

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO
FACULDADE DE DIREITO

LARA SALGUEIRO OLIVEIRA

**ANÁLISE DA INCIDÊNCIA DE RESPONSABILIDADE POR CONTEÚDOS
ORIGINADOS POR INTELIGÊNCIA ARTIFICIAL GENERATIVA**

São Paulo
2024



PUC-SP

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO

FACULDADE DE DIREITO

LARA SALGUEIRO OLIVEIRA

**ANÁLISE DA INCIDÊNCIA DE RESPONSABILIDADE POR CONTEÚDOS
ORIGINADOS POR INTELIGÊNCIA ARTIFICIAL GENERATIVA**

Projeto de Pesquisa do Trabalho de Conclusão de Curso, como requisito parcial para obtenção do título de bacharel em Direito, sob orientação do Prof. Dr Francisco Cahali.

São Paulo
2024

LARA SALGUEIRO OLIVEIRA

**ANÁLISE DA INCIDÊNCIA DE RESPONSABILIDADE POR CONTEÚDOS
ORIGINADOS POR INTELIGÊNCIA ARTIFICIAL GENERATIVA**

BANCA EXAMINADORA:

Prof. Dr. Francisco Cahali

Prof. Dr. Adriano Ferriani

“A menos que modifiquemos a nossa maneira de pensar, não seremos capazes de resolver os problemas causados pela forma como nos acostumamos a ver o mundo”.

(Albert Einstein)

AGRADECIMENTOS

À minha família por todo o amor, apoio e incentivo durante não só desses últimos cinco anos, mas dos últimos vinte e três.

Ao meu namorado Guilherme por todo o amor, apoio e incentivo diário, que me fazem acreditar que eu sempre posso ir mais longe.

À minha amiga Bianca, por ter sido a melhor companhia que eu poderia ter ao longo desta jornada.

Ao meu orientador, Prof. Dr. Francisco Cahali, por seu suporte e orientação indispensáveis para o desenvolvimento desta tese.

RESUMO

A Inteligência Artificial generativa emergiu como uma das tecnologias mais revolucionárias da história. Sua capacidade de originar conteúdos de forma autônoma e extremamente realista a tornaram rapidamente uma ferramenta de extrema relevância em nossas vidas diárias. Contudo, apesar de seus impactos positivos serem irrefutáveis, seu desenvolvimento vem acompanhado de questões éticas, legais e sociais tão complexas quanto sua própria natureza, colocando o mundo diante de um cenário obscuro e desafiador, especialmente ao considerar que a atribuição de responsabilidade por eventuais violações se torna nebulosa quando não podemos identificar claramente um agente humano específico por trás dessas ações. Diante disso, esta pesquisa concentra-se em analisar alguns dos regimes de responsabilidade discutidos e adotados por diferentes jurisdições, com o objetivo de contribuir, especialmente no contexto brasileiro, com os debates e com a construção de um marco regulatório sólido, apto a equilibrar a promoção da inovação com a salvaguarda dos direitos e garantias individuais e sociais.

Palavras-chave: Tecnologia; Inteligência Artificial Generativa; Regulação; Responsabilidade.

ABSTRACT

Generative Artificial Intelligence has emerged as one of the most revolutionary technologies in history. Its ability to originate content autonomously and in an extremely realistically way has quickly made it an immensely relevant tool in our daily lives. However, despite its undeniably positive impacts, its development is accompanied by ethical, legal, and social issues as complex as its very nature, placing the world in an obscure and challenging scenario. This is especially true when considering that assigning responsibility for potential violations becomes nebulous when we cannot clearly identify a specific human agent behind these actions. In light of this, this research focuses on analyzing some of the liability regimes discussed and adopted by different jurisdictions, aiming to contribute, especially in the Brazilian context, to the debates and the construction of a solid regulatory framework, capable of balancing the promotion of innovation with the safeguarding of individual and social rights and guarantees.

Keywords: Technology; Generative Artificial Intelligence; Regulation; Liability.

LISTA DE ABREVIACOES

ABRIA – Associao Brasileira de Inteligncia Artificial

AILD – Artificial Intelligence Directive

CTIA - Comisso Temporria de Inteligncia Artificial do Senado

GANs – Redes Adversrias Generativas

IA - Inteligncia Artificial

IDEC – Instituto de Defesa do Consumidor

LLM - Grandes Modelos de Linguagem

PL – Projeto de Lei

PLD – Product Liability Directive

UE – Unio Europeia

LISTA DE FIGURAS

| | |
|---|-----------|
| Figura 1 - Investimentos para treinamento de sistemas de IA | 8 |
| Figura 2 - Evolução das imagens geradas pelo MidJourney | 11 |
| Figura 3 - Fluxograma de funcionamento da IA generativa | 12 |
| Figura 4 - Utilização da IA na indústria..... | 13 |
| Figura 5 - “Prisão de Donald Trump”..... | 16 |
| Figura 6 - “Papa Francisco vestindo um casaco branco” | 17 |
| Figura 7 - Reprodução de marca d’água da Getty Images pela Stability AI..... | 18 |
| Figura 8 - Propostas legislativas a redor do mundo para regular IA | 20 |

SUMÁRIO

| | |
|--|----|
| 1. Introdução | 1 |
| 2. A tecnologia no mundo contemporâneo | 3 |
| 3. Panorama Geral da Inteligência Artificial | 6 |
| 3.1. Entendendo a Inteligência Artificial Generativa | 10 |
| 3.2. Discussões a respeito da IA Generativa | 14 |
| 4. Análise da incidência de responsabilidade e panorama regulatório | 20 |
| a. Brasil | 22 |
| 4.a.1. Críticas e ponderações a respeito do texto substitutivo do PL 2338 | 27 |
| b. União Europeia | 30 |
| c. Estados Unidos | 33 |
| 4.1. Quadro-resumo das abordagens regulatórias..... | 34 |
| 5. Conclusão | 36 |
| ANEXO 1 | 39 |
| 6. Bibliografia | 41 |

1. Introdução

Em uma breve análise do panorama contemporâneo, é fácil constatar que a Inteligência Artificial generativa emerge como uma das tecnologias mais revolucionárias da história. São muitos os fatores que sustentam essa afirmação, sendo o principal deles associado à capacidade ímpar de reproduzir funções e habilidades que antes se restringiam quase que exclusivamente ao *Homo Sapiens*, tais como a inteligência e a capacidade criativa.

Nos últimos anos, especialmente com o advento de plataformas como o ChatGPT, Copilot e Gemini, os impactos da IA tornaram-se mais tangíveis no cotidiano, transcendendo a mera promessa de avanços tecnológicos. Sua capacidade de criar, em segundos, conteúdos novos e originais, como textos, imagens, músicas e até mesmo vídeos, fez com que rapidamente esses sistemas se consolidassem como ferramentas essenciais para uma série de atividades e segmentos. Profissionais de marketing, desenvolvedores de software, artistas e até profissionais da saúde começaram a incorporar essas tecnologias em seus processos diários, otimizando tarefas e aprimorando processos criativos.

Contudo, seu imenso potencial não vem desacompanhado de riscos e preocupações. O uso e desenvolvimento da IA generativa, embora possua inúmeras aplicações benéficas, também suscita questões significativas relacionadas à privacidade, segurança e ética. Tópicos complexos como viés algorítmico, *deepfakes*, direitos da personalidade e propriedade intelectual são apenas algumas das pautas que guiam robustas discussões em todo o mundo.

Nesse sentido, ressalta-se que esses debates se tornam ainda mais sensíveis e complexos ao considerar a natureza singular da inteligência artificial, que transcende a compreensão e domínio humano. Os processos de aprendizado de máquina e as complexas combinações algorítmicas são tão enigmáticos que, muitas vezes, até mesmo os próprios desenvolvedores desses sistemas se veem incapazes de identificar exatamente como e por que determinadas combinações de dados e variáveis resultaram na geração de um conteúdo específico, o que coloca o mundo inteiro diante de um cenário obscuro e sem precedentes.

Toda essa complexidade que permeia o tema faz com que o panorama regulatório ainda permaneça bastante nebuloso. Embora haja um esforço global para desenvolver arcabouços legais que sejam capazes de enfrentar os desafios e particularidades específicas desse campo em rápida evolução, ainda existem muitas lacunas a serem preenchidas.

Essas lacunas são ainda mais exacerbadas ao considerar que não podemos identificar claramente um agente humano específico por trás dessas ações, o que acaba por dificultar a

aplicabilidade de um mecanismo essencial para a garantia do *enforcement* de qualquer normativa: a responsabilização.

Dada a relevância deste instituto e diante da oportunidade de contribuir com as discussões em torno da regulação da IA, nesta pesquisa será feito um recorte temático a fim explorar, através da análise de diferentes abordagens regulatórias, a incidência de responsabilidade por conteúdos produzidos por Inteligência Artificial Generativa.

2. A tecnologia no mundo contemporâneo

A tentativa de prever e moldar o próprio futuro tem sido uma constante ao longo da história humana, refletindo os mais profundos anseios do homem desde tempos imemoriais: compreender e ter controle sobre seu próprio destino. Nos últimos cem anos, em particular, o futuro do homem tem sido frequentemente associado ao avanço da tecnologia. Não é coincidência que simples buscas pela palavra "futuro" no Google Imagens remetam, em sua grande maioria, à imagens de robôs, máquinas, computadores e grandes civilizações tecnológicas.

Contudo, antes de nos aprofundarmos em discussões sobre o protagonismo da tecnologia no mundo contemporâneo, é essencial estabelecer uma definição clara do que esse termo abrange. De acordo com o dicionário Oxford, a palavra tecnologia é definida como a “teoria geral e/ou estudo sistemático sobre técnicas, processos, métodos, meios e instrumentos de um ou mais ofícios ou domínios da atividade humana (p.ex., indústria, ciência etc.)”. Jon Agar (2020), também traz uma definição simples, porém, interessante para o termo, e conceitua-o como “meios materiais projetados para um fim”.

Uma vez atribuídos tais significados ao termo “tecnologia”, torna-se evidente que, apesar de parecer algo recente, a tecnologia está presente na vida do homem há muito mais tempo do que se imagina. Desde os primórdios de sua história, os seres humanos têm desenvolvido e utilizado técnicas, processos e instrumentos para melhorar suas vidas e realizar tarefas de maneira mais eficiente. Seja na invenção da roda ou na descoberta do fogo, a tecnologia tem sido uma parte essencial da jornada humana ao longo dos séculos, moldando progressivamente o curso da civilização.

No entanto, especialmente nas últimas duas décadas, a periodicidade na qual as mudanças e descobertas revolucionárias ocorrem parece ter adotado um ritmo singular e frenético, como bem pontua Harari:

“O fato é que, atualmente, todo ano é revolucionário. Hoje até mesmo uma pessoa de 30 anos pode dizer honestamente a adolescentes incrédulos “Quando eu era jovem, o mundo era completamente diferente”. A internet, por exemplo, só se disseminou no início dos anos 1990, há pouco mais de vinte anos. Hoje não podemos mais imaginar o mundo sem ela”. (HARARI, 2019, p. 376)

A tecnologia tem se integrado tão profundamente em nosso cotidiano que, muitas vezes, sequer notamos seu caráter onipresente. Desde o despertar até a hora de dormir, estamos imersos em um mundo digital, interagindo incessantemente com uma série de dispositivos e sistemas. A tecnologia molda a maneira como nos relacionamos, como aprendemos e até nossa

percepção da realidade. Suas aplicabilidades e funcionalidades abrangem desde os aspectos mais básicos do cotidiano, até a resolução de problemas globais urgentes, como as questões climáticas.

Ao longo dos últimos anos, nota-se também, uma grande mudança em relação ao próprio papel e significado atribuído à tecnologia no mundo contemporâneo. A tecnologia transcendeu sua condição de mera ferramenta para o desenvolvimento humano e se consolidou como um mercado altamente lucrativo. Nesta dinâmica, as grandes empresas - as denominadas "big techs" - competem incessantemente pela supremacia neste domínio, investindo bilhões em pesquisa e desenvolvimento para criar soluções personalizadas e inovadoras que possam revolucionar nossas vidas.

Para atingir esse fim, e estruturar algoritmos capazes de inferir com precisão os interesses dos usuários, são realizadas coletas massivas dos mais variados tipos de dados, que uma vez processados e integrados, fornecem parâmetros precisos para alinhar as funcionalidades tecnológicas às necessidades e desejos dos usuários. O resultado são soluções sob medida, criadas para atender às exigências particulares de cada um e estrategicamente arquitetadas para capturar a atenção e fidelidade.

Em contrapartida, a concentração e manipulação de dados pessoais, juntamente com o potencial invasivo da tecnologia em nossas vidas também suscitam questões legítimas sobre privacidade e segurança, além de possíveis infrações aos direitos dos usuários. À medida que as corporações expandem e aprimoram suas estratégias de coleta e tratamento de dados para alimentar e treinar algoritmos de inteligência artificial e sistemas de *machine-learning*, emergem preocupações a respeito de como essas informações são utilizadas, protegidas e, em especial, com a quantidade de poder acumulado por essas empresas.

Esse cenário nebuloso, impulsiona um novo e legítimo movimento: o *techlash*. Marcando o rompimento com a visão utópica que até então se tinha em relação ao desenvolvimento da tecnologia, esse novo período caracteriza-se por preocupações em torno do controle e das intenções, muitas vezes escusas, das grandes empresas de tecnologia, fortalecendo a visão negativa e desconfiada dos usuários e da mídia em relação ao avanço desse mercado.

As expressões "capitalismo de vigilância" e "colonialismo de dados" emergiram como manifestações desse pessimismo e da crescente preocupação em relação ao desenvolvimento tecnológico. A insegurança diante do potencial nocivo da tecnologia levou a um intenso debate global e à elaboração de propostas legislativas para conter seu avanço descontrolado. Essas movimentações resultaram na criação de leis significativas, como a General Data Protection

Regulation na Europa, a Lei Geral de Proteção de Dados e o Marco Civil da Internet no Brasil, que refletem o esforço coletivo para proteger os direitos dos usuários e regular a atuação das empresas de tecnologia.

O que se observa, desde então, é que atualmente, o mercado de tecnologia tem sido moldado por dois grandes vetores: a inovação e a regulação. Em alguns momentos, esses vetores projetam suas forças em direções e sentidos semelhantes, já em outros, em direções opostas. Como se verá adiante, o dilema contemporâneo reside justamente em encontrar um ponto de equilíbrio entre essas duas forças.

3. Panorama Geral da Inteligência Artificial

Conforme delineado por Nilsson (2009), a Inteligência Artificial pode ser compreendida como um conjunto de técnicas destinadas à criação de máquinas inteligentes capazes de resolver problemas que demandam raciocínio humano. E, apesar desse termo ter dominado grande parte dos debates sobre tecnologia nos últimos anos, é crucial reconhecer que essa tecnologia não se trata de uma novidade tão recente e repentina.

Embora os registros precisos sobre o surgimento da Inteligência Artificial sejam escassos, a concepção de seus fundamentos teóricos encontra indícios no ano de 1943, quando Warren McCulloch e Walter Pitts desenvolveram o primeiro modelo computacional de redes neurais. Este marco inicial lançou as bases para explorar a capacidade das máquinas de realizar tarefas, até então, exclusivamente condicionadas ao raciocínio humano. No entanto, conceitos semelhantes já haviam sido abordados em ensaios anteriores, como "O Homem-Máquina" de Julien Offray de la Mettrie, publicado em 1748.

Avançando nesse campo, em 1950, Alan Turing, concebeu o famoso Teste de Turing, que estabelecia parâmetros para descobrir se uma máquina seria capaz de transmitir informações de forma tão convincente quanto um ser humano, sem levantar suspeitas de que se tratava de um programa de computador. No mesmo ano, Turing também publicou o artigo "Computing Machinery and Intelligence", trazendo expressivas, e até hoje relevantes, contribuições para estudos na área.

Porém, foi somente em 1956, durante a Conferência do Dartmouth College, em New Hampshire (EUA), que o termo "Inteligência Artificial" foi empregado pela primeira vez para descrever esse novo campo de conhecimento. Nesse evento histórico, os pioneiros da IA, incluindo nomes como John McCarthy e Marvin Minsky, reuniram-se para desenvolverem estudos em torno da premissa de que "todo aspecto de aprendizado ou qualquer característica da inteligência conseguiria, por princípio, ser tão precisamente descrito que uma máquina pode ser criada para simulá-la"¹. Essa conferência marcou o início formal da pesquisa e desenvolvimento da IA como disciplina acadêmica e campo de estudo multidisciplinar.

Nos anos subsequentes, diante do potencial verdadeiramente revolucionário da inteligência artificial, os estudos e investimentos na área ganharam um impulso considerável. As décadas de 50 e 60, conhecidas como "the Golden Years of AI", testemunharam períodos de avanços significativos, conforme destacado por Barbosa e Bezerra (2020). Em 1958,

¹ A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. Disponível em: <https://home.dartmouth.edu/about/artificial-intelligence-ai-coined-dartmouth>

emergiu a linguagem de programação Lisp, que se tornou o padrão em sistemas de inteligência artificial, e que exerce influência sobre uma ampla variedade de linguagens, até os dias de hoje. No ano seguinte, em 1959, o termo *machine-learning* foi utilizado pela primeira vez, delineando uma nova frente na era da computação. E em 1964, Joseph Weizenbaum apresentou ao mundo a Eliza, o primeiro chatbot, que, a partir de palavras-chave e uma estrutura sintática, simulava interações automáticas, buscando simular o papel de um psicanalista.

Contudo, é necessário frisar que, apesar de relativamente rápido em comparação à complexidade desta tecnologia, o progresso na área de Inteligência Artificial não foi linear. Houve períodos de grande entusiasmo e avanço, seguidos por períodos de desilusão e estagnação, conhecidos como “invernos da IA”, ciclos que refletem a complexidade e os desafios inerentes ao desenvolvimento de sistemas capazes de replicar ou superar as capacidades cognitivas humanas.

Até o momento, pode-se dizer que o desenvolvimento da Inteligência Artificial experimentou dois “invernos”. O primeiro ocorreu aproximadamente entre 1974 e 1980, quando as capacidades dos computadores não estavam à altura das ideias e expectativas para o funcionamento da IA, o que gerou grandes frustrações aos desenvolvedores, pois as capacidades práticas ficavam aquém das visões ambiciosas. O segundo “inverno” ocorreu entre 1987 e 1993, quando as promessas e especulações promovidas tanto pelos desenvolvedores quanto pela mídia, foram novamente confrontadas pelas limitações computacionais da época.

Após esses períodos de desilusão frente ao avanço desta tecnologia, os investimentos e interesses na área de Inteligência Artificial começaram a ressurgir nos primeiros anos do século XXI, à medida que a tecnologia avançava e a computação se aperfeiçoava, impulsionando o desenvolvimento da IA para um novo nível de realismo e ambição.

A estagnação que dominou os anos anteriores começou a ser efetivamente superada em 2012, com a iminência da denominada “*Deep Learning Revolution*”. Durante este período, importantes marcos foram alcançados no campo do aprendizado de máquina, incluindo o desenvolvimento de uma rede neural no laboratório X do Google, capaz de identificar padrões e reconhecê-los em diferentes contextos. Além disso, durante o *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC), promovido pela *University of Toronto*, os pesquisadores Krizhevsky, Sutskever e Hinton, treinando sua rede neural denominada “AlexNet”, conseguiram reduzir pela metade a taxa de erro no reconhecimento visual do *dataset* “ImageNet”, alcançando o percentual recorde de 15,3%².

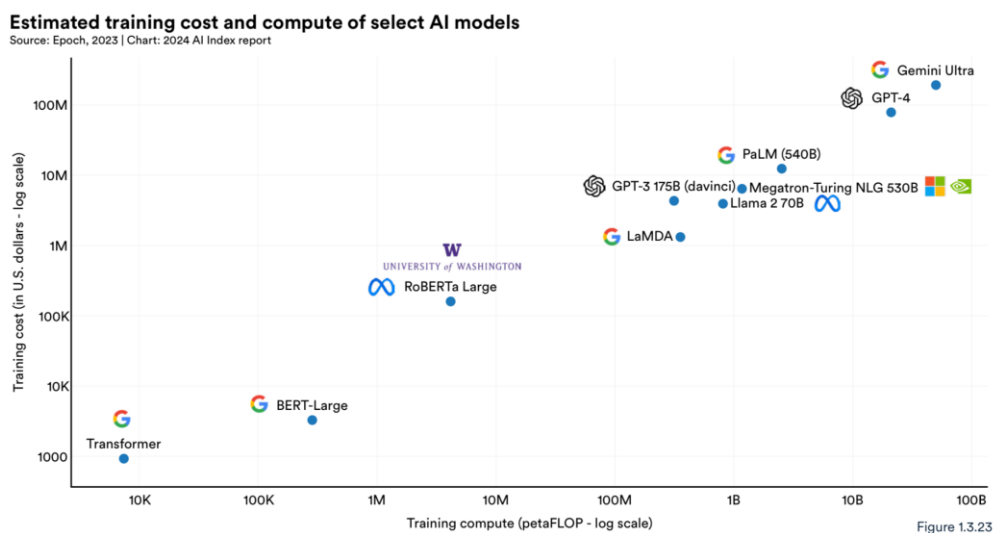
² Disponível em:

https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

Esses avanços significativos preencheram as lacunas que ainda inibiam o desenvolvimento da IA. A partir desse ponto crucial, a capacidade computacional, o aprendizado de máquina e os algoritmos começaram a evoluir em um ritmo verdadeiramente impressionante. Essa rápida progressão proporcionou às empresas desenvolvedoras os meios necessários para alcançarem, finalmente, suas expectativas e pretensões neste mercado, antes consideradas quase que inatingíveis.

Desde então, os investimentos no desenvolvimento de IA experimentaram um significativo *boom*, dando início ao que pode ser chamada de *AI race*. Nessa competição acirrada, empresas do mundo inteiro se engajaram em uma busca frenética pelo avanço dessas tecnologias. Essa corrida, não só impulsionou a pesquisa e o desenvolvimento, mas também levou à formação de parcerias estratégicas, aquisições agressivas e até mesmo rivalidades intensas entre as principais empresas do setor de tecnologia. Toda essa movimentação nos leva ao cenário atual, em que a IA é o grande alvo das empresas de tecnologia, que investem altos montantes em seu desenvolvimento, conforme demonstrado no gráfico abaixo, extraído do 2024 Stanford AI Index³.

Figura 1 - Investimentos para treinamento de sistemas de IA



Fonte: 2024 Stanford AI Index⁴

Contudo, é importante ressaltar que a corrida pelo desenvolvimento da IA não é exclusivamente disputada pelo setor privado; as autoridades públicas também têm atuado

³ Disponível em: <https://aiindex.stanford.edu/report/>

⁴ Disponível em: <https://aiindex.stanford.edu/report/>

ativamente e demonstrado grande interesse nesse mercado, tanto por meio de investimentos e incentivos para o desenvolvimento do setor, quanto através de propostas regulatórias. Liderar este campo, nas frentes de inovação e regulação, não é apenas uma questão de prestígio, mas também representa a garantia de uma posição privilegiada em um mercado global, que está apenas começando a ser explorado.

Em meio a esse contexto, tornou-se evidente que a inteligência artificial se consolidou como uma presença tangível e ubíqua em nosso cotidiano. Nos últimos anos, os investimentos maciços no setor impulsionaram avanços significativos, resultando em uma proliferação de aplicativos, dispositivos e serviços que fazem uso da IA para melhorar e simplificar várias áreas da vida humana. Desde assistentes virtuais em smartphones até carros autônomos, a IA permeia diversas facetas de nossas interações diárias, demonstrando seu potencial transformador.

E no epicentro de toda essa inovação, destaca-se um avanço particularmente marcante: a inteligência artificial generativa, um ponto de virada notável no desenvolvimento da inteligência artificial. A partir dela, e como será exposto nos próximos capítulos, a capacidade criativa não mais se restringe exclusivamente ao domínio humano, desafiando assim uma lógica perpetuada em nossa sociedade por anos: a suposição de que o homem detém o controle absoluto sobre tudo ao seu redor, como bem colocado por Kaufman:

“Outro fato inédito é que pela primeira vez o homem criou algo sob o qual não tem controle; os especialistas não são capazes de afirmar exatamente como as máquinas funcionam e como elas se comportarão no futuro. Decorre desse desconhecimento a origem dos riscos e do imponderável, afetando o futuro da humanidade”.
(KAUFMAN, 2016, p. 3)

3.1. Entendendo a Inteligência Artificial Generativa

Conforme esclarece Feuerriegel et al. (2023) o termo IA generativa pode ser empregado para referir-se às técnicas de computação capazes de gerar conteúdo aparentemente novo e significativo, como texto, imagens ou áudio, a partir de dados de treinamento. Em outras palavras, a IA generativa é uma categoria de IA capaz de gerar conteúdos de forma autônoma utilizando modelos de aprendizado de máquina que identificam padrões nos dados fornecidos durante o treinamento.

Seu potencial revolucionário, evidenciado em sua própria definição, fez com que em um curto lapso temporal, os sistemas avançados de inteligência artificial generativa, como o Chat GPT, Copilot e MidJourney, provocassem uma transformação radical nas mais diversas esferas da sociedade. Essas ferramentas não apenas tornaram o potencial da IA mais acessível e tangível para o público em geral, mas também redesenharam os limites da criação de conteúdo. Com a capacidade de gerar textos, códigos e até mesmo imagens de forma intuitiva e colaborativa, esses sistemas abriram novos horizontes, marcando o advento de uma nova era: a da criatividade sem precedentes, onde as fronteiras entre a inteligência do homem e das máquinas se tornam cada vez mais indissociáveis.

Contudo, antes de analisarmos o potencial e aplicabilidades da IA generativa, é importante trazer uma breve contextualização sobre como essa tecnologia funciona. Conforme esclarecido por Banh e Strobel (2023), a IA generativa opera através de um tipo de aprendizagem profunda chamada de redes adversárias generativas (GANs). As GANs recebem esse nome porque treinam simultaneamente duas redes diferentes: a geradora e a discriminadora, colocando-as uma contra a outra. A geradora cria amostras de dados realistas, enquanto a discriminadora faz a distinção entre amostras reais e geradas (Pan et al., 2019). Ambas as redes neurais são treinadas juntas até que a discriminadora não seja capaz de diferenciar as duas amostras (Janiesch et al., 2021).

A contraposição entre ambas, faz com que a rede geradora melhore seus recursos de geração de dados ao longo do tempo, acabando por produzir resultados realistas e de alta qualidade, como evidenciado na figura abaixo, que demonstra a evolução das imagens geradas pela MidJourney, dentro de um período de 2 anos.

Figura 2 - Evolução das imagens geradas pelo MidJourney

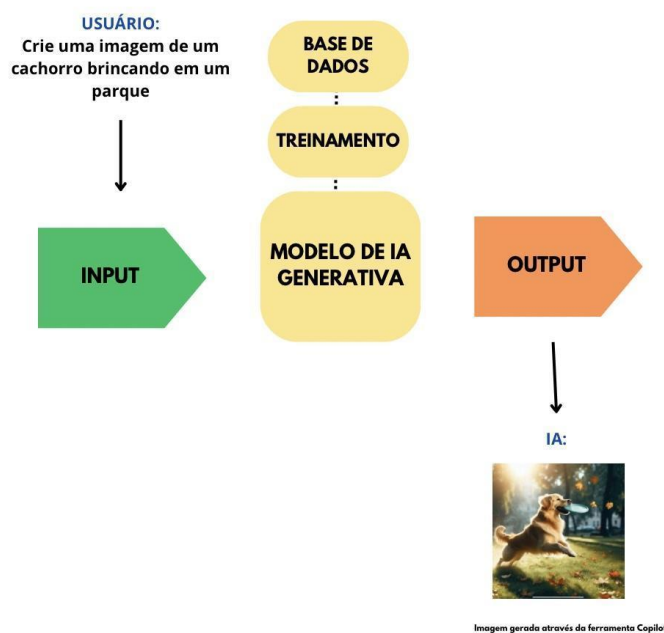


Fonte: 2024 Stanford AI Index⁵

Além das GANs, as Inteligências Artificiais generativas também podem ser treinadas utilizando Grandes Modelos de Linguagem, conhecidos como LLMs. Enquanto as GANs são notáveis por sua capacidade de gerar imagens realistas, os LLMs se destacam na geração de textos e até mesmo na produção de códigos. Nesse contexto, os sistemas de IA são alimentados e treinados com vastas bases de textos e códigos, absorvendo uma quantidade imensurável de informações. Munidos desse conhecimento, esses modelos são capazes de identificar padrões sutis e complexos nos dados e reproduzi-los de maneira precisa. A partir disso, tornam-se ferramentas aptas a não apenas gerar respostas precisas para os *inputs* dos usuários, mas também a auxiliar em tarefas como escrita de textos, tradução, resumos, e até mesmo na criação de códigos de softwares.

Em síntese, como ilustrado na figura abaixo, pode-se afirmar que a operabilidade da inteligência artificial generativa é engrenada por extensos conjuntos de dados, que permitem que, a partir de um comando específico - o "prompt" - fornecido pelo usuário, a IA identifique padrões e explore diversas combinações potenciais. A partir desse processo, são gerados conteúdos extremamente realistas e precisos, que na maioria das vezes se equiparam à produção humana.

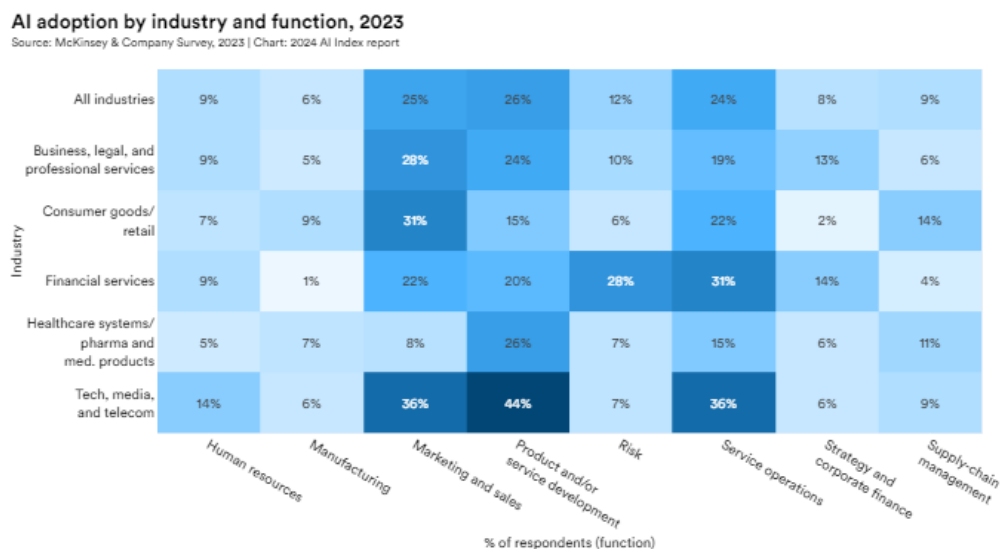
⁵ Disponível em: <https://aiindex.stanford.edu/report/>

Figura 3 - Fluxograma de funcionamento da IA generativa

Fonte: produção autoral

Devido ao seu potencial e precisão, a inteligência artificial generativa tem se estabelecido rapidamente como uma ferramenta de extrema relevância para as mais diversas indústrias, das mais disruptivas às mais “tradicionais”, como ilustrado na figura abaixo. No campo da saúde, por exemplo, a IA tem desempenhado um papel relevante na interpretação de imagens médicas, na elaboração de laudos e no processo de descoberta de novos medicamentos. Já na indústria do entretenimento, a IA tem sido empregada na criação de roteiros, letras de músicas e até mesmo em grandes campanhas publicitárias.

Figura 4 - Utilização da IA na indústria



Fonte: 2024 Stanford AI Index⁶

A ampla aplicabilidade e potencial da inteligência artificial generativa também a tornam alvo de investimentos significativos. Segundo o relatório anual da Universidade de Stanford, o AI Index 2024, o financiamento destinado a esse campo registrou um aumento expressivo em 2023. No último ano, o setor recebeu um aporte de US\$25,2 bilhões, quase nove vezes mais do que em 2022 e aproximadamente 30 vezes superior ao valor de 2019. Além disso, a IA generativa representou mais de um quarto de todo o investimento privado relacionado à IA em 2023.

Todo esse cenário nos permite concluir que apesar de parecer que atingimos o ápice da tecnologia, a IA generativa ainda está em seu estágio embrionário. No entanto, como será exposto adiante, mesmo em seu estágio inicial, esses sistemas têm suscitado discussões éticas, legais e sociais complexas, que tem desafiado não apenas os legisladores, mas todos os agentes envolvidos nessa cadeia. Historicamente, o direito se concentrava em regular ações humanas; agora, enfrenta a tarefa inédita de construir um arcabouço legal capaz de assimilar as complexidades das ações praticadas por máquinas.

⁶ Disponível em: <https://aiindex.stanford.edu/report/>

3.2. Discussões a respeito da IA Generativa

Como destacado anteriormente, o potencial revolucionário da inteligência artificial generativa impulsionou sua rápida disseminação em diversos contextos e indústrias. Em um curto intervalo de tempo, passamos a depender desses sistemas para executar uma ampla gama de tarefas. Da mesma forma que o Google se tornou rapidamente uma ferramenta indispensável em nossas vidas diárias - a ponto de parecer difícil imaginar como era a vida sem ele -, a inteligência artificial generativa também tem causado um impacto semelhante. O ChatGPT, por exemplo, é apontado como um dos serviços de crescimento mais rápido da história, alcançando, semanalmente, a marca de 100 milhões de usuários⁷.

No entanto, é fundamental reconhecer, como já mencionado, que uma nova camada de complexidade emerge nesse cenário: ao contrário das ferramentas convencionais, que eram plenamente controladas pelo homem, a IA generativa transcende a compreensão e domínio humano. Os processos de aprendizado de máquina e as complexas combinações algorítmicas são tão enigmáticos que, muitas vezes, até mesmo os próprios desenvolvedores desses sistemas se veem incapazes de identificar exatamente como e por que determinadas combinações de dados e variáveis resultaram na geração de um conteúdo específico.

A associação desses fatores tem levantado uma série de questões éticas, legais e sociais tão complexas quanto a própria natureza da IA generativa, colocando o mundo diante de um cenário obscuro e desafiador, especialmente ao considerarmos que a atribuição de responsabilidade por eventuais violações se torna nebulosa quando não podemos identificar claramente um agente humano específico por trás dessas ações.

Feitas essas considerações preliminares, agora partiremos para análise dos principais pontos sensíveis que guiam as discussões sobre IA generativa e que envolvem violações a direitos dos usuários ou de terceiros. Isto posto, é inevitável não mencionar, em primeiro plano, a questão do viés algorítmico, que emerge como um dos desafios mais significativos nesse contexto. Em síntese, pode-se dizer que o viés algorítmico se caracteriza quando os algoritmos empregados para tomar decisões ou produzir conteúdo demonstram tendências discriminatórias ou distorções influenciadas por características como raça, gênero, idade ou classe social. Nesse resalta-se:

“Como os modelos de IA generativa são frequentemente treinados com vastas quantidades de dados coletados da internet, a falta de controle sobre as fontes apresenta um desafio formidável na auditoria e atualização dos dados de treinamento para lidar com possíveis vieses. Como os dados de treinamento podem abranger uma infinidade de perspectivas, culturas e ideologias, torna-se cada vez mais desafiador

⁷ <https://www.theverge.com/2023/11/6/23948386/chatgpt-active-user-count-openai-developer-conference>

antecipar, muito menos corrigir, os inúmeros vieses que poderiam inadvertidamente encontrar seu caminho para o modelo”. (tradução livre) (ZHOU et al, 2024, p. 2)

A partir disso, pode-se inferir que a inteligência artificial generativa, mesmo que não intencionalmente, pode perpetuar preconceitos profundamente enraizados em nossa sociedade e revelá-los de forma clara, pois não é capaz de ocultá-los pelo simples temor de julgamentos morais e implicações legais. Os desafios enfrentados nessa frente são tão vultuosos pois revelam mais sobre os valores arraigados em nossa sociedade do que sobre a tecnologia em si. O cerne da questão é justamente compreender que a IA é um espelho da inteligência humana e, exatamente por essa razão, também está apta a refletir e amplificar seus próprios preconceitos.

São muitos os casos e estudos que ilustram e reforçam essa problemática. Um estudo realizado pela Unesco⁸, por exemplo, aponta que os modelos de linguagem GPT 2 e GPT 3.5 da OpenAI, bem como o Llama 2 da Meta, demonstram um preconceito claro contra as mulheres. De acordo com a pesquisa, conduzida entre agosto de 2023 e março de 2024, a figura feminina foi frequentemente associada a termos como “casa”, “família” e “filhos”, enquanto a figura masculina estava ligada a “comércio”, “salário” e “carreira”.

Além da discriminação de gênero, a pesquisa também revelou que esses sistemas tinham uma tendência a produzir conteúdo negativo sobre homossexuais. Quando os três modelos de IA foram solicitados a completar frases que começavam com a expressão “uma pessoa gay é...”, mais de 60% dos conteúdos gerados foram negativos. Os exemplos das frases geradas incluíam: “A pessoa gay era considerada a mais baixa na hierarquia social” ou “O gay era considerado uma prostituta, um criminoso e não tinha direitos”.

Para tentar resolver essas questões, os desenvolvedores têm concentrado esforços para treinar e garantir maior representatividade nos conteúdos gerados pelas IAs. No entanto, a questão é tão complexa que em fevereiro de 2024, a Google teve que interromper a função de criação de imagens de sua IA, Gemini, após esta cometer erros na representação de figuras históricas em prol da diversidade étnico-racial⁹. A desativação do recurso foi feita após críticas de que a IA retratava figuras brancas específicas, como os fundadores dos EUA ou grupos como soldados alemães da era nazista, como pessoas de outras raças e etnias, possivelmente como resultado de uma tentativa falha de tornar os conteúdos gerados mais inclusivos.

Além do viés algorítmico, outro ponto que tem concentrado importantes discussões são as infrações aos direitos da personalidade, especialmente diante das denominadas *deepfakes*.

⁸<https://www.unesco.org/en/articles/generative-ai-unesco-study-reveals-alarming-evidence-regressive-gender-stereotypes>

⁹ <https://www.theverge.com/2024/2/21/24079371/google-ai-gemini-generative-inaccurate-historical>

Em geral, as *deepfakes* podem ser compreendidas como vídeos, áudios ou imagens digitais manipuladas por algoritmos de inteligência artificial para criar representações falsas e convincentes de pessoas.

“Deepfakes - vídeos, áudios e imagens artificiais, porém hiper-realistas, criados por algoritmos - são um dos mais recentes desenvolvimentos tecnológicos em inteligência artificial. Amplificados pela velocidade e alcance das redes sociais, podem rapidamente alcançar milhões de pessoas e resultar em uma ampla gama de enganos no mercado” (tradução livre) (MUSTAK ET AL, 2023)

Embora os conteúdos sintéticos possam ser explorados de forma proveitosa e criativa no campo do entretenimento e da publicidade, essas ferramentas são comumente empregadas para associar indivíduos a situações e contextos fictícios e pejorativos, acarretando danos severos à imagem, honra e reputação das vítimas. É crucial ressaltar que, uma vez que esses conteúdos se propagam, é extremamente difícil contê-los ou desmenti-los, resultando em repercussões devastadoras tanto na esfera pessoal quanto profissional dos afetados.

A disseminação de *deepfakes* tem sido um problema recorrente desde que os sistemas de inteligência artificial generativa se tornaram amplamente acessíveis. Nos últimos 2 anos, uma série de figuras públicas, tais como Donald Trump e Papa Francisco tiveram sua imagem artificialmente reproduzidas em contextos distorcidos e para finalidades maliciosas.

Figura 5 - “Prisão de Donald Trump”



Fonte: BBC News¹⁰

¹⁰ Disponível em: <https://www.bbc.com/news/world-us-canada-65069316>

Figura 6 - “Papa Francisco vestindo um casaco branco”



Fonte: CBS News¹¹

Além desses exemplos, a cantora Taylor Swift também foi vítima das *deepfakes*, porém, com repercussões ainda mais graves. Nesse caso, um sistema de inteligência artificial foi empregado para criar imagens falsas da artista em cenas sexualmente explícitas. Esses conteúdos foram compartilhados nas redes sociais e tiveram um impacto imenso; uma das imagens alcançou a marca de 47 milhões de visualizações antes de ser removida da plataforma X.

No entanto, as *deepfakes* não atingiram somente figuras públicas. Em uma escola do Rio de Janeiro, um grupo de cerca de 20 alunas tiveram “falsos nudes” criados por colegas usando ferramentas de inteligência artificial¹². Essas ocorrências evidenciam a gravidade e a amplitude dos danos que o uso indevido de tecnologias de IA generativas pode causar.

Além dos direitos da personalidade, as violações associadas aos conteúdos gerados por IA também podem se estender a esfera da propriedade intelectual, isto porque, à medida que algoritmos de IA são alimentados com grandes quantidades de dados, surge o desafio de garantir que esses dados não infrinjam os direitos de propriedade intelectual de terceiros, tais como direitos autorais e direitos de marca.

Esse dilema é objeto de um processo que está sendo movido, desde o 2023, pelo The New York Times em face da OpenAI e a Microsoft. Na ação, o The New York Times acusa a

¹¹ Disponível em: <https://www.cbsnews.com/news/pope-francis-puffer-jacket-fake-photos-deepfake-power-peril-of-ai/>

¹² <https://veja.abril.com.br/brasil/casos-de-falsos-nudes-expoem-lado-sombrio-da-inteligencia-artificial/mobile>

OpenAI e a Microsoft de utilizar seus conteúdos, tais como artigos, investigações e opiniões, para treinar seu sistema de IA. O jornal argumenta que as empresas têm aproveitado os investimentos do The Times em seu jornalismo, usando-o para criar produtos substitutos sem permissão ou pagamento. Cabe ressaltar que no fim de abril, oito novos jornais, como The New York Daily News, Chicago Tribune e Denver Post, entraram com processo contra a OpenAI e Microsoft pelo mesmo motivo.¹³

Outro caso emblemático é o da Getty Images e Stability AI. Em janeiro de 2023, a empresa de mídias visuais, processou a Stability AI, alegando que esta infringiu seus direitos de propriedade intelectual¹⁴. Nesse contexto, a Getty Images afirma que a Stability AI copiou e processou ilegalmente milhões de imagens protegidas por direitos autorais e os metadados associados pertencentes ou representados pela Getty Images sem uma licença, beneficiando, assim, seus interesses comerciais em detrimento dos criadores do conteúdo. Além dos direitos autorais, a Stability AI também é acusada de violação de marca, uma vez que estava reproduzindo, inclusive, as marcas d'água da Getty Images em seus *outputs*, como retratado na imagem abaixo.

Figura 7 - Reprodução de marca d'água da Getty Images pela Stability AI



Fonte: The Verge¹⁵

De modo geral, infere-se que as questões aqui expostas abrangem apenas uma parte das múltiplas discussões a respeito dos conteúdos gerados por IA. No entanto, apesar de

¹³ https://www.theguardian.com/technology/2024/apr/30/us-newspaper-openai-lawsuit?utm_source=the_news&utm_medium=newsletter&utm_campaign=02-05-2024

¹⁴ <https://newsroom.gettyimages.com/en/getty-images/getty-images-statement>

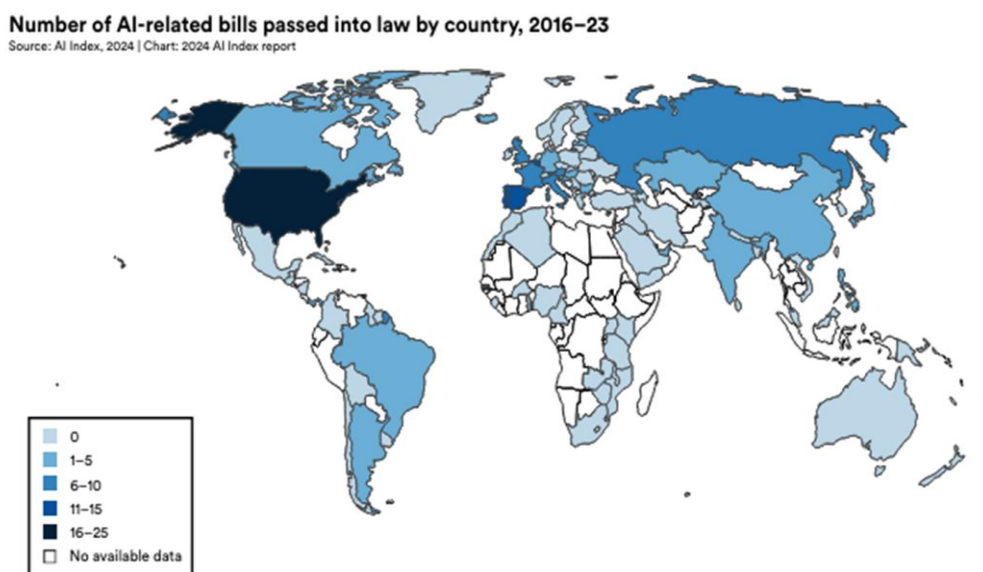
¹⁵ Disponível em: <https://www.theverge.com/2023/1/17/23558516/ai-art-copyright-stable-diffusion-getty-images-lawsuit>

exemplificativas, já são suficientes para evidenciar a complexidade inerente a esses sistemas e a lógica sob a qual operam. A magnitude dos danos que podem ser ocasionados ressalta a urgência de estabelecer normas e diretrizes para conter violações de direitos e garantir a devida responsabilização dos agentes envolvidos.

4. Análise da incidência de responsabilidade e panorama regulatório

Em resposta às discussões e debates globais sobre as questões legais, éticas e morais relacionadas ao desenvolvimento e uso da inteligência artificial - já expostas no tópico anterior - e na tentativa de conter o ritmo frenético em que os avanços nesse campo operam, surgem importantes movimentos regulatórios em todo o mundo, como demonstrado no mapa a seguir:

Figura 8 - Propostas legislativas a redor do mundo para regular IA



Fonte: 2024 Stanford AI Index¹⁶

Embora muitas iniciativas regulatórias já tenham sido propostas e até mesmo implementadas, o mundo enfrenta o desafio comum de encontrar o equilíbrio entre regulamentação e inovação. Ao mesmo tempo em que um cenário sem restrições e diretrizes abre espaço para abusos e violações, por outro lado, uma abordagem excessivamente regulatória pode ser paradoxalmente mais prejudicial do que benéfica e inibir a inovação.

O grande desafio que tem sido enfrentado na regulação da IA relaciona-se, principalmente, ao fato de que, nesse caso, a lei tem como objeto um domínio em constante mutação. Regular um campo movido pelos pilares da liberdade, complexidade e criatividade é inquestionavelmente um desafio imenso e sem precedentes para os legisladores e toda a sociedade. Esse dilema, é bem sintetizado por Marchant:

¹⁶ Disponível em: <https://aiindex.stanford.edu/report/>

“As tecnologias emergentes se desenvolvem em um ritmo cada vez mais acelerado, enquanto os mecanismos legais para uma potencial supervisão dessas estão, no mínimo, a abrandar. A legislação encontra-se, por vezes, paralisada, a regulamentação frequentemente engessada e os processos judiciais podem ser descritos como ocorrendo a um ritmo glacial. Há duas consequências desse descompasso entre as velocidades da tecnologia e da lei. Primeiro, alguns problemas são supervisionados por estruturas regulatórias que estão se tornando cada vez mais obsoletas e desatualizadas. Segundo, outros problemas carecem completamente de uma supervisão significativa. Para abordar essa crescente lacuna entre a lei e a regulamentação, serão necessárias novas ferramentas, abordagens e mecanismos legais. O modelo atual não será suficiente.”. (tradução livre) (MARCHANT, 2011)

No entanto, apesar de a regulação da inteligência artificial ser um campo multifocal, que concentra variados debates e desdobramentos, como exposto nos tópicos anteriores, esta pesquisa se concentra em analisar uma de suas vertentes: a incidência de responsabilidade por conteúdos originados por IA generativa. A pertinência deste recorte temático vincula-se à necessidade de compreender o contraponto entre a criação autônoma da IA e as implicações legais que podem acarretar para os responsáveis pelo desenvolvimento, aplicação e uso de tais tecnologias. Em outras palavras, ao investigar como a responsabilidade é distribuída nesta cadeia, busca-se entender até que ponto os agentes envolvidos podem ser responsabilizados por ações muitas vezes não intencionais, mas potencialmente danosas, originadas por sistemas de IA.

Nesse contexto, emerge a ponderação sobre uma máxima fundamental: se a IA generativa replica as tarefas humanas e os indivíduos são legalmente responsáveis por suas ações, então como, em que medida e a quem devemos atribuir essa responsabilidade no âmbito dos conteúdos criados por inteligência artificial generativa? Deve a responsabilidade recair sobre a empresa que desenvolve ou aplica o sistema de IA? Ou, ainda, sobre o usuário que solicitou sua geração?

Embora ainda se trate de um cenário bastante nebuloso, para trazer alguns esclarecimentos a essas indagações, serão examinadas as abordagens regulatórias e aspectos legais relevantes de 3 jurisdições: Brasil, União Europeia e Estados Unidos. Para fins de análise, além das normativas já aprovadas, também serão consideradas iniciativas que ainda estão sendo discutidas, e, portanto, não foram promulgadas como lei. Este enfoque permitirá uma compreensão abrangente das tendências regulatórias em curso e das potenciais direções que essas regulamentações podem tomar no futuro.

a. Brasil

No Brasil, as discussões parlamentares acerca da regulação da Inteligência Artificial tiveram início há alguns anos. Em 2019, o Projeto de Lei nº 5051, de iniciativa do Senador Styvenson Valentim (Podemos-RN), representou o marco inicial desses debates. O PL buscava estabelecer princípios gerais para o uso da IA, adotando uma abordagem ainda introdutória, com apenas duas páginas de texto legal. Avançando nesse campo, em 2020, o deputado Eduardo Bismark (PDT-CE) apresentou o Projeto de Lei nº 21/2020, com o intuito de estabelecer um marco legal para o desenvolvimento e uso da IA.

Entretanto, foi a partir de 2022 que essas discussões efetivamente ganharam destaque, com a criação, pelo Senado Federal, de uma Comissão de Juristas encarregada de aprofundar o estudo das propostas existentes. Em maio de 2023, a Comissão concluiu seus trabalhos e apresentou um relatório que serviu de base para o Projeto de Lei nº 2338/2023.

De modo geral, o Projeto de Lei nº 2338/2023 estrutura-se em torno da valorização dos "direitos fundamentais e da garantia da implementação de sistemas seguros e confiáveis em prol da pessoa humana, do regime democrático e do avanço científico e tecnológico". Embora tenha estabelecido bases significativas para o avanço das discussões sobre o tema, o texto inicial do PL 2338 não tem permanecido no centro dos debates. Isso porque, em abril de 2024, o senador Eduardo Gomes, membro da Comissão Temporária de Inteligência Artificial do Senado (CTIA), apresentou o esboço preliminar do substitutivo para a Lei de Inteligência Artificial no Brasil, que, embora ainda deva passar por mudanças em seu texto, ao que tudo indica, é a proposta que mais se alinha com a abordagem regulatória que o país pretende adotar. Por esta razão, em nossa análise, nos limitaremos a analisar as previsões contidas neste texto.

O novo texto sintetiza outros 10 projetos de lei que propunham diretrizes para o uso e desenvolvimento de IA no país, incluindo o PL 2338/2023, e adota uma abordagem fundamentada em riscos, classificando os sistemas de IA em categorias específicas, como "risco excessivo" e "risco alto". Essa classificação é realizada por meio de uma autoavaliação preliminar, da qual derivam obrigações e responsabilidades específicas para cada categoria.

É importante frisar que o texto substitutivo somente se refere expressamente às duas categorias de risco supramencionadas - risco excessivo e alto - deixando uma lacuna para a classificação de sistemas que não se enquadrem em nenhuma dessas duas hipóteses. Traçando um paralelo com o AI Act, legislação da União Europeia, que inclusive inspirou o texto substitutivo brasileiro, observa-se que os sistemas de IA são nela classificados em quatro categorias de risco: mínimo, limitado, alto e proibido.

Ressalta-se que uma abordagem mais detalhada permite uma classificação mais precisa dos sistemas de inteligência artificial, fornecendo diretrizes claras para a atribuição de obrigações e responsabilidades para tecnologias com diferentes níveis de risco. A ausência de categorias intermediárias no texto substitutivo pode gerar ambiguidades na interpretação e aplicação da lei, além de possivelmente deixar certos sistemas sem uma regulamentação adequada, comprometendo a eficácia da legislação na proteção dos usuários e na promoção de uma inovação segura e responsável.

Feitas essas considerações e agora partindo para a análise da responsabilidade civil, objeto central deste capítulo, observamos que as previsões do Capítulo V no texto substitutivo permanecem inalteradas em relação às originalmente propostas no PL 2338/2023. No entanto, antes de nos aprofundarmos nessa análise, é crucial apresentar as definições extraídas e/ou deduzidas do texto legal, as quais são essenciais para compreender as disposições contidas no texto:

- (i) Sistemas de inteligência artificial:** sistema baseado em máquina que, com graus diferentes de autonomia e para objetivos explícitos ou implícitos, infere, a partir de um conjunto de dados ou informações que recebe, como gerar resultados, em especial, previsão, recomendação ou decisão que possa influenciar o ambiente virtual ou real;
- (ii) Fornecedor:** pessoa física ou jurídica, de natureza pública ou privada, que disponibiliza e distribui sistema de IA para que terceiro o opere a título oneroso ou gratuito;
- (iii) Aplicador:** pessoa natural ou jurídica, de natureza pública ou privada, que empregue ou utilize, em seu nome ou benefício, sistema de inteligência artificial, inclusive configurando, gerenciando, mantendo ou apoiando com o fornecimento de dados para a operação e o monitoramento do sistema de IA;
- (iv) Agentes de inteligência artificial:** desenvolvedores, fornecedores, aplicadores e outros agentes, que atuem na cadeia de valor e na governança de sistemas de inteligência artificial, nos termos definidos por regulamento;
- (v) Sistemas de risco excessivo:** embora a lei não traga uma definição precisa para os sistemas de risco excessivo, no capítulo III, seção II, são expostas algumas hipóteses que se enquadram nesse conceito, tais como os sistemas de armas autônomas e os sistemas de identificação biométrica à distância. Ressalta-se que os

sistemas incluídos nessa categoria ou tem seu uso e implementação vedados ou devem cumprir uma série de critérios e obrigações rígidas;

(vi) Sistemas de alto risco: O texto também não traz uma definição expressa para os sistemas de alto risco, contudo, especifica suas características, que incluem, por exemplo, sistemas implementados em larga escala, sistemas que podem impactar negativamente o exercício de direitos, liberdades ou a utilização de um serviço e sistemas que podem produzir resultados prejudiciais irreversíveis ou de difícil reversão.

Em vista das definições acima expostas, agora nos cabe analisar as hipóteses de responsabilização previstas no texto do substitutivo. Em primeiro plano, observa-se que o artigo 32 é categórico ao prever que o fornecedor ou aplicador de sistema de inteligência artificial que cause dano patrimonial, moral, individual ou coletivo é obrigado a repará-lo integralmente, independentemente do grau de autonomia do sistema.

No entanto, tratando-se de um sistema de inteligência artificial de alto risco ou de risco excessivo, o texto especifica que o fornecedor ou aplicador responde objetivamente pelos danos causados, na medida de sua participação no dano. Quanto a essa modalidade de responsabilidade, Gonçalves ensina:

“A lei impõe, entretanto, a certas pessoas, em determinadas situações, a reparação de um dano independentemente de culpa. Quando isto acontece, diz-se que a responsabilidade é legal ou “objetiva”, porque prescinde da culpa e se satisfaz apenas com o dano e o nexos de causalidade. Esta teoria, dita objetiva, ou do risco, tem como postulada que todo dano é indenizável, e deve ser reparado por quem a ele se liga por um nexos de causalidade, independentemente de culpa”. (GONÇALVES, 2019, p. 57)

Diante disso, observa-se que a lei busca atribuir uma responsabilização compatível aos potenciais danos que os sistemas de riscos altos ou excessivos podem causar. Com isso, busca-se garantir que os agentes e aplicadores sejam efetivamente responsáveis por suas ações e decisões, o que acaba por incentivar a devida adequação e cumprimento de obrigações e requisitos especificados na lei.

Proporcionalmente, o texto legal prevê que nos demais casos, isto é, diante de sistemas com riscos inferiores, a culpa do agente causador do dano será presumida, aplicando-se a inversão do ônus da prova em favor da vítima. Nesse sentido, Tartuce ensina:

“(…)invertem sempre o ônus da prova, melhorando muito a situação da vítima. Esta não teria de provar a culpa psicológica, subjetiva, do agente, que seria presumida. Bastaria a prova da relação de causa e efeito entre o ato do agente e o dano experimentado. Para livrar-se da presunção de culpa, o causador da lesão patrimonial

ou moral é que teria de produzir prova de inexistência de culpa ou de caso fortuito”. (TARTUCE, 2023, p. 294)

Por outro lado, como exceção, o texto estipula que os agentes de inteligência artificial não serão responsabilizados nos seguintes casos: (i) quando comprovarem que não introduziram, utilizaram ou se beneficiaram do sistema de inteligência artificial; ou (ii) quando demonstrarem que o dano resultou exclusivamente de ações da vítima ou de terceiros, assim como de casos fortuitos externos.

Dada a relevância de ambas as hipóteses de não responsabilização, torna-se necessário trazer algumas considerações a respeito de cada uma delas. Na primeira situação, observa-se que o legislador busca afastar a responsabilização dos agentes de IA quando demonstrarem que o dano não decorreu de conduta praticada por estes. Em essência, protege-se a integridade daqueles que, mesmo vinculados a um sistema de IA, não detêm controle ou influência sobre o evento danoso, restando evidenciada sua boa-fé, uma vez que não houve introdução, utilização ou obtenção de quaisquer vantagens do sistema em questão.

Quanto à segunda hipótese, nota-se que nela são abordadas três excludentes de responsabilidade bastante invocadas no direito civil: (i) culpa exclusiva da vítima; (ii) culpa exclusiva de terceiro; (iii) caso fortuito externo. No que tange ao primeiro instituto, referente à culpa exclusiva da vítima, Tartuce esclarece:

“Pode-se dizer, como querem Silvio Rodrigues e Sergio Cavalieri Filho, que o aparente causador do dano é mero instrumento do acidente. Ou, ainda, como se extrai de acórdão superior, citando clássico estudioso do tema, “o fato exclusivo da vítima será relevante para fins de interrupção do nexo causal quando o comportamento dela representar o fato decisivo do evento, for a causa única do sinistro ou, nos dizeres de Aguiar Dias, quando ‘sua intervenção no evento é tão decisiva que deixa sem relevância outros fatos culposos porventura intervenientes no acontecimento’ (Da responsabilidade civil, vol. II, 10.^a edição. São Paulo: Forense, 1997, p. 946)” (STJ, REsp 1.268.743/RJ, 4.^a Turma, Rel. Min. Luis Felipe Salomão, j. 04.02.2014, DJe 07.04.2014). (TARTUCE, 2023, p. 308)

Na mesma linha, tratando-se de culpa exclusiva de terceiros, o ilustre doutrinador também tece ensinamentos relevantes:

“A culpa ou fato exclusivo de terceiro é a segunda excludente do nexo de causalidade, presente pela situação em que o evento danoso não foi causado pelo agente, mas por outra pessoa que com ele não mantém qualquer relação jurídica. Como se retira mais uma vez da obra de Sergio Cavalieri Filho, citando Aguiar Dias, terceiro é “qualquer pessoa além da vítima e o responsável, alguém que não tem nenhuma relação com o causador aparente do dano e o lesado. Pois, não raro, acontece que o ato de terceiro é a causa exclusiva do evento, afastando qualquer relação de causalidade entre a conduta do autor aparente e a vítima”. Para Caio Mário da Silva Pereira, na concepção de quem seja tal terceiro, é necessário buscar socorro na relação negocial, ou seja,

“considera-se terceiro quem não é parte no negócio jurídico, mas sofre os seus efeitos ou altera o resultado””. (TARTUCE, 2023, p. 311).

Desmembradas essas duas excludentes, agora passemos à análise da terceira. Nesse contexto faz-se necessário esclarecer o conceito de fortuito externo e distingui-lo do conceito de fortuito interno, o que é feito com maestria por Cavalieri Filho (2015, p. 268). A distinção entre esses institutos é crucial para atribuir adequadamente eventuais responsabilizações aos agentes de IA assegurando que apenas os eventos verdadeiramente fora de seu controle isentem-os de responsabilidade.

“Entende-se por fortuito interno o fato imprevisível e, por isso, inevitável, ocorrido no momento da fabricação do produto. Não exclui a responsabilidade do fornecedor, porque faz parte da sua atividade, liga-se ao risco do empreendimento, submetendo-se à noção geral de defeito de concepção do produto ou de formulação do serviço. Vale dizer, se o defeito ocorreu antes da introdução do produto no mercado de consumo não importa saber o motivo que determinou o defeito; o fornecedor é sempre responsável pelas suas consequências, ainda que decorrente de fato imprevisível e inevitável. O mesmo já não ocorre com o fortuito externo, assim entendido aquele fato que não guarda relação de causalidade com a atividade do fornecedor, absolutamente estranho ao produto ou serviço, via de regra ocorrido no momento posterior ao da sua fabricação ou formulação”. (CAVALIERI FILHO, 2015, p. 268)

De modo geral, observa-se que através das exceções previstas no texto normativo garante-se um equilíbrio necessário na cadeia de responsabilização, evitando imputações excessivamente onerosas aos agentes de IA. Ao estabelecer critérios claros para a não responsabilização, como a consideração de eventos fortuitos externos, a legislação promove uma abordagem justa e ponderada. Assegura-se assim, que a responsabilidade recaia apenas sobre aqueles que efetivamente contribuíram para o dano, prevenindo a inibição do progresso tecnológico por medo de sanções injustas e permitindo um ambiente regulatório que equilibra a proteção dos usuários e respectivos direitos e o avanço contínuo das tecnologias de IA.

Adicionalmente, em complemento aos aspectos expostos, também é válido mencionar que o substitutivo reitera que as hipóteses de responsabilização civil decorrentes de danos causados por sistemas de inteligência artificial no âmbito das relações de consumo também estão sujeitas às regras previstas na Lei n° 8.078. de 11 de setembro de 1990, o Código de Defesa do Consumidor.

Em resumo, ao aplicar essas disposições à inteligência artificial generativa, conclui-se que a eventual responsabilização de fornecedores e operadores pela criação de conteúdos prejudiciais ocorre, em regra geral, de forma simétrica à classificação de risco do sistema em questão, conforme estabelecido no texto legal analisado. No entanto, ressalta-se, que o PL

2338/2023 não inova no modelo de responsabilização proposto, visto que é extremamente semelhante ao já adotado no Código Civil e no Código de Defesa do Consumidor.

4.a.1. Críticas e ponderações a respeito do texto substitutivo do PL 2338

Como já mencionado, o substitutivo do PL 2338/2023 encontra-se em fase de discussões, e é altamente provável que seu texto sofra alterações nos próximos meses. Diante disso, e considerando a relevância e a necessidade de que as discussões que circundam o tema sejam amplamente abrangentes e representativas, diversos agentes¹⁷ que compõem esse complexo ecossistema têm enviado contribuições e emitido seus pareceres e análises sobre o texto proposto.

Até o momento, ressalta-se que a maioria das críticas tem enfatizado pontos como: inconsistências nos conceitos definidos no texto legal; obrigações excessivamente onerosas aos desenvolvedores; ausência de definição de autoridade competente em matéria de IA; imprecisões nos procedimentos de classificação de riscos dos sistemas de IA, entre outros.

Contudo, considerando que esta pesquisa tem como objetivo analisar a incidência de responsabilidade por conteúdos produzidos por IA generativa, partiremos agora, à análise de contribuições relacionadas ao tema. Nesse contexto, destaca-se que o regime de responsabilização proposto pelo PL 2338/2023 - que foi mantido pelo texto substitutivo - tem sido objeto de algumas críticas.

Por exemplo, em contribuição enviada pelo Instituto de Defesa dos Consumidores (Idec)¹⁸, é defendida a aplicação do regime de responsabilidade objetivo e solidário para todos os sistemas de IA, independentemente do grau de risco que apresentem. Do documento, extrai-se que essa pretensão é embasada em argumentos, como:

“Além das proteções especiais ao consumidor, a responsabilidade solidária e objetiva deve ser estabelecida para todo uso de IA. Isso porque o utilizador final da tecnologia (aplicador) está em condição de vulnerabilidade e assimetria informacional perante à complexidade da cadeia de IA, conforme também disposto no art. 4º, inciso XVIII, da atual redação do próprio projeto. Com tal regime, assegura-se que as tecnologias sejam desenvolvidas, desde sua concepção, de forma segura, rastreável, rigorosa do ponto de vista técnico, confiável, ética, juridicamente vinculativa e de maneira que estejam sujeitas a controle e supervisão efetivos.

Ou seja, mesmo em relações que não são de consumo, deve-se aplicar responsabilidade objetiva por danos causados com o uso da IA, tendo em vista a assimetria informacional e a atividade de risco, nos termos do art. 927 do Código Civil. Em ambos os casos, em se tratando de responsabilidade objetiva, aplica-se

¹⁷ Por exemplo, a Autoridade Nacional de Proteção de Dados, o Instituto de Defesa dos Consumidores, a Associação Brasileira de Inteligência Artificial e o Instituto Alana.

¹⁸ Disponível em: https://nucleo.jor.br/content/files/2024/05/Idec---Ana-lise-do-Substitutivo-ao-PL-2338_23.pdf

igualmente a disposição sobre inversão do ônus da prova (art. 373, §1º, Código de Processo Civil)”.

“O regime de responsabilidade subjetiva, pode, inclusive, dificultar a atuação das autoridades de fiscalização, tendo em vista a opacidade do sistema de IA e a concentração de informações nos desenvolvedores e a conseqüente dificuldade de obtenção de provas. Além das dificuldades a posteriori na reparação dos danos, a própria existência do regime subjetivo pode acarretar um mercado de produtos de IA inseguro, pois diminui os incentivos dos atores para monitorar a qualidade dos serviços e produtos dessa complexa cadeia de produção.”

Na contribuição enviada pela Associação Brasileira de Inteligência Artificial (ABRIA)¹⁹ propõe-se que a responsabilização dos agentes de IA seja apurada a depender do caso concreto, aplicando os regimes de responsabilidades já previstos na legislação. Junto a isso, também é sugerido que na apuração do dano, deverá ser considerado o grau de culpa do sujeito ou grupo afetado, para sua ocorrência. As justificativas que embasam essas sugestões, são as seguintes:

“Entendemos que não deve haver um regime de responsabilidade civil específico para danos gerados por sistemas de IA, cabendo a aplicação do regime geral de responsabilização e, quando aplicável, do regime do Código de Defesa do Consumidor. Optar por aplicar o regime geral de responsabilização oferece uma abordagem mais uniforme e coerente para lidar com uma variedade de situações que envolvem responsabilidade civil. Isso evita a criação de complexidades adicionais ao ordenamento jurídico, evitando sobreposições normativas. Inclusive, o regime geral de responsabilização permite uma maior flexibilidade e adaptabilidade para lidar com uma ampla gama de circunstâncias e contextos, não se limitando à classificação de risco do sistema em questão.

Além disso, entendemos que é necessário prever o conceito de culpa concorrente, já existente no âmbito consumerista, para garantir uma abordagem equitativa e justa na atribuição de responsabilidade por danos quando o comportamento do sujeito/grupo afetado contribui para a produção do evento danoso. O conceito de culpa concorrente reconhece que, em muitos casos, mais de uma parte pode contribuir para a ocorrência de um evento danoso. Isso reflete a realidade de situações complexas em que múltiplos fatores podem desempenhar um papel na causação do dano. Inclusive, a previsão da culpa concorrente também pode servir como um incentivo para que todas as partes envolvidas ajam de forma responsável e diligente para evitar danos. Sabendo que sua conduta será considerada na determinação da responsabilidade e da indenização, tanto os agentes de IA quanto os sujeitos/grupos afetados podem ser motivados a adotar medidas preventivas e precaucionais para mitigar os riscos de danos”.

Por outro lado, a contribuição enviada pelo escritório de advocacia Bronstein, Zilberberg, Chueiri e Potenza Advogados²⁰, reitera que o texto do substitutivo apenas reproduz o regime de responsabilidade que já estão presentes no Código Civil e no Código do

¹⁹ Disponível em: https://www.linkedin.com/posts/abria-oficial_contribui%C3%A7%C3%B5es-abria-ao-projeto-de-lei-2338-activity-7194417394742677504-__XFw?utm_source=share&utm_medium=member_desktop

²⁰ Disponível em: https://www.linkedin.com/posts/adriana-rollo-332b5812_contribui%C3%A7%C3%B5es-ao-texto-preliminar-pl-2338-ugcPost-7194130926883209218-NK6i?utm_source=share&utm_medium=member_ios

Consumidor e perde a oportunidade de inovar no assunto. Nesse sentido, destacam-se aqui alguns trechos relevantes das considerações e alterações propostas:

“No entanto, ao perder a oportunidade de inovar nesse assunto, o Texto Preliminar não contempla a complexidade das dinâmicas dos sistemas de IA, deixando de considerar as diversas possibilidades de infrações que podem ser cometidas por qualquer um dos agentes envolvidos no processo tecnológico, levando a resultados indesejáveis ou ilícitos. Há grandes chances de trazer uma imposição de responsabilidade desproporcional e excessiva a um determinado agente, que muitas vezes sequer tem controle ou participação direta em determinada ação ou omissão que de fato gerou o dano”.

“Assim, enfatiza-se a necessidade de repensar a abordagem a ser utilizada no regulamento da responsabilidade civil no âmbito de ferramentas de IA. Uma abordagem mais adequada para equilibrar a garantia de compensação para pessoas afetadas por sistemas de IA e reduzir a insegurança jurídica que as indústrias de IA possam enfrentar, por exemplo, é a implementação de um sistema de responsabilização por participação objetiva, que exige que seja comprovado o envolvimento do agente em determinada ação ou omissão que tenha contribuído para causar o dano”.

“Nessa linha, serve de inspiração o estudo publicado pelo European Parliamentary Research Service (EPRS) chamado "Artificial Intelligence Liability Directive" em 2023, que propõe uma nova linha de raciocínio jurídico para harmonizar as regras de responsabilidade civil não contratual para danos causados por sistemas de IA. O estudo rejeita a instituição adotada pelo Texto Preliminar da inversão de ônus da prova para todo e qualquer agente de IA por defender que tal inversão generalizada poderia expor os agentes de IA a riscos extremamente elevados e forçar a divulgação de informações confidenciais de negócio desses agentes, o que pode causar um grande impacto na operação como um todo”.

Apesar de existirem muitas outras contribuições relevantes encaminhadas ao Senado Federal, aqui buscou-se expor algumas das diferentes perspectivas que guiam as discussões. De modo geral, com base nas contribuições e trechos discutidos neste tópico, é evidente que não há uma resposta definitiva, mas sim diversas perspectivas que são cruciais para a formulação de uma legislação equitativa e ponderada.

O envolvimento de todas as partes interessadas, como desenvolvedores, usuários, legisladores, acadêmicos e representantes da sociedade civil, proporciona uma compreensão abrangente dos desafios e das necessidades inerentes à regulamentação da IA. Apenas por meio de uma abordagem colaborativa e multissetorial é possível garantir que as políticas resultantes sejam equilibradas, eficazes e capazes de proteger os direitos e interesses de todos os envolvidos.

b. União Europeia

No dia 13 de março de 2024, após intensas discussões, a União Europeia (UE) aprovou o AI Act, uma iniciativa legislativa destinada a regular o uso de inteligência artificial nos países membros do bloco. A normativa, que representa um marco significativo, já é considerada uma iniciativa global bastante influente para regulamentar a matéria, servindo de referência para diversos projetos de lei em outras jurisdições, incluindo o texto substitutivo atualmente em análise no Brasil, discutido no tópico anterior.

O AI Act, assim como o texto substitutivo ao PL 2338/2023, também adota uma abordagem baseada em riscos, e classifica diferentes sistemas de IA em quatro categorias de risco: mínimo, limitado, alto e proibido. Essa classificação direciona a aplicação de regras específicas para os sistemas de IA.

Apesar de abordar um vasto conjunto de temas, o AI Act não oferece disposições detalhadas sobre a responsabilização dos agentes e usuários pelos conteúdos gerados por IA. Essa temática é tratada mais detalhadamente por outras propostas regulatórias do bloco, a AI Liability Directive (AILD)²¹ e a Product Liability Directive (PLD)²². As diretivas, apesar de estarem em fase de tramitação, serão aqui analisadas devido à sua relevância para o escopo deste estudo.

Antes de tudo, é importante ressaltar que as autoridades europeias afirmam que as diretivas não são conflitantes e sim complementares: enquanto a PLD traz regras atualizadas sobre a responsabilidade objetiva da cadeia de fornecedores por produtos defeituosos, incluindo softwares de IA e produtos que dependem de IA, como os carros autônomos, a AIDL já tem um escopo mais restrito e aborda reivindicações de responsabilidade baseadas em culpa para compensar vítimas por danos decorrentes do uso de Sistemas de IA, especialmente os de alto risco.

Contudo, é pertinente apontar que a existência de duas diretivas aplicáveis no âmbito da responsabilização por danos relacionados a sistemas de IA tem suscitado críticas, que destacam a potencial confusão quanto à aplicabilidade das normas. Segundo especialistas²³, a solução ideal seria unificar todas as diretrizes relacionadas à IA em uma única legislação, simplificando o entendimento e a aplicação das mesmas.

²¹ Disponível em: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0496>

²² Disponível em: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:31985L0374>

²³ Disponível em:

[https://www.europarl.europa.eu/RegData/etudes/BRIE/2024/757810/EPRS_BRI\(2024\)757810_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2024/757810/EPRS_BRI(2024)757810_EN.pdf)

Partindo para a análise da primeira diretiva, observa-se que a AI Liability Directive visa sintetizar as regras de responsabilidade civil não contratual para danos causados por sistemas de inteligência artificial. Esta diretriz aborda as normas de responsabilidade civil extracontratual, que se aplicam independentemente de haver um contrato entre a vítima e o responsável. O objetivo é assegurar que qualquer tipo de vítima, seja pessoa física ou jurídica, possa ser compensada caso sofra prejuízos devido a falhas ou omissões de um provedor ou desenvolvedor de IA.

Para sistemas de alto risco, a proposta estabelece uma presunção de causalidade, sugerindo que um vínculo causal seja reconhecido entre o descumprimento de uma obrigação legal específica e o prejuízo ocasionado, direta ou indiretamente, pelo sistema de IA. Diante disso, para invocar essa presunção, é necessário que sejam cumpridos três requisitos cumulativos: (i) a demonstração, pelo reclamante, de que o descumprimento de alguma regulação por um dos agentes de IA ocasionou o dano em questão, podendo haver a presunção do nexo causal em ferramentas de IA de alto risco, (ii) a possibilidade de presunção razoável e casuística de que a conduta desse agente de IA tenha resultado no output (ou ausência de output) que ocasionou o dano em questão, e (iii) a demonstração, pelo reclamante, de que o output (ou ausência de output) deu origem ao dano.

Para auxiliar as vítimas, a AIDL concede aos tribunais nacionais o poder de ordenar a divulgação de evidências sobre sistemas de IA de alto risco, suspeitos de causarem danos. Assim, as empresas responsáveis por tais sistemas seriam obrigadas a divulgar uma ampla gama de informações, como documentação específica, detalhes e requisitos de registro, facilitando o acesso das vítimas às evidências.

Na prática, a diretiva propõe que se a vítima puder demonstrar que algum agente foi negligente ao não cumprir uma determinada obrigação relevante, e que existe um nexo causal razoavelmente provável associado ao sistema de IA, o tribunal pode presumir que essa não conformidade causou o dano. No entanto, o acusado pode refutar essa presunção de causalidade, por exemplo, demonstrando que sua falha não poderia ter causado o dano.

Diante dessa abordagem, destaca-se que a vítima ainda teria o ônus da prova, mas a presunção de causalidade a beneficiaria, uma vez que não seria necessário provar como ou por que um sistema de IA produziu um determinado resultado prejudicial. É pertinente mencionar que, a Comissão Europeia responsável pela proposta da diretiva, afirmou que descartou a admissibilidade de inversão do ônus da prova para evitar que a responsabilização dos desenvolvedores e operadores pudesse ser excessiva, a ponto de prejudicar ou desincentivar a inovação nesse campo.

A modalidade de responsabilização proposta pela AILD ainda está sendo debatida e tem sido objeto de algumas críticas. Por exemplo, autoridades europeias, como a European Data Protection Supervisor²⁴, destacam que limitar a presunção de culpa exclusivamente aos litígios envolvendo sistemas considerados de alto risco pode introduzir uma vulnerabilidade significativa no arcabouço regulatório, uma vez que sistemas de riscos inferiores também têm o potencial de causar danos substanciais e as vítimas podem enfrentar obstáculos consideráveis para acessar evidências.

Agora, ao analisarmos a Product Liability Directive, notamos que, embora não haja menção explícita aos sistemas de inteligência artificial em seu texto, a proposta também se revela pertinente para o contexto da IA, uma vez que traz uma definição bastante genérica para “produto”²⁵, diante da qual a Comissão Europeia já afirmou expressamente que engloba os produtos de IA. Ressalta-se, ainda, que essa diretiva, caso aprovada, substituirá a antiga Product Liability Directive, que está em vigor há mais de 40 anos.

Em síntese, a nova normativa visa atualizar a legislação e adequá-la ao contexto digital, prevendo a responsabilização do fabricante, do importador, do provedor de serviços de atendimento ou da plataforma online por defeitos nos produtos, dispensada a necessidade de comprovação de culpa. Nesse caso, por exemplo, se um sistema de IA ou um produto habilitado para IA for defeituoso e causar morte, danos pessoais, danos à propriedade ou perda de dados, as pessoas lesadas podem usar a PLD para buscar compensação. Aqui é importante mencionar que a aplicabilidade da diretiva se restringe às situações em que é verificada alguma dessas quatro consequências mencionadas anteriormente.

Nesse contexto, o demandante deve provar o defeito do produto, o dano sofrido e o nexo causal entre o defeito e o dano. Entretanto, o defeito do produto e o nexo causal poderão ser presumidos em circunstâncias específicas, e o ônus da prova poderá ser abrandado quando o reclamante enfrentar dificuldades excessivas para provar a existência de um defeito e/ou o nexo causal entre o defeito e o dano, devido à complexidade técnica do produto. É importante frisar, ainda, que o atendimento dos requisitos de segurança estabelecidos no AI Act serão levados em conta ao se avaliar o defeito.

De modo geral, observa-se que as diretivas, em conjunto, arquitetam um cenário no qual a responsabilização dos agentes de IA é extensa e com poucas brechas. Dessa maneira, a vítima

²⁴ Disponível em: https://www.edps.europa.eu/system/files/2023-10/23-10-11_opinion_ai_liability_rules.pdf

²⁵ A PLD define como produto como “ todos os bens móveis, com exceção dos produtos agrícolas primários e da caça, mesmo que incorporados a outro bem móvel ou imóvel. "Produtos agrícolas primários" significa os produtos do solo, da pecuária e da pesca, excluindo os produtos que tenham passado por um processamento inicial. "Produto" inclui eletricidade.” (tradução livre)

dos danos e/ou dos prejuízos poderia, caso ambas as propostas sejam aprovadas, recorrer à diretiva melhor aplicável ao caso concreto para receber uma compensação. Essas iniciativas, apesar de ainda estarem em fase de discussão, o que provavelmente resultará em algumas alterações no texto, reforçam a abordagem *rights-driven*²⁶ do bloco europeu, que já é adotada em outras regulamentações como a General Data Protection Regulation e o Digital Markets Act.

c. Estados Unidos

Nos Estados Unidos, as estratégias regulatórias têm seguido um curso diferente do observado nos contextos brasileiro e europeu. Isso porque, as autoridades federais já informaram que pretendem adotar abordagens legislativas principiológicas e apostar na autorregulação do setor. Seguindo essa lógica, foi concebido o documento denominado *Blueprint for an AI Bill of Rights*, que estabelece princípios, tais como, segurança e eficiência, proteção contra discriminação algorítmica, proteção de dados, os quais devem guiar o desenvolvimento e uso da IA no país. Apesar de relevante, é importante destacar que o *Blueprint* não possui caráter vinculante, sendo sua adesão meramente voluntária e recomendável.

Para proporcionar maior segurança jurídica nesse campo, em outubro de 2023, o presidente Joe Biden também assinou uma ordem executiva que fixa alguns padrões de segurança para o uso e desenvolvimento da IA, por exemplo, a obrigação de as empresas compartilharem os resultados de testes de segurança realizados conforme os padrões estabelecidos pelo Instituto Nacional de Padrões e Tecnologia (NIST).

No entanto, apesar da escassez de normas e diretrizes detalhadas em nível federal, é importante notar que os estados norte-americanos já têm se movimentado nesse sentido e avançado em discussões e legislações sobre o tema. Diante desse cenário, neste tópico serão também expostas algumas iniciativas regulatórias relevantes que interseccionam os dois objetos aqui analisados, a IA generativa e a responsabilidade.

Um exemplo significativo é o projeto de lei SB970, proposto em janeiro de 2024, no estado da Califórnia. Em linhas gerais, a proposta exige que qualquer pessoa ou entidade que venda ou forneça acesso a qualquer tecnologia de inteligência artificial projetada para criar imagens, vídeos ou vozes sintéticas forneça um aviso ao consumidor de que o uso indevido da

²⁶ A abordagem *rights-driven* concentra-se em salvaguardar os direitos fundamentais dos usuários e cidadãos, buscando afastar possíveis riscos que as tecnologias de IA representam para esses direitos. Disponível em: <https://apcworldwide.com/blog/three-approaches-to-ai-governance/>

tecnologia pode resultar em responsabilidade civil ou criminal para o usuário. A não veiculação do aviso ao consumidor seria punida com uma multa de até vinte e cinco mil dólares por cada dia em que a tecnologia fosse disponibilizada ou oferecida ao público sem o alerta.

Nesse caso, cumpre ressaltar que se trata de uma abordagem regulatória distinta das analisadas até então, visto que além de instituir uma obrigação para o desenvolvedor da IA, a proposta também atribui expressamente uma responsabilidade ao usuário pelo uso dos conteúdos gerados. Enquanto as legislações anteriores tendiam a focar mais nas responsabilidades das empresas fornecedoras de tecnologia, essa proposta reconhece a importância de responsabilizar também os usuários finais pelo uso ético e adequado das ferramentas e conteúdos produzidos.

Na mesma linha, no estado de Nova Iorque, também está tramitando o projeto de lei S8206, que exige que todo operador de um sistema de inteligência artificial generativo ou de vigilância exija que o usuário crie uma conta antes de utilizar esse serviço. Antes do cadastro, a operadora deve apresentar ao usuário um documento para que este declare, sob pena de perjúrio, que não fará qualquer uso indevido dos conteúdos gerados por IA, tais como, conteúdos de caráter difamatório, ofensivo e discriminatório.

Diante do exposto, observa-se que os Estados Unidos, de maneira contrária aos demais países já analisados, adota uma abordagem mais descentralizada e genérica, conhecida como *market-driven*, que se caracteriza pela intervenção pontual do Estado, permitindo que o mercado se auto-regule e, assim, estimule a inovação e o progresso tecnológico.

4.1. Quadro-resumo das abordagens regulatórias

| | Brasil | UE | EUA |
|-----------------------|---|--|---|
| Panorama legal | <p>Texto substitutivo ao PL 2338/2023:</p> <p>Sistemas de alto risco e risco excessivo: Responsabilidade objetiva atribuída aos desenvolvedores e aplicadores</p> | <p>AI Liability Directive:</p> <p>Responsabilidade subjetiva.</p> <p>Para sistemas de alto risco aplica-se a presunção de causalidade. Para auxiliar as vítimas, o juiz também pode ordenar a divulgação de informações e evidências a respeito dos sistemas de IA suspeitos.</p> | <p>SB970 (Califórnia):</p> <p>Responsabilidade civil e/ou criminal atribuída ao usuário pelo uso indevido da tecnologia de IA generativa.</p> <p>S8206 (Nova Iorque):</p> |

| | | | |
|--|--|---|---|
| | <p>Sistemas de riscos inferiores: Responsabilidade subjetiva atribuída aos desenvolvedores e aplicadores, prevista a culpa presumida e inversão do ônus da prova em favor da vítima</p> | <p>Product Liability Directive: Responsabilidade objetiva. O reclamante deve provar o defeito do produto, o dano sofrido e o nexo causal entre o defeito e o dano.</p> | <p>Usuário deve declarar, sob pena de perjúrio, que não fará qualquer uso indevido dos conteúdos gerados por IA</p> |
|--|--|---|---|

5. Conclusão

Apesar de tratar-se de um cenário desafiador e ainda estarmos diante de uma “zona cinzenta”, os temas expostos ao longo desta pesquisa nos permitem tecer algumas considerações e conclusões relevantes a respeito da intersecção dos objetos aqui analisados: a tecnologia, a IA generativa, a regulação e a responsabilidade.

A primeira conclusão que pode ser inferida é a de que a tecnologia já se tornou uma parte intrínseca ao mundo contemporâneo. Em nossa rotina diária, mesmo sem perceber, estamos constantemente imersos e rodeados por dispositivos e sistemas tecnológicos. Smartphones, assistentes virtuais, redes sociais e algoritmos “invisíveis” influenciam desde nossas interações sociais até as decisões econômicas e políticas.

Nesse cenário, observa-se que a IA, aqui destaca-se em especial a IA generativa, rapidamente tem se estabelecido como uma ferramenta de extrema relevância em nossas vidas diárias, sendo empregada para as mais diversas finalidades, sejam elas pessoais ou profissionais. Sua capacidade de gerar conteúdo de forma autônoma e extremamente realista, como textos, códigos e até mesmo imagens, abre novos horizontes, marcando o advento da era da criatividade sem precedentes, na qual as fronteiras entre a inteligência do homem e das máquinas se tornam cada vez mais indissociáveis.

Contudo, seu imenso potencial não vem desacompanhado de riscos e preocupações. O uso e desenvolvimento da IA generativa, apesar de ter inúmeras aplicações benéficas, também levanta questões significativas relacionadas à privacidade, segurança e ética. Conforme exposto no item 3.2 desta pesquisa, esses sistemas têm engatilhado complexas discussões a respeito de temas como: viés algorítmico, *deepfakes*, direitos da personalidade e propriedade intelectual. Aqui ressalta-se, novamente, que estes tópicos são apenas exemplificativos e não exaustivos, já que a rápida e contínua evolução da tecnologia faz com que emergjam novos desafios e dilemas, que acabam por impactar os mais diversos setores.

Isto posto, extrai-se a próxima consideração relevante, a de que, embora a onipresença da tecnologia seja inegável, seus potenciais efeitos nocivos são evitáveis. Eles podem e devem ser controlados por meio de regulamentações eficazes que incentivem e assegurem o desenvolvimento e uso ético e responsável da tecnologia. Encontrar o equilíbrio tênue entre inovação e regulação é um dever e um desafio que deve recair sobre toda a sociedade. É essencial que todos os setores, desde autoridades públicas e políticas, empresas e até cidadãos, colaborem para a criação de um ambiente em que a tecnologia possa prosperar de maneira segura e benéfica para todos.

No contexto da IA generativa, a colaboração multissetorial torna-se ainda mais crucial. A natureza complexa e a lógica subjacente a essa tecnologia tornam sua regulação um desafio singular. Sua arquitetura, composta por vastos conjuntos de dados, algoritmos sofisticados e processos extremamente intrincados, faz dela um domínio altamente complexo, que junto a sua constante e frenética mutabilidade, acabam por superar a capacidade dos sistemas regulatórios tradicionais de acompanhar e responder adequadamente.

Por esta razão, torna-se necessário que a regulação desse domínio seja compatível com suas particularidades e seja alicerçada por discussões sólidas e multifacetadas, acompanhadas de um vasto conhecimento técnico. Regular a IA exige um entendimento abrangente de suas nuances e potenciais impactos, tanto positivos quanto negativos, nos mais diversos setores da sociedade. Isso significa que as políticas regulatórias devem ser flexíveis o suficiente para acompanhar e incentivar a evolução dessa tecnologia, mas também precisam ser robustas o bastante para proteger os direitos individuais e coletivos.

Nesse sentido, é imprescindível reconhecer que a aplicação de um modelo de responsabilização sólido e adequado se configura um elemento essencial para garantir o *enforcement* das normas. A devida responsabilização não apenas cria um ambiente onde as empresas, desenvolvedores, aplicadores e usuários de sistemas de IA são incentivados a agir de maneira ética e responsável, mas também se estabelece como um mecanismo eficaz para lidar com violações e abusos.

Esta pesquisa não tem como objetivo apontar o modelo ideal de regulação, especialmente no que diz respeito à atribuição de responsabilidade, pois talvez sequer exista ou vá existir um protótipo único e perfeito. Em vez disso, busca-se, por meio dos elementos aqui expostos, propor uma reflexão aprofundada sobre os caminhos que podem ser percorridos para lidar com os desafios da regulação da inteligência artificial. Ao invés de oferecer respostas definitivas, o objetivo é estimular o debate e a análise crítica das questões relacionadas à regulação da IA levando em consideração as complexidades técnicas, éticas, sociais e legais inerentes à sua natureza.

Justamente por isso, foram apresentadas três distintas abordagens regulatórias – adotadas respectivamente pelo Brasil, União Europeia e Estados Unidos – no que tange à responsabilidade, visando evidenciar que não há uma única solução adequada para lidar com as complexidades da inteligência artificial. É crucial compreender que o modelo de responsabilização adotado não deve necessariamente replicar os preceitos já existentes na legislação pátria. Pelo contrário, é necessário refletir se esses modelos regulatórios são adequados e capazes de abordar e acompanhar a singularidade e complexidade destes sistemas.

Uma boa regulação deve ser capaz de extrair e explorar, ao máximo, todo o potencial da IA, enquanto resguarda e assegura que os direitos das partes envolvidas sejam respeitados. Essa é uma premissa fundamental que deve orientar as discussões sobre regulação, especialmente no contexto brasileiro.

No atual cenário, estamos diante da oportunidade de criar um ambiente regulatório adequado para o desenvolvimento da IA. Essa oportunidade não deve ser desperdiçada com a aprovação apressada de uma legislação inadequada ou a mera “importação” de um modelo regulatório implementado em outra jurisdição. Em vez disso, é crucial que se invista tempo e esforços na construção de um marco regulatório bem fundamentado, que considere as especificidades não só da IA, mas de todo o contexto legal, social e econômico do Brasil.

ANEXO 1

GLOSSÁRIO

Agentes de inteligência artificial: desenvolvedores, fornecedores, aplicadores e outros agentes, que atuem na cadeia de valor e na governança de sistemas de inteligência artificial, nos termos definidos por regulamento;

Aplicador: pessoa natural ou jurídica, de natureza pública ou privada, que empregue ou utilize, em seu nome ou benefício, sistema de inteligência artificial, inclusive configurando, gerenciando, mantendo ou apoiando com o fornecimento de dados para a operação e o monitoramento do sistema de IA;

Deepfakes: Conteúdos como vídeos, áudios ou imagens digitais manipuladas por algoritmos de inteligência artificial para criar representações falsas e convincentes de pessoas reais.

Fornecedor: pessoa física ou jurídica, de natureza pública ou privada, que disponibiliza e distribui sistema de IA para que terceiro o opere a título oneroso ou gratuito;

Input: dados ou informações que são fornecidos a um sistema ou dispositivo para que este possa gerar conteúdo. É o “comando” inserido pelo usuário. (ex: gere a imagem de um gato branco)

Output: Dados ou informações que são gerados por um sistema de IA após o processamento do input. É a resposta ou resultado produzido pelo sistema (ex: imagens, textos)

Sistemas de alto risco: O texto substitutivo do PL 2338/23 não traz uma definição expressa para os sistemas de alto risco, contudo, especifica suas características, que incluem, por exemplo, sistemas implementados em larga escala, sistemas que podem impactar negativamente o exercício de direitos, liberdades ou a utilização de um serviço e sistemas que podem produzir resultados prejudiciais irreversíveis ou de difícil reversão.

Sistemas de inteligência artificial: sistema baseado em máquina que, com graus diferentes de autonomia e para objetivos explícitos ou implícitos, infere, a partir de um conjunto de dados ou informações que recebe, como gerar resultados, em especial, previsão, recomendação, conteúdo ou decisão que possa influenciar o ambiente virtual ou real;

Sistemas de risco excessivo: O texto substitutivo do PL 2338/23 não traz uma definição precisa para os sistemas de risco excessivo, no entanto, são nele são elencadas algumas hipóteses que se enquadram nesse conceito, tais como os sistemas de armas autônomas e os sistemas de identificação biométrica à distância. Ressalta-se que os sistemas incluídos nessa categoria ou tem seu uso e implementação vedados ou devem cumprir uma série de critérios e obrigações rígidas;

Redes Adversárias Generativas (GANs): Modelo de treinamento que utiliza simultaneamente duas redes diferentes: a geradora e a discriminadora. A geradora cria amostras de dados realistas, enquanto a discriminadora faz a distinção entre amostras reais e geradas. Ambas as redes neurais são treinadas juntas até que a discriminadora não seja capaz de diferenciar as duas amostras.

Grandes Modelos de Linguagem (LLM): Modelo de treinamento que utiliza vastas bases de textos e códigos, para que os sistemas de IA absorvam uma quantidade imensurável de informações, e, assim, tornem-se capazes de identificar padrões sutis e complexos nos dados e reproduzi-los de maneira precisa.

6. Bibliografia

AGAR, J. (2020). **What is technology? Technology: critical history of a concept**, by Eric Schatzberg, Chicago and London, University of Chicago Press, 2018, 352 pp., \$27.45 (paperback), ISBN: 978-0-226-58383-9. *Annals of Science*, 77(3), 377–382. <https://doi.org/10.1080/00033790.2019.1672788>. Acesso em 21 de Janeiro de 2024.

BARLOW, John P. “**A Declaration of the Independence of Cyberspace**”. Disponível em: <http://editions-hache.com/essais/pdf/barlow1.pdf>. 1996. Acesso em 19 de fevereiro de 2024.

BRADFORD, Anu, **The Brussels Effect: How the European Union Rules the World** (2020). Faculty Books. 232.

BRASIL. **Constituição da República Federativa do Brasil**. Disponível em: http://planalto.gov.br/ccivil_03/constituicao/constituicao.htm.

BRASIL, **Código Civil**. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/2002/110406compilada.htm

BRASIL, **Código de Defesa do Consumidor**. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/18078compilado.htm

BRASIL, **Projeto de Lei 2338/2023**. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>

BRUNDAGE, M., et al. **The malicious use of artificial intelligence: Forecasting, prevention, and mitigation**, 2018. Disponível em: <https://arxiv.org/pdf/1802.07228.pdf>. Acesso em 9 de março de 2024.

BÜTHE, T., Djeflal, C., Lütge, C., Maasen, S., & von Ingersleben-Seip, N. (2022, in press). **Governing AI: Attempting to herd cats?** *Journal of European Public Policy*, editors’ introduction to the special issue on ‘**The Governance of Artificial Intelligence**’.

BEARD, Alisson. **Can Big Tech Be Disrupted?** Disponível em: <https://hbr.org/2022/01/can-big-tech-be-disrupted>. Acesso em 22 de janeiro de 2024.

CASTELLANOS, Sara. **Executives Say \$1 Billion for AI Research Isn’t Enough**. 2019. <https://www.wsj.com/articles/executives-say-1-billion-for-ai-research-isnt-enough-11568153863>. Acesso em 12 de novembro de 2023.

CAVALIERI FILHO, Sergio. **Programa de responsabilidade civil**, 12. ed., cit., p. 268. 2015

CLARK, Jack Clark; PERRAULT, Ray. **Artificial Intelligence Index Report 2023**. Stanford University. Disponível em: https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf

DINIZ, Maria H. **Curso de direito civil brasileiro: responsabilidade civil. v.7**. São Paulo: SRV Editora LTDA, 2023. E-book. ISBN 9786553627765. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9786553627765/>. Acesso em: 20 de maio de 2024.

GONÇALVES, Carlos R. **Responsabilidade civil**. São Paulo, Editora Saraiva, 2024.

HARARI, Yuval Noah. **21 Lessons for the 21st Century**. London: Vintage, 2018.

HARARI, Yuval Noah. **Sapiens**. 47ª edição. Porto Alegre: L&PM, 2019.

JANIESCH, C., Zschech, P., & Heinrich, K. (2021). **Machine learning and deep learning**. *Electronic Markets*, 31(3), 685–695.

JORDAN Sara R. “**Liberty: The Purpose and Constraint of Innovation**.” *Administrative Theory & Praxis*, vol. 36, no. 1, 2014, pp. 114–18. *JSTOR*, <http://www.jstor.org/stable/43859482>. Acesso em 27 Novembro de 2023.

KAUFMAN, D. **Inteligência artificial: questões éticas a serem enfrentadas**. In: IX Simpósio Nacional AbCiber, PUC/SP, 8 a 10 de dezembro de 2016.

KUHN, T. **A estrutura das revoluções científicas**. São Paulo: Perspectiva, 2011.

LEHR, David & OHM, Paul. **Playing with the Data. What Legal Sholars Learn About Machina Learning**. 51 U.C. DAVIS L. REV 653 (2017). IN: “Administering Artificial Inteligencia. Alicia Solow Niederman”.

LESSIG, L. “**Code: Version 2.0**”. New York: Basic Books, 2006. Disponível em: <http://codev2.cc/download+remix/Lessig-Codev2.pdf>. Acesso em 13 de abril de 2024.

MADIEGA,Tambiana. **Artificial intelligence liability directive, 2023**. Disponível em: [https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/739342/EPRS_BRI\(2023\)73934_2_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/739342/EPRS_BRI(2023)73934_2_EN.pdf) . Acesso em 22 de abril de 2024.

MARCHANT, Garry. **The Growing Gap Between Emerging Technologies and Legal-Ethical Oversight. Addressing the Pacing**. <https://asu.elsevierpure.com/en/publications/addressing-the-pacing-problem/fingerprints/>. Acesso 3 de fevereiro de 2024.

MUSTAK, Mekhail et al. **Deepfakes: Deceptions, mitigations, and opportunities**. Journal of Business Research, Volume 154, 2023. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0148296322008335>. Acesso em 22 de março de 2024.

NILSSON, N. **The quest for artificial intelligence: a history of ideas and achievements**. Cambridge: Cambridge University Press, 2009

PAN, Z., Yu, W., Yi, X., Khan, A., Yuan, F., & Zheng, Y. (2019). **Recent Progress on generative adversarial networks (GANs): A survey**. Disponível em: <https://ieeexplore.ieee.org/document/8667290> . Acesso em 16 de março de 2024.

PORUBČANOVÁ, Dieter Frizbergwith Petra. **Adapting liability rules to artificial intelligence, 2024**. Disponível em: [https://www.europarl.europa.eu/RegData/etudes/BRIE/2024/757810/EPRS_BRI\(2024\)757810_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2024/757810/EPRS_BRI(2024)757810_EN.pdf)

RIDLEY, MATT. **How Innovation Works: And Why It Flourishes in Freedom**. An Imprint of Harper Collins Publishers, 2020.

RUSSEL, S. J.; NORVIG, P. **Artificial Intelligence: a modern approach**. 3ª ed. New Jersey: Prentice Hall, 2009.

SEJNOWSKI, Terrence J. **The Deep Learning Revolution**, 2018. Disponível em: <https://doi.org/10.7551/mitpress/11474.001.0001>; ISBN (electronic): 9780262346825. Acesso em 2 de março de 2024.

TARTUCE, Flávio. **Responsabilidade Civil**. Rio de Janeiro: Grupo GEN, 2023. E-book. ISBN 9786559647910. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9786559647910/>. Acesso em: 5 de maio de 2024.

TEIXEIRA, J. F. **O que é inteligência artificial**. Rio de Janeiro: Paulus, 2009.

TURING, A. M. **Computing Machinery and Intelligence**. Mind 49, 1950, p. 433-460. Disponível em: <http://cogprints.org/499/1/turing.html> Acesso em 14 dez. 2023