudos que têm sido produzidos na temática central do Grupo de Trabalho.

As temáticas abordadas decorrem de intensas e numerosas discussões que acontecem pelo Brasil, com temas que reforçam a diversidade cultural brasileira e as preocupações que abrangem problemas relevantes e interessantes, a exemplo do direito digital, proteção da privacidade, crise da verdade, regulamentação de tecnologias, transformação digital e Inteligência artificial, bem como políticas públicas e tecnologia.

Espera-se, então, que o leitor possa vivenciar parcela destas discussões por meio da leitura dos textos. Agradecemos a todos os pesquisadores, colaboradores e pessoas envolvidas nos debates e organização do evento pela sua inestimável contribuição e desejamos uma proveitosa leitura!

Danielle Jacon Ayres Pinto - Universidade Federal de Santa Catarina

Yuri Nathan da Costa Lannes - Faculdade de Direito de Franca

Laura Inés Nahabetián Brunet - Universidad Mayor de la República Oriental del Uruguay

O CHATGPT É PRECONCEITUOSO? REFLEXÕES EM TORNO DOS SISTEMAS DE INTELIGÊNCIA ARTIFICIAL E SEUS VIESES À LUZ DA LEI GERAL DE PROTEÇÃO DE DADOS

IS CHATGPT BIASED? REFLECTIONS ON ARTIFICIAL INTELLIGENCE SYSTEMS AND THEIR BIASES IN LIGHT OF THE GENERAL DATA PROTECTION LAW

Camilla Gomes Pinheiro Costa [1]

Resumo

A popularização das IAs, como o ChatGPT, levantou preocupações sobre vieses algorítmicos e a proteção de dados pessoais. O presente artigo tem como objetivo explorar as implicações éticas e jurídicas dos sistemas de inteligência artificial e seus vieses à luz da Lei Geral de Proteção de Dados (Lei 13.709/2018), oferecendo um contributo para as pesquisas interdisciplinares que envolvem direito e tecnologia. Elenca-se como objetivos específicos: a) compreender o funcionamento da IA e sua sujeição a vieses; b) analisar disposições da LGPD que possam ser aplicadas a situações envolvendo inteligência artificial. Parte-se da premissa de que as leis de proteção de dados possuem proximidade técnica com possíveis mecanismos de regulação de IA. Para o desenvolvimento da pesquisa, foram feitas revisão de literatura, análise e síntese de obras, artigos científicos, legislação e relatórios existentes sobre inteligência artificial e proteção de dados pessoais. Foi feita também a análise dos principais pontos levantados pelo Centro de Inteligência Artificial e Política Digital em reclamação à Federal Trade Comission (EUA) alegando que a OpenAI estaria ciente dos vieses perpetuados pelo ChatGPT e os riscos a violações a direitos. Reafirma-se, no artigo, a importância de uma governança eficaz de inteligência artificial, fundamentada em princípios éticos e legais, para maximizar os benefícios das IAs generativas enquanto se protege os direitos dos indivíduos. Por fim, o artigo indica a necessidade de aprofundamento do tema em futuras agendas de pesquisa.

Palavras-chave: Inteligência artificial, Proteção de dados pessoais, Vieses algorítmicos, Ia generativa, Lei geral de proteção de dados

Abstract/Resumen/Résumé

The popularization of AI, such as ChatGPT, has raised concerns about algorithmic biases and the protection of personal data. This article aims to explore the ethical and legal implications of artificial intelligence systems and their biases in light of the General Data Protection Law (Law 13.709/2018), offering a contribution to interdisciplinary research involving law and technology. The specific objectives are: a) to understand the functioning of AI and its susceptibility to biases; b) to analyze the provisions of the LGPD that can be applied to

[1] Mestranda em Comunicação Digital. Especialização lato sensu em Lei Geral de Proteção de Dados. Especialização lato sensu em Docência no Ensino Superior. Bacharela em Direito.

situations involving artificial intelligence. It is based on the premise that data protection laws have technical proximity to possible AI regulation mechanisms. For the development of this research, a comprehensive literature review, analysis, and synthesis of existing works, scientific articles, legislation, and reports on artificial intelligence and personal data protection were conducted. An analysis of the main points raised by the Center for Artificial Intelligence and Digital Policy in a complaint to the Federal Trade Commission (USA) was also conducted, alleging that OpenAI was aware of the biases perpetuated by ChatGPT and the risks of rights violations. The article reaffirms the importance of effective artificial intelligence governance, grounded in ethical and legal principles, to maximize the benefits of generative AIs while protecting individuals' rights. Finally, the article indicates the need for further research on this topic in future research agendas.

Keywords/Palabras-claves/Mots-clés: Artificial intelligence, Personal data protection, Algorithmic biases, Generative ai, General data protection law

1.INTRODUÇÃO

Uma matéria publicada em 31 de março de 2023 na Exame[1] apresenta o intrigante título: *"Centro de ética diz que ChatGPT[2] é "preconceituoso e enganoso" em denúncia nos EUA"*. No subtítulo, consta a seguinte descrição: *"Organização sem fins lucrativos afirma que a inteligência artificial pode acabar reforçando preconceitos e imagens negativas na sociedade"*.

A matéria em questão, refere-se a uma queixa apresentada pelo Centro de Inteligência Artificial e Política Digital[3] (CAIDP) à Comissão Federal de Comércio dos EUA (FTC)[4] contra a OpenAI Inc[5]., destacando preocupações com o lançamento do produto GPT-4, sob o argumento de que se trataria de solução enviesada e que representaria riscos à privacidade e à segurança pública. Em termos de estratégia de comunicação, certamente, o título da matéria cumpre o objetivo de capturar a atenção dos leitores. É cediço, todavia, que o assunto demanda um olhar técnico para apurada compreensão. O ChatGPT, assim como qualquer outra ferramenta de inteligência artificial, não é dotado de emoções, crenças ou capacidade de realizar juízo de valor. No entanto, como alertam Bioni e Luciano (2019), acreditar que os algoritmos de IA sejam imparciais ou isentos de erros, manipulações e vieses é uma ficção.

Como asseveram Martins, Longhi e Mocellin (2024), assim como os potenciais benefícios das aplicações da IA para a sociedade ainda não são completamente conhecidos, da mesma forma os riscos decorrentes do uso de tais ferramentas ainda estão sendo identificados. O desconhecimento em torno das decorrências da era da inteligência artificial ainda é grande, e surgem diversas implicações éticas e jurídicas sem que advenham imediatas soluções.

As indagações na esfera da proteção de dados pessoais, sobretudo, preocupam, tendo em vista que, para o funcionamento dos modelos de IA, "é necessário coletar e tratar dados em larga

[1] Centro de ética diz que ChatGPT é "preconceituoso e enganoso" em denúncia nos EUA | Exame, acesso em 10/06/2024, às 14h41.
[2] Ferramenta de IA generativa desenvolvida pela OpenIA.
[3] O CAIDP é uma organização dedicada ao estudo e à promoção de políticas públicas relacionadas ao desenvolvimento e à implementação de inteligência artificial (IA) e tecnologias digitais. CAIDP - Center for AI and Digital Policy, acesso em 10/06/2024, às 15h49.
[4] A FTC é uma agência independente do governo dos Estados Unidos que tem como principal missão promover a proteção dos consumidores e a eliminação e prevenção de práticas anticompetitivas nos mercados. Federal Trade Commission | Protecting America's Consumers (ftc.gov), acesso em 10/06/2024, às 15h49.
[5] A Open AI é uma organização de pesquisa e implementação de inteligência artificial, About | OpenAI, acesso em 10/06/2024, às 15h57.

escala, inclusive aqueles que identificam ou permitem identificar uma pessoa natural: os dados pessoais" (MARTINS, LONGHI, MUCELIN, p. 830, 2024). Ademais, a forma como a base de dados de uma IA é montada pode ser um dos fatores provocadores dos vieses algorítmicos. Nesse sentido, questiona-se: como a Lei Geral de Proteção de Dados (LGPD) pode contribuir para as reflexões em torno dos sistemas de inteligência artificial e seus vieses?

A presente pesquisa tem como proposta refletir sobre os sistemas de inteligência artificial e seus vieses à luz da Lei Geral de Proteção de Dados, oferecendo um contributo acadêmico para o desenvolvimento do debate interdisciplinar que engloba direito e tecnologia.

Tendo em vista que a popularização da inteligência artificial, potencializada pelas ferramentas de IA generativa, está revolucionando práticas nos mais diversos setores e evolui de forma exponencial, razoável prever um aumento expressivo de questionamentos acerca da temática nos próximos anos, o que faz com que toda pesquisa e contribuição acadêmica que se dedique a investigar o assunto seja considerada de extrema valia.

Para a elaboração deste artigo, foram realizadas revisão de literatura, análise e síntese de informações provenientes de diversas fontes bibliográficas, incluindo livros, artigos científicos, legislação e relatórios relevantes para a compreensão do tema pesquisado. Parte-se da premissa de que vários dos temas tratados em leis de proteção de dados estão intimamente relacionados a possíveis mecanismos regulatórios para o uso da IA.

Aponta-se, como objetivo geral, desenvolver reflexões acerca de como a as disposições da LGPD podem contribuir para análise de questões jurídicas decorrentes do desenvolvimento e uso de ferramentas de IA, especialmente no que tange a possíveis vieses embarcados nos modelos e suas decorrências.

Os objetivos específicos são: a) compreender o funcionamento de uma inteligência artificial e como ela pode estar sujeita a vieses; b) estudar os princípios e demais disposições da LGPD que possam ser aplicados a situações envolvendo IA. Tais objetivos dão origem aos dois tópicos apresentados a seguir nos quais se desdobra a pesquisa.

2. INTELIGÊNCIA ARTIFICIAL: FALHAS E VIESES

2.1 Compreendendo a inteligência artificial

Uma parte considerável das atividades humanas é regida por algoritmos de inteligência artificial: do entretenimento, por meio das sugestões de filmes feitas pelas plataformas de streaming, a questões mais complexas e potencialmente lesivas a direitos, como as decisões referentes à concessão ou não concessão de crédito bancário a um determinado indivíduo, ou se determinado suspeito ou acusado de crime representa alto risco para a sociedade.

Em que pese a inteligência artificial já esteja integrada ao cotidiano de grande parte da população, especialmente com o avanço da IA generativa, uma pequena parcela compreende minimamente o que ela é, qual é seu propósito e como ela funciona. Em parte, o desconhecimento se dá em virtude da própria terminologia que, na visão de estudiosos como Johansen (2020), pode ser prejudicial para a compreensão do fenômeno. Segundo o autor, a adoção de terminologias irrefletidas para definir coisas novas influencia negativamente na forma como indivíduos e organizações se relacionam com as inovações. A denominação "inteligência artificial", na visão do autor, conduziria a um pensamento equivocado de que essa inovação representaria uma espécie de "carruagem sem cavalos", algo do qual o ser humano não possui controle algum. Além de não expressar de forma assertiva seu sentido, a expressão contribui para um temor generalizado de que as máquinas poderiam, a qualquer momento, suplantar os seres humanos.

A partir desse raciocínio, uma forma eficaz de se iniciar a compreensão do que é a IA, é definir o que ela não é. A inteligência artificial não é uma nova tecnologia. Tampouco trata-se de uma entidade sobre humana com capacidades ilimitadas. Ela é uma decorrência de décadas de estudos e pesquisas envolvendo os campos da estatística, da computação, lógica, análise de dados e modelos de aprendizagem, e seus avanços só se concretizaram em razão da evolução de um conjunto de tecnologias que são utilizadas para o desenvolvimento de uma IA.

Entende-se por inteligência artificial "todo sistema computacional que simula a capacidade humana de raciocinar e resolver problemas, por meio de tomadas de decisão baseadas em análises probabilísticas" (FRAZÃO, MULHOLLAND, p. 05, 2019).

Essa emulação das capacidades humanas é possível graças à "habilidade de um sistema de interpretar corretamente dados externos, aprender a partir desses dados para alcançar objetivos

e tarefas específicos por meio da adaptação flexível" (STEIBEL, VICENTE, JESUS, p. 54, 2019).

A inteligência artificial, portanto, aprende a partir da análise de grandes bases de dados que lhes são fornecidas depois uma etapa inicial de coleta e pré-processamento dos dados para que se tornem compreensíveis para a IA. Após fases de treinamento, validação e testes, um sistema de IA estabelecerá uma regra própria – um algoritmo - para tomar determinadas decisões ou executar determinadas tarefas, podendo, para tal, percorrer caminhos que poderão fugir do controle dos programadores e superar processos cognitivos humanos tanto no quesito tempo de execução quanto no quesito complexidade da tarefa. (ROSSETTI, GARCIA, 2023).

Com efeito, os dados são a base sobre a qual é construído um sistema de IA, por viabilizarem treinamento, validação, personalização e melhoria contínua dos modelos de IA:

> "A Inteligência Artificial usa a informação externa obtida por esses meios como um *input* para a identificação de regras e modelos subjacentes ao confiar em perspectivas, como o aprendizado das máquinas, o qual descreve métodos que auxiliam os computadores a aprenderem sem serem explicitamente programados. (STEIBEL, VICENTE, JESUS, p. 55, 2019)

A oferta crescente de dados em formato de vídeos imagens, áudios, textos e números é, por certo, elemento fundamental na recente popularização das ferramentas de IA. Sem dados de alta qualidade e em quantidade suficiente, os sistemas de inteligência artificial não seriam capazes de alcançar o desempenho necessário para resolver problemas complexos e fornecer valor significativo em suas soluções. Não se deve olvidar, contudo, que a difusão da IA é um fenômeno multifatorial, sendo a redução dos custos computacionais e o vertiginoso aumento da capacidade de processamento dos dispositivos modernos, fatores primordiais para sua ocorrência (FRAZÃO, MULHOLLAND, 2019).

Em seu propósito, a IA foi idealizada com o intuito de "otimizar processos que podem ou poderiam ser feitos pela inteligência humana". (ROSSETTI, GARCIA, p. 255, 2023). Com a propagação das novas ferramentas de IA generativa, vemos esse propósito se concretizar de forma democrática, em diversos setores da sociedade.

A IA generativa é explicada pelo Fórum Econômico Mundial como

> uma categoria de algoritmos de inteligência artificial (IA) que geram novos resultados com base nos dados nos quais foram treinados. Diferentemente dos sistemas tradicionais de IA, que são projetados para reconhecer padrões e fazer previsões, a IA generativa cria novos conteúdos na forma de imagens, textos, áudios e muito mais.[6] (Tradução livre)

Para complementar o entendimento do que seriam as IAs generativas e porque elas estão tão difundidas, tem-se, ainda, que:

> As IAs generativas distinguem-se das tradicionais porque são programadas apenas para realizar inferências, predições e classificações em razão de uma funcionalidade particular, senão aprendem padrões para criar novos exemplos similares aos dados de treinamento e para performar diversas tarefas inespecíficas. Seu propósito é, pois, gerar novos dados (...) os quais são adjetivados como sintéticos, bem como expandir as possíveis aplicações de um modelo em outras tecnologias para novas tarefas, independentemente de predefinição acerca do uso e do setor empregado. (MARTINS, MOCELLIN, p. 381,2024)

Alguns exemplos de soluções que foram amplamente popularizadas nos últimos dois anos e que usam IA generativa são Dall-E[7], Midjourney[8] e o próprio ChatGPT.

Feitas as considerações sobre o funcionamento da inteligência artificial, as distinções entre as IAs tradicionais a as IAs generativas, e compreendidos seus principais aspectos, passa-se à seguinte investigação: o ChatGPT e outras ferramentas de IA generativa podem, de fato, ser preconceituosas?

2.2 Uma IA pode ser preconceituosa?

O que uma inteligência artificial generativa responderia diante dessa pergunta? Seria possível que as próprias IAs colaborassem na compreensão desse questionamento?

Acreditamos que a tentativa é válida e perguntamos ao ChatGPT o seguinte: “ChatGPT, você é preconceituoso?”

Antes de trazer à baila a explicação fornecida pela IA para essa pergunta, vale fazer a seguinte ressalva: é provável que não seja possível que o leitor replique esse comando e obtenha

[6] Disponível em What is Generative AI? Artificial intelligence explains | World Economic Forum (weforum.org), acesso em 10/06/2024, às 12h22.
[7] DALL·E 2 | OpenAI: sistema de inteligência artificial capaz de criar imagens realistas a partir de uma descrição feita em linguagem natural.
[8] Midjourney Grátis - Gerador de Imagens IA: Gerador de imagens gratuito de alta qualidade a partir de simples descrições.

exatamente a mesma resposta que aqui será apresentada. O *output* fornecido depende de fatores como a versão utilizada, o contexto do chat aberto e a data em que o comando foi dado. Para esse comando, foi usada a versão 4.0 do ChatGPT e o comando foi efetuado em junho de 2024. Colaciona-se, abaixo, a resposta articulada pela ferramenta:

> Como uma inteligência artificial desenvolvida pela OpenAI, não possuo emoções, crenças ou preconceitos próprios. No entanto, o modelo pode refletir preconceitos presentes nos dados nos quais foi treinado. Esses dados incluem uma vasta quantidade de texto da internet e outras fontes, que podem conter vieses humanos. A equipe da OpenAI trabalha continuamente para mitigar esses vieses e garantir que o modelo forneça respostas justas e imparciais. Se em algum momento você perceber que uma resposta pode estar enviesada ou ser inadequada, por favor, avise-me para que eu possa melhorar e fornecer a assistência mais adequada possível. (CHATGPT. Resposta à pergunta sobre preconceito do ChatGPT. OpenAI, 2024. Disponível em: https://chat.openai.com/. Acesso em: 10 jun. 2024).

Como mencionado na introdução, a afirmação de que uma IA generativa é "preconceituosa e enganosa" não é tecnicamente correta, podendo ser realizada apenas para fins de estratégias de comunicação e obtenção de atenção e cliques.

O fato é que, para o CAIDP, as práticas de negócios da OpenAI são injustas e enganosas, violam declarações da FTC, relatórios e *guidelines* para práticas de IA, bem como normas legais para governança da inteligência artificial.

Ao analisar a reclamação[9] do CAIDP à FTC contra a OpenAI, destacam-se, dentre os argumentos apresentados pelo CAIDP, o de que o GPT-4[10] teria sido lançado sem uma avaliação independente, contendo vieses que podem reforçar estereótipos prejudiciais; que ele não atenderia aos requisitos de transparência, justiça e responsabilidade declarados pela FTC; que representaria riscos para a segurança digital das crianças, apresentando potencial para encorajar comportamentos inseguros; que impactaria os direitos dos consumidores; que poderia ser explorado por cibercriminosos causando riscos à privacidade dos usuários; que processaria dados pessoais de cidadãos europeus sem a devida autorização legal pelo General Data

[9] The Center for Artificial Intelligence and Digital Policy (CAIDP), In the matter of OpenAI, Inc. FEDERAL TRADE COMMISSION, Washington, DC 20580, March 30, 2023.
[10] "O GPT-4, ou Generative Pre-trained Transformer 4, é a quarta versão do modelo de linguagem desenvolvido pela OpenAI. Ele representa um avanço significativo em relação às versões anteriores, oferecendo melhorias em precisão, coerência e capacidade de compreensão de texto". OPENAI. ChatGPT: O que é o GPT-4. Disponível em: https://www.openai.com/chatgpt. Acesso em: 11 jun. 2024.

Protection Regualtion (GDPR); que não teria divulgado detalhes fundamentais sobre sua arquitetura, técnicas de treinamento ou construção do conjunto de dados, dentre outros.
No documento consta ainda que a OpenAI emitiu declarações que validam vários dos argumentos tecidos:

> A OpenAI reconheceu especificamente o risco de viés e, mais precisamente, "associações estereotipadas e depreciativas prejudiciais para determinados grupos marginalizados" (ênfase adicionada). No documento GPT-4 System Card, a OpenAI declara: "O processo de avaliação que realizamos ajudou a gerar evidências qualitativas adicionais de vieses sociais em várias versões do modelo GPT-4. Descobrimos que o modelo tem o potencial de reforçar e reproduzir vieses e visões de mundo específicos, incluindo associações estereotipadas e depreciativas prejudiciais para determinados grupos marginalizados". (The Center for Artificial Intelligence and Digital Policy (CAIDP), In the matter of OpenAI, Inc. FEDERAL TRADE COMMISSION, Washington, DC 20580, March 30, 2023. Tradução livre)

Nesse sentido, solicitou a CAIDP que a FTC estabelecesse imediatas medidas para investigação da OpenAI, determinando, ainda, a suspensão de novos lançamentos comerciais do GPT-4 e estabelecendo rigorosas diretrizes para governança da IA.

Não é surpresa alguma que o GPT-4 ou outros modelos de IA sejam investigados por oferecem riscos e problemas decorrentes de algoritmos enviesados, e isso acontece por diversos motivos como "falta de regulação, monopólios no setor de IA, estruturas de governança insuficientes dentro de empresas de tecnologia, assimetrias de poder entre empresas e usuários, a distância cultural entre os responsáveis por pesquisas em tecnologia e a diversidade das populações nas quais essa tecnologia é utilizada" (BIONI, LUCIANO, p. 208, 2019).

Todavia, para fins da presente pesquisa, atemo-nos à montagem da base de dados que alimenta um sistema de IA como o principal motivo para a existência dos vieses algorítmicos. É necessário questionar as escolhas de quais dados importam e o porquê de eles importarem em um processamento por inteligência artificial, e isso é feito a partir de suposições previamente prescritas aos sistemas de IA. (BIONI, LUCIANO, 2019).

O problema é que esses questionamentos, além de nem sempre serem feitos, quase sempre, mesmo quando realizados, ficam sem respostas adequadas. Cathy O'neil (2020), critica contundentemente os modelos de IA responsáveis por operar atividades como vasculhar milhares de currículos ou pedidos de empréstimo em míseros segundos e ordená-los em listas de candidatos mais promissores, apontando a obscuridade de seus mecanismos que, como

deuses, não revelam sua natureza aos meros mortais, mas apenas aos matemáticos e cientistas da computação, altos sacerdotes de seus domínios.
Acrescenta a autora que tais soluções de IA que otimizam as tarefas cotidianas, não obstante sejam vendidas como imparciais e objetivas, em contraposição à análise humana preconceituosa, não passam de aplicações baseadas em escolhas feitas por seres humanos infalíveis, o que faz com que vários desses modelos programem preconceitos, equívocos e vises humanos, criando algoritmos que vão tomar decisões igualmente preconceituosas, equivocadas e enviesadas, além de potencialmente lesivas e – o pior de tudo – acima de quaisquer possibilidades de contestação. Não por acaso, O'neil nomeia esses modelos nocivos de "Armas de Destruição Matemáticas" (ADMs).

Ante o exposto, observa-se que, resta claro que uma IA não pode ser "preconceituosa" ou "enganosa", no sentido de exercer intencionalmente juízo de valor ou manipulação, faltando-lhe, para tal, a consciência e demais atributos próprios da existência humana.

Nesse ínterim, seguindo a linha de pensamento de Johansen (2020) de que a denominação importa para a correta compreensão de um fenômeno, válido questionar a razoabilidade do uso do termo "discriminação algorítmica", ainda que muito propagado, dada sua insuficiência para exprimir o processo de decisões enviesadas tomadas por algoritmos de inteligência artificial, uma vez que pode levar à interpretação distorcida de que a discriminação seria uma escolha feita de forma intencional pela própria IA que, como visto, não possui atributos humanos ou capacidade para empregar tal intencionalidade em seus processos decisórios, sendo as decisões, boas ou ruins, meros resultados da análise de um *input* de dados.

Todavia, ao mesmo tempo em que não podemos rotular a IA de preconceituosa, é patente que os preconceitos e manipulações não só podem como já se perpetuam nas inteligências artificiais, levando a sociedade a refletir sobre uma série de questões éticas e jurídicas referentes ao combate à propagação dos vieses algorítmicos.

Países de todo o mundo encontram-se imbuídos na busca pelas melhores soluções para governança da inteligência artificial, bem como modelos regulatórios capazes de minimizar os potenciais riscos decorrentes do desenvolvimento e uso de sistemas de IA e promover adequada responsabilização de agentes ocasionadores de eventuais danos.

Contudo, a missão de combater vieses potencialmente lesivos a direitos nos modelos de IA é árdua, especialmente ante a falta de regulação específica em vários países, a exemplo do Brasil.[11] O cumprimento dessa missão demanda, antes de tudo, o estabelecimento de questionamentos assertivos, que atinjam de forma certeira os principais desdobramentos éticos e jurídicas do desenvolvimento e do uso da inteligência artificial, possibilitando reflexões essenciais acerca de tais desdobramentos.

Em que pese o Brasil não conte com regulação destinada especificamente ao tema inteligência artificial, não se pode olvidar que, por certo, todo o supedâneo encontrado no ordenamento jurídico brasileiro deve ser aplicado a questões que envolvem desenvolvimento e uso de inteligência artificial, a exemplo da proteção do consumidor e da teoria da responsabilidade civil, ainda que a aplicação seja cercada de dúvidas e necessidades interpretativas.

O Brasil conta também com a Lei Geral de Proteção de Dados – Lei 13.709/2018, que possui diversas disposições passíveis de serem aplicadas em contextos decorrentes do uso de IA trazendo, inclusive, de forma expressa, determinações sobre decisões automatizadas. Passa-se, dessa forma, ao estudo dessa lei.

3. LGPD, INTELIGÊNCIA ARTIFICIAL E VIESES

3.1 Lei Geral de Proteção de Dados

Algumas das mais relevantes questões éticas e jurídicas decorrentes da IA gravitam em torno do direito à proteção de dados pessoais, dada a essencialidade do uso de dados para o funcionamento da IA. Bioni e Luciano (2019), propõem que as leis de proteção de dados, possam representar o portal de entrada para uma regulação efetiva da inteligência artificial, considerando que as indagações que circundam o processamento de dados pessoais já têm sido endereçadas em leis de proteção de dados, a exemplo do *General Data Protection Regulation* na União Europeia - GDPR e da Lei Geral de Proteção de Dados no Brasil – LGPD.

A Lei Geral de Proteção de Dados dispõe sobre o tratamento de dados pessoais, inclusive nos meios digitais, por pessoa natural ou por pessoa jurídica de direito público ou privado. O

[11] Está em tramitação no Brasil o PL 2338/203, PL 2338/2023 - Senado Federal, acesso em 11/06/2024.

objetivo da LPGD é proteger os direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa natural.
Para a lei, dado pessoal deve ser entendido como toda informação relacionada a pessoa natural diretamente identificada, como também da pessoa natural identificável – aquela que, embora não seja diretamente identificada de imediato, poderá ser eventualmente identificada a partir de um conjunto de informações. Ademais, é delineada na legislação uma outra categoria – os dados pessoais sensíveis, taxativamente descritos no art. 5º, II como aqueles dados pessoais sobre origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de caráter religioso, filosófico ou político, dado referente à saúde ou à vida sexual, dado genético ou biométrico, quando vinculado a uma pessoa natural. Tratam-se de dados pessoais que podem acarretar algum tipo de discriminação quando do seu tratamento, implicando potenciais riscos ao direitos e liberdades fundamentais de seus titulares.

Para que um tratamento de dados pessoais seja considerado legal, deverão ser respeitados os princípios para as atividades de tratamento, (como necessidade, adequação, transparência, precaução etc.), sendo necessário também que cada operação possa ser justificada com base em uma das hipóteses taxativas de tratamento, previstas nos arts. 7º (dados pessoais) e 11 (dados pessoais sensíveis) da LGPD.

Dentre outras disposições acerca dos direitos dos titulares de dados pessoais, das medidas de segurança e governança, das sanções e fiscalização, a LGPD estabelece a definição dos agentes de tratamento de dados (controlador e operador), bem como sua responsabilização por violações à lei e realização de tratamentos indevidos.

Martins e Mucelin (2024) propõem uma sistematização da abordagem do tema IA e sua relação com a Lei Geral de Proteção de Dados, apontando que o treinamento de IAs generativas desafia o direito à proteção de dados em pelo menos seis níveis, sendo que, para fins desse estudo, os mais relevantes dizem respeito aos parâmetros adotados para o treinamento de sistemas algorítmicos e de medidas para mitigação de vieses, discriminações, ilicitudes e abusividades, bem como a disponibilização de informações sobre as categorias e fontes de dados pessoais adotadas para desenvolvimento e aplicação da IA.

Interessante ressaltar que, no que tange à Lei Geral de Proteção de Dados, argumentam Bioni, Garrote e Guedes (2023) que se trata de legislação transversal, cuja abordagem baseia-se em

risco e direitos, guardando intensa proximidade com possíveis técnicas regulatórias para o cenário da inteligência artificial.

3.2 Princípios da proteção de dados no Brasil e sistemas de IA

Uma breve análise dos princípios para o tratamento de dados pessoais estabelecidos na LGPD, demonstra a aplicação da lei a questões fundamentais ao campo da inteligência artificial.
A análise parte do princípio da não discriminação, previsto no art. 6º, IX, consubstanciado na impossibilidade de realização do tratamento de dados para fins discriminatórios, ilícitos ou abusivos.

Conforme lembram Maldonado e Blum (2021)

> proteger dados pessoais supera a tutela da privacidade, abarcando também os direitos de personalidade, diante da possibilidade de estigmatização do ser humano em razão da sua classificação e segmentação baseada no tratamento de suas informações (MALDONADO, BLUM, p. 168, 2021)

Ademais, o próprio reconhecimento do direito à proteção de dados pessoais como um direito fundamental constante no rol do art. 5º da Constituição Federal, reforça esse posicionamento.
A garantia do princípio da não discriminação é um dos caminhos para a mitigação de vieses, discriminações, ilicitudes e abusividades não só de forma geral no tratamento de dados pessoais, mas também e talvez especialmente quando esse tratamento acontece por meio de sistemas de IA. Nesse sentido, temos que

> (...) o princípio da não discriminação (artigo 6º, IX) deve ser refletido em todas as circunstâncias em que o uso de dados, sejam sensíveis ou não, gere algum tipo de desvalor ou indução a resultados que seriam inequitativos. Esse princípio deve servir como base de sustentação da tutela dos dados sensíveis, especialmente quando estamos diante do exercício democrático e do acesso a direitos sociais, tais como o direito ao trabalho, à saúde e à moradia" (MULHOLLAND, FRAJHOF, P. 269, 2019)

O princípio da não discriminação abrange as mais diversas práticas discriminatórias que tenham finalidade não permitida, desde o impedimento à vaga de emprego em razão de religião ou doença do titular de dados (com ou sem emprego de IA), até técnicas de perfilamento nas quais o controlador utiliza ferramentas de inteligência artificial para servir como parâmetro avaliativo de aspectos de sua personalidade ou vida financeira, para conceder ou negar acesso a crédito. (MENDES, DONEDA, SARLET, RODRIGUES JR., 2021)

Além da discriminação por atos abusivos ou ilícitos, aponta-se a existência da discriminação pela generalização injusta ou correlação abusiva, pode ser constatada em qualquer contexto de discriminação estatística, ou seja, a atribuição de um tratamento diferenciado com base na média, sem levar em conta as características e condições individuais. (MENDES, MATTIUZZO, FUJIMOTO, 2021).

A não discriminação encontra complemento e fortalecimento nos demais princípios da LGPD, a exemplo do princípio da transparência.

O princípio da transparência, previsto no artigo 6º, VI da LGPD é um dos princípios norteadores das atividades de tratamento de dados pessoais no Brasil, é a garantia de fornecimento aos titulares de dados de informações claras, precisas e facilmente acessíveis sobre a realização do tratamento e os respectivos agentes de tratamento, observados os segredos comercial e industrial.

Todavia, em situações nas quais exista um tratamento automatizado e decisões alcançadas por processos algorítmicos, garantir ao titular dos dados a transparência de forma plena pode ser uma tarefa dificultosa. A falta de clareza nos processos decisórios de IA pode representar obstáculos à efetivação do princípio a transparência, seja dificultando o exercício do direito do titular dos dados à explicação, seja trazendo aos agentes de tratamento de dados pessoais um excesso de onerosidade ou aumentado grau de dificuldade no cumprimento de suas obrigações de prestar informações ao titular.

Por tais motivos, a necessidade de garantia do princípio da responsabilização e prestação de contas merece destaque junto ao princípio da transparência. A responsabilização e prestação de contas está presente no art. 6º, X e determina que seja dada demonstração, pelo agente de tratamento, da adoção de medidas eficazes e capazes de comprovar a observância e o cumprimento das normas de proteção de dados, bem como eficácia de tais medidas. A prestação de contas ou *accountability* é de grande importância na proteção de dados, para garantia da transparência, e é entendida como medida fundamental na governança da inteligência artificial e celebrada nos debates regulatórios de IA ao redor do mundo.

Não somente a falta de transparência tensiona as questões éticas e jurídicas decorrentes do uso desses algoritmos opacos, como também a qualidade da própria base de dados utilizados para alimentar uma IA.

O princípio da qualidade dos dados, disposto no art. 6º, V da lei, determina a garantia, aos titulares de dados pessoais, de exatidão, clareza, relevância e atualização de dados, em conformidade com a necessidade e para o cumprimento da finalidade do tratamento.

A observância do princípio da qualidade dos dados ajudaria a minimizar eventuais vícios em bases de dados utilizadas para treinar modelos de IA reduzindo, assim, os riscos e potenciais prejuízos decorrentes. Alertam Maldonado e Blum que, embora dados pessoais isoladamente vistos ou coletados pareçam ter pouco potencial de afetação de seus titulares,

> quando colocados em conjunto e processados por mecanismos altamente capacitados, formam um compilado da personalidade de cada pessoa, passando a representar, perante terceiros, inúmeras características, como se o indivíduo é rico ou pobre, bom pagador ou caloteiro, sem antecedentes criminais ou criminoso, apartidário ou filiado a um determinado partido, ateu ou religioso, HIV positivo ou negativo, homossexual ou bissexual, alcoólatra ou abstêmio, entre outras incontáveis hipóteses que podem interferir de forma contundente na vida da pessoa (MALDONADO, BLUM, p. 156, 2021)

Adicionalmente, no âmbito individual também é possível identificar riscos significativos se a imprecisão em um dado pessoal, qualquer que seja, ocasionar, por exemplo, recusa de concessão de crédito, erro em tratamento médico, eliminações em processos seletivos ou injusto cerceamento de liberdade. Situações ainda mais lesivas podem ocorrer, ato contínuo, se a imprecisão não for devidamente sanada na fonte, pois os dados viciados permanecerão indefinidamente em tratamento, cabendo, portanto, aos controladores a adoção de medidas suficientes para garantia da precisão e da qualidade dos dados, bem como de sua atualização mediante necessidade (MALDONADO, BLUM, 2021).

O'neil (2020) argumenta, ainda, que a qualidade dos dados pode ser um fator de grande risco pela altíssima possibilidade de promover perpetuação de discriminações em sistemas de IA. Ela justifica seu argumento alertando que a perpetuação do racismo por meio de modelos de inteligência artificial pode ocorrer de forma quase que "desleixada", pois o racismo estrutural encrustado nas sociedades seria suficiente para que a base de dados que alimenta uma IA seja incompleta, defeituosa ou generalizada, gerando uma previsão binária de que todas as pessoas

de determinada raça iriam se comportar da mesma forma, ou fazendo com que a IA realize correlações espúrias.

Ainda, no cabedal das leis de proteção de dados pessoais, temos a existência do direito à revisão de decisões automatizadas.

3.3 Direito à revisão de decisões automatizadas

A LGPD, em seu art. 20, trata sobre o direito do titular de solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de seus dados pessoais que afetem seus interesses. São listadas em rol não taxativo as decisões destinadas a definir o perfil pessoal, profissional, de consumo e de crédito do titular, ou aspectos de sua personalidade. Essas decisões comuns no nosso dia a dia, podem trazer prejuízos e desencadear abusos.

Caberá ao controlador, nos termos do § 1º do art. 20 o fornecimento de informações claras e adequadas a respeito dos critérios e dos procedimentos utilizados para a tomada de decisões automatizadas, quando solicitado, observados os segredos comercial e industrial. Em caso de não fornecimento das informações sob a alegação de que se trataria de segredo comercial e industrial, a ANPD – Autoridade Nacional de Proteção de Dados, poderá realizar auditoria para verificação de aspectos discriminatórios em tratamento automatizado de dados pessoais.

A previsão do direito do titular em solicitar a revisão de tais decisões, coaduna com o princípio da transparência (art. 6º, VI, LGPD), que preconiza que o titular receba informações claras sobre o tratamento de seus dados e determina a obrigação aos agentes de tratamento de não repassar informações imprecisas ou confusas.

Asseveram Muholland e Frajhof (2019) que, em se tratando de decisões baseadas em tratamento automatizado, necessário investigar como será feita a segurança dos dados coletados e como serão tutelados os titulares cujos dados pessoais serviram de insumo para que o sistema de inteligência artificial alcançasse a tomada de decisão, haja vista que tal decisão teria o potencial de negar ou não promover direitos, ou até mesmo de gerar processos discriminatórios abusivos ou ilegítimos (2019).

Casos diversos de prejuízos a direitos fundamentais, discriminação e abusos decorrentes do uso de sistemas de inteligência artificial para tomada de decisões têm sido notificados ao redor do mundo e merecem destaque nesse ponto. Um dos casos mais impactantes diz respeito ao uso de decisões automatizadas no âmbito da justiça criminal, nos Estados Unidos.

Nos EUA, ferramentas de IA são adotadas no âmbito do Sistema Criminal Norte-Amerciano para práticas de *risk assessement* - os algoritmos são empregados para determinar se indivíduos submetidos ao Sistema Criminal representa potenciais riscos, para fins de se estabelecer a dosimetria da pena, a fixação de regimes de cumprimento e execução de suas penas, a concessão ou não de liberdade condicional e a efetivação de prisões provisórias.

> (...) submetidos os dados da pessoa a ser avaliada ao(s) programa(s) de IA, que calculará, segundo os algoritmos nele implementados, o risco que tal indivíduo traz à sociedade, os profissionais envolvidos (Juízes, Agentes penitenciários, Agentes de *probation,* Psicólogos etc.) acessam, em geral por um link no qual tenham autorização de acesso, o score obtido na avaliação. (SULOKI, 653, 2019)

Essa ideia de justiça preditiva, por si só, de que um algoritmo possa ser capaz de prever a reincidência de um indivíduo, constitui grave violação de direitos fundamentais e aponta para a necessidade de se refletir com mais critérios sobre o uso de tratamento automatizado de dados pessoais, também do ponto de vista ético, além do jurídico, sob pena, nas palavras de Suloki, de reproduzirmos exatamente aquilo que pretendemos combater.

O potencial lesivo do tratamento de dados pessoais realizado pelos processos algorítmicos aqui descritos clarifica a necessidade de garantia de uma espécie de direito à explicação, que pode ser traduzido no "direito de receber informações suficientes e inteligíveis que permita o titular dos dados entender a lógica e os critérios utilizados para tratar seus dados pessoais para uma ou várias finalidades" (MONTEIRO, p. 4, 2018). Embora não previsto expressamente na LGPD, o direito à explicação pode ser depreendido da mesma, bem como também do GDPR, como um instrumento para concretização do princípio da transparência e mitigação de possíveis riscos de potenciais discriminações, e abuso no tratamento de dados pessoais, em observância também a princípios como o livre acesso (art. 6º, IV, LGPD) e a não discriminação (art. 6º, XI, LGPD).

3.3 *Privacy by design*

Sem dúvidas, um dos pontos trazidos pela Lei Geral de Proteção de Dados de maior potencial de aplicação eficaz para questões referentes ao uso de sistemas de IA e combate a seus vieses, é a ideia do *privacy by design*.

O *privacy by design*, previsto no art. 46, § 2º da LGPD é refere-se à incorporação de medidas técnicas e administrativas aptas à proteção dos dados pessoais desde a fase de conceção de um novo produto ou serviço. Quando aplicado à governança de sistemas de inteligência artificial o *privacy by design* tem o condão de gerar significativas contribuições.

A metodologia do *privacy by design*, se aplicada no desenvolvimento de sistemas de IA, implicaria na privacidade sendo priorizada como configuração padrão da mesma forma que seria embutida no design de tais sistemas. Da mesma forma, já se estabeleceria a segurança de ponta a ponta nessas novas soluções e a transparência aos usuários e titulares de dados pessoais acerca da coleta e tratamento de seus dados.

A aplicação do *privacy by design* é quase que "um mundo ideal" para agentes de tratamento que desenvolvam IAs, tendo em vista que poderiam, desde o início, ser implementados critérios para se garantir a qualidade dos dados (desde sua coleta e pré-processamento), bem como mecanismos para se garantir a transparência do tratamento, dentro daquilo que for possível auditar em tratamentos automatizados. Além disso, poderiam ser adotadas as chamadas *Privacy Enhancing Technologies* – PETs, tecnologias que reforçam e melhoram a privacidade e que sejam capazes de abrigar sob seu guarda-chuvas toda e qualquer tecnologia que seja minimamente amigável e facilitadora da privacidade (LEMOS, BRANCO, 2021).

O *privacy by design* não é opcional – é de observância obrigatória a todos os agentes de tratamento de dados pessoais, inclusive os que desenvolvem e usam inteligência artificial, e não se trata de obrigatoriedade presente apenas na Lei Geral de Proteção de Dados, com também no General Data Protection Regulation, da União Europeia.

CONSIDERAÇÕES FINAIS

A ascensão das inteligências artificiais generativas trouxe uma série de benefícios, mas também inúmeros desafios, configurando, dentre os desafios mais proeminentes, os vieses algorítmicos e a proteção dos dados pessoais.

Não é tecnicamente correto afirmar que o ChatGPT é preconceituoso, mas ele pode se travestir em vetor de perpetuação de preconceitos, discriminações a abusos, assim como outras

ferramentas de IA generativa. Não obstante sua reputação de imparcialidade, modelos de IA acabam por refletir objetivos, modelos e ideologias. Nas sábias palavras de O'neil (2020) "modelos são opiniões embutidas em matemática".

É cediço, todavia, que o combustível que alimenta esses modelos são os dados, dentre os quais dados pessoais em larga escala, de modo que se faz essencial a análise das implicações do desenvolvimento e uso de sistemas de inteligência artificial à luz do direito à proteção dos dados pessoais, mais especificamente, no contexto brasileiro, da Lei Geral de Proteção de Dados. A LGPD, com seus princípios e diretrizes, oferece uma base sólida para enfrentar esses desafios. Princípios como a não discriminação, a transparência e a qualidade dos dados são fundamentais para mitigar os riscos associados ao uso de IA.

Essa análise à luz da Lei Geral de Proteção de Dados revela, em última instância, a complexidade de se equilibrar inovação tecnológica e proteção de dados pessoais e de direitos fundamentais dos indivíduos, e aponta para a necessidade urgente de se estabelecer regras e mecanismos de governança da inteligência artificial, bem como de se avaliarem as questões regulatórias.

O tema é atual e os *cases* vão se multiplicar com a mesma velocidade em que o uso da inteligência artificial avança e a cultura preconizada pela LGPD passa a ser internalizada pelos cidadãos e instituições brasileiras.

É evidente que o Brasil possui na LGPD um instrumento poderoso para endereçar muitas das questões éticas e jurídicas levantadas pelo uso de IA. No entanto, propõe-se que as agendas de pesquisa sigam se debruçando sobre o assunto, a fim de oferecer suporte para o desenvolvimento de normas voltadas à regulação de questões decorrentes do desenvolvimento e uso da inteligência artificial que sejam, ao mesmo tempo, suficientemente coercitivas e suficientemente flexíveis para se adaptar às evoluções da IA.

REFERÊNCIAS

BRASIL. **Lei n.° 13.709 de 14 de agosto de 2018. Lei Geral de Proteção de Dados.** Brasília, DF.

BIONI, Bruno Ricardo; LUCIANO, Maria. **O Princípio da Precaução na Regulação de Inteligência Artificial: seriam as leis de proteção de dados o seu portal de entrada?** In:

FRAZÃO, Ana; MULHOLLAND, Caitlin (Org.). Inteligência artificial e direito: ética, regulação e responsabilidade. São Paulo: Thomson Reuters Brasil, 2019.

BIONI, Bruno; GARROTE, Marina; GUEDES, Paula. **Temas centrais na Regulação de IA: O local, o regional e o global na busca da interoperabilidade regulatória.** São Paulo: Associação Data Privacy Brasil de Pesquisa, 2023.

FRAZÃO, Ana; MULHOLLAND, Caitlin. **Inteligência artificial e direito: ética, regulação e responsabilidade.** São Paulo: Thomson Reuters Brasil, 2019.

JOHANSEN, Robert. **Full-spectrum thinking: how to escape boxes in a post-categorical future.** Oakland, CA: Berret-Koehler Publishers, 2020.

LEMOS, Ronaldo; BRANCO, Sérgio. ***Privacy by design:* conceito, fundamentos e aplicabilidade na LGPD.** In: MENDES, Laura Schertel; DONEDA, Danilo; SARLET, Ingo Wolfgang. RODRIGUES JR. Otávio Luiz; BIONI; Bruno. Tratado de Proteção de Dados Pessoais. Rio de Janeiro: Forense, 2021.

MALDONADO, Viviane Nóbrega; BLUM, Renato Ópice (coord). **LGPD - Lei Geral de Proteção de Dados pessoais: comentada.** São Paulo: Thomson Reuters Brasil, 2021.
MARTINS, Guilherme Magalhães; LONGHI, João Victor Rozatti (coord.); MUCELIN, Guilherme (org.). **Direito Digital: direito privado e internet.** 5a. ed. Indaiatuba: Editora Foco, 2024.

MASSENO, Manuel David. **Entre dados e algoritmos: como a União Europeia procura proteger os cidadãos em tempos de inteligência artificial e big data.** In: MARTINS, Guilherme Magalhães; LONGHI, João Victor Rozatti (coord.); MUCELIN, Guilherme (org.). Direito Digital: direito privado e internet. 5a. ed. Indaiatuba: Editora Foco, 2024.

MENDES, Laura Schertel; DONEDA, Danilo; SARLET, Ingo Wolfgang. RODRIGUES JR. Otávio Luiz; BIONI; Bruno. **Tratado de Proteção de Dados Pessoais**. Rio de Janeiro: Forense, 2021.

MENDES, Laura Schertel; MATTIUZZO, Marcela; FUJIMOTO, Mônica Tiemy. **Discriminação algorítmica à luz da Lei Geral de Proteção de Dados.** In: MENDES, Laura Schertel; DONEDA, Danilo; SARLET, Ingo Wolfgang. RODRIGUES JR. Otávio Luiz; BIONI; Bruno. Tratado de Proteção de Dados Pessoais. Rio de Janeiro: Forense, 2021.

MONTEIRO, Renato Leite. **Existe um direito à explicação na Lei Geral de Proteção de Dados no Brasil?** Rio de Janeiro: Instituto Igarapé, 2018.

O'NEIL, Cathy. **Algoritmos de destruição em massa: como o *big data* aumenta a desigualdade e ameaça a democracia**. Trad. Rafael Abraham, 1ª ed, Santo André, Editora Rua do Sabão, 2020.

ROSSETTI, Regina; GARCIA, Kethly. **Inteligência artificial generativa: questões jurídicas e éticas em torno do ChatGPT.** VirtuaJus, Belo Horizonte, v. 8, n. 15, p. 253-264, 2º sem. 2023 – ISSN 1678-3425

SULOKI, Victoria. **Novas tecnologias, velhas discriminações: ou da falta de reflexão sobre o sistema de algoritmos na justiça criminal.** In: FRAZÃO, Ana; MULHOLLAND, Caitlin (Org.). Inteligência artificial e direito: ética, regulação e responsabilidade. São Paulo: Thomson Reuters Brasil, 2019.

The Center for Artificial Intelligence and Digital Policy (CAIDP), **In the matter of OpenAI, Inc.** FEDERAL TRADE COMMISSION, Washington, DC 20580, March 30, 2023.

UNIÃO EUROPEIA. Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho de 27 de abril de 2016 relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (**Regulamento Geral sobre a Proteção de Dados**). Jornal Oficial da União Europeia, L 119, 4 maio 2016. Disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX%3A32016R0679. Acesso em: 10 jun. 2024.