UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE DIREITO
THAÍS VICTORETTI

MODELOS GENERATIVOS E MECANISMOS DE TUTELA DA PESSOA HUMANA: regimes de responsabilização de agentes autônomos perante a regulamentação da IA de propósito geral no ordenamento jurídico brasileiro

Juiz de Fora
2025

THAÍS VICTORETTI

MODELOS GENERATIVOS E MECANISMOS DE TUTELA DA PESSOA HUMANA: regimes de responsabilização de agentes autônomos perante a regulamentação da IA de propósito geral no ordenamento jurídico brasileiro

Dissertação apresentada ao Programa de Pós-Graduação em Direito da Faculdade de Direito da Universidade Federal de Juiz de Fora, como requisito parcial para obtenção do grau de Mestra no Mestrado em Direito e Inovação, sob orientação do Prof. Dr. Sergio Marcos Carvalho de Ávila Negri.

Juiz de Fora
2025

Ficha catalográfica elaborada através do programa de geração
automática da Biblioteca Universitária da UFJF,
com os dados fornecidos pelo(a) autor(a)

Victoretti, Thaís.
Modelos generativos e mecanismos de tutela da pessoa humana: regimes de responsabilização de agentes autônomos perante a regulamentação da IA de propósito geral no ordenamento jurídico brasileiro/ Thaís Victoretti. -- 2025.
110 f.

Orientador: Sergio Marcos Carvalho de Ávila Negri
Dissertação (mestrado acadêmico) - Universidade Federal de Juiz de Fora, Faculdade de Direito. Programa de Pós-Graduação em Direito, 2025.

1. Inteligência artificial. 2. Modelos generativos. 3. Tutela da pessoa humana. 4. Personalidade digital. 5. Responsabilização de agentes autônomos. I. Negri, Sergio Marcos Carvalho de Ávila, orient. II. Título.

FOLHA DE APROVAÇÃO

THAÍS VICTORETTI

MODELOS GENERATIVOS E MECANISMOS DE TUTELA DA PESSOA HUMANA: regimes de responsabilização de agentes autônomos a regulamentação da IA de propósito geral no ordenamento jurídico brasileiro

Dissertação apresentada ao Programa de Pós-Graduação em Direito da Faculdade de Direito da Universidade Federal de Juiz de Fora, como requisito parcial para obtenção do grau de Mestra no Mestrado em Direito e Inovação, submetida à Banca Examinadora composta pelos membros:

Orientador: Prof. Dr. Sergio Marcos Carvalho de Ávila Negri
Universidade Federal de Juiz de Fora

Prof. Dr. Bruno Farage da Costa Felipe
Faculdade Doctum / Faculdade Metodista Granbery

Profa. Dra. Maria Regina Detoni Cavalcanti Rigolon Korkmaz
Universidade Federal de Juiz de Fora

PARECER DA BANCA

(X) APROVADA

( ) REPROVADA

Juiz de Fora, 8 de julho de 2025.

Processo: 23071.926725/2025-58 Documento:2452885

Thaís Victoretti

Modelos Generativos e Mecanismos de Tutela da Pessoa Humana: regimes de responsabilização de agentes autônomos perante a regulamentação da IA de propósito geral no ordenamento jurídico brasileiro

Dissertação apresentada ao Programa de Mestrado em Direito da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Direito. Área de concentração: Direito e Inovação

Aprovada em 16 de junho de 2025.

BANCA EXAMINADORA

Sergio Marcos Carvalho de Ávila Negri - Orientador
Universidade Federal de Juiz de Fora

Bruno Farage da Costa Felipe
DOCTUM e Granbery

Maria Regina Detoni Cavalcanti Rigolon Korkmaz
Universidade Federal de Juiz de Fora

Juiz de Fora, 11/06/2025.

Documento assinado eletronicamente por **Sergio Marcos Carvalho de Avila Negri, Professor(a)**, em 30/06/2025, às 22:19, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Thaís Victoretti, Usuário Externo**, em 01/07/2025, às 18:10, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Maria Regina Rigolon Korkmaz, Professor(a)**, em 02/07/2025, às 03:12, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **BRUNO FARAGE DA COSTA FELIPE, Usuário Externo**, em 06/07/2025, às 17:44, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no Portal do SEI-Ufjf (www2.ufjf.br/SEI) através do ícone Conferência de Documentos, informando o código verificador 2452885 e o código CRC 88C361F7.

AGRADECIMENTOS

Aprendi, primeiramente, a agradecer a Deus. Ele que sempre me amparou nos momentos mais difíceis da minha caminhada, que presenciou minhas súplicas e meus choros mais íntimos, me mostrando o verdadeiro sentido da palavra resiliência para crer que tudo acontece em seu momento. Em especial, agradeço a oportunidade de ter me trazido de volta ao meu lugar.

Agradeço à minha família, pilar fundamental da minha existência. Com eles aprendi a importância de se ter valores, de ser uma pessoa verdadeiramente boa e de cultivar nossa relação para que estejamos sempre unidos. Agradeço, em especial, à minha mãe, Eliane, que sempre se fez presente e resiliente para que nada me faltasse – apesar dela achar que ela é a força e eu o cérebro, eu nada seria sem o cérebro dela –; ao meu avô, Hélio (*in memoriam*), por, ao lado dos meus tios e padrinho, ter sido a referência de pai que eu sempre quis ter, por ter feito eu me sentir a neta/filha mais amada do mundo e sempre ter apoiado minha mãe a subsidiar meus estudos; e à minha avó, Cidinha, por ter me mostrado que ser uma "grande mulher" diz muito mais sobre o bem que você faz ao próximo do que ao sucesso material.

Agradeço ao meu companheiro, Joaquim, sinônimo de inspiração, cumplicidade, amor, carinho e companheirismo diários. Aquele que não só torna a vida menos atribulada, mas que faz com que ela seja ainda mais doce de ser vivida. Agradeço, ainda, a toda sua família, que também aprendi a chamar de minha.

Agradeço ao meu orientador, Sergio Negri, pela oportunidade e por todo o aprendizado e orientação que foram fundamentais para a minha trajetória acadêmica, desde a graduação.

Agradeço à Cata, pelo carinho e pela gentileza de ser a revisora deste trabalho. Ao João, pela amizade que encontrei nesta jornada acadêmica.

Agradeço aos meus gestores e ao meu time do Machado Meyer, pela oportunidade de viver essa jornada acadêmica e por toda a força, me cobrindo na nossa rotina do escritório para que eu pudesse me dedicar aos meus estudos.

Agradeço à Universidade Federal de Juiz de Fora e à Faculdade de Direito da UFJF pela oportunidade de desenvolvimento acadêmico desde a minha graduação em Direito.

Por fim, agradeço a mim mesma, principalmente à Thaís dos 20 e poucos, pela coragem e dedicação de fazer com que a Thaís dos 30 e poucos chegasse até aqui.

“Seja você quem for, seja qual for a posição social que você tenha na vida, a mais alta ou a mais baixa, tenha sempre como meta muita força, muita determinação e sempre faça tudo com muito amor e com muita fé em Deus, que um dia você chega lá.
De alguma maneira você chega lá.”

Ayrton Senna

RESUMO

O trabalho aborda a necessidade de uma regulação clara e eficaz da inteligência artificial generativa no Brasil, especialmente em relação ao Projeto de Lei nº 2.338/2023. O estudo se propõe a analisar as implicações jurídicas da IA generativa, destacando a importância de adequações legislativas que respeitem os direitos humanos e garantam uma IA segura e confiável. A pesquisa também explora a viabilidade de aplicação dos conceitos de personalidade jurídica e responsabilidade civil, conforme teoria de Beckers e Teubner, aos modelos generativos. Ademais, respeitadas as devidas peculiaridades que diferenciam o direito privado alemão do brasileiro, pretende investigar o cabimento dessa teoria no contexto jurídico do Brasil. Apesar da adoção dos citados autores como referencial teórico, a presente dissertação não pretende utilizar-se do direito comparativo. A metodologia utilizada é o método dedutivo com uma análise qualitativa, buscando entender a dinâmica das relações sociais e os preceitos legais que estão sendo definidos. A pesquisa é exploratória, dada a natureza ainda em construção da legislação sobre IA no Brasil, e se foca nas implicações específicas da IA generativa, que traz questões únicas e complexas. Por fim, o trabalho explora a necessidade de um arranjo legislativo que permita ao ordenamento jurídico brasileiro adaptar-se às inovações e aos conflitos impostos pela IA generativa, evitando ambiguidades que possam comprometer a segurança jurídica e a ética no desenvolvimento e na utilização dessa tecnologia. O presente estudo foi desenvolvido no âmbito da participação da autora no Núcleo de Estudos Avançados em Pessoa, Inovação e Direito (NEAPID), da Faculdade de Direito da Universidade Federal de Juiz de Fora, e no projeto de pesquisa “Inovação e Direito na IA: mapeamento normativo e análise do exercício de direitos fundamentais” (CNPq-universal), coordenado pelo professor Sergio Negri, orientador deste trabalho.

Palavras-chave: inteligência artificial generativa; modelos generativos; tutela da pessoa humana; personalidade digital; responsabilização de agentes autônomos.

ABSTRACT

The essay addresses the need for clear and effective regulation of generative artificial intelligence in Brazil, especially in relation to Bill No. 2,338/2023. The study aims to analyze the legal implications of generative AI, highlighting the importance of legislative adjustments that respect human rights and ensure safe and reliable AI. The research also explores the feasibility of applying the concepts of legal personality and civil liability, according to the theories of Beckers and Teubner, to generative models. Furthermore, while respecting the particularities that differentiate German private law from Brazilian law, it seeks to investigate the applicability of this theory within the Brazilian legal context. Despite adopting the authors as theoretical references, this dissertation does not intend to use comparative law. The methodology employed is the deductive method with a qualitative analysis, aiming to understand the dynamics of social relations and the legal precepts that are being defined. The research is exploratory, given the still-developing nature of AI legislation in Brazil, and focuses on the specific implications of generative AI, which brings unique and complex issues. Finally, the work explores the need for a legislative framework that allows the Brazilian legal system to adapt to the innovations and conflicts imposed by generative AI, avoiding ambiguities that could compromise legal certainty and ethics in the development and use of this technology. The present study was developed within the scope of the author's participation in the Núcleo de Estudos Avançados em Pessoa, Inovação e Direito (NEAPID) at the Law School of the Federal University of Juiz de Fora, as well as in the research project "Inovação e Direito na IA: mapeamento normative e análise do exercício de direitos fundamentais" (CNPq-universal), coordinated by Professor Sergio Negri, the supervisor of this work.

Keywords: *generative artificial intelligence; generative models; protection of the human person; digital personality; responsibility of autonomous agents.*

LISTA DE ABREVIATURAS E SIGLAS

AI	*Artificial intelligence*
AI Act	*European Artificial Intelligence Act*
A&O Shearman	Allen & Overy Shearman
APIs	*Application Programming Interfaces*
Apud	Junto de (outra obra)
AEPD	*Agencia Española Protección Datos*
AGI	*Artificial General Intelligence*
AM	Aprendizado de Máquina
ANPD	Autoridade Nacional de Proteção de Dados
Art./Arts.	Artigo/Artigos
CC	Código Civil
CDC	Código de Defesa do Consumidor
CIn-UFPE	Centro de Informática da Universidade Federal de Pernambuco
CRFB	Constituição da República Federativa do Brasil
CTIA	Comissão Temporária Interna sobre Inteligência Artificial
DP	*Deep Learning*
EBIA	Estratégia Brasileira de Inteligência Artificial
Ed.	Edição
EUA	Estados Unidos da América
GANs	*Generative Adversarial Networks*
GDPR	*General Data Protection Regulation*
GPT ou GPTs	*Generative Pre-training Transformer* ou *Transformers*
HITL	*Human In The Loop*
IA ou IAs	Inteligência Artificial ou Artificiais
LLM ou LLMs	*Large Language Models*
LGPD	Lei Geral de Proteção de Dados Pessoais
MCTI	Ministério da Ciência, Tecnologia e Inovações

MIT	*Massachusetts Institute of Technology*
ML	*Machine Learning*
MLLE	Modelos de Linguagem de Grande Escala
NEAPID	Núcleo de Estudos Avançados em Pessoa, Inovação e Direito
NLP	*Natural Language Processing*
n.º	Número
p./pp.	Página/Páginas
PL	Projeto de Lei
PLN	Processamento de Linguagem Natural
TechLab	Laboratório de Tecnologia e Direito da Faculdade de Direito da Universidade de São Paulo
TI	Tecnologia da Informação
EU	União Europeia
Vol.	Volume

LISTA DE SÍMBOLOS

§	Parágrafo
%	Porcentagem
US$	Dólar americano

SUMÁRIO

1 INTRODUÇÃO

Ao longo dos últimos anos, a inclusão da inteligência artificial (IA) no cotidiano das pessoas ganhou ascensão exponencial. Esse tipo de tecnologia, até então não materializada para o público não especialista[1], emerge do cenário científico e passa a ser introduzida em atividades rotineiras, a exemplo do atendimento por inteligências artificiais em serviços essenciais, como bancos, telefonia e internet; assistentes virtuais, como Siri e Alexa; portarias eletrônicas por reconhecimento virtual; e sistemas de recomendação de conteúdos de plataformas como Netflix, YouTube e Spotify, além das mídias sociais, como as do grupo Meta.

Para Barroso e Mello (2024, p. 5):

> "A quarta revolução industrial, que começa a invadir nossas vidas, vem com a combinação da Inteligência Artificial, da Biotecnologia e a expansão do uso da Internet, criando um ecossistema de interconexão que abrange pessoas, objetos e mesmo animais de estimação, numa Internet de coisas e de sentidos".

As novas tecnologias, em certa medida, implicam liberar as pessoas humanas de atividades cotidianas mais simples, mas também possuem potencial para desempenhar tarefas altamente complexas.

Estimativas recentes da consultoria McKinsey indicam que, já em 2025, ferramentas baseadas em IA generativa poderão acrescentar entre 2,6 e 4,4 trilhões de dólares ao PIB global, sobretudo a partir de ganhos de produtividade em setores intensivos de informação, como serviços financeiros, saúde e comunicação (Mckinsey, 2023). No plano laboral, o Fórum Econômico Mundial projeta que aproximadamente 44% das habilidades exigidas para o desempenho de funções humanas serão transformadas, substituídas ou requalificadas até 2030 em razão da adoção massiva de sistemas de IA generativa[2].

Como discorre Romano (2011, p. 120) acerca do desenvolvimento da linguagem Church[3], que subsidiou, por exemplo, o surgimento de sistemas que realizam tradução automatizada de textos:

> No início, o pensamento era visto pela inteligência artificial como uma inferência lógica. Por isto, as primeiras linguagens de programação eram baseadas em linguagem

[1] Público ou pessoa não especialista, no contexto deste trabalho, deverá ser entendido como o usuário que não detém o conhecimento técnico e específico do funcionamento da tecnologia da IA generativa.

[2] Fórum Econômico Mundial. Comunicados à imprensa: Relatório sobre o Futuro dos Empregos 2025: 78 milhões de novas oportunidades de emprego até 2030, mas é preciso melhorar a qualificação das forças de trabalho urgentemente. 2025. Disponível em: https://reports.weforum.org/docs/WEF_Future_of_Jobs_2025_Press_Release_PTBR.pdf. Acesso em 5 mai. 2025.

[3] Explica Romano (2021) que a linguagem Church, desenvolvida em 2009 por pesquisadores do MIT, é uma linguagem probabilística que combina conceitos de linguagens de programação tradicionais com modelos probabilísticos, permitindo a criação de programas que podem aprender e inferir informações a partir de dados.

> matemática e utilizavam frases afirmativas, como por exemplo: “As aves põem ovos, logo, as galinhas são aves”. Essas linguagens conseguiam mostrar conceitos distintos, porém, nem todos, como por exemplo, “o ornitorrinco, que não é uma ave, põe ovos”.
>
> Nos últimos tempos, as pesquisas vêm avançando na inteligência artificial, com o uso da inteligência probabilística, no qual os computadores podem aprender com a análise dos dados de treinamento (grandes conjuntos de dados) utilizando os padrões estatísticos.

O aprendizado de máquina (ou *machine learning*) representa uma metodologia que engloba um conjunto de técnicas, suportadas por métodos, que permitem que as máquinas automatizem a criação de modelos a partir da identificação sistemática de padrões estatísticos em um conjunto intensivo de dados (Chowdhury *et al*., 2017). No caso da inteligência artificial, o aprendizado de máquina profundo subsidia a criação dos modelos com base na extração dos dados brutos e conversão em múltiplas camadas de neurônios artificiais (Gartner, 2024). Essas camadas incrementalmente obtêm características de alto nível dos dados brutos, permitindo a solução de problemas mais complexos com uma alta acurácia e uma menor dependência de ajustes manuais (Gartner, 2024).

Os modelos generativos, dentre eles as *Generative Adversarial Networks* (GANs[4]), ganharam notável destaque e ascensão exponencial nos últimos anos. A IA generativa, conforme definida por Goodfellow *et al.* (2014), é um processo de modelagem que utiliza um jogo adversarial entre dois modelos para gerar amostras que imitam a distribuição de dados reais. Ou, em definição mais simples, é um subconjunto de inteligências artificiais de *machine learning* que possuem a capacidade de criar conteúdo rapidamente em resposta a solicitações dos usuários (o que chamamos de *prompt*), criando conteúdo sintético, que pode variar entre texto, áudio, imagem e vídeo (Pavlik, 2023).

O grande diferencial da IA generativa em comparação à IA de maneira geral é que aquela é mais ampla, de modo que o conteúdo gerado em resposta aos *prompts* não será encontrado no conjunto de treinamento da máquina exatamente como produzido. Não bastasse isso, os modelos de IA generativa são treinados a partir de um considerável e diverso volume de dados, enquanto os sistemas tradicionais de *machine learning* são treinados com foco em dados específicos para a função pretendida[5].

Recentemente, com maior precisão em novembro de 2022, através do lançamento do ChatGPT, o uso da inteligência artificial tornou-se, de fato, acessível ao público em geral[6]. Essa solução de IA generativa é um assistente virtual de linguagem natural, desenvolvido pela

---

[4] GANs, ou Redes Adversariais Generativas, em português, são modelos de aprendizado de máquina que utilizam a abordagem adversarial para a geração de conteúdo sintético (ANPD, 2024).

[5] Conceitos não referenciados refletem o entendimento da própria autora desta dissertação que, além de Bacharel em Direito pela Universidade Federal de Juiz de Fora, é também Especialista em Big Data Analytics pela Universidade Presbiteriana Mackenzie.

[6] Amaral e Xavier (2023, p. 22) destacam que o ChatGPT “talvez seja a primeira encarnação de uma inteligência artificial realmente capaz de convencer membros do grande público de sua inteligência”.

OpenAI, que utiliza esse tipo de tecnologia para conversar com usuários e responder perguntas (OpenAI, 2024), facilitando a obtenção de insights e automatização de tarefas. Estima-se que, do seu lançamento até janeiro de 2023, a ferramenta tenha atingido 100 milhões de usuários ativos, ao passo que, comparativamente, para chegar a esta marca, o TikTok levou nove meses após seu lançamento global, e o Instagram pouco mais de dois anos (Hu, 2023 *apud* Amaral; Xavier, 2023).

Assim, o que de fato popularizou esse tipo de tecnologia é que, além do acesso gratuito via internet, as pessoas puderam aprender a conversar com a IA de modo a obter conhecimento instantaneamente, de forma resumida e mais direcionada, o que acaba corroborando para as diversas atividades do dia a dia, seja no âmbito profissional, acadêmico ou pessoal. Como resultado disso, temos vislumbrado o crescimento vultuoso de partes interessadas nessa tecnologia, sem que tenham, contudo, dimensão de como lidar com o impacto da IA generativa nos negócios e na sociedade.

A corrida pelo desenvolvimento da IA generativa tem evidenciado uma crescente hegemonia do poder econômico, refletindo a intensa competição entre as grandes potências globais, especialmente os Estados Unidos e a China. À medida em que empresas de tecnologia buscam incessantemente aprimorar seus modelos de linguagem e sistemas de IA, o que se observa é uma centralização do acesso a essas tecnologias em mãos de poucos *players* dominantes.

O acesso facilitado à IA generativa, embora promova inovações e simplifique tarefas cotidianas, ascende preocupações sobre a monopolização da informação e o controle que essas empresas exercem sobre os dados. Essa dinâmica não apenas acentua a desigualdade no acesso à tecnologia, mas também gera um ambiente propício à exploração de recursos naturais e humanos, especialmente em países em desenvolvimento que dependem das inovações tecnológicas das potências econômicas.

Embora as projeções sobre IA generativa sejam promissoras, elas expõem uma assimetria fundamental: o mesmo impulso econômico que concentra capitais e dados em torno de poucos conglomerados tecnológicos aprofunda a dependência dos países periféricos, gerando novas formas de colonialismo digital e reforçando exigências de tutela jurídica da pessoa humana diante do poder computacional e econômico dessas entidades privadas.

“A personalidade, enquanto valor da pessoa que caracteriza o ordenamento jurídico e garante a sua unidade” (Perlingeri, 2005 *apud* Korkmaz, 2019, p. 11), passou a compor o objeto de estudo nas pesquisas acerca da IA generativa, sobretudo nos últimos oito anos, à vista dos debates sobre a responsabilização de agentes digitais. Isso porque, por prerrogativa inicial do Parlamento Europeu, em 2017, o entendimento caminhava para o sentido de assunção de subjetividade jurídica plena por esses agentes, categorizando como pessoa um agente virtual

que não era alcançado pela definição jurídica de ordenamentos jurídicos ao redor do mundo, inclusive o do Brasil[7]. Nas palavras de Rodotà (2004), “a unidade da pessoa somente pode ser reconstituída com a ampliação ao corpo eletrônico das garantias elaboradas para o corpo físico”.

Em meio ao cenário de plena expansão da interação humano-máquina, o estudo da IA generativa progressivamente aproxima-se de uma abordagem de tutela da pessoa humana, no sentido de que o processo evolutivo de desenvolvimento da própria tecnologia tem se associado a um processo de fragmentação da pessoa e de seus direitos fundamentais. Isto é, suas vivências, suas experiências e seus pertences são vistos essencialmente como dados para subsidiar a evolução tecnológica dos modelos generativos, ainda porque o uso dessa tecnologia por pessoas humanas corrobora para um potencial lesivo aos seus direitos fundamentais.

Nesse contexto, o processo de construção regulatória da IA generativa deve ser capaz não só de garantir a evolução tecnológica, entendida no contexto de se fazer útil ao próprio desenvolvimento das sociedades, mas, ao mesmo tempo, de tutelar os direitos fundamentais, seja no processo de construção ou de utilização da tecnologia pelas pessoas.

A transmutação dessa tecnologia implica não apenas os desafios de políticas públicas – como programas de requalificação e redes de proteção – mas também a reconfiguração da própria noção de responsabilidade civil. Afinal, quando um agente autônomo assume tarefas outrora desempenhadas por seres humanos, a pergunta que se impõe é: quem responde, e em que medida, por eventuais danos decorrentes de erros, vieses ou falhas sistêmicas intrínsecas ao modelo?

Em virtude disso, é necessário que a discussão acerca dos modelos generativos, bem como sobre os regimes jurídicos da personalidade digital e da responsabilidade jurídica aplicáveis a danos em decorrência do uso da tecnologia da IA generativa, receba ênfase na composição regulatória.

Nesse sentido, infere-se que a dignidade da pessoa humana é o núcleo dos direitos fundamentais e deve orientar toda a análise jurídica. Não basta que a discussão acerca desses direitos se volte tão somente aos riscos aos quais o indivíduo é exposto em relação à própria tecnologia, mas que também haja a consideração deles no momento de sopesamento da responsabilidade atribuída em relação a um dano. Isso significa que, ao avaliar casos de danos causados por modelos e/ou sistemas de IA generativa, é preciso a avaliação de afronta à dignidade da vítima e a busca de soluções que promovam sua proteção integral.

Entende-se que a linha entre os direitos fundamentais, em razão do risco a que a tecnologia pode expor a vítima quanto ao uso (análise preventiva) ou ao dano que a ela pode

[7] Atualmente, o Parlamento Europeu não apoia mais a ideia de conceder personalidade jurídica plena a agentes de IA, como sugerido anteriormente com o conceito de "*e-person*". Em vez disso, a abordagem atual foca na regulamentação e no controle dos sistemas de IA com base em seu potencial de risco (European Parliament, 2024).

causar (análise corretiva), é tênue. Contudo, como será demonstrado ao longo deste trabalho, existem situações em que, não necessariamente, o desenvolvedor de um sistema vai expor o consumidor da tecnologia a risco, embora o emprego do modelo de IA utilizado para a construção do sistema possa ocasionar um dano a outrem, em situação específica.

Sob a ótica regulatória, observa-se um movimento pendular entre abordagens mais prescritivas, tal qual a estrutura de "*risk-based approach*", consagrada pelo aprovado AI Act, e propostas principiológicas focadas em "*soft law*", como as "*OECD AI Principles*" e a "*Recommendation on the Ethics of Artificial Intelligence*" da UNESCO. No Brasil, o Projeto de Lei n.º 2.338/2023 caminha para internalizar, ainda que de forma híbrida, esse confronto de paradigmas: de um lado, a ambição de fomentar inovação e competitividade internacional; de outro, a necessidade de preservar o núcleo essencial dos direitos fundamentais insculpido na Constituição de 1988.

A discussão sobre responsabilidade de agentes autônomos, nesse contexto, demanda o resgate de categorias consolidadas no direito privado brasileiro, ao mesmo tempo em que exige a construção de novos enquadramentos dogmáticos capazes de acomodar fenômenos inéditos, como a opacidade algorítmica ("*black box problem*") e a capacidade de aprendizado contínuo dos modelos.

Assim, lança-se mão do seguinte questionamento: o projeto de lei proposto para regular a IA generativa em solo brasileiro, bem como demais legislações já em vigor, estão preparados para uma aplicação adequada de responsabilidade civil e coibição de práticas abusivas em relação ao dano à pessoa humana?

Parte-se da hipótese de que o ordenamento jurídico ainda não está preparado para tanto, sendo viável delinear os seguintes pressupostos para adequação: em primeira análise, é necessária a correta diferenciação entre os conceitos de modelos e de sistemas de IA no PL nº 2.338/2023; segundo, é necessário que se revisite o conceito de personalidade jurídica para que a responsabilidade civil em decorrência de danos por IA generativa consiga ser distribuída entre toda a cadeia produtiva da tecnologia; por fim, apesar dos desafios impostos pelo progresso das novas tecnologias e pelos interesses mercadológicos decorrentes de sua exploração, é necessário que o ordenamento jurídico brasileiro fomente normativas com o propósito não só de coibir a lacuna de responsabilidade decorrente do dano causado pela IA generativa à pessoa humana, como também da violação aos direitos fundamentais em virtude disso.

Traçadas as linhas introdutórias, no Capítulo 2 deste estudo serão apresentadas as diretrizes metodológicas que nortearão o desenvolvimento da pesquisa. No Capítulo 3, por sua vez, os conceitos dos modelos generativos e de IA de propósito geral serão mais bem aprofundados, demonstrando, ainda, a tecnicidade conceitual trazida pelo AI Act quanto à diferenciação entre modelo e sistema de IA.

No Capítulo 4, serão introduzidas as bases teóricas adotadas, a fim de analisar a proposição do conceito jurídico de personalidade digital e os regimes de responsabilização na perspectiva de Beckers e Teubner (2021) e o porquê deste estudo associá-la aos modelos gerativos.

Já no Capítulo 5, a proposta será demonstrar o cenário atual brasileiro e fazer uma análise acerca da fusão conceitual vista no Projeto de Lei nº 2.338/2023, além de se verificar os conceitos de personalidade digital e responsabilização, abordados por Beckers e Teubner (2021), são aplicáveis ou não à realidade jurídica brasileira. Tudo isso com foco em entender a proposta de responsabilização dos agentes autônomos em face da violação de direitos fundamentais da pessoa humana.

Por fim, no Capítulo 6, conclui-se o trabalho com uma retomada de seu conteúdo e a apresentação das considerações finais.

2 METODOLOGIA

Para a produção da presente pesquisa, utilizou-se o método dedutivo com predominância de uma análise qualitativa, na medida em que o objeto específico em estudo é o Projeto de Lei nº 2.338/2023, que tem o intuito de regular a inteligência artificial no Brasil a partir da centralidade da pessoa humana. Visto isso, a pesquisa qualitativa busca "explicar o porquê das coisas, exprimindo o que convém ser feito, mas não quantificam valores e as trocas simbólicas nem se submetem à prova de fatos" (Gerhardt; Silveira, 2009, p. 34), preocupando-se, portanto, em explicar a dinâmica das relações sociais (Gerhardt; Silveira, 2009).

Considerando que a legislação citada ainda está sendo construída no âmbito brasileiro, é racional a adoção de uma abordagem exploratória, no sentido de entender se os preceitos legais que estão sendo definidos vão ao encontro da expectativa de respeito aos direitos humanos e aos valores democráticos, garantindo a implementação de uma IA segura e confiável, beneficiando a sociedade como um todo.

É importante destacar que a presente pesquisa realizará um recorte temático específico, concentrando-se exclusivamente nas implicações da inteligência artificial generativa, mais detidamente aos modelos que a sustentam. Essa delimitação é fundamental, uma vez que a IA generativa representa uma vertente inovadora e complexa dentro do campo mais amplo dessa tecnologia, trazendo à tona questões únicas que demandam uma análise aprofundada.

Ao focar nesse segmento, busca-se compreender não apenas as particularidades técnicas e operacionais desses modelos, mas também as implicações éticas e jurídicas que emergem de sua utilização. Dessa forma, a pesquisa se propõe a investigar como os preceitos legais brasileiros em construção podem efetivamente abordar os desafios e as oportunidades apresentados pela inteligência artificial generativa, assegurando que os direitos humanos e os valores democráticos sejam respeitados e promovidos em um contexto de rápida evolução tecnológica.

A interseção entre proteção de dados e inteligência artificial – mais especificamente, a generativa – constitui um dos temas mais relevantes e desafiadores do cenário jurídico contemporâneo, especialmente diante do volume e da sensibilidade das informações pessoais utilizadas para o treinamento desses modelos e funcionamento desses sistemas. Embora o presente trabalho não se proponha a aprofundar a análise sobre as complexas questões envolvendo a proteção de dados no contexto da IA generativa, parte-se da premissa de que os dados pessoais são o alicerce fundamental de toda a cadeia de desenvolvimento, operação e aprimoramento dessas tecnologias. Ainda porque, a LGPD, diante da ausência de legislação específica sobre IA no cenário brasileiro, serve atualmente como o principal parâmetro para a análise de questões relacionadas à essa tecnologia.

A LGPD estabelece, por exemplo, diretrizes para o tratamento de dados pessoais, o que inclui a coleta, o armazenamento e o uso desses dados. As duas primeiras etapas são cruciais para o desenvolvimento dos modelos generativos. Assim, como a inteligência artificial generativa depende fortemente de dados para operar, as regras da LGPD se aplicam diretamente ao uso de IA.

Além disso, a LGPD exige transparência e responsabilidade no uso de dados pessoais, o que inclui fornecer explicações claras sobre decisões automatizadas tomadas por sistemas de IA. À vista disso, a LGPD garante a proteção de dados pessoais e a privacidade dos indivíduos no contexto de tecnologias emergentes, como a inteligência artificial.

Reconhece-se, assim, a importância crucial da observância dos princípios e das normas de proteção de dados como forma de resguardar os direitos fundamentais e garantir a confiança social no uso da inteligência artificial, ainda que a abordagem detalhada desse tema extrapole os objetivos deste estudo.

Como suporte ao desenvolvimento do presente trabalho, tem-se o Projeto de Lei nº 2.338/2023 (ainda em votação pela Câmara dos Deputados) como ponto focal de análise, abarcando, ainda, documentos relevantes que permeiam a própria formação do projeto, bem como sua ampla alteração através do substitutivo. A análise do PL nº 2.338/2023 terá foco tão somente ao que trespasse os modelos generativos, com a finalidade de identificar os mecanismos de responsabilização prescritos, em coordenação com o direito privado brasileiro.

Nesse sentido, verificou-se a necessidade de combinar a análise do PL nº 2.338/2023 aos institutos jurídicos de personalidade e responsabilidade civil regulados pelo Código Civil brasileiro e, também, pelo Código do Consumidor, na medida em que o estudo se propõe a analisar os danos decorrentes da utilização desses modelos.

O presente trabalho reconhece a relevância e a imprescindibilidade da análise e categorização dos riscos previstos nas legislações que regulam e pretendem regular a inteligência artificial em âmbito global. Contudo, optou-se, neste estudo, por direcionar o foco investigativo para as minúcias inerentes à própria configuração da IA, especialmente no que tange à responsabilização pelos danos eventualmente causados por seus modelos e sistemas.

Assim, a pesquisa se debruça sobre a identificação e a delimitação proporcional da responsabilidade dos agentes responsáveis, buscando aprofundar a compreensão acerca das fronteiras e dos critérios que devem nortear a atribuição de responsabilidade civil no contexto da inteligência artificial generativa. Ao privilegiar essa abordagem, o presente estudo visa contribuir para o debate acadêmico e jurídico, oferecendo subsídios para a construção de soluções mais precisas e adequadas à complexidade dos desafios impostos pela autonomia e opacidade dos modelos e sistemas de IA.

Em nível internacional, o *European Artificial Intelligence Act*, ou, tão somente, AI Act,

normativa que entrou em vigor no dia 1 de agosto de 2024, ora utilizado como base para o referido projeto de lei, será explorado no sentido de comparação quanto ao entendimento dos termos técnicos relacionados aos modelos generativos e suas implicações legais, bem como o tratamento dado à proteção da pessoa humana face ao uso dessa tecnologia. Frisa-se que, pelo fato do AI Act ter subsidiado a construção do PL nº 2.338/2023, é inevitável que a pesquisa assuma um caráter majoritariamente comparativo, mas que estará restrito aos aspectos ora mencionados. Nesse sentido, o estudo exploratório proporciona uma visão geral do problema, desenvolvendo uma investigação mais ampla (Gil, 2008 *apud* Korkmaz, 2019).

A distinção conceitual entre modelo e sistema de IA, nesta medida, é crucial para a correta atribuição de responsabilidade e para a tutela dos indivíduos afetados. Sem essa diferenciação, a legislação pode falhar em identificar corretamente os pontos de controle e as obrigações de cada parte envolvida, de maneira proporcional, resultando em lacunas na proteção jurídica dos direitos fundamentais e da própria pessoa humana.

O presente trabalho não pretende, contudo, enrijecer a legislação em construção, levando-a a um existencialismo meramente técnico que impossibilite a imputação de responsabilidade e submissão do indivíduo afetado a uma condição ainda mais vulnerável (ante a ausência de responsabilização do agente do dano). Ao contrário, o que se pretende é que a tecnicidade, do ponto de vista da computação, seja considerada nas discussões jurídicas para que a checagem e determinação da responsabilidade sejam sopesadas proporcionalmente a partir do quanto do dano, de fato, cada agente contribuiu.

Ademais, a interpretação do Projeto de Lei nº 2.338/2023 à luz da Constituição Federal é de suma importância, pois garante que a regulamentação da inteligência artificial esteja em consonância, sobretudo, com os princípios fundamentais da dignidade da pessoa humana, da igualdade, da transparência e da eficiência. A dignidade do indivíduo assegura que os direitos fundamentais sejam respeitados, evitando discriminações e protegendo a sua privacidade. O princípio da igualdade promove a aplicação equitativa das normas, garantindo a inclusão digital e o acesso justo às tecnologias. A transparência é essencial para permitir o controle social e a fiscalização, assegurando que o uso da inteligência artificial seja claro e compreensível para todos. Por fim, a eficiência incentiva o uso inovador e sustentável da tecnologia, promovendo o desenvolvimento econômico e social. Dessa forma, interpretar constitucionalmente o referido projeto de lei assegura que a regulamentação da inteligência artificial, de forma ampla ou mais restrita, a exemplo da IA generativa, ora em estudo, contribua positivamente para a sociedade, respeitando os valores e direitos estabelecidos pela CRFB.

É comum que no método dedutivo o pesquisador use, como base, referências teóricas, buscando a aplicação da teoria aos objetivos da investigação (Gerhardt; Silveira, 2009). À vista de tal entendimento, o trabalho de Beckers e Teubner (2021) servirá como referência teórica ao

presente estudo, buscando compreender a extensão da aplicabilidade aos modelos generativos dos regimes de responsabilidade elucidados pelos autores.

Ressalta-se que a teoria de Beckers e Teubner (2021) não foi moldada propriamente para o enquadramento de modelos generativos. Porém, à vista dos estudos empenhados neste trabalho, entendeu-se pelo possível enquadramento do comportamento, do risco e da responsabilidade definidos pelos autores à realidade dos modelos generativos.

Para Beckers e Teubner (2021), a lei precisa ser capaz de coibir riscos específicos, reconhecendo esses algoritmos personificados como agentes vicários, as associações humano-máquina como empresas coletivas e os sistemas interconectados como *pools* de risco. Assim, a partir da identificação dessas instituições sociodigitais, os autores partem da análise de um comportamento, ao qual atribuem um risco e, por fim, inferem a responsabilidade aplicável.

Especialmente em face de novas tecnologias e suas implicações sociais, a análise sociológica do direito combinada à dogmática jurídica comparativa é oferecida como solução para os desafios contemporâneos da responsabilidade legal (Beckers; Teubner, 2021). Nesse sentido, Beckers e Teubner (2021) discorrem, ainda, sobre a aplicação da "jurisprudência sociológica comparada" de Collins (2011) à sua obra, utilizando-se do método para analisar as instituições sociodigitais e seus riscos inerentes ao enquadramento das categorias jurídicas relevantes em diferentes sistemas jurídicos, comparando-as às especificidades das doutrinas nacionais.

Em relação ao risco analisado pelos autores, é necessário destacar que se difere dos riscos abordados pelas legislações em construção e pelo próprio AI Act. Isso porque, para Beckers e Teubner (2021), a análise do risco é pautada no próprio dano, enquanto o risco categorizado na regulação da IA ao redor do mundo tem uma abordagem anterior ao dano e, portanto, é mais abstrata.

Apesar da obra de Beckers e Teubner (2021) traçar uma análise jurídica com base em uma dimensão comparativa, utilizando a teoria civilista alemã para o entendimento do status dos algoritmos na lei combinada ao direito consuetudinário, particularmente americano e inglês, é relevante destacar que a problemática também está em discussão na formatação do ordenamento jurídico brasileiro para lidar com a matéria de IA, incluindo a IA generativa.

No entanto, este trabalho reconhece as particularidades de cada ordenamento jurídico e a impossibilidade de transição imediata dos institutos propostos pelos autores para a realidade brasileira. Exemplificativamente: as instituições sociodigitais sugeridas como sujeitos do dano não são imediatamente aplicáveis ao ordenamento jurídico brasileiro, para o qual somente existe personalidade jurídica para pessoa física ou jurídica; já com relação ao dano, no direito alemão ele é mais definido e analisado, em certa medida, pela contribuição da própria vítima para o dano sofrido, ao passo que no direito privado brasileiro o dano é mais generalista.

Nesse sentido, o presente estudo não se propõe a fazer uma análise comparativa entre os dispositivos legais dos ordenamentos jurídicos elucidados por Beckers e Teubner (2021) e a realidade legislativa brasileira. Mas, entende que, a investigação e a articulação jurídica de legislações e doutrinas de países que já estão demasiadamente avançados nas discussões sobre o tema emergem como uma estratégia eficaz para enfrentar as lacunas que se apresentam de forma concreta no contexto atual brasileiro. Essa abordagem não apenas possibilita uma compreensão mais aprofundada das dinâmicas legais em jogo, mas também propicia a identificação de soluções que possam ser adaptadas à realidade brasileira, contribuindo para um arcabouço normativo mais robusto e eficaz.

A adaptação, nesse sentido, deve respeitar a própria configuração do ordenamento jurídico brasileiro. Exemplificativamente, em matéria de responsabilidade, o direito privado brasileiro privilegia a análise do dano em relação ao sujeito. Isso, contudo, não se observa na teoria de Beckers e Teubner (2021), que partem da análise do sujeito (no caso, das instituições sociodigitais) para a identificação do dano.

Entretanto, ainda que esta dissertação entenda pela necessidade de adequação e equalização de atuais e futuras normas jurídicas brasileiras ao contexto global da inteligência artificial generativa, não pretende ir de encontro a uma ideia de autorregulação da IA. Valendo-se do entendimento de Frazão (2021), "o fato de ainda sabermos muito pouco sobre a inteligência artificial"[8] e, por isso, "diante de incertezas significativas", não é justificativa para que o governo não intervenha ou intervenha muito pouco na regulação da IA, ao passo que as empresas desenvolvedoras da tecnologia, enquanto supostas autoridades no assunto, intervenha de maneira exclusiva e excessiva[9].

Ainda que a pesquisa tome Beckers e Teubner (2021) como referencial teórico, é necessário esclarecer que ela se ancora nos estudos desenvolvidos no NEAPID, da Faculdade de Direito da Universidade Federal de Juiz de Fora, e no projeto de pesquisa "Inovação e Direito na IA: mapeamento normativo e análise do exercício de direitos fundamentais" (CNPq-universal), coordenado pelo professor Sergio Negri, orientador deste trabalho, dos quais a autora faz parte.

Nesse sentido, o posicionamento central de Negri (2020) é crítico quanto ao uso do antropomorfismo na abordagem jurídica e social da inteligência artificial (IA) e da robótica. Para ele, a tendência de projetar características humanas — como autonomia, consciência e

[8] Segundo Frazão (2021, pp. 1-2), "muito da nossa ignorância em relação à inteligência artificial decorre da atitude deliberada das empresas que atuam no setor que, por meio de uma série de estratégias empresariais e de instrumentos legais, como o segredo de negócios, criaram uma arquitetura de extrema opacidade, em que sistemas algorítmicos funcionam como verdadeiras caixas pretas (*black boxes*)".

[9] No mesmo sentido também se pronunciam Barroso e Mello (2024), afirmando que a "assimetria de informação e de poder entre empresas e reguladores" (Barroso; Mello, 2024, p. 26) é saliente, possibilitando que essas empresas desfrutem "de um poder econômico que é facilmente transformável em poder político" (Barroso; Mello, 2024, p. 26).

agência moral — em robôs e sistemas de IA é uma metáfora perigosa, especialmente quando utilizada como base para propostas regulatórias ou para a criação de uma personalidade jurídica eletrônica.

Negri (2020), argumenta que o antropomorfismo, ou seja, a atribuição de qualidades humanas a máquinas influencia diretamente o modo como o Direito e a sociedade pensam e regulam a IA, levando, inclusive, a uma confusão conceitual, apagando as diferenças fundamentais entre humanos e artefatos tecnológicos.

O professor ressalta que a adoção de metáforas antropomórficas pode comprometer a elaboração de normas jurídicas adequadas, pois cria a ilusão de que robôs e sistemas de IA possuem agência e autonomia comparáveis às humanas. O que o autor enfatiza, nesse sentido, é o risco de se transferir responsabilidades de programadores, engenheiros e operadores para entidades artificiais, obscurecendo quem são os verdadeiros responsáveis por eventuais danos.

É nesse momento que este estudo enxerga a proposta de Beckers e Teubner (2021) como interessante de ser aplicável aos modelos generativos. Ao tratar a responsabilidade do dano a partir de instituições sociodigitais, os autores são capazes de distribuir proporcionalmente a responsabilidade entre os usuários que usufruem da tecnologia, os intermediadores que desenvolvem sistemas com base em modelos generativos e, em última medida, os desenvolvedores e financiadores do modelo em concreto.

Na medida em que a pessoa humana se destaca como núcleo central da fundamentação e da axiologia do sistema jurídico, a elaboração de normativas para a proteção da pessoa no contexto tecnológico deve buscar um equilíbrio que evite a submissão à lógica de eficiência do mercado, que tende a instrumentalizar o indivíduo para seus próprios fins, ao mesmo tempo em que deve ser pertinente à realidade e à evolução tecnológica, sob pena de se tornar hermética e ineficaz.

Por fim, destaca-se que a proeminência de exemplos provenientes da seara jurídica ao longo desta dissertação não constitui mera coincidência, mas reflete, a priori, a trajetória acadêmica e profissional da autora: bacharela em Direito pela Universidade Federal de Juiz de Fora, construiu sua identidade profissional, a partir da imersão constante nos fundamentos dogmáticos e metodológicos do Direito, para a exploração das zonas de confluência entre direito, inovação e tecnologia.

Assim, ao selecionar casos práticos, jurisprudências paradigmáticas e referenciais normativos que ilustram a aplicação de metodologias tecnológicas no ambiente jurídico, a autora possui a plena consciência de que seu repertório técnico-científico se ancora predominantemente no ordenamento jurídico. Nesse sentido, a interseção entre direito e tecnologia não é mero plano de fundo, mas a própria arena na qual se desdobra a investigação.

A construção de uma legislação que tutele efetivamente a centralidade da pessoa frente

à evolução tecnológica é impreterível ante uma corrida mercadológica que vislumbra coleções de dados como infraestrutura, e não mais como materiais pessoais (Crowford, 2021). Nas lições de Rodotà (2004), nem tudo que é tecnicamente possível é socialmente desejável, eticamente aceitável e juridicamente admissível quando o objeto é analisado a partir da perspectiva de constitucionalização da pessoa humana.

3 MODELOS GENERATIVOS E IA DE PROPÓSITO GERAL

A compreensão de modelos generativos é subsidiada pela própria evolução dos modelos de *machine learning*, os quais, segundo Mitchell (1997), são o campo da inteligência artificial que utiliza algoritmos e técnicas estatísticas para permitir que os computadores aprendam[10] a partir de dados, fazendo previsões ou, até mesmo, tomando decisões sem serem explicitamente programados para tal.

Na concepção de Goodfellow *et al.* (2016), a máquina aprende sem necessariamente ter sido programada para aquela finalidade, apenas identificando padrões a partir do conjunto de dados que subsidiaram o seu treinamento. Esse processo é chamado de aprendizado supervisionado, sendo aquele em que é possível definir qual é o *output* adequado, a exemplo dos sistemas de IA voltados para categorizações (em gatos, cachorros e meios de transporte) (Barroso; Mello, 2024).

Kaufman (2022, p. 43), sobre o entendimento de aprendizado de máquina, expõe:

> A técnica de aprendizado de máquina (*machine learning*), que permeia a maior parte das implementações atuais de inteligência artificial, redes neurais profundas (*deep learning*), consiste em extrair padrões de grandes conjuntos de dados, origem de seu sucesso, mas igualmente de sua fragilidade. Durante anos diversas bases de dados tendenciosas foram usadas para desenvolver e treinar algoritmos de IA, sem escrutínio. 'Reunir dados de qualidade em grande escala é caro e difícil. Criar grandes conjuntos de dados logo se tornou a versão da IA para a fabricação de salsichas: tedioso e difícil, com alto risco de usar ingredientes ruins', pondera Hutson.

No entendimento de Hastie *et al.* (2009), a partir do momento em que a máquina identifica padrões, ela será capaz de tomar decisões, seja pela combinação do aprendizado supervisionado ou pelas técnicas de aprendizado não supervisionado, em que o modelo é treinado usando dados não rotulados. Ademais, esse processo também envolve o aprendizado por reforço (ou *reinforcement*), cujo objetivo é treinar e aprimorar máquinas por meio de um sistema de tentativa e erro (Brown, 2021 *apud* Barroso; Mello, 2024), por meio do qual a máquina aprende a tomar decisões interagindo com o ambiente (Sutton *et al*., 2018).

A partir dessa decisão, o modelo é capaz de executar tarefas, tais como classificação, que consiste na categorização dos dados em classes previamente conhecidas, e regressão, que prediz o valor numérico de uma variável contínua de acordo com os valores dos dados (ANPD, 2024). Isso porque a implementação de modelos de ML envolve um processo de quatro etapas, basicamente, sendo elas: a coleta e preparação dos dados; o treinamento do modelo; a validação do modelo; e, finalmente, a interpretação dos resultados (Géron, 2019).

Em virtude disso, técnicas como o pré-processamento e a seleção dos recursos de

[10] Este Capítulo 3 se propõem a trazer uma abordagem mais técnica da IA, sob o prisma da Ciência da Computação. A análise acerca das ações que acabam por humanizar a IA, trazendo uma postura antropomórfica, será mais bem delineada no Capítulo 4, em que será abordado o tema da personalidade jurídica em matéria de IA.

entrada do algoritmo podem ser essenciais para aprimorar o seu desempenho, garantindo que ele capture as informações relevantes e ignore ruídos irrelevantes (ANPD, 2024), além de prevenir quaisquer tipos de discriminação em relação à pessoa humana, sobretudo aos grupos minoritários[11]. Ademais, é importante frisar que os desenvolvedores[12] possuem uma responsabilidade crucial para a performance ética e não discriminatória do modelo.

Suplantando os modelos de *machine learning*, pelo entendimento de Goodfellow (2016), os algoritmos de aprendizado profundo (ou *deep learning*) são capazes de identificar padrões complexos nos dados e melhorar suas previsões à medida em que são expostos a mais informações. Assim, esses modelos melhoram sua acurácia à semelhança com que humanos evoluem seu nível de aprendizagem. Isso porque os modelos computacionais inspirados no cérebro humano, chamados de redes neurais, são compostos por múltiplas camadas de neurônios artificiais, sendo que, no *deep learning*, essas redes são profundas, permitindo a modelagem de dados complexos e a extração de características em múltiplos níveis de abstração (Goodfellow *et al*., 2016). Essas arquiteturas construídas em DL subsidiam, por exemplo, o reconhecimento de imagens e o processamento de linguagem natural (*natural language processing*[13]).

É justamente a evolução do campo do processamento de linguagem natural que permitiu o aprimoramento e o crescimento exponencial dos modelos generativos, imputando ao legislador brasileiro a urgência de tratá-los ainda no texto propositivo de uma regulação para a inteligência artificial ampla.

Traçados os conceitos iniciais, os subcapítulos, a seguir, darão ênfase ao entendimento da tecnicidade da matéria em face do tratamento legislativo proposto, considerando, ainda, a evolução hierarquizada desse modelo de IA e a potencialidade lesiva à pessoa humana a partir do uso da tecnologia generativa.

3.1 Dicotomia entre o conceito técnico de modelos e de sistemas generativos

Em 2014, Ian Goodfellow publicou um trabalho sobre GANs, que se tornou o ponto de

[11] Entendidos como segmentos da população que, devido a características distintas como etnia, religião, orientação sexual ou outras, encontram-se em uma posição de desvantagem ou marginalização em relação ao grupo majoritário dominante, enfrentando, frequentemente, discriminação e tendo menos acesso a recursos e oportunidades (Machado, 2021).

[12] Segundo o texto do Projeto de Lei nº 2.338/2023, o conceito de desenvolvedor perfaz-se na "pessoa natural ou jurídica, de natureza pública ou privada, que desenvolva sistema de IA, diretamente ou por encomenda, com vistas a sua colocação no mercado ou a sua aplicação em serviço por ela fornecido, sob seu próprio nome ou marca, a título oneroso ou gratuito".

[13] Para Liddy (2021, p. 2, traduzido pela autora): "Processamento de Linguagem Natural (PLN) é a abordagem computacional para analisar textos que se baseia tanto em um conjunto de teorias quanto em um conjunto de tecnologias. E, por ser uma área de pesquisa e desenvolvimento muito ativa, não existe uma definição única e consensual que satisfaça a todos, mas há alguns aspectos que fariam parte da definição de qualquer pessoa conhecedora do assunto".

partida para o desenvolvimento de um novo tipo de modelo de ML, denominado generativo. No estudo, o autor sintetiza a diferenciação das GANs a partir de dois componentes principais, sendo o primeiro deles o gerador e, o segundo, o discriminador.

No processo iterativo de treinamento desses componentes, o gerador tem o papel de criar instâncias a partir de um espaço latente[14], enquanto o discriminador avalia se essas instâncias são reais (retiradas do conjunto de treinamento) ou falsas (criadas pelo gerador) (Goodfellow *et al*., 2014). A iteração se repete até que o gerador se torne bom o suficiente para "enganar" o discriminador, criando instâncias indistinguíveis das reais.

O grande diferencial que sustenta os modelos generativos é que, além do aprendizado a partir da associação dos dados do conjunto de treinamento (ou seja, dos modelos discriminativos convencionais de ML), essa nova modalidade é capaz de capturar características estatísticas subjacentes dos dados para gerar novos exemplos que se assemelham aos dados reais (ANPD, 2024).

À vista disso, o modelo generativo amolda-se a partir da combinação de uma rede neural discriminativa e de outra generativa. A rede generativa, por sua vez, será responsável por produzir dados sintéticos[15] a partir de ruídos aleatórios, enquanto a rede discriminativa é responsável por determinar se os dados apresentados a ela são reais ou foram gerados pela rede generativa. Assim sendo, o que ocorre é que a rede generativa estará em constante evolução para que seus dados sejam os mais próximos possíveis de dados reais, enquanto a rede discriminativa deverá desenvolver-se na mesma medida, assegurando a distinção do que parece real para o que realmente é. Segundo a ANPD (2024), "idealmente, os dados sintéticos gerados devem ser tão próximos dos dados reais, que a distinção entres eles não devem ser possíveis".

Atualmente, além dos modelos baseados em GANs, outros baseados em *Generative Pre-training Transformer* (GPT[16]) tornaram-se peças importantes na atual revolução da IA, na medida em que os modelos de aprendizado profundo disponíveis até então incorriam em uma série de dificuldades diante das exigências para uma IA geral (Amaral; Xavier, 2023). Isso porque os modelos anteriores operavam sequencialmente, processando palavra por palavra, o que trazia dificuldades em termos de eficiência e de tempo empregados no treinamento desses modelos, além de ser custoso do ponto de vista computacional (Amaral; Xavier, 2023).

Assim como nas GANs, os GPTs também utilizam a combinação de duas técnicas para

---

[14] Espaço latente refere-se a um espaço de menor dimensão no qual os dados de alta dimensão são incorporados. Esse espaço captura a estrutura ou as características subjacentes dos dados e é tipicamente aprendido por um modelo de aprendizado de máquina, como um modelo de aprendizado profundo ou um autocodificador (TedAI, 2024).

[15] Segundo o Radar Tecnológico nº 3, da Autoridade Nacional de Proteção de Dados (2024), o termo "dado sintético" é utilizado tanto no campo da TI quanto na área de privacidade e proteção de dados para se referir aos dados gerados a partir de modelos generativos, "(...) gerados artificialmente, em contraste com os dados reais que são oriundos da realidade" (AEPD, 2023).

[16] GPTs, ou Transformadores Generativos Pré-Treinados, em português, são modelos generativos de aprendizado de máquina que utilizam a abordagem codificador-decodificador para a geração de conteúdo sintético (ANPD, 2024).

o desenvolvimento de seus modelos. A primeira delas, a arquitetura de transformadores, é composta por camadas de codificadores e decodificadores, sendo utilizadas com o objetivo de gerar dados (ANPD, 2024). As camadas codificantes são responsáveis pela transformação do dado de entrada em uma representação numérica (espaço de representação), ao passo que os decodificadores são responsáveis pelas saídas do modelo, com base nas informações contidas no espaço de representação e em um contexto anterior (ANPD, 2024).

A segunda arquitetura, o mecanismo de atenção, é tida por Amaral e Xavier (2023) como o agente de revolução do transformador, que procura prever o próximo elemento (por exemplo, uma palavra) dentro de uma sequência. Pela abordagem do Radar Tecnológico da ANPD (2024), nesse caso, valendo-se do conceito elucidado por Vaswani (2017), essa é a técnica que permite ao modelo generativo concentrar-se nas especificidades do dado real, em vez de tratá-lo de maneira uniforme. Esse foco seletivo atribui pesos distintos às partes do dado de entrada com base na sua importância para a execução da tarefa, permitindo que sejam priorizadas em detrimento de outras de menor importância (ANPD, 2024).

A partir do mecanismo de atenção, o modelo se vale de uma sequência de palavras como *input* e, através da iteração entre camadas de codificação, procura prever a próxima palavra numa sequência de *output*. Essa técnica, conhecida como processamento paralelo, permite que os modelos sejam treinados com uma quantidade gigantesca de dados em um tempo relativamente curto, se comparado aos modelos anteriores (Amaral; Xavier, 2023).

Os GPTs possibilitaram o avanço dos modelos generativos, sobretudo a partir de 2018, quando a empresa OpenAI criou o primeiro GPT (denominado GPT-1). Nos últimos dois anos, esse avanço se intensificou com o lançamento do ChatGPT, um *chatbot* desenvolvido pela mesma empresa, capaz de oferecer respostas contextualmente relevantes a partir de entradas fornecidas pelos usuários. Até então arquitetado com base no GPT-1, esse sistema impulsionou a IA generativa a uma proporção de crescimento exponencial. Com isso, a OpenAI tornou essa tecnologia amplamente acessível em todo o mundo, sobretudo às pessoas não especialistas.

O funcionamento do ChatGPT, por exemplo, é propiciado a partir da experiência que o agente de IA adquire com a interação humana. Tecnicamente, o modelo quebra os textos que compõem a sua base de treinamento em pequenas unidades chamadas "*tokens*", atribuindo a cada uma delas um número que varia de acordo com o contexto, podendo um mesmo atributo (por exemplo, uma palavra) receber números distintos (Amaral; Xavier, 2023). Feito isso, o modelo observa a relação entre os atributos dentro do mesmo conjunto de treinamento, com foco em identificar quais deles aparecem juntos com mais frequência, calculando a probabilidade de um aparecer junto a outro e atribuindo pesos (Amaral; Xavier, 2023).

A partir da composição do atributo de entrada em número (*tokens*) e da associação probabilística dentro de um contexto, o modelo torna-se capaz de gerar, como saída, sequências

cada vez mais parecidas com os atributos que estão na sua base de treinamento (Amaral; Xavier, 2023). Explicam os autores:

> Por exemplo, diante da sequência “machado de”, o modelo consegue saber que, num contexto em que o texto trata de ferramentas, é maior a probabilidade de ocorrer a palavra “mão”. Fosse um contexto sobre literatura e cultura, o modelo saberia que a palavra com maior probabilidade de ocorrência seria “Assis”. De forma geral, esta é uma operação semelhante ao recurso autocompletar que já funciona em nossos motores de busca na internet ou nos aplicativos de mensagens disponíveis nos celulares. A diferença é que o modelo de linguagem, no caso do ChatGPT, seria uma ferramenta de autocompletar muito mais poderosa e complexa do que qualquer recurso com qual tivemos contato até o momento no mundo informático (Amaral; Xavier, 2023, p. 29).

Os mecanismos de atenção também se tornam importantes ao entendimento dos modelos GPTs, pois permitem que a máquina, por exemplo, diferencie o que é um dado de entrada de texto (e, portanto, estruturado) daquele não estruturado (como imagens e vídeos). O modelo generativo que opera somente com um único tipo de dado é dito unimodal (a exemplo do GPT-1, GPT-3, LaMDA e LLaMa, que geram textos como saída a partir de entradas também textuais), enquanto o modelo multimodal é aquele que processa e gera múltiplos tipos de dados simultaneamente, integrando informações de entrada de diferentes tipos, como texto, imagem e vídeo (Skrlj, 2024). O GPT-4, por sua vez, é um exemplo de modelo multimodal.

Através dos mecanismos de atenção desses modelos, a máquina tanto pode valer-se da visão computacional, a partir da priorização de partes mais importantes de uma imagem para sua completa identificação, quanto do processamento de linguagem natural, que identifica as palavras mais importantes para entender o contexto.

Com a junção da arquitetura de transformadores e dos mecanismos de atenção (combinação da visão computacional com o processamento de linguagem natural), os modelos GPTs servem de base para o desenvolvimento de uma gama de sistemas de IA generativa.

O DALL-E é um sistema de inteligência artificial que cria imagens (dados de saída) a partir de descrições textuais (dados de entrada) do usuário, usando uma versão de 12 bilhões de parâmetros do modelo GPT-3, modelo generativo construído pela OpenAI (Johson, 2021). O próprio ChatGPT, como demonstrado acima, é um sistema de IA que, em sua primeira versão, teve o GPT-1 como suporte arquitetônico e, agora, chega à sua nova versão amparado pelo GPT-4o (OpenAI, 2025). Gemini, do Google, construído através do modelo de linguagem LaMDA (Google AI for Developers, 2025); Meta AI, baseando-se na arquitetura do modelo LLaMa (Meta AI, 2025); e, mais recentemente, o DeepSeek, sistema de IA generativa chinês, que utiliza o DeepSeek-V3 (modelo mais avançado) (DeepSeek, 2025), encerram esta lista exemplificativa.

Nessa medida, o que se pretende demonstrar é que existe uma diferença técnica entre os

conceitos de modelos e sistemas de inteligência artificial.

Um modelo, segundo Sommerville (2015), é uma representação abstrata que busca simplificar e entender a complexidade do que está sendo estudado, sem a necessidade de implementá-lo, ao passo que um sistema é um conjunto de componentes interconectados, incluindo modelos, que trabalham juntos para realizar uma tarefa específica.

Em literaturas recentes, o termo "modelo fundacional" ou "modelo fundamental" (ou *foundation model*, em inglês) passou a ser comumente associado ao estudo da inteligência artificial generativa. Pelo entendimento do Gartner (2024), o modelo fundacional é aquele com grandes parâmetros treinados a partir de uma gama de dados de maneira autossupervisionada, sendo que o termo, literalmente, faz referência à sua importância e aplicabilidade de servir de base para uma coleção diversificada de sistemas e aplicações.

É comum encontrar na literatura a associação de modelos fundacionais apenas a tarefas generativas, uma vez que são amplamente utilizados no processamento de linguagem natural para a tarefa de geração de texto (ANPD, 2024). No entanto, é válido frisar que não é correto assumir que modelos generativos e modelos fundacionais são sinônimos. Um modelo generativo, a bem da verdade, pode ser entendido como um tipo de modelo fundacional, mas, nos ensinamentos de Jones (2023 *apud* Barroso e Mello 2024, p. 10), "nem toda IA generativa constitui um modelo fundacional". Capacidades generativas para manipulação e análise de texto, de imagem e vídeo, e a produção de discurso, bem como sistemas de recomendação de conteúdo em plataformas de streaming, destinam-se a uma finalidade bastante delimitada (Barroso; Mello, 2024).

Do mesmo modo, o AI Act, considerado uma das legislações mais avançadas e abrangentes sobre regulação de IA no mundo, define *foundation models* como modelos de inteligência artificial de propósito geral, que são treinados em grandes quantidades de dados e podem ser adaptados para uma ampla variedade de tarefas, servindo de base para muitas aplicações de IA. Aliás, o texto do Considerando 97, do AI Act, é reluzente ao tratar da diferenciação entre modelo de IA e sistema de IA:

> **A noção de modelos de IA de propósito geral deve ser claramente definida e diferenciada da noção de sistemas de IA para permitir a segurança jurídica.** A definição deve ser baseada nas **principais características funcionais de um modelo de IA de propósito geral, em particular a generalidade e a capacidade de realizar, de forma competente, uma ampla gama de tarefas distintas.** Esses modelos são tipicamente treinados em grandes quantidades de dados, por meio de vários métodos, como aprendizado autossupervisionado, não supervisionado ou por reforço. **Modelos de IA de propósito geral podem ser disponibilizados no mercado de várias maneiras, incluindo por meio de bibliotecas, interfaces de programação de aplicativos (APIs),** download direto ou como cópia física. Esses modelos podem ser ainda modificados ou ajustados em novos. **Embora sejam componentes essenciais dos sistemas de IA, os modelos não constituem sistemas de IA por si só. Eles requerem a adição de outros componentes, como, por exemplo, uma interface de usuário, para se tornarem sistemas de IA.** Os modelos de IA são tipicamente

> integrados e fazem parte dos sistemas de IA (...) (European Parliament, 2024, traduzido e realçado pela autora).

Nesse sentido, em síntese, um modelo generativo é um algoritmo de *machine learning* que aprende a capturar os dados de entrada para criar amostras, com vista a atender a uma ampla gama de tarefas distintas. Um sistema de inteligência artificial generativa, por sua vez, integra esses modelos generativos a outras entidades (como *hardware*, banco de dados e interface de usuário) para gerar um contexto prático, a exemplo do DALL-E, explicitado acima.

Nos últimos meses, tribunais ao redor do mundo precisaram lidar com questões envolvendo IA generativa de forma acessória à demanda levada ao conhecimento do juízo. Em linhas gerais, os magistrados voltaram-se para a avaliação da responsabilização de advogados utilizando essa tecnologia para revisão ou elaboração do seu trabalho jurídico, o qual integrou os autos, sem que, contudo, tenham se atentado a uma revisão quanto à veracidade e autenticidade da informação gerada.

Quanto a isso, exemplifica-se a partir de casos trazidos a conhecimento público entre fevereiro e março de 2025. No primeiro deles, um tribunal federal americano, em Wyoming, multou três advogados do escritório Morgan & Morgan, em uma ação contra a Walmart, por terem citado, na petição inicial, oito precedentes que não existem[17]. O erro deu-se em virtude da utilização de um sistema de IA generativa, levando o juiz ao entendimento de que os advogados faltaram com o dever ético de checar as citações usadas em suas petições.

No segundo caso, um juiz de um tribunal federal em Indiana foi um pouco mais repreensivo, aplicando uma multa de US$ 15 mil em virtude de citações falsas encontradas na petição, entendendo, também, que faltou ao advogado a checagem de autenticidade das informações[18].

No Brasil, a retórica também se aplica – não só pelo viés de profissionais patrocinando causas judiciais, mas também dos próprios magistrados colocando-se nessa situação. Em janeiro de 2025, o Portal Jota noticiou um caso em que um advogado solicitou a anulação de uma sentença proferida pela 4ª Vara Cível de Osasco, sob o argumento de que ela teria sido elaborada por inteligência artificial, ferindo, assim, o princípio do juiz natural[19].

O que se quer chamar atenção não é simplesmente o fato do cabimento de responsabilidade civil em decorrência da ausência de zelo profissional, mas sim de entender se

---

[17] MELO, J. O. de. Cortes dos EUA recorrem a multas para conter precedentes falsos gerados por IA. Consultor Jurídico. 2025. Disponível em: https://www.conjur.com.br/2025-mar-04/cortes-dos-eua-recorrem-a-multas-para-conter-precedentes-falsos-gerados-por-ia/. Acesso em: 16 mar. 2025.

[18] MELO, J. O. de. Cortes dos EUA recorrem a multas para conter precedentes falsos gerados por IA. Consultor Jurídico. 2025. Disponível em: https://www.conjur.com.br/2025-mar-04/cortes-dos-eua-recorrem-a-multas-para-conter-precedentes-falsos-gerados-por-ia/. Acesso em: 16 mar. 2025.

[19] CARVALHO, M. Advogado usa ChatGPT para identificar uso de IA em sentença e requer anulação. JOTA. 2025. Disponível em: https://www.jota.info/justica/advogado-usa-chatgpt-para-identificar-uso-de-ia-em-sentenca-e-requer-anulacao. Acesso em: 16 mar. 2025.

recai somente à pessoa que utilizou a IA, que alucinou, a totalidade dessa responsabilidade. No caso em análise, é essencial que os aplicadores do Direito estejam confortáveis com a conceituação técnica que propicia a própria existência da IA generativa.

Hipoteticamente, imagine um magistrado brasileiro que precisa prolatar uma sentença em uma demanda judicial que envolva dano a um grupo vulnerável decorrente da utilização de um sistema de IA, desenvolvido por uma empresa brasileira. Ao longo da ação, no entanto, a empresa brasileira reforça que seu sistema se utilizou de um modelo generativo desenvolvido nos Estados Unidos, por uma empresa norte-americana, e que, na realidade, o dano foi ocasionado em virtude desse modelo ter se valido de dados pessoais para o treinamento, fato que não era de conhecimento da empresa brasileira. Supondo que o entendimento norte-americano seja o de que dados pessoais que estejam on-line são de domínio público e, portanto, sem óbice a ser usado como subsídio para o treinamento de uma IA, como o tribunal deverá se manifestar? A empresa brasileira também poderia estar na condição de consumidora do modelo e, então, exposta a certa vulnerabilidade que deveria ser considerada pelo magistrado?

A IA generativa caminha para um viés de agentes personalizados, que, para além da incorporação de um modelo para operacionalizar um sistema, a técnica do *fine-tuning*, ou ajuste fino, tem contribuído para essa tendência. O processo de *fine-tuning* consiste em valer-se de um modelo pré-treinado com uma grande quantidade de dados genéricos para, posteriormente, ajustá-lo a um conjunto de dados menor e mais específico, de acordo com a necessidade de uma aplicação particular (Goodfellow *et al*., 2016). Retomando a situação hipotética acima, caso a empresa brasileira tenha realizado um *fine-tuning* do modelo da empresa norte-americana, a responsabilidade a ela cabível na primeira ilustração deveria ser a mesma desta segunda?

Como abordado no Capítulo 2, a distinção conceitual é crucial para a correta atribuição de responsabilidade e para a tutela dos indivíduos afetados. Como na hipótese acima, um modelo generativo, enquanto representação abstrata que identifica padrões a partir dos dados de entrada, pode ser desenvolvido e treinado em um contexto internacional, enquanto o sistema de IA, que integra esse modelo a outros componentes para realizar tarefas específicas, pode ser implementado e operado localmente.

Ao inserir ferramentas de IA na elaboração de documentos, na pesquisa jurídica e na análise de contratos, os profissionais do Direito podem otimizar suas atividades e reduzir seus custos operacionais. Nesses casos, a imputação de responsabilidade exclusiva ao advogado que utiliza IA generativa como meio para realização do serviço jurídico atrai a atenção, pois atribui a um único sujeito toda a cadeia de responsabilidade que engloba a construção e o desenvolvimento de um sistema generativo. Não que ao advogado não caiba o dever de supervisionar o conteúdo produzido por essas ferramentas, em conformidade com os princípios éticos e legais que regem a profissão, mas é necessária uma abordagem mais ampla dos atores

envolvidos no desenvolvimento desse agente digital, sob pena de uma injusta atribuição de responsabilidade exclusiva.

Reforça-se, ademais, que o presente trabalho não pretende enrijecer a legislação em construção, simplificando-a a uma discussão essencialmente técnica da tecnologia, que impossibilite a imputação de responsabilidade e submeta o indivíduo afetado a uma condição ainda mais vulnerável.

Schuett (2021), exemplificativamente, ao tratar da perspectiva comparativa de regulamentação da IA, com base em conceitos técnicos versus a abordagem baseada em risco, frisa que a simples definição do termo "inteligência artificial" para definir o escopo material das normas não é suficiente. Ele aponta que as definições existentes de IA são excessivamente amplas, vagas e, frequentemente, não atendem aos requisitos necessários para definições legais e eficazes, como precisão, compreensibilidade, praticidade e flexibilidade. Isso ocorre porque o termo IA abrange uma vasta gama de sistemas com perfis de risco muito diferentes, tornando impossível tratá-los de forma adequada sob uma única definição.

Ao invés disso, o autor sugere que, em vez de usar o termo genérico "IA", os legisladores deveriam considerar a inclusão de abordagens técnicas específicas (por exemplo, aprendizado por reforço, aprendizado supervisionado, aprendizado não supervisionado) na definição do escopo das regulações. Cada abordagem técnica pode apresentar riscos próprios, como efeitos colaterais negativos, exploração de falhas na função de recompensa ou dificuldades de interrupção segura em sistemas de aprendizado por reforço.

Definir o escopo com base em abordagens técnicas é útil, especialmente em níveis mais detalhados de regulamentação. Este estudo entende que a tecnicidade, do ponto de vista da computação, agrega às discussões e às próprias definições jurídicas, trazendo um lastro mais seguro para a eficácia legal, sobretudo no que tange à responsabilização em virtude de dano decorrente do uso de IA generativa.

A definição formalmente técnica, ainda que pareça meramente um ajuste de conduta, perfaz-se para além da significância semântica. Ela implica em consequências e debates ainda não vislumbrados, sobretudo em uma seara em estágio exploratório pela própria comunidade jurídica. Ainda que não deva se fazer como elemento central da regulamentação, ela é um atributo acessório de importância significativa (e, até mesmo, decisiva em certas situações).

Veja-se o seguinte caso concreto: lançado oficialmente em janeiro de 2025, o DeepSeek emergiu como um sistema de IA generativa com potencial para sobrepujar a hegemonia do ChatGPT, da OpenAI. Em entrevista ao G1, dias após o lançamento do DeepSeek, Cleber Zanchettin, professor do CIn-UFPE e um dos líderes em pesquisa em inteligência artificial na América Latina, destacou quatro principais características associadas ao *boom* do DeepSeek (Mota, 2025).

A empresa chinesa utiliza seu modelo de linguagem de grande escala (*large language models*) em código aberto, o que permite à comunidade científica entender a cadeia de raciocínio para se chegar aos modelos mais avançados de IA. Até então, essa prática não era adotada pelas demais empresas, que disponibilizavam seu LLM em código fechado, como no caso do ChatGPT, ou parcialmente fechado, como no caso do LLaMa, da Meta, que divulgava somente alguns parâmetros.

O raciocínio explícito, que detalha o passo a passo lógico para se chegar às respostas, também é um diferencial, segundo Zanchettin, haja vista que, até o lançamento do DeepSeek, "a maioria das empresas não queria que a gente entendesse direito [como o modelo raciocina], porque isso pode levar você a perceber que ele está fazendo as coisas direito ou que não entendeu nada, e que o resultado é mais ou menos aleatório" (Mota, 2025). Essa característica é essencial para entender a perspectiva centralizadora e hegemônica que poucas empresas de tecnologia comumente usam para se manterem como supremas no mercado de tecnologia de IA, ditando o modelo de funcionamento do mercado e da economia de modo global, quesito que será mais bem examinado no próximo subcapítulo.

A terceira característica, de acordo com o pesquisador, diz respeito à aprendizagem por reforço do modelo, que apresenta uma dependência muito menor da supervisão humana, ao passo que os modelos até então disponíveis no mercado demandavam bastante intervenção do indivíduo, utilizando a estratégia conhecida como "humano no *loop*" (HITL). Tal estratégia possui uma abordagem na qual humanos são integrados ao ciclo de treinamento, ajuste e teste de modelos de aprendizado de máquina, contribuindo para melhorar a precisão e a eficácia dos sistemas de IA (Amershi *et al*., 2014). Segundo Zanchettin, essa é a característica que permite ao modelo ser mais barato computacionalmente, à medida em que passa por uma autoevolução (Mota, 2025).

Por fim, Zanchettin elucida como quarta e última característica a questão da restrição à inovação, uma vez que a companhia chinesa foi capaz de desenvolver um modelo extremamente potente, do ponto de vista técnico, sem os melhores chips disponíveis no mercado, visto que os Estados Unidos, em 2022, impuseram restrições à China para a importação desses insumos (Mota, 2025).

Dias após o lançamento do sistema DeepSeek, fruto do desenvolvimento conjunto da Hangzhou DeepSeek Artificial Intelligence Co. e da Beijing DeepSeek Artificial Intelligence Co., a OpenAI as acusou de terem estruturado o seu modelo a partir do modelo generativo GPT-4, da própria OpenAI (Financial Times, 2025). Em análise mais técnica, a acusação da empresa é amparada pelos seus "Termos de Serviço para Negócios" que, além de dispor que é proibido a empresas ou desenvolvedores terceiros usar o conteúdo do modelo proprietário da empresa para desenvolver quaisquer outros modelos de IA que concorram com seus produtos e serviços,

também reforça que é vedada a extração de dados diferentes dos permitidos pelas APIs fornecidas pela OpenAI.

Em entrevista ao Financial Times (2025), a OpenAI afirma ter evidências de que o modelo do DeepSeek se utilizou de uma técnica conhecida como "destilação" (*distillation*, em inglês), por meio da qual os desenvolvedores extraem dados de grandes modelos de IA para treinar modelos menores ainda em desenvolvimento, para construção do próprio modelo de IA dessa empresa (Financial Times, 2025).

Como reforçado por Zanchettin, um modelo de linguagem computacional pode configurar-se como de código aberto ou fechado (Mota, 2025). Sendo um modelo de código aberto, seu código-fonte é disponibilizado publicamente, permitindo que qualquer pessoa possa visualizá-lo, modificá-lo, distribuí-lo ou implementá-lo em qualquer sistema (Raymond, 2001). Isso não se confirma em um modelo de código fechado, no qual apenas os desenvolvedores autorizados pela empresa detêm direitos de acessar, modificar ou distribuir o código (Stallman, 2009), preservando, assim, os direitos de autor do desenvolvedor. Sendo o modelo da OpenAI de código fechado, isso implica que as empresas chinesas não poderiam tê-lo utilizado como arcabouço para desenvolvimento do seu modelo, que serviria de base para o seu sistema de IA generativa DeepSeek.

Diante desse cenário, confirma-se que a tecnicidade relacionada à seara da Ciência da Computação apresenta importância considerável para resoluções de disputas jurídicas em matéria de IA generativa.

3.2 Desenvolvimento hegemônico da IA de propósito geral a partir de uma perspectiva geopolítica e não sustentável

Segundo o Gartner (2024), inteligência artificial de propósito geral, ou *general purpose AI*, é uma forma de IA com capacidade para entender, aprender e aplicar o conhecimento em uma gama de tarefas e domínios, podendo ser dedicada a um conjunto muito mais amplo de casos de uso, à medida em que incorpora flexibilidade cognitiva, adaptabilidade e habilidades gerais de resolução de problemas[20].

Os modelos de inteligência artificial de propósito geral (*general-purpose AI models*), segundo a ANPD (2024), caracterizam-se por ser um tipo de modelo de IA com habilidades gerais de resolução de problemas, com o objetivo de criação de sistemas mais capazes de se adaptarem a uma variedade de cenários, estando no horizonte do avanço científico.

Sob o ponto de vista de Russell e Norvig (2021), a IA de propósito geral se aproxima

[20] Como será demonstrado no Capítulo 4, sobretudo no Subcapítulo 4.1., este estudo não entende a IA sob essa ótica, que acaba por antropomorfizar a tecnologia.

do conceito da IA forte, em contraposição ao de uma IA fraca[21]:

> Definição de IA fraca: "é um sistema de IA que é projetado para realizar tarefas específicas e limitadas, com base em um conjunto de regras predefinidas e modelos estatísticos. Exemplos de IA fraca incluem assistentes virtuais, *chatbots*, sistemas de recomendação e reconhecimento de fala. Embora esses sistemas possam ser muito eficazes em suas tarefas específicas, eles geralmente não possuem a capacidade de aprender e adaptar-se a novas situações ou contextos" (Russell e Norvig, 2021, p. 27, traduzido pela autora).
>
> Definição de IA forte: "é um sistema de IA que é projetado para ter a capacidade de pensar, aprender e resolver problemas como um ser humano. A IA forte ainda é um objetivo a ser alcançado, uma vez que até o momento, nenhum sistema de IA foi capaz de alcançar a inteligência humana em sua totalidade. No entanto, pesquisadores continuam a trabalhar em direção a esse objetivo, utilizando técnicas como aprendizado profundo, redes neurais e processamento de linguagem natural" (Russell e Norvig, 2021, p. 27, traduzido pela autora).

Com a disputa pela supremacia da inteligência artificial generativa de forma global, sobretudo após o lançamento do DeepSeek, que acirrou os conflitos entre os Estados Unidos e a China, empresas desenvolvedoras dessa tecnologia estão constantemente lançando novas versões de seus modelos e sistemas com o intuito de aprimorar incessantemente seus LLMs e se manterem competitivas no mercado.

Para além do anseio pela liderança tecnológica, o que se observa, sobremaneira, é uma disputa geopolítica, refletindo a crescente rivalidade entre as duas maiores economias do mundo. A título de exemplo, a primeira versão do GPT, em 2018, utilizou 110 milhões de parâmetros para o aprendizado do modelo (Floridi; Chiriatti, 2020, p. 684). Um ano depois, o GPT-2 utilizou 1,5 bilhão (Floridi; Chiriatti, 2020, p. 684). Para a versão do GPT-3, 175 bilhões de parâmetros, com custo estimado em US$ 12 milhões, foram treinados em 2020, com a dedicação de um dos supercomputadores mais potentes da Microsoft à época, através do Azure[22] (Floridi; Chiriatti, 2020, p. 684).

É válido reforçar que o serviço do Azure OpenAI pode implicar em um crescimento monopolizado da IA generativa em escala mundial, na medida em que qualquer cliente ou parceiro da Microsoft pode ter acesso ao modelo de IA da OpenAI para subsidiar a criação de seus próprios sistemas de inteligência artificial generativa. Exemplificativamente, elencam-se algumas ferramentas de IA criadas a partir do Azure OpenAI: Harvey, plataforma de IA voltada para profissionais do Direito, oferecendo assistentes virtuais, automação de fluxos de trabalho

---

[21] Barroso e Mello (2024) também entendem como modelos de propósito determinado ou estrito (*fixed-purpose AI ou narrow AI models*), que, como o próprio nome sugere, são treinados com uma base mais direcionada de dados e destinam-se a um propósito específico.

[22] Microsoft Azure é uma plataforma de computação em nuvem da Microsoft (Microsoft, 2025). Em 2019, a empresa firmou parceria com a OpenAI, investindo bilhões de dólares e fornecendo acesso aos recursos de computação do Azure para a OpenAI, o que permitiu, em 2021, o lançamento do Azure OpenAI Service, que possibilita que os clientes da Microsoft acessem os modelos avançados da OpenAI diretamente, através da plataforma Azure (Microsoft Corporate Blogs, 2025).

e pesquisa (OpenAI, 2025); ContractMatrix, plataforma de IA jurídica do escritório inglês Allen & Overy Shearman, criada a partir do Harvey, com foco em revisão, redação e gestão de contratos para os advogados do escritório (A&O Shearman, 2025); Axios Local, ferramenta criada em parceria entre Axios e OpenAI com foco na expansão de cobertura de notícias locais em cidades dos Estados Unidos (OpenAI, 2025); entre muitas outras.

A facilidade com que grandes empresas têm proporcionado acesso à tecnologia de IA generativa deve se tornar uma preocupação constante. À primeira vista, esse tipo de tecnologia brilha aos olhos, sobretudo dos indivíduos não especialistas, que hoje são os principais consumidores da tecnologia. Mas, em análise mais crítica e detida, o que se observa é uma monopolização do próprio acesso à informação, na medida em que os sistemas de IA generativa têm sido estruturados a partir de modelos restritos, fornecidos por poucas e grandes empresas de tecnologia, que dominam o cenário tecnológico global há anos (ou não, a exemplo das empresas chinesas desenvolvedoras do DeepSeek).

Os discursos relacionados à "assistente pessoal", "realizar uma ampla gama de tarefas" e "informação instantânea" têm conquistado e motivado o uso da IA generativa de maneira exponencial na maioria dos países do mundo. Entretanto, é válido reforçar que, em conjunto com uma evolução tecnológica que promete simplificar e melhorar a vida das pessoas, observa-se, ainda mais, um controle de dados e informações. As próprias empresas desenvolvedoras exercem governança acerca dos dados e da forma como arquitetam esses modelos, gerando conflitos a direitos fundamentais, sobretudo no que diz respeito à violação da privacidade humana, à propagação de informações inverídicas e, não menos importante, ao esgotamento de recursos naturais.

Causam ainda mais preocupação ao cenário político global, sobretudo ao desenvolvimento da inteligência artificial, enquanto tema do presente estudo, as atuais medidas tomadas pelo governo Trump em tão pouco tempo de gestão. Em 20 de janeiro de 2025, Donald Trump assume a presidência dos Estados Unidos pela segunda vez. Em 22 de janeiro de 2025, apenas dois dias após tomar posse, o político derruba regras de antidiscriminação em contratações do governo, dando fim a programas de diversidade e ocasionando uma série de demissões em massa[23].

Segundo o presidente americano, em ordem executiva:

> As políticas ilegais de DEI [Diversidade, Equidade e Inclusão] e DEIA [Diversidade, Equidade, Inclusão e Acessibilidade] não apenas violam o texto e o espírito de nossas antigas leis federais de direitos civis, mas também minam nossa unidade nacional, pois negam, desacreditam e minam os valores americanos tradicionais de trabalho duro, excelência e realização individual em favor de um sistema de espólios baseado em identidade ilegal, corrosivo e pernicioso. Os americanos trabalhadores que

[23]O trecho refere-se à derrubada, por Trump, de uma ordem executiva de 1965, que impedia a discriminação por cor, raça, sexo e nacionalidade em contratações do governo federal.

> merecem uma chance no Sonho Americano não devem ser estigmatizados, humilhados ou excluídos de oportunidades por causa de sua raça, ou sexo (Bonfim, 2025).

Nesse sentido, é imperioso reforçar que empresas de tecnologia como Meta e OpenAI, que se destacam no cenário global de desenvolvimento de IA generativa, estavam alinhadas, à época, às políticas de Trump. Isso porque não só adotaram algumas diretrizes do governo, como a redução de iniciativas de diversidade e inclusão nos seus ambientes corporativos[24], como também buscaram apoio político, para contornar políticas que beneficiassem seus negócios ao redor do mundo[25], e apoio financeiro, para firmar os Estados Unidos como potência hegemônica em matéria de IA generativa[26].

Em síntese, e com extrema preocupação, esses são os movimentos de gigantes da tecnologia que ascendem exponencialmente, em âmbito global, no desenvolvimento da IA. Ressaltam-se movimentos amplamente distantes da relevância de se promover um desenvolvimento ético e não discriminatório dessa tecnologia. Reforça-se, por fim, que essas são as empresas que estão construindo os modelos generativos que estão sendo utilizados para subsidiar a construção de diversos sistemas ao redor do mundo.

O efeito da competição pela liderança da evolução da IA generativa acaba não só se refletindo na alteração do mercado financeiro[27], mas também na sustentabilidade global em face dos recursos necessários para manter o crescimento desenfreado dessa tecnologia.

No estudo intitulado "*The Environmental Impacts of AI and Digital Technologies*", Naeeni e Nouhi (2023) coletaram dados de entrevistas com 27 profissionais do setor de tecnologia, de pesquisas ambientais, do legislativo (no texto original, "*policy-makers*") e da academia, com o objetivo principal de entender os impactos ambientais da IA e das tecnologias digitais. Através dele, os autores ressaltam que, apesar do potencial da IA para promover o

---

[24] No decreto assinado por Trump a respeito da extinção das políticas de diversidade e inclusão, o presidente incentiva o setor privado a seguir o exemplo do governo federal, que foi prontamente adotado pela Meta e pelo McDonalds, por exemplo (O Globo, 2025).

[25] "Meta promete se aliar a Trump contra países que regulam redes sociais" (León, 2025).

[26] "Na véspera da posse de Donald Trump, Sam Altman fez uma ligação estratégica para o presidente eleito. Em um telefonema de 25 minutos, o CEO da OpenAI apresentou uma visão ambiciosa: os Estados Unidos estavam prestes a alcançar a inteligência artificial no nível humano, e isso aconteceria ainda no governo Trump. Mas para garantir essa liderança e superar a China, seria necessário um investimento massivo em infraestrutura. Altman então revelou o Stargate, um projeto de US$ 100 bilhões em data centers espalhados pelo país, financiado por Oracle, SoftBank e investidores do Oriente Médio. Trump, sempre atraído por grandes promessas e projetos de impacto, se entusiasmou. No dia seguinte à posse, anunciou o Stargate na Casa Branca como o 'maior projeto de IA da história', com Altman ao seu lado. A cena marcava não apenas a oficialização do acordo, mas também um feito inesperado: o CEO da OpenAI havia superado Elon Musk, que investira pesadamente na campanha de Trump e tentava controlar a agenda do novo governo" (Lopes, 2025).

[27] Com a alta do DeepSeek, a empresa norte-americana Nvidia, que faz chips potentes e de alto custo para o mercado de inteligência artificial, registrou uma perda de quase US$ 600 bilhões dias após o lançamento do sistema pelas empresas chinesas (Financial Times, 2025). O empreendimento chinês afirma ter conseguido treinar seu modelo de IA generativa a um custo infinitamente menor do que a concorrente norte-americana, mesmo com as restrições de semicondutores impostas a ela pelos Estados Unidos. Na mesma data, em 27 de janeiro de 2025, o índice Nasdaq, na Bolsa de Valores de Nova York, que é focado em empresas de tecnologia, teve uma queda de 3% (Lara, 2025).

desenvolvimento sustentável, é fundamental gerenciar sua implementação para mitigar os impactos negativos no meio ambiente, como a pegada de carbono e o uso excessivo de recursos. Isso porque envolve etapas complexas e recursos significativos, especialmente o treinamento dos grandes modelos de linguagem (LLMs), que consomem vastas quantidades de energia elétrica, grande parte dela proveniente de fontes não renováveis, como carvão e gás natural (Strubell *et al*., 2019).

A pesquisa também aponta uma mudança crescente nas discussões sobre IA, que historicamente se concentraram em questões éticas, para uma abordagem que inclui a sustentabilidade ambiental. A proposta de "*greening AI*" sugere a integração de considerações sustentáveis no desenvolvimento e na aplicação da IA. Strubell *et al*. (2019), exemplificativamente, apontam de forma analógica que, para o treinamento desses modelos, a emissão de gases de efeito estufa no meio ambiente pode ser comparada à de um carro em toda a sua vida útil, incluindo a fabricação do veículo e o combustível consumido.

Na mesma linha, um outro estudo realizado por Luccioni *et al*. (2024), conduzido a partir de uma análise sistemática sobre o consumo de energia e as emissões de carbono durante a fase de inferência de modelos de aprendizado de máquina, comparando modelos específicos para tarefas e modelos multiuso (IA de propósito geral), aponta que arquiteturas de IA generativa, especialmente aquelas multiuso, podem ser significativamente mais caras em termos de energia do que sistemas específicos para tarefas. Em particular, o estudo aponta que essas arquiteturas podem consumir até 33 vezes mais energia para realizar inferências em comparação com sistemas específicos para tarefas, mesmo quando o número de parâmetros do modelo é controlado.

Smith *et al*. (2022), em última análise, no artigo em que discutem o impacto ambiental das arquiteturas de IA generativa, reforçam que o consumo energético dos data centers é responsável por, aproximadamente, 2% do consumo global de energia, o que tende a aumentar a procura por computação em nuvem e armazenamento de dados para o treinamento dessa tecnologia. Comprovadamente, entre 2017 e 2021, a eletricidade usada pela Meta, Amazon, Microsoft e Google, principais provedoras de computação em nuvem, mais que dobrou (IEA, 2023 *apud* Luccioni *et al*., 2024).

Isso destaca um importante *trade-off* entre a utilidade dos modelos multiuso e seus custos aumentados em termos de energia e emissões de carbono. Nos Estados Unidos, país cujas empresas desenvolvedoras e fornecedoras de modelos de IA são as mais globalmente emblemáticas do ponto de vista comercial, a Constituição não garante explicitamente um direito ao meio ambiente saudável[28]. Ainda que se vislumbre uma interpretação crescente de que certas

[28] No Brasil, o artigo 225, da CRFB, estabelece que todos têm direito a um meio ambiente ecologicamente equilibrado, que é um bem de uso comum do povo e essencial para uma qualidade de vida saudável, impondo ao Poder Público e

cláusulas constitucionais podem ser aplicadas para proteger interesses ambientais, a exemplo da *Due Process Clause*, da 14ª Emenda[29], a regulamentação de IA americana, até o momento, tem sido voltada para a segurança e a liderança nacional no setor, abordagem que, espera-se, se mantenha e até se intensifique sob o governo Trump (López, 2025).

Em 13 de janeiro de 2025, o governo dos Estados Unidos publicou um documento com diretrizes e restrições a respeito do uso da IA, com o objetivo de coibir o advento da tecnologia na China (López, 2025). O contexto se dá justamente no período de rumores sobre o lançamento do DeepSeek no mercado de tecnologia, com o próprio governo reforçando que a medida tem em vista fortalecer a segurança e a força econômica nacional, direcionando uma não terceirização da tecnologia.

O fortalecimento das políticas de IA nos Estados Unidos, país líder em inovações tecnológicas, pode ditar a maneira de consumo do restante do globo. O Brasil, como já reforçado anteriormente, é dependente dos LLMs norte-americanos para o desenvolvimento dos seus sistemas de IA, como também da importação de semicondutores para a arquitetura dos seus modelos.

Crowford (2021) também reforça a questão da própria exploração humana em virtude da mineração de dados para a inteligência artificial, tendo em vista que a extração de materiais necessários para a construção da inteligência artificial depende, sobremaneira, da exploração da mão de obra minerária em locais como Nevada, Bolívia, Congo, Mongólia, Indonésia e Austrália, reforçando a violência local e geopolítica associada à extração desses minerais. A autora chama atenção, também, ao ciclo de obsolescência dos dispositivos eletrônicos, ressaltando o impacto do descarte inadequado de e-lixo em locais como Gana e Paquistão (Crowford, 2021).

A título de ilustração, Crowford (2021) destaca a história da cidade de Blair, que foi criada em 1906 devido à atividade de mineração de ouro e prata, e abandonada 12 anos depois devido à contaminação por cianeto e ao esgotamento das reservas minerais. Em sentido análogo, infere que Silver Peak, cidade norte-americana que vem sofrendo com a exploração de minas de lítio, matéria-prima essencial para a construção da IA, também se tornará uma cidade fantasma em breve, em virtude da alta demanda pelo mineral e do impacto ambiental da extração desenfreada.

Nessa medida, para além de uma regulamentação da IA voltada à proteção do indivíduo em relação ao seu uso, se faz notória uma análise precursora da exploração da pessoa humana

---

à coletividade o dever de defendê-lo e preservá-lo para as presentes e futuras gerações. Contudo, no cenário atual, mesmo que o Brasil tenha uma presença significativa no uso de IA generativa (em estudo recente, 54% dos brasileiros afirmaram terem usado sistemas de IA generativa em 2024 (ROSA, 2025)), ainda é dependente da importação de modelos desse tipo de tecnologia para a construção dos seus sistemas.

[29] Que pode ser invocada para argumentar que um meio ambiente saudável é essencial para a vida, liberdade e busca da felicidade (Cometti, 2024)

em detrimento da própria formação da tecnologia, que implica não só na exploração do trabalho de grupos essencialmente minoritários, mas também no esgotamento de recursos naturais essenciais para a continuidade da vida humana. É necessário, assim, que o legislativo também seja capaz de coibir a exploração humana dentro da cadeia de suprimentos dessa tecnologia. Entretanto, ao que parece, esse tipo de tratativa ainda é tímida dentro das legislações que vêm sendo construídas, de maneira global, para a regulamentação da IA.

O AI Act, em seu Considerando 48, menciona o direito fundamental a um elevado nível de proteção ambiental em virtude da avaliação da gravidade dos danos que um sistema de IA pode causar em relação à saúde e à segurança das pessoas. Trata, aqui, o texto de forma remediativa, cuidando dos impactos já causados, sem explicitamente coibir toda a cadeia ambiental exploratória no estágio inicial da tecnologia. Do mesmo modo, o Considerando 176 reforça que o objetivo da lei está centrado na melhoria do funcionamento do mercado interno para promover a adoção de uma IA centrada no ser humano, garantindo um elevado nível de proteção da saúde e do ambiente, bem como dos direitos fundamentais. Novamente, o legislador europeu centra sua preocupação na proteção humana quando o produto já está no mercado.

Por fim, o PL nº 2.338/2023, construção da lei a que se atenta o presente estudo, cita o incentivo, em instituições de ensino, à inclusão de disciplinas voltadas ao impacto ambiental e à sustentabilidade no desenvolvimento e na operação de sistemas de IA (Brasil, 2024, art. 70, V). Prevê, também, o fomento à pesquisa e ao desenvolvimento de programas de certificação para redução do impacto ambiental de sistemas de IA (Brasil, 2024, art. 61). Veja-se que o texto intenta regular o impacto ambiental quanto ao desenvolvimento do sistema, sem que, contudo, seja capaz de coibir o impacto na construção do modelo. Até porque, como já reforçado, o Brasil se mantém na posição majoritária de importador desses modelos generativos.

3.3 Potencialidade lesiva aos direitos fundamentais a partir do uso de modelos generativos

No primeiro subcapítulo do presente tópico, foi abordado a conceitualidade técnica dos modelos e dos sistemas de IA generativa, com o objetivo de trazer mais clareza às discussões que viriam a seguir.

O segundo subcapítulo, por sua vez, teve o propósito de demonstrar a ascensão da IA generativa sob o prisma de uma intensa competição global por liderança tecnológica, sobretudo entre os Estados Unidos e a China, o que culmina em um mercado dominado por grandes empresas de tecnologia. É imperioso que essa monopolização levante preocupações não só sobre o controle da informação e da privacidade, como também das implicações ambientais, uma vez que o treinamento desses modelos, cada vez mais intensificados com vista a manter a hegemonia tecnológica, tem consumido exageradamente e esgotado grandes quantidades de

energia e recursos naturais. Ademais, tal exploração acaba por violar, consequentemente, direitos fundamentais da pessoa humana, que deve ter um meio ambiente sustentável assegurado, prezando pela qualidade de vida e preservação das presentes e futuras gerações.

Em razão disso, e em continuidade à contextualização até aqui realizada, este subcapítulo terá a finalidade de concluir o estudo dos modelos generativos, a partir do potencial lesivo ocasionado ao indivíduo com sua disponibilização ao mercado. Chega-se, aqui, à abordagem que tem determinado a construção das legislações de IA de maneira global: a centralidade da pessoa humana face ao desenvolvimento, fomento e uso não só da IA de propósito geral, mas também da IA de maneira ampla.

A IA de propósito geral lança-se ao mercado com foco em facilitar o desenvolvimento de atividades pelos seres humanos ao integrá-los a esse tipo de tecnologia. Esse é o discurso que fomenta o desenvolvimento acelerado dessa tecnologia, ainda que, como já demonstrado, os interesses econômicos hegemônicos das grandes potências tenham enviesado o posicionamento do ser humano no centro desses avanços tecnológicos.

Barroso e Mello (2024) reforçam que é preciso ter atenção aos efeitos adversos da IA, procurando neutralizá-los ou mitigá-los, na medida em que podem implicar negativamente em impactos sociais, econômicos, políticos e, até mesmo, na paz mundial. Em seu artigo, intitulado "Inteligência artificial: promessas, riscos e regulação. Algo de novo debaixo do sol", os autores abordam alguns riscos e efeitos que entendem como desfavoráveis aos indivíduos e aos seus direitos fundamentais.

Em primeira análise, Barroso e Mello (2024) discutem o impacto da IA e da automação no mercado de trabalho, destacando que a introdução de novas tecnologias altera profundamente a paisagem laboral, exigindo que trabalhadores de diversas áreas se adaptem a novas funções. Essa transição pode ser desafiadora, afetando não apenas empregos mecânicos, como também funções qualificadas e criativas.

Embora novas tecnologias possam gerar novos mercados e empregos, há um descompasso entre a criação de novas oportunidades e a eliminação de postos existentes, o que representa um desafio significativo, sendo necessário que os governos invistam em proteção social e capacitação profissional, uma vez que a vulnerabilidade econômica pode ameaçar a estabilidade democrática (Barroso; Mello, 2024).

De acordo com um levantamento recente da McKinsey (2023), a IA generativa tem o potencial de mudar a anatomia do trabalho, automatizando atividades que ocupam de 60 a 70% do tempo dos trabalhadores hoje, permitindo um crescimento da produtividade do trabalho de 0,1 a 0,6% ao ano até 2040. No contexto da IA generativa, diferentemente do que ocorre com IAs de modo geral, o rápido avanço do potencial de automação se deve, em grande parte, à sua maior capacidade de entender a linguagem natural, que é essencial para atividades que

respondem a 25% do tempo total de trabalho, fazendo-se propício o maior impacto em ocupações com salários e requisitos educacionais mais elevados (McKinsey, 2023).

Com igualdade, um estudo científico recente, intitulado "*GPT are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models*", redigido, em conjunto, por pesquisadores da OpenAI, Open Research e University of Pennsylvania, retrata potenciais implicações de modelos generativos no mercado de trabalho norte-americano. Segundo a avaliação, cerca de 80% da força de trabalho estadunidense terá, ao menos, 10% de suas rotinas de trabalho alteradas pela IA, além de, possivelmente, um em cada cinco trabalhadores ter metade da sua rotina impactada por essa tecnologia (Eloudoun *et al*., 2023). À visto disso, e em consonância ao estudo da McKinsey, esse artigo também aponta uma maior exposição de ocupações com salários mais elevados aos impactos dos GPTs.

Peixoto e Bonat (2023) pressupõem que será possível presenciar uma maior probabilidade de impactos de GPTs em trabalhos jurídicos, sobretudo naqueles que demandam maior tempo para o desenvolvimento de habilidades. De acordo com a classificação adotada pelos autores[30], 42,2% das atividades jurídicas altamente impactadas por sensibilidade aos dados não possuem atenuantes de impacto[31], a exemplo de atividades como despachar, preparar informações, assessorar, apurar liquidez e certeza de dívidas, editar enunciados etc. (Peixoto; Bonat, 2023). A área jurídica, enquanto geradora de uma enorme quantidade de dados não estruturados e demandante de apoio constante, tem aberto espaço significativo para o uso de soluções de IA. Entretanto, atividades que envolvam conhecimentos científicos sofisticados ou que utilizem predominantemente análises críticas não estariam contempladas nesse rol (Peixoto; Bonat, 2023).

Segundo a pesquisa da Thomson Reuters Institute[32], escritórios de advocacia têm inclinações positivas para o uso da IA generativa na prestação do serviço jurídico, ainda que de forma tímida. Embora 82% dos entrevistados tenham sinalizado positivamente pela aplicação da tecnologia ao trabalho jurídico, apenas 3% afirmaram estar utilizando a tecnologia em seus escritórios (Thomson Reuters, 2023).

Em fevereiro de 2023, sendo um dos precursores na área jurídica, o escritório de advocacia britânico Allen & Overy Shearman anunciou a adoção do seu sistema de IA

[30] Os autores classificam as atividades inerentes à prática jurídica de acordo com o índice à sensibilidade e composição de dados de uma determinada tarefa. "Assim, por exemplo, baixa exposição/dependência/utilização a/ de dados de uma determinada tarefa receberá o índice [0], uma média exposição [1] e uma alta, [2]" (Peixoto; Bonat, 2023).

[31] Os autores também aplicam um índice para medir as atenuantes de impacto nas atividades impactadas por sensibilidade aos dados: "a) concentrada aplicação de conhecimentos científicos, b) elevada aplicação de abordagem crítica, atribuindo-se [-1] para cada incidência. Aqui, não havendo atenuante, será aplicado índice [0]" (Peixoto; Bonat, 2023).

[32] A pesquisa foi realizada com 433 profissionais de escritórios de advocacia de médio e grande porte, juntamente com escritórios de advocacia membros do painel da Thomson Reuters Influencer Coalition, localizados nos Estados Unidos, Reino Unido e Canadá, entre 21 e 31 de março de 2023.

generativa, ContractMatrix, desenvolvido através do modelo generativo da Counsel AI (A&O Shearman, 2023). Essa solução intenta auxiliar em pesquisas jurídicas, a partir da base de conhecimento interna do escritório, bem como na redação e revisão de contratos (A&O Shearman, 2023). A banca iniciou os testes em novembro de 2022, já tendo implementado o sistema para mais de 3.500 advogados, em seus 43 escritórios (A&O Shearman, 2023).

A elaboração, a revisão e o resumo de documentos, a tradução, a correção ortográfica, a pesquisa jurídica e a geração de insights para tomada de decisão estratégica são algumas das aplicabilidades da IA generativa cabíveis às atividades jurídicas. Com efeito, o que se tem presenciado é que a aplicação dessa tecnologia à prática do Direito não só otimiza o desempenho das atividades, como também culmina em significativa redução de custo[33], principalmente em rotinas que demandam um elevado investimento de tempo e uso não efetivo do intelectual técnico exigido pela profissão.

A massificação da desinformação é outro efeito negativo do mau uso da IA, na medida em que a disseminação de informações falsas por plataformas digitais e aplicativos de mensagens já vem se tornando um problema grave para a democracia e os processos eleitorais desde 2016 (Barroso; Mello, 2024), ano marcante para a observação do fenômeno em virtude da eleição presidencial nos Estados Unidos. Durante essa eleição, houve um aumento notável na disseminação de notícias falsas e desinformação nas redes sociais, muitas vezes amplificadas por *bots* e modelos de IA, demonstrando o potencial dessa tecnologia em influenciar a opinião pública. De igual modo, a desinformação também esteve diretamente ligada ao rompimento do Reino Unido com a União Europeia (Brexit) (Barroso; Mello, 2024), na medida em que a amplificação da inveracidade por meio da IA moldou a opinião pública com informações exageradas ou falsas sobre os benefícios econômicos com a saída da UE.

Nessa mesma época, presenciou-se, também, o surgimento das *deepfakes*, que utilizam IA para criar vídeos falsos realistas, evoluindo rápida e significativamente, tornando-se mais sofisticadas e difíceis de detectar. A eleição presidencial brasileira de 2018 é um exemplo evidente de desinformação e de *deepfakes*, com o agravante de ter sido exponencialmente ampliada através do compartilhamento instantâneo de mensagens via aplicativo WhatsApp. Fake news, robôs e contas automatizadas e *microtargeting* (técnica que utiliza dados pessoais para direcionar anúncios políticos específicos a grupos de eleitores[34]) contribuíram

[33] Com a utilização de seu sistema de IA, A&O Shearman estimou uma economia de até sete horas nas negociações contratuais (A&O Shearman, 2023).

[34] Farage e Mulholland (2022), em artigo que analisa a relação entre a democracia e as tecnologias emergentes, como IA, *big data* e *machine learning*, destacam como essas ferramentas influenciam a participação política. Dois conceitos centrais são discutidos. O primeiro deles, o "filtro bolha", que limita a exposição a informações diversas, os autores argumentam que a aplicação de algoritmos na política resulta na customização da democracia, semelhante a estratégias de *neuromarketing*, o que pode criar bolhas informativas que isolam os cidadãos de opiniões divergentes. O segundo conceito, "*Big Nudging*", utiliza *big data* para moldar comportamentos e políticas públicas, combinando *nudging* com *big data* para influenciar decisões de forma sutil. Exemplos de políticas públicas no Reino Unido mostram como

significativamente não só para influenciar o resultado das eleições em virtude da intervenção desinformativa na opção de escolha das pessoas, como também polarizaram politicamente o país (Ruediger, 2019).

Cabe reforçar a manipulação de conteúdo audiovisual ultrarrealista por meio da IA generativa, que permite a criação de imagens ou vídeos sintéticos que são capazes de burlar verificações de identidade, além da sua utilização para a prática de crimes contra a honra, por meio da adulteração de registros.

A violação da privacidade também é um efeito abordado por Barroso e Mello (2024). Citando o brilhante trabalho de Morozov (2018, p. 36), os autores reforçam que "o modelo de negócio das plataformas que se valem da IA se baseia na coleta da maior quantidade possível de dados pessoais dos indivíduos, o que transforma a privacidade em mercadoria" (Barroso; Mello, 2024, p. 21). Em sua obra, a principal crítica de Morozov (2018) ao que chama de "solucionismo" tecnológico é o fato de a tecnologia passar a ser empregada com o subterfúgio de resolução de problemas que instituições falharam em resolver. Nesse sentido, plataformas tecnológicas baseadas em dados pessoais (como Airbnb, Uber e WhatsApp), para além do que se costuma acreditar em termos de facilitadores do dia a dia humano, podem, na verdade, fazerem-se contrárias à democracia (Morozov, 2018).

O *input* abrangente e acelerado de dados pessoais em sistemas de IA tem proporcionado às grandes empresas de tecnologia uma perfilização humana cada vez mais assertiva. Comportamento de consumo, posição política, sexualidade, religião, vulnerabilidades, para citar algumas questões intimistas, estão se tornando ativos econômicos valiosos no mercado digital. Isso porque a IA, através do acesso aos dados privados, "é capaz de realizar predições, recomendações, manipular interesses e produzir os resultados almejados pelo algoritmo" (Zuboff, 2022, pp. 1-79). Aqui vê-se a centralidade da pessoa humana no desenvolvimento da IA, entretanto, em cenário antagônico ao modelo fomentado pelas legislações em construção, tal como o PL nº 2.338/2023.

Para Barroso e Mello (2024), existem três aspectos, relativos à privacidade, que exigem atenção, sendo eles a ausência de consentimento[35] de usuários para obtenção de seus dados na internet; a vigilância e o rastreamento pelo governo e por autoridades policiais, mediante tecnologias de reconhecimento facial e ferramentas de localização; e, por último, o risco de

nudges podem ser aplicados para promover comportamentos desejáveis, mas o texto também destaca os riscos de manipulação da autonomia individual. A análise de dados em tempo real e o uso de *bots* são mencionados como ferramentas de manipulação política, que podem disseminar informações verdadeiras e falsas.

[35] A legislação brasileira, especialmente a LGPD, define o consentimento como uma manifestação livre, informada e inequívoca, mas existem desafios significativos para garantir que essa permissão seja realmente informada e não coercitiva (Korkmaz, 2019). A autora, ademais, a partir do entendimento de Rodotà (2008), reforça a impossibilidade de o consentimento ser manejado em todas as situações, sobretudo em relação aos dados de proteção dos valores individuais, devendo ser postos para concessão "todos aqueles que poderiam sofrer uma 'perda de dignidade' ou de autonomia através do seu consentimento para a coleta, tratamento e difusão das informações" (Rodotà, 2008).

vazamento e os ataques cibernéticos por atores maliciosos, que não raro alimentam práticas de assédio, violência política, *malinformation* e desinformação, em virtude da significativa quantidade de dados necessários para o treinamento do modelo.

A complexidade do consentimento, no contexto da privacidade em sentido amplo e da proteção de dados em sentido estrito, merece especial atenção em virtude da sua importância como um instrumento de autonomia e autodeterminação do indivíduo. O consentimento é visto como essencial para a utilização de dados pessoais, mas sua aplicação pode ser problemática, especialmente quando se considera a comoditização dos dados e a possibilidade de manipulação por interesses de mercado (Korkmaz, 2019). Isso porque a mineração dos dados pessoais tem ensejado sua venda para direcionamento de informações e publicidades, bem como para a manipulação da vontade dos usuários (Barroso; Mello, 2024). No entendimento de Korkmaz (2019, n.p.), segundo Doneda (2006):

> A utilização do consentimento pode ser "instrumentalizada pelos interesses que pretendem que a sua disciplina não seja mais que uma via para legitimar a inserção dos dados pessoais no mercado" (Doneda, 2006, p. 375). Com o verniz de juridicidade, o consentimento poderia, em última análise, legitimar uma apropriação do corpo eletrônico da pessoa por parte do mercado com diversas repercussões em desfavor da própria pessoa e da coletividade. A partir de Rodotà, Doneda (2006) evidencia que uma falsa premissa de conceder o consentimento como instrumento para determinar livremente a utilização dos dados pessoais poderia, por parte do Estado, representar um "falso álibi" para não interferir na situação que, em realidade, demandaria a sua atuação positiva na defesa de direitos fundamentais.

Ainda, em seu trabalho de 2022, Korkmaz destaca as limitações e os desafios específicos do consentimento no contexto da inteligência artificial. Mesmo que esse conceito tenha sido fundamental para a construção das legislações de proteção de dados, Korkmaz (2022) entende que ele, por si só, não é suficiente para garantir a proteção efetiva dos titulares de dados diante das novas tecnologias, especialmente da IA.

No contexto dessa tecnologia, para a autora, o consentimento enfrenta alguns desafios adicionais, os quais passam-se a analisar.

Quanto à assimetria informacional: a autora destaca a profunda assimetria entre os agentes de tratamento (empresas, plataformas e governos) e os titulares de dados, os quais não têm condições reais de compreender o alcance do tratamento de seus dados, especialmente quando envolvem sistemas automatizados e algoritmos complexos (Korkmaz, 2022, pp. 41-43).

O consentimento, ademais, pode dar ao titular uma falsa impressão de controle ou propriedade sobre seus dados, quando, na prática, ele não tem meios de entender ou controlar a forma como serão processados por sistemas de IA (Korkmaz, 2022, p. 42). Por isso, entende-se o desafio do consentimento enquanto falsa garantia.

O consentimento, pela ótica da discriminação algorítmica, não impede que sistemas de

IA produzam resultados discriminatórios ou violem outros direitos fundamentais[36], especialmente quando se trata de dados sensíveis ou de decisões automatizadas com impacto significativo, mesmo que o titular consinta com o uso de seus dados (Korkmaz, 2022, pp. 47-63).

Por fim, a autora observa que, no cenário do *big data* e da IA, o tratamento de dados muitas vezes ocorre de forma massiva, agregada e para finalidades não previstas originalmente, tornando o consentimento inicial insuficiente ou até irrelevante para proteger o titular (Korkmaz, 2022, pp. 48-57).

Dessa forma, o que Korkmaz (2022) sistematiza e enfatiza é que a proteção de dados pessoais, especialmente diante da IA, não pode depender exclusivamente do consentimento do titular[37], prezando por uma abordagem mais ampla, que combine consentimento, princípios normativos, controles institucionais e mecanismos coletivos de proteção, para garantir efetivamente os direitos dos titulares de dados em um cenário de decisões automatizadas e uso intensivo de IA.

Ademais, analisando a proteção de dados a partir da sua posição enquanto direito fundamental, impõe-se ao Estado e aos agentes de tratamento deveres de proteção ativa, fiscalização e controle, inclusive para garantir a igualdade material e prevenir discriminações.

À vista disso, a análise do consentimento deve ser contextualizada dentro do paradigma da proteção da personalidade, não da propriedade (Rodotà, 2008), sobretudo quando o titular do dado se encontra em uma posição vulnerável (Korkmaz, 2019).

Conte (2023), em extensão, reforça o contexto da patrimonialização dos dados pessoais em relação ao mercado, sob o prisma de que as novas tecnologias impactam a produção, a circulação e o consumo de informações, fazendo com que a produção descentralizada de informações e o caráter aberto e global dos fluxos de comunicação caracterizem a nova era da informação.

Nessa medida, as novas tecnologias têm impacto não apenas na esfera privada física, mas também na dimensão imaterial das informações que contribuem para constituí-la, encontrando-se os dados pessoais no cruzamento entre sujeito e objeto: por um lado são atraídos para a esfera subjetiva, constituindo uma representação da pessoa, descrevendo seu

---

[36] A violação e restrição aos direitos (sobretudo os direitos fundamentais) estão diretamente relacionadas a "danos físicos, reputacionais, relacionais, psicológicos (emocionais), econômicos, discriminatórios e relacionados à autonomia humana (coerção, manipulação, desinformação, deformação de expectativas, perda de controle entre outros)" (Citron; Solove, 2022, em especial p. 831; Huq, 2020, em especial pp. 35- 41 *apud* Barroso; Mello, 2024, p. 22).

[37] Inclusive, a autora aponta que as legislações modernas, como a LGPD e o GDPR, já preveem hipóteses de tratamento de dados independentemente do consentimento, desde que fundamentadas em bases legais e princípios normativos: "Nessa direção, verificam-se, em regulações de proteção de dados pessoais, hipóteses que autorizam o tratamento de dados independentemente da vontade de seu titular, podendo mesmo ser, contrariamente àquela, desde que lastreado nas balizas normativas, em sentido amplo, pertinentes, superando-se, portanto, um paradigma de equivalência entre sigilo e proteção" (Korkmaz, 2022, p. 47).

comportamento e relacionamentos; e, por outro, são objetos de uso por parte de atores públicos e privados (Conte, 2023).

Com relação à discriminação algorítmica, segundo Barroso e Mello (2024), esta é entendida pelo fato de os algoritmos tenderem a reproduzir estruturas sociais atuais e pretéritas de inclusão e exclusão, repletas de vieses[38] e preconceitos em razão de circunstâncias históricas, culturais e sociais.

Já para Frazão (2021), em seu artigo sobre discriminação algorítmica, analisando-a a partir do crescente uso de algoritmos em processos decisórios que afetam diretamente a vida das pessoas, a autora enfatiza que os julgamentos algorítmicos não são neutros. Eles são implementados com objetivos específicos, podendo negar acesso a serviços, oportunidades ou direitos com base em critérios muitas vezes opacos[39] e questionáveis[40].

---

[38] A bem da verdade, não se fala em "viés" enquanto uma "distorção" algorítmica. No entendimento de Crowford (2021), os conjuntos de dados de treinamento da inteligência artificial apenas naturalizam hierarquias e ampliam as desigualdades que já existem na sociedade, ressaltando que a tendência atual de focar apenas na questão do viés na IA acaba por afastar a avaliação das principais práticas de classificação em IA, juntamente com suas políticas concomitantes. Nesse sentido, as discussões sobre preconceitos na IA muitas vezes se limitam à busca por paridade matemática para produzir "sistemas mais justos", sem competir com as próprias estruturas sociais subjacentes (Crowford, 2021).

[39] Para Burrell (2016, p. 1, traduzido pela autora), os algoritmos são "opacos no sentido de que, se alguém é um receptor da saída do algoritmo (a decisão de classificação), raramente se tem qualquer noção concreta de como ou por que uma determinada classificação foi obtida a partir de entradas". Ou seja, a opacidade está intimamente vinculada à ausência de transparência quanto ao treinamento e à classificação do modelo.

Ao abordar a questão da opacidade como um problema para mecanismos de classificação e decisões socialmente consequentes, Burrell (2016) menciona três formas distintas: como segredo corporativo ou estatal intencional, como analfabetismo técnico e como uma característica dos algoritmos de aprendizado de máquina e da escala necessária para aplicá-los de forma útil, destacando a importância de reconhecer as formas distintas de opacidade para determinar soluções técnicas e não técnicas que possam ajudar a prevenir danos.

A opacidade como segredo corporativo ou estatal intencional está em linha com o que foi analisado no Capítulo 3, a respeito da hegemonia de grandes empresas de tecnologia dominando o desenvolvimento, a evolução e a distribuição de modelos de IA de maneira global. Essa forma de opacidade refere-se à proteção deliberada de informações por empresas ou instituições governamentais, com o intuito de preservar suas vantagens competitivas e proteger suas inovações (Burrell, 2016).

A opacidade como analfabetismo técnico surge da complexidade técnica dos algoritmos e da programação necessária para sua implementação, cuja habilidade de escrever e ler o código mostra-se como uma competência especializada que não está amplamente disponível à população em geral (Burrell, 2016) que, ao fim, é a destinatária e usuária dessas tecnologias. Como as pessoas não entendem o funcionamento dos algoritmos, sua capacidade de questionar ou contestar as decisões tomadas por esses modelos e sistemas acaba sendo limitada.

Por fim, a opacidade como uma característica dos algoritmos de aprendizado de máquina está relacionada à discrepância entre os procedimentos matemáticos utilizados nos algoritmos e à maneira como os seres humanos interpretam informações semanticamente. Os algoritmos de aprendizado de máquina operam em alta dimensionalidade e utilizam otimizações matemáticas que podem não se alinhar com o raciocínio humano tradicional, gerando dificuldades na compreensão dos resultados produzidos pelos algoritmos, pois as decisões podem parecer arbitrárias ou incompreensíveis para os usuários (Burrell, 2016).

A relevância dos mecanismos de classificação está intimamente atrelada a questões de desigualdade econômica e mobilidade social. Burrell (2016) entende que a mudança na apresentação pública do termo "algoritmo" é hoje influenciada por narrativas midiáticas e esforços de branding corporativo. Nesse sentido, infere-se que a transparência e a interpretabilidade dos algoritmos, expondo a lógica por trás das classificações, refuta o subjetivismo da existência de uma personalidade digital aos agentes autônomos.

Do mesmo modo, discorre Brostom (2017 *apud* Beckers; Teubner, 2021, p. 2) sobre os riscos de exposição dos seres humanos a um ambiente algorítmico opaco, elucidando que o algoritmo individual pode até satisfazer o objetivo estabelecido pelo participante humano, mas pode ser que escolha um meio que viole as intenções do humano.

[40] *Ipsis litteris*: "Alimentados por bases de dados cada vez maiores, no atual contexto do *big data*, os algoritmos têm sido utilizados para decisões e tarefas que envolvem análises qualitativas e subjetivas, comumente marcadas por alta carga valorativa, tal como acontece nos julgamentos para classificação, ranqueamento e criação de perfis das pessoas. Mais do que isso, os algoritmos têm sido vistos como chaves para se compreender o passado, se diagnosticar o presente

Frisa a autora que a questão da discriminação algorítmica vai além do mercado, atingindo dimensões fundamentais da autonomia privada e pública dos cidadãos. Exemplos incluem decisões sobre quem ingressa em universidades, quem obtém empregos ou, até mesmo, quem é poupado em situações de risco por carros autônomos. No âmbito estatal, Frazão (2021) chama atenção para o uso de algoritmos em classificações políticas, o que pode levar a perseguições e à manipulação de processos democráticos.

Kaufman (2022, p. 114), adicionalmente, enfatiza a discriminação algorítmica a partir do gênero:

> Como a técnica de inteligência artificial que permeia a maior parte das aplicações atuais é baseada em dados (*machine learning/deep learning*), a sociedade está tomando decisões enviesadas por gênero em número maior que o percebido. Na Inglaterra, por exemplo, as mulheres têm 50% mais chances de serem diagnosticadas erroneamente após um ataque cardíaco, em função da predominância dos homens nos estudos científicos sobre insuficiência cardíaca.

O entendimento central de Ana Frazão é que, apesar dos avanços e benefícios dos julgamentos algorítmicos, é imprescindível discutir a validade e os limites éticos e jurídicos desses julgamentos. Ela questiona se é legítimo submeter seres humanos a qualquer tipo de julgamento algorítmico apenas porque a tecnologia permite, defendendo que devem existir critérios claros para distinguir julgamentos legítimos daqueles que não o são.

Assim, para Frazão (2021), a discussão sobre discriminação algorítmica deve partir da análise da própria legitimidade desses julgamentos, e não apenas de suas eventuais falhas ou acertos, ressaltando a necessidade de reflexão ética e jurídica antes da adoção indiscriminada desses sistemas.

Silva (2022), ademais, aborda a discriminação algorítmica como uma atualização do racismo estrutural, argumentando que os algoritmos, ao serem desenvolvidos e implementados sem uma análise crítica, podem perpetuar e até intensificar desigualdades existentes. Ele destaca, por exemplo, que tecnologias como reconhecimento facial, filtros de *selfies*, moderação de conteúdo e policiamento preditivo frequentemente apresentam vieses que desfavorecem minorias raciais.

Assim como Frazão (2021), Silva (2022) também afirma que a crença na neutralidade das tecnologias digitais é equivocada, tendo em vista que estas são alimentadas por dados históricos que refletem preconceitos sociais. Nesse sentido, os algoritmos podem reforçar hierarquias raciais e contribuir para a manutenção de privilégios de grupos hegemônicos.

De forma mais atual, fala-se, ainda, no efeito da violação dos direitos da propriedade intelectual, sobretudo dos direitos autorais. Ao que parece, o desafio central das complexas

---

e se antever o futuro, por meio de prognósticos e análises preditivas a respeito das pessoas, tanto individual como coletivamente" (Frazão, 2021).

questões de propriedade intelectual relacionadas ao uso de conteúdos protegidos por IA generativa é justamente o fato de a tecnologia “criar” obras objetivamente inéditas, o que gera um vácuo legislativo quanto à proteção dos direitos autorais dessas criações.

A legislação brasileira de direitos autorais (Lei nº 9.610/1998), fundamentada na Constituição Federal e na Convenção de Berna, exige que a obra seja uma "criação do espírito" e detenha originalidade subjetiva, requisitos tradicionalmente ligados à intervenção e criatividade humanas. Como as máquinas não possuem personalidade jurídica nem subjetividade, surge a dúvida sobre a possibilidade de proteção autoral para obras criadas de forma autônoma por IA.

Especialmente no contexto das *Big Techs*, desenvolvedoras da tecnologia generativa, a resposta que se busca é justamente a quem pertencem os direitos dos conteúdos utilizados para treinar essas IAs: se aos criadores originais ou às empresas que os exploram (Barroso; Mello, 2024).

Como já ilustrado anteriormente, os LLMs foram construídos a partir do consumo incessante de dados disponíveis on-line. Ainda que esses dados possam ser considerados públicos, o princípio fundamental das discussões sobre privacidade, no qual as informações dos indivíduos não sejam reveladas fora do contexto em que foram originalmente produzidas, deve ser utilizado como regra.

Por meio da Lei nº 9.610/1998, tem-se a compreensão da divisão entre o direito de autor e o direito patrimonial, imputado àquele que possui o direito econômico sobre a obra (Brasil, 1998, arts. 28 a 45), e o direito moral, cabível àquele que, de fato, cria a obra (Brasil, 1998, arts. 24 a 27). Ao autor da obra são conferidos tanto o direito patrimonial quanto o moral, podendo, inclusive, transferir o direito patrimonial proveniente de sua criação a outra pessoa (Brasil, 1998, arts. 49 ao 52).

Em termos de direito do autor, no âmbito dos modelos e dos sistemas generativos, a problemática, além de buscar compreender a titularidade dos direitos autorais sobre obras produzidas pela própria inteligência artificial, também busca identificar quais e como os direitos autorais acabam sendo violados no processo de desenvolvimento da própria tecnologia.

Sobre o primeiro ponto, ainda não há consenso nem regramento específico que discorra sobre o sujeito sobre o qual recaem os direitos autorais de obras geradas a partir de um modelo generativo. A doutrina está dividida em teorias que entendem a autoria cabível individualmente aos seguintes supostos atores: ao desenvolvedor (no sentido de financiador) e/ou ao programador da solução, ao usuário, à própria IA ou ao domínio público (Ehrhardt Júnior; Milhazes Neto, 2024[41]).

[41] O artigo apresenta e discute quatro principais teorias sobre a titularidade das obras criadas por IA:
1. Domínio Público: a teoria mais aceita atualmente defende que tais obras devem ser consideradas de domínio

Quanto ao segundo ponto, ações judiciais já começaram a discutir a matéria sob a ótica de a quem cabe a responsabilidade sobre a violação desses dados[42]: se às provedoras dos sistemas generativos ou às intermediadoras responsáveis pela compilação dos dados (entendido, aqui, como os modelos generativos). A principal linha de defesa de empresas provedoras dos sistemas generativos é que não existe manipulação do conjunto de dados por parte delas, apenas consumo do modelo desenvolvido por uma interventora[43]. Nessa perspectiva, a violação da propriedade intelectual ocorreria no tratamento dos dados pela empresa intermediadora (ou seja, na construção do modelo da IA), e não quando as provedoras utilizam esse conjunto de dados para arquitetar seus sistemas.

Por fim, à luz dos desafios emergentes impostos pela inteligência artificial generativa, a responsabilidade civil assume papel de salvaguarda concretizadora dos direitos fundamentais, pois a sofisticação algorítmica dos modelos, aliada à opacidade de seus processos decisórios, potencializa violações à dignidade, à privacidade, à honra e à liberdade informacional dos indivíduos, enquanto fragiliza os tradicionais mecanismos de imputação subjetiva de culpa.

Nesse cenário, este estudo pretende não apenas investigar a insuficiência normativa em coibir condutas danosas, mas também a eventual lacuna de efetiva punibilidade dos agentes que, direta ou indiretamente, participam da cadeia causal do prejuízo. Assim, o que está em análise é a lacuna de responsabilidade civil em detrimento do uso de modelos e sistemas generativos como mais um dos efeitos lesivos aos direitos fundamentais.

A partir da premissa desse déficit protetivo, este estudo vai investigar um sistema de imputação de responsabilidade proporcional que, partindo da premissa de que o usuário lesado ocupa uma posição hipossuficiente no mercado digital, imponha, de um lado, deveres de

---

público, pois não preenchem os requisitos subjetivos exigidos pela legislação brasileira, como a originalidade subjetiva e a criação do espírito;

2. Direitos ao Operador da IA: outra corrente sugere que os direitos autorais deveriam pertencer ao operador da IA, ou seja, à pessoa que insere os comandos e prompts, já que a máquina seria apenas uma ferramenta;

3. Direitos ao Desenvolvedor do Algoritmo: uma terceira posição atribui os direitos ao criador do algoritmo ou da plataforma de IA, sob o argumento de que este teria maior controle criativo sobre o resultado; e

4. Personalidade Jurídica para Máquinas: por fim, há quem defenda a criação de uma personalidade jurídica para as máquinas, permitindo que estas sejam titulares de direitos autorais. Contudo, essa solução é vista como pouco prática e ontologicamente problemática, pois a personalidade jurídica tradicionalmente se fundamenta em valores antropológicos, como a dignidade humana.

[42] Exemplo disso é uma ação em curso em um tribunal da Califórnia, movida por alguns artistas, em face das empresas Stability AI Ltd., Stability AI Inc., Midjourney Inc. e DeviantArt Inc., na qual se discute a violação de direitos autorais de obras dos demandantes. Em breve análise, Stability AI Ltd. e Stability AI Inc. criaram, treinaram e mantêm em conjunto o Stable Diffusion, um produto de imagem de IA usado para derivar as imagens de saída do produto DreamStudio, da Stability. A Midjourney criou, vende, comercializa e distribui o produto Midjourney, e a DeviantArt faz o mesmo com o produto DreamUP, os quais, assim como o Stable Diffusion, produzem imagens em resposta a solicitações de texto. É válido reforçar que um subconjunto de imagens do Stable Diffusion foi usado para treinar tanto o produto da Midjourney quanto o da DeviantArt. Os requerentes alegam que suas obras foram utilizadas para habilitar os produtos dos réus sem permissão e buscam reparação em virtude da violação de seus direitos autorais. O processo ainda está em discussão, e o que é alegado, em sede de defesa pelas rés, é que não compilaram o conjunto de dados, mas consumiram os dados de uma empresa terceira. Nesse sentido, segundo elas, a violação da propriedade intelectual teria ocorrido no tratamento dos dados pela terceira, e não quando utilizaram esse conjunto de dados para o treinamento de seus modelos.

[43] Com base na leitura demanda judicial citada na nota de rodapé anterior.

diligência reforçados ao fornecedor do modelo generativo — titular do poder de concepção e treinamento da IA — e, de outro, distribua responsabilidade àqueles que, como desenvolvedores ou operadores de sistemas baseados nesse modelo, contribuam para a concretização do dano.

A proteção dos consumidores, sejam os usuários finais ou os desenvolvedores dos sistemas generativos, em relação a esses produtos digitais complexos, exige, portanto, uma interpretação finalística do artigo 7º do Código de Defesa do Consumidor, conjugada com a principiologia constitucional, a fim de assegurar tutela inibitória e compensatória efetiva, a responsabilização solidária ou subsidiária conforme o grau de culpabilidade e a internalização dos custos sociais da atividade, de modo a prevenir a externalização de riscos e a perpetuação de práticas lesivas sem resposta sancionatória adequada.

4. PERSONALIDADE DIGITAL E REGIMES DE RESPONSABILIDADE JURÍDICA

O capítulo anterior trouxe um recorte acerca da corrida polarizada pela supremacia do desenvolvimento da IA generativa, detalhando os impactos negativos para a sustentabilidade e a modelagem da economia global pelas *Big Techs*. Fato é que, além da motivação dessas empresas de tecnologia em impulsionarem seus produtos no mercado, a ascensão desse campo inovador também tem sido subsidiada pela ambição dos líderes empresariais internacionais em aumentar suas receitas em virtude da implementação de soluções de IA, posicionando-se, também, como líderes na transformação digital ocasionada por essa tecnologia.

Segundo Mayer *et al*. (2025), aproximadamente 31% dos executivos de alto escalão global acreditam que a IA poderá gerar um crescimento de receita superior a 10% nos próximos três anos, um otimismo que se destaca ainda mais entre os líderes indianos, dos quais 55% preveem esse aumento, em contraste com os líderes nos Estados Unidos, que apenas compartilham dessa mesma expectativa, evidenciando uma disparidade nas percepções sobre o potencial transformador da inteligência artificial. Essa dinâmica não apenas acentua a competitividade entre nações como também ressalta a urgência de uma adaptação rápida e eficaz às inovações tecnológicas, configurando um cenário em que a supremacia em IA se torna um determinante crucial para o sucesso econômico e estratégico no futuro próximo.

De forma antagônica ao que, por anos, manteve-se como objeto central de estudo da pesquisa científica e da doutrina acerca da temática da inteligência artificial (ou seja, o desenvolvimento da tecnologia a partir de pilares éticos bem determinados), o que se observa, sobremaneira, é que a atual configuração do mercado global de liderança, em razão do pioneirismo, tem posto a discussão ética como quesito secundário em razão da preeminência acelerada do desenvolvimento da IA, com a justificativa de facilitação do cotidiano empresarial. Sobre isso, reforçam Amaral e Xavier (2023, p. 40, realçado pela autora):

> No final de março de 2023, importantes cientistas, pesquisadores e empresários do campo assinaram uma carta pedindo uma "pausa" na pesquisa/desenvolvimento de tecnologia de IA. Inaudita cautela. Seria a primeira vez que isso seria feito nas modernas condições de produção e desenvolvimento tecnológico. **Não temos experiência praticamente nenhuma em sermos cautelosos quando lidamos com a base produtiva de nossas sociedades e as forças produtivas liberadas por nossas tecnologias. A regra que vale é: "primeiro inventa e implementa, depois lidamos com as externalidades negativas"**. Essa regra pressupõe que isso que estamos chamando lógica concorrencial operaria para resolver os problemas, restituir um equilíbrio inicial perdido. Balela. Uma lógica puramente competitiva nos afasta de condições que favoreçam, por exemplo, que fossem implementadas legislações internas nos países e acordos e tratados internacionais para controle das IAs.

Diante disso, torna-se imperativo que os institutos de personalidade e responsabilidade

civil sejam considerados no enfrentamento das complexas consequências que essa tecnologia emergente pode acarretar. A rápida evolução da IA não apenas transforma a dinâmica das relações sociais e comerciais, mas também levanta questões éticas e legais que desafiam os paradigmas tradicionais de responsabilidade.

A adoção desenfreada da IA pode resultar em impactos significativos, como a violação de direitos de personalidade, a geração de conteúdos prejudiciais e a responsabilização por decisões automatizadas. Assim, é fundamental que o arcabouço jurídico esteja apto a garantir a proteção dos indivíduos e a responsabilização adequada das entidades, promovendo um ambiente que equilibre inovação e segurança jurídica. Isso não apenas assegurará a proteção dos direitos fundamentais, mas também fomentará um desenvolvimento sustentável e ético da inteligência artificial na sociedade.

Na perspectiva atual, a IA ainda carece de autoconsciência, discernimento moral e emocional e de um senso comum. Essa tecnologia permanece totalmente dependente da inteligência humana para sua alimentação, incluindo a incorporação de valores éticos. Assim, é importante ressaltar que os computadores não possuem vontade própria, conforme discutido por Barroso e Mello (2024, p. 7).

> No estágio atual[44], a Inteligência Artificial não tem consciência de si mesma, não tem discernimento do que é certo ou errado, nem tampouco possui emoções, sentimentos, moralidade ou mesmo senso comum. Vale dizer: ela é inteiramente dependente da inteligência humana para alimentá-la, inclusive com valores éticos. Computadores não têm vontade própria[45].

Ou, ainda, “a IA não é artificial nem inteligente” (Crowford, 2021, p. 8, traduzido pela autora), uma vez que, além de depender da exploração do trabalho humano ao longo de toda a cadeia de suprimentos de extração, os sistemas automatizados também aparentam realizar tarefas anteriormente feitas por humanos que, na verdade, são apenas uma transferência de carga de trabalho para os consumidores ou outros trabalhadores não remunerados (Crowford, 2021).

Ao longo dos anos, duas visões distintas emergiram nas pesquisas sobre o funcionamento da IA. A primeira abordagem, que prevaleceu até a década de 1980, buscou imitar o funcionamento da mente humana, focando na forma como questões são elaboradas e raciocínios lógicos são desenvolvidos. Em contraste, a segunda perspectiva se baseou nas

---

[44] “A ressalva se impõe tendo em vista que não se descarta que a IA do futuro conceda, às máquinas, doses intensas de autonomia e de consciência, em um panorama que as aplicações inteligentes adquiram uma racionalidade própria, perseguindo objetivos não previstos (Degli-Esposti, 2023, p. 10; Rebollo Delgado, 2023, p. 24)” (Barroso; Mello, 2024, p. 7).

[45] “Minuta de 30 set. 2018, gentilmente enviada pelo autor: ‘Eles (os programas) não percebem como nós e não pensam como nós; na verdade, eles não pensam nada’ (Winston, 2018, p. 2). Sobre o tema cf.: Lenhado, 2023” (Barroso; Mello, 2024, p. 7).

estruturas do cérebro humano, propondo a conexão de unidades de processamento de informações que se assemelham a neurônios, com o objetivo de simular seu funcionamento (Dreyfus, L.; Dreyfus, E., 1988, pp. 15-44 *apud* Barroso; Mello, 2024, p. 7).

Essa última abordagem, conhecida como "abordagem conexionista", tornou-se predominante no campo da IA, não se limitando a reproduzir a racionalidade humana, mas sim em estabelecer correlações e padrões entre vastas quantidades de dados e resultados específicos, fundamentando-se em princípios da estatística e da neurociência.

Para além da ilusão de uma inteligência artificial autônoma e eficiente, sustentada fortemente pelo trabalho humano repetitivo e não reconhecido[46], os modelos e sistemas de IA generativa dependem, primordialmente, da inteligência da pessoa que os desenvolvem, sob financiamento de grandes empresas que estão à frente das inovações tecnológicas.

Nessa toada, Beckers e Teubner (2021), como também Negri (2020) e Negri e Lopes (2021), apresentam uma posição crítica em relação à atribuição de personalidade digital a agentes autônomos. Eles argumentam que a mera existência social desses agentes não é suficiente para justificar sua personificação legal. A discussão é polarizada, e os críticos da personalidade legal insistem que a autonomia atribuída socialmente não deve obrigar o direito a reconhecer a personalidade legal, pois a legislação pode decidir sobre a autonomia com base em critérios próprios, sem depender exclusivamente das características tecnológicas.

Beckers e Teubner destacam a complexidade das interações entre a atribuição social de assistência digital (ou agência) e a personalidade legal, sugerindo que a insistência na liberdade do direito positivo para conceder a personalidade ignora as exigências normativas do contexto social. Eles também ressaltam que a atribuição de responsabilidade legal deve ser cuidadosamente considerada, levando em conta a capacidade dos agentes digitais de tomar decisões sob incerteza, e não apenas a capacidade de autoaprendizado.

Negri (2020) expressa preocupações sobre os riscos e problemas associados à ampliação da subjetividade jurídica de robôs e agentes de IA em diferentes sistemas legais, criticando a tendência tratá-los de forma antropomórfica sem considerar as complexidades do processo de personificação e o significado do termo "pessoa jurídica" no Direito. Assim, a tendência de naturalizar conceitos como autonomia e consciência na robótica e na IA, sem uma definição clara do que esses termos significam no contexto tecnológico, também é uma preocupação para o autor.

Autonomia "forte", na visão de Negri (2020), está associada à consciência e ao livre-arbítrio, típica de agentes morais humanos, ao passo que uma autonomia "fraca" ou aparente

[46] Crowford (2021, p. 108) exemplifica a ascensão do ImageNet a partir de um exército de trabalhadores dedicados a classificar imagens manualmente, de modo que cada trabalhador chegava a executar, em um minuto, a classificação de 50 imagens.

diz respeito à capacidade de operar sem supervisão constante, mas ainda dentro dos limites programados. Com isso, o autor conclui que nenhum artefato robótico atual possui autonomia no sentido forte, e a confusão entre esses conceitos reforça o erro do antropomorfismo (Negri, 2020, p. 5).

Assim, atribuir personalidade jurídica eletrônica a agentes autônomos de IA é desconsiderar as limitações dessa ideia em relação ao conceito tradicional de pessoa jurídica (Negri; Lopes, 2021).

O presente capítulo se debruçará em analisar o humanismo digital a partir da interação humano-máquina, com o objetivo de entender a proposta de instituições digitais de Beckers e Teubner (2021) e como a responsabilidade de agentes autônomos dá-se a partir dela.

4.1 Humanismo digital

Na atual configuração de mercado da sociedade do século XXI, a inovação tem sido tratada como um imperativo para a competitividade. As empresas que não se reinventam estão fadadas ao declínio, o que destaca a necessidade de práticas inovadoras para se manterem relevantes. A transformação digital, nesse sentido, atua como um catalisador para essa inovação, exigindo que as empresas explorem novas tecnologias e novos modelos de negócios, além de realizar uma análise contínua das tendências e dos comportamentos do consumidor.

A promoção de uma cultura de inovação é fundamental, mas não deve ficar restrita apenas à adoção de novas tecnologias[47]. É indispensável que também se crie um ambiente que estimule a criatividade e a experimentação, integrando o ser humano a uma nova rotina facilitada pela tecnologia. A interação entre humanos e máquinas, facilitada pela IA, traz ainda desafios quanto à dependência da tecnologia, como a complexidade na definição de objetivos que alinhem as máquinas aos valores e às necessidades humanas[48].

[47] A opinião refletida pela autora deste trabalho aproxima-se, ainda, do que foi disposto no Relatório Especial da União Europeia quanto a ambições para a inteligência artificial (União Europeia, 2024). Logo na introdução e nos pontos iniciais do relatório, destaca-se que a IA não é apenas uma tecnologia a ser adotada, mas um vetor de transformação econômica e social, com potencial para impulsionar o crescimento e responder aos desafios societais. O relatório afirma que "os decisores das políticas públicas desempenham um papel importante na organização do ecossistema da IA" (União Europeia, 2024, p. 9), e que a recomendação da OCDE para a IA inclui, entre outros pontos, "reforçar as capacidades humanas e fazer a preparação para a transformação do mercado de trabalho" (União Europeia, 2024, p. 10) e "criar um ambiente político propício à inovação e à concorrência, visando uma IA de confiança e o apoio à transição da investigação para a implantação" (União Europeia, 2024, p. 10).

[48] O relatório citado na nota acima também aborda a necessidade de garantir que a IA esteja "ao serviço das pessoas e seja uma força positiva para a sociedade" (União Europeia, 2024, p. 14, figura 4), um dos quatro pilares do plano da UE para a IA. O documento enfatiza que a IA deve ser desenvolvida e utilizada de forma ética, segura e centrada no ser humano, reconhecendo os desafios da dependência tecnológica e da necessidade de alinhar os objetivos das máquinas com os valores e as necessidades humanas. Menciona, ainda, a importância de "uma IA de confiança" (União Europeia, 2024, p. 5), e que a UE tem buscado criar um "quadro regulamentar previsível" (União Europeia, 2024, p. 14, figura 4) para garantir que a IA respeite os direitos fundamentais e valores da União. Além disso, o relatório destaca que a promoção de uma IA ética envolve não apenas a regulamentação, mas também a criação de orientações e ambientes de experimentação, como instalações de ensaio e experimentação de inteligência artificial, que permitem testar soluções em ambientes reais, promovendo a criatividade e a experimentação.

Christian Fuchs, cientista social austríaco, em sua obra "*Digital Humanism: A Philosophy for 21st Century Digital Society*", aborda criticamente os impactos das novas tecnologias digitais na sociedade, argumentando que a digitalização tem exacerbado problemas como o autoritarismo, a polarização política, a disseminação de notícias falsas e as desigualdades econômicas. O autor defende que o humanismo digital pode ajudar a entender criticamente como essas tecnologias moldam a sociedade e a humanidade, propondo uma abordagem radical que inclui a decolonização da academia e a análise do papel de robôs e da inteligência artificial no capitalismo digital (Fuchs, 2022). Assim, Fuchs (2022, pp. 45-46) entende que a desumanidade figura como o problema central das sociedades digitais contemporâneas, fazendo-se necessário que a abordagem filosófica do humanismo permita o enfretamento dos problemas globais dessas sociedades.

Nesse sentido, a corrente filosófica do humanismo digital emerge como uma conduta que prioriza a adaptação da tecnologia às necessidades humanas, promovendo um equilíbrio entre eficiência e resiliência. Ou, ainda, segundo Fuchs (2022, p. 50, traduzido pela autora), “humanismo digital é uma abordagem filosófica que estressa as capacidades ativas e transformativas dos seres humanos na era digital”.

Assim, como ilustrado anteriormente com o entendimento de Barroso e Mello (2024), Fuchs (2022, p. 50) também reforça a inexistência de características epistemológicas, ontológicas e axiológicas nas máquinas e nas tecnologias computacionais, ao passo que são naturais nos seres humanos. Nesse sentido, respectivamente, o autor destaca que i) as máquinas e as tecnologias computacionais carecem de razão, consciência, moralidade e pensamento crítico, não estando aptas a substituírem as pessoas humanas na sociedade; ii) as tecnologias digitais, na contemporaneidade, moldam e são moldadas por seres humanos, não devendo ser analisadas como se fossem humanos, assim como os humanos não devem ser analisados como se fossem máquinas; iii) e, por fim, como máquinas não são humanos e humanos não são máquinas, é imperioso que máquinas não recebam um tratamento humano (Fuchs, 2022, pp. 50-51, traduzido pela autora).

Em sua obra de 1997, Lucia Santaella entende a relação entre humanos e máquinas a partir da classificação em três níveis distintos: muscular, sensório e cerebral. No que toca às máquinas musculares, surgidas durante a Revolução Industrial e projetadas para substituir ou amplificar a força física humana, entende Santaella (1997, pp. 34-36) que são caracterizadas por sua capacidade de transformar energia em trabalho mecânico, substituindo os músculos humanos e aumentando a eficiência das tarefas, uma vez que absorvem tarefas físicas e repetitivas, e aceleram o ritmo do trabalho humano.

Já em relação às máquinas sensórias, Santaella (1997, pp. 37-38) destaca que funcionam como extensões dos sentidos humanos, especialmente a visão e a audição. Desenvolvidas a

partir da Revolução Industrial, essas máquinas não apenas imitam os nossos sentidos, como também registram e reproduzem signos (imagens e sons), ampliando a capacidade de percepção sensorial.

Por fim, com relação às máquinas cerebrais, a autora explica que são essas que simulam os processos mentais e intelectuais humanos. Desenvolvidas com a invenção dos computadores, evoluíram para dispositivos capazes de processar símbolos e realizar operações complexas (Santaella, 1997, pp. 38-39). Essas máquinas ampliam habilidades mentais, como o processamento de informações e memória, representando uma nova forma de humanidade, integrando sistemas biológicos e eletrônicos em um ecossistema híbrido (Santaella, 1997, p. 39).

O que destaca Santaella (1997) é que a interação entre humanos e máquinas se dá em níveis históricos e evolutivos que coexistem e colaboram, formando um ecossistema híbrido, sugerindo que cada nível de máquina não apenas substitua, mas também amplie e complemente as funções das interações humanas, promovendo uma integração crescente entre o homem e a tecnologia. Esse pensamento de substituição, no entanto, precisa ser combinado com a exclusividade de características epistemológicas, ontológicas e axiológicas que diferenciam máquinas e tecnologias computacionais dos seres humanos. Portanto, é imperativo reconhecer que, apesar da crescente capacidade das máquinas de absorver certas atividades humanas, a essência da humanidade e a complexidade das interações sociais não podem ser reduzidas a meras operações tecnológicas.

Ainda no sentido de complementação das funções humanas, Santaella (2022) aborda a hipótese da extrassomatização[49] do cérebro, que sugere que a inteligência humana tem crescido fora do corpo biológico ao longo da história. Desde os primórdios, o ser humano buscou superar a fragilidade do cérebro mortal e a efemeridade da fala, começando com representações de imagens nas grutas e a invenção de formas de escrita como pictográficas, ideográficas e hieroglíficas.

Com o desenvolvimento da escrita alfabética no mundo grego e a invenção de Gutenberg[50], a propagação dos livros impulsionou a exossomatização da inteligência. A Revolução Industrial trouxe tecnologias de linguagem, como a máquina fotográfica, o fonógrafo e o cinematógrafo, seguidas pela revolução eletroeletrônica, como o rádio e a TV, ampliando ainda mais a inteligência humana (Santaella, 2022).

---

[49] Expressão utilizada por Santaella (2022) para abordar o crescimento da inteligência humana fora do corpo biológico, um processo que começou nas imagens das cavernas e evoluiu significativamente ao longo do tempo, o que a fez discernir sobre a hipótese do neo-humano, uma transformação radical da ontologia do humano impulsionada pelas linguagens e tecnologias. Ela argumenta que a inteligência artificial está dando força à hipótese do crescimento da inteligência fora do corpo biológico, o que justifica a hipótese do neo-humano.

[50] A autora refere-se à prensa de tipos móveis, que revolucionou a impressão e a disseminação do conhecimento no século XV. Antes dessa invenção de Gutenberg, os livros eram copiados à mão, um processo lento e caro. Com a prensa, foi possível produzir livros em massa de forma mais rápida e acessível.

O ponto culminante dessa evolução, segundo a autora, foi a criação dos computadores, que permitiram o desenvolvimento da inteligência fora do corpo humano, integrando-se ao corpo biológico através das linguagens. Assim, Santaella (2022) entende que a IA representa a amplificação dessa inteligência, simulando e emulando os atributos constitutivos da inteligência humana, marcando um novo estágio nessa evolução, tanto dentro quanto fora do corpo biológico.

Entretanto, é válido reforçar que a imitação se configura como um princípio fundamental na modificação que um instrumento ou uma máquina exerce em relação a uma habilidade-alvo, operando em um nível abstrato de semelhança. Amaral e Xavier (2023) corroboram que, ainda que seja possível mobilizar matéria não orgânica para armazenar informações, emulando capacidades do cérebro humano, a base material das redes neurais que sustentam a revolução da inteligência artificial é substancialmente distinta da rede de neurônios que inspira esse modelo.

Nesse sentido, os próximos subcapítulos abordarão o processo de imitação de uma habilidade-alvo por um agente autônomo sob a ótica do planejamento e da regulação, enfatizando a necessidade de uma distinção clara entre máquinas e humanos. Essa análise é fundamental, pois, embora a inteligência artificial busque conquistar um espaço de autonomia ao substituir atividades humanas, é imperativo delimitar que, na atual configuração jurídica, dificilmente a IA conquistará uma autonomia que se assemelhe à autonomia humana.

Ainda que os institutos jurídicos tradicionais, em certos momentos, careçam de ser reinterpretados e moldados para atender às intensas transformações sociais, sobretudo as provocadas pelo crescimento exponencial da IA, reforça-se a importância de tratar as máquinas como máquinas e os humanos como humanos, garantindo que as interações e responsabilidades sejam adequadamente definidas em respeito às próprias diretrizes bases do ordenamento.

4.2 Interação humano-máquina e personalidade jurídica de agentes autônomos

A trajetória do desenvolvimento da IA ao longo das últimas sete décadas foi marcada por uma série de altos e baixos, refletindo as tentativas de reproduzir capacidades intelectivas humanas em máquinas. Esse campo de pesquisa se organiza em subdisciplinas que visam emular habilidades específicas, como o raciocínio, a percepção e o processamento de linguagem natural, com o objetivo de permitir que um agente de IA possa operar à semelhança da performance humana, o que, como analisado neste estudo, entende-se não ser cabível.

Embora os LLMs sejam capazes de gerar conteúdo textual com base em grandes volumes de dados, esses modelos operam de maneira probabilística, gerando texto com base na máxima probabilidade condicional das palavras do conjunto de treinamento, sem realmente

compreender o significado profundo ou o contexto de maneira holística. Ou seja, os LLMs não possuem a habilidade de generalizar o conhecimento adquirido para além dos padrões que foram previamente aprendidos. Exemplo disso é a capacidade dos modelos generativos de gerarem conteúdos em diversos idiomas. Ainda que o modelo possa gerar respostas assertivas em um idioma, não significa que a máquina realmente o entenda.

Nesse ponto, é forçoso fazer-se a necessária distinção entre os conceitos de automação e de autonomia.

Para Korkmaz (2022), em sua tese sobre revisão de decisões automatizadas na LGPD, a automação em IA é apresentada como um fenômeno que, ao processar grandes volumes de dados e tomar decisões com base em algoritmos, tende a substituir ou apoiar o julgamento humano em diversas esferas sociais, econômicas e jurídicas. A autora destaca que, embora sistemas automatizados[51] possam superar vieses e ruídos cognitivos humanos, eles também carregam riscos próprios, como a opacidade, a falta de explicabilidade, o enviesamento algorítmico e a dificuldade de responsabilização.

Ela enfatiza que a automação, especialmente quando baseada em *machine learning* e *deep learning*, opera predominantemente em uma lógica formal, matemática e estatística, reduzindo a complexidade da realidade a parâmetros quantificáveis e, muitas vezes, inacessíveis à linguagem e compreensão humanas (Korkmaz, 2022, p. 13).

A autonomia de não humanos, segundo Beckers e Teubner (2021), refere-se à atribuição de capacidades de ação e decisão a entidades não humanas, como algoritmos ou robôs. Beckers e Teubner (2021) discutem que essa autonomia não é um fato tecnológico em si, mas uma construção social. A sociedade, através de diferentes contextos sociais (como economia, política e direito), atribui características de autonomia a esses sistemas, reconhecendo-os como agentes que podem interagir, afetar mudanças e adaptar estratégias.

Para os autores, essa autonomia é frequentemente vista em sistemas de IA que operam de forma independente, como veículos autônomos ou assistentes virtuais, que tomam decisões baseadas em dados e aprendizado de máquina. Nesse sentido, a discussão sobre autonomia de não humanos destaca a complexidade das decisões algorítmicas, que podem ser imprevisíveis e, portanto, desafiadoras em termos de responsabilidade legal. O que deverá, então, a sociedade considerar quanto à percepção dessas decisões e quais responsabilidades devem ser atribuídas quando ocorre um erro ou uma falha?

Na visão de Korkmaz (2022), a autonomia é entendida como a capacidade de

[51] A autora recorre a exemplos filosóficos, como o experimento do "quarto chinês", de John Searle, para ilustrar que sistemas automatizados podem manipular símbolos e produzir respostas indistinguíveis das humanas, mas sem qualquer compreensão semântica ou intencionalidade genuína. Assim, a IA, por mais avançada que seja, permanece restrita à manipulação sintática de dados, sem acessar o conteúdo hermenêutico, existencial e valorativo que caracteriza a cognição e a autonomia humanas (Korkmaz, 2022, p. 24).

autodeterminação, de construção da própria identidade e de participação consciente e crítica nos processos decisórios que afetam a vida do indivíduo. Nesse sentido, a automação, quando não controlada, pode ameaçar essa autonomia ao transformar pessoas em objetos de classificação, perfis e decisões tomadas por sistemas opacos e potencialmente discriminatórios[52].

Em virtude disso, no entendimento de Korkmaz (2022), a intervenção humana é fundamental para garantir a autonomia diante da automação. A autora defende que a revisão de decisões automatizadas deve ser substancial, permitindo que um ser humano compreenda, altere e justifique o resultado do processo decisório (Korkmaz, 2022, pp. 200-206). A mera presença nominal de um humano no processo não é suficiente: é necessário que haja efetiva capacidade de influência e contestação, sob pena de a automação esvaziar a proteção da pessoa e sua dignidade (Korkmaz, 2022, pp. 200-206).

Amaral e Xavier (2023), por outro lado, entendem que o avanço na autonomia e eficiência da IA é um reflexo da crescente distância entre a máquina e o programador, evidenciando a capacidade das máquinas de aprender e operar com maior independência, reforçando que a evolução da IA representa não apenas um desafio técnico, mas também uma nova fase na relação entre humanos e máquinas, em que a transferência de habilidades humanas para as máquinas se torna cada vez mais evidente.

Ao contrário, Beckers e Teubner (2021) entendem que, em vez de tratar os agentes digitais como meras extensões dos humanos ou conferir-lhes plena personalidade legal, é necessário encontrar um equilíbrio que reconheça a autonomia dos agentes digitais e que também leve em consideração os riscos e as lacunas de responsabilidade que essa autonomia pode gerar.

Negri (2016), por sua vez, ainda que tenha uma visão antropocêntrica da IA que se assemelha ao pensamento de Beckers e Teubner (2021), é crítico em relação ao modelo de análise de Teubner (1996) sob a perspectiva dos processos de personificação do ser humano e das pessoas jurídicas. Segundo ele, os conceitos de personalidade e de capacidade jurídica são aplicados de forma indistinta a ambos, o que negligencia as razões subjacentes à personificação do ser humano.

A análise proposta por Beckers e Teubner (2021) desafia a concepção tradicional de autonomia, que está intimamente ligada à autodeterminação humana, liberdade e moralidade. Eles argumentam que a autonomia digital não precisa necessariamente estar atrelada a características como inteligência artificial, empatia ou autoconsciência. Essa perspectiva é

[52] A autora alerta para o risco de uma "ditadura dos algoritmos", em que a pessoa é reduzida a um perfil ou classificação, perdendo sua individualidade e capacidade de contestação. Ela ressalta que a automação pode promover uma desapropriação tecnológica de prerrogativas humanas, criando assimetrias de poder entre quem controla os sistemas e quem é afetado por eles (Korkmaz, 2022, pp. 16-17).

crucial, especialmente quando se considera a ação irracional de agentes digitais em casos de violação da lei, que podem ter implicações significativas para a responsabilidade legal.

A crítica à ideia de que apenas agentes digitais com autoconsciência poderiam ser considerados portadores de personalidade jurídica é um ponto central no debate. Beckers e Teubner (2021) sugerem que essa visão é insustentável, pois ignora o contexto social e institucional em que esses agentes operam. A autonomia jurídica deve ser entendida de forma mais ampla, considerando a capacidade de autoaprendizagem e a adaptação dos agentes digitais às normas e expectativas sociais, sem a necessidade de atribuir-lhes características humanas ou morais.

Em janeiro de 2017, o Parlamento Europeu adotou uma resolução que propôs a criação de um estatuto jurídico especial para robôs, sugerindo que robôs autônomos mais avançados fossem reconhecidos como "pessoas eletrônicas" (*e-person*)[53], dotadas de direitos e obrigações específicas, incluindo a reparação de danos que causam. Para os autores, entretanto, essa proposta é vista como uma simplificação que não leva em conta as complexidades das relações sociais contemporâneas, à medida em que os agentes digitais, ao serem inseridos em um contexto socioeconômico, não atuam como maximizadores de utilidade de forma independente, mas sim como participantes de um sistema mais amplo que influencia suas ações e decisões (Beckers; Teubner, 2021).

Assim, a proposta de "pessoa eletrônica", segundo Beckers e Teubner (2021), ignora o contexto social e as dinâmicas de interação entre humanos e algoritmos, não sendo a personalidade jurídica plena uma solução adequada para todos os casos, pois os entes digitais não atuam como entidades autônomas, mas sim como assistentes em decisões humanas ou como partes de sistemas interconectados, onde suas decisões não são acessíveis à consciência humana. Ademais, a crítica dos autores também se estende à ideia de que a personalidade jurídica deve ser uniformizada, haja vista que a compreensão do papel socioeconômico dos agentes digitais é fundamental para determinar seu *status* legal.

A análise dos motivos pelos quais os sistemas sociais personificam não humanos, como organizações e algoritmos, revela uma complexidade nas relações contemporâneas entre humanos e não humanos, emergindo como uma estratégia em resposta à incerteza e à imprevisibilidade associadas ao comportamento[54] desses agentes. Ao atribuir características humanas a esses autores, Beckers e Teubner (2021, pp. 26-27) explicam que a relação passa de

[53] A proposta do Parlamento Europeu foi amplamente criticada, tendo a Resolução 2020/2014 (INL), de 20 de outubro de 2020, também do Parlamento, afastado a criação de uma personalidade jurídica própria aos sistemas comandados por IA (União Europeia, 2021).

[54] Negri (2020) discorre sobre o conceito de "comportamento emergente" referindo-se a comportamentos complexos que surgem da interação de componentes simples em sistemas adaptativos, o que levanta desafios na programação e utilização de agentes digitais (no caso analisado pelo artigo, sistemas de IA e robôs), especialmente em contextos críticos.

uma dinâmica de sujeito-objeto para uma de Ego-Alter, onde o "Ego" (humano) interage com o "Alter" (não humano) de maneira mais complexa, permitindo que o primeiro reaja de forma mais adaptativa às respostas do segundo, facilitando a comunicação e a coordenação.

Assim, na visão dos autores, a capacidade de um agente digital de agir, comunicar e decidir deve ser avaliada em função do contexto social específico em que ele opera, e não apenas dos critérios abstratos de autonomia, considerando a interdependência entre os agentes digitais e as instituições sociais que os cercam, permitindo uma compreensão mais rica e contextualizada da autonomia e da personalidade jurídica no mundo digital.

Beckers e Teubner (2021), ao discutirem sobre os processos de comunicação entre humanos e não humanos, os entendem como complexos e multifacetados, reforçando que essa comunicação pode ser entendida através da teoria dos sistemas[55], que oferece uma estrutura para analisar as interações entre humanos e máquinas. A comunicação, segundo os autores, envolve três componentes principais: enunciado, informação e compreensão.

Para que haja uma comunicação genuína entre humanos e algoritmos, é necessário que os encontros entre eles produzam eventos que possam ser reconhecidos como "enunciados" que contêm "informações". A verdadeira comunicação, portanto, só ocorre quando há uma compreensão mútua, o que é mais difícil de alcançar no contexto das interações homem-máquina. Os algoritmos, por sua natureza, operam de maneira diferente dos humanos. Enquanto os humanos têm uma vida interior complexa e subjetiva, os algoritmos funcionam com base em operações matemáticas e lógicas, o que leva a uma assimetria na comunicação.

Essa assimetria é tripla: primeiro, as operações internas dos algoritmos não se equiparam às operações mentais humanas; segundo, a comunicação entre humanos e algoritmos não é simétrica, pois os humanos interpretam as respostas dos algoritmos com base em suas próprias experiências; e, por último, a capacidade dos algoritmos de se comunicarem como "actantes" (entidades que podem agir) é limitada e depende do contexto institucional em que estão inseridos (Beckers; Teubner, 2021, pp. 38-40).

Beckers e Teubner (2021) também ressaltam que, mesmo que os algoritmos não possuam qualidades ontológicas que lhes permitam se envolver em relações sociais como os humanos, sua utilização em contextos institucionais pode levar à atribuição de capacidades comunicativas e *status* de ator, exatamente o que foi exemplificado acima sobre a interação

[55] A visão de Beckers e Teubner (2021) sobre algoritmos como atores antropomórficos ou corporativos se insere em um contexto mais amplo de análise das interações entre humanos e não humanos, especialmente no que diz respeito à atribuição de agência e capacidade de ação a entidades não humanas, como algoritmos. Essa discussão se relaciona com a teoria dos sistemas sociais de Luhmann, que propõe que os sistemas sociais são construídos por meio de comunicações e interações, e não apenas por indivíduos. Teubner (2017), em sua obra sobre corporativismo empresarial, argumenta que a personificação de entidades não humanas, como algoritmos, ocorre dentro de um quadro institucional que atribui a essas entidades uma identidade social e a capacidade de agir. Isso é reforçado pela ideia de que as práticas sociais institucionalizadas permitem que esses processos de comunicação sejam vistos como agentes autônomos, capazes de influenciar decisões e comportamentos humanos.

entre humanos e IA generativa.

Por conseguinte, a diferenciação entre o *status* de ator social e a personificação jurídica é crucial, na medida em que um algoritmo pode atuar como um agente dentro de um contexto social, sendo reconhecido por suas ações e interações, sem implicar que ele tenha direitos ou deveres legais da mesma forma que um ser humano ou uma pessoa jurídica (Beckers; Teubner, 2021 p. 41).

Negri (2016), analisando a assimetria entre a personificação do ser humano e das pessoas jurídicas e reforçando que personalidade e capacidade de direito são aplicadas indistintamente a ambos, propõe uma reavaliação do direito privado brasileiro a partir de um novo modelo de classificação capaz de distinguir as razões da personificação das relações normativas e das formas de uso. Exemplificativamente, o autor menciona o caso de Dartmouth College[56], que estabeleceu a autonomia das corporações privadas e a impossibilidade de interferência estatal (Negri; Lopes, 2021, pp. 1-2).

Aqui, encontra-se o principal ponto de divergência entre Negri (2016) e Teubner (1996). No citado caso concreto, para Teubner (1996), do ponto de vista da relação jurídica, não haveria problema em aplicar a lógica da tutela jurídica do ser humano visível a esse "ser artificial". Negri (2016), entretanto, aponta o risco dessa aproximação metafórica, que tende a mascarar as diferenças fundamentais entre pessoas naturais e jurídicas. Portanto, o entendimento de Teubner (1996), criticado por Negri (2016), é a aceitação acrítica da equiparação entre pessoas naturais e jurídicas, especialmente no que diz respeito à titularidade de direitos e à aplicação dos conceitos de personalidade e subjetividade, sem considerar as razões e os fundamentos distintos que justificam a personificação de cada um desses sujeitos no Direito.

O que Negri (2016) concentra-se em criticar é o modelo de análise que aplica indistintamente os conceitos de personalidade[57] e de capacidade de direito tanto à pessoa natural

---

[56] Segundo o autor: "O caso, julgado em 1819, é considerado um marco nos Estados Unidos para aplicação da chamada *contract clause*, que impõe o respeito à autonomia contratual, nas situações envolvendo corporações privadas. Na ocasião, o reitor da Dartmouth College foi deposto pelo conselho de administração da faculdade, o que levou o estado de New Hampshire a tentar transformar a faculdade em uma instituição pública, para reconduzi-lo ao cargo. A Suprema Corte entendeu que o Estado não poderia interferir nas atividades da faculdade, preservando, assim, a autonomia da instituição. Na ocasião Marshall afirmou: '*A corporation is an artificial being, invisible, intangible, and existing only in contemplation of law. Being the mere creature of law, it possesses only those properties which the charter of its creation confers upon it, either expressly or as incidental to its very existence*' (HALL, Kermit L. The Oxford Companion to the Supreme Court of the United States. Oxford: Oxford University Press, 2005, p. 248)" (Negri; Lopes, 2021. pp. 1-2).

[57] Korkmaz (2019), citando, respectivamente, Tepedino (2004) e Schreiber (2014), destaca que a categoria dos direitos da personalidade emerge como resultado de uma construção teórica desenvolvida na Alemanha e na França durante a segunda metade do século XIX. Evolução essa impulsionada pela necessidade de estabelecer um conjunto de direitos essenciais ao ser humano, que transcendesse a mera liberdade formal e protegesse a vulnerabilidade da vontade individual. Nesse sentido, o reconhecimento jurídico dos direitos da personalidade está intrinsecamente atrelado à proteção da dignidade e da integridade da pessoa, refletindo conquistas históricas significativas. Entretanto, a autora reforça que o mecanismo disponível para a tutela da pessoa à época era o direito subjetivo, legado da Revolução Francesa, que se fundamentava em uma lógica patrimonial e estabelecia uma dualidade entre sujeito e objeto, alinhando-se a um substrato ideológico liberal (Korkmaz, 2019). Embora alguns aspectos da personalidade, como o direito moral de autor e a proteção da imagem, tenham sido desenvolvidos entre o século

quanto à pessoa jurídica, como se fosse possível equiparar as razões que fundamentam a personificação do ser humano às que justificam a atribuição de personalidade jurídica a sociedades, associações e fundações. Para Negri (2016), essa equiparação desconsidera as particularidades do processo de imputação de direitos e deveres à pessoa jurídica, promovendo uma naturalização indevida e ocultando as diferenças essenciais entre os dois tipos de sujeitos.

Com isso, Negri e Lopes (2021) destacam, criticamente, a falta de consideração das razões que fundamentam a personificação do ser humano em comparação às que justificam a atribuição de personalidade jurídica a entidades como sociedades e associações, sendo incisivos quanto ao risco de uma "expropriação da subjetividade", em que a proteção dos direitos individuais pode ser comprometida em nome de uma abordagem mais ampla.

Assim, infere-se que as operações mentais dos modelos generativos diferem significativamente das operações mentais humanas, principalmente em função da natureza dos processos que cada um utiliza para gerar resultados. Os modelos generativos, como os descritos no contexto da inteligência artificial, operam por meio de algoritmos que analisam grandes volumes de dados e aprendem a reconhecer padrões e correlações entre eles.

Essa abordagem é essencialmente estatística e probabilística, permitindo que o modelo gere novos dados com base nas informações previamente processadas. Por outro lado, a mente humana é capaz de raciocínios complexos que envolvem não apenas a análise de dados, mas também a incorporação de emoções, de experiências pessoais, dos contextos sociais e culturais, e a capacidade de formular conceitos abstratos. Enquanto os modelos generativos, como os LLMs, utilizam técnicas como a amostragem probabilística para criar texto ou imagens, os humanos utilizam uma combinação de raciocínio lógico, intuição e criatividade, que não pode ser completamente replicada por algoritmos.

A popularização da tecnologia de IA generativa demonstrou como essas máquinas podem produzir resultados que, em muitos casos, parecem indistinguíveis da produção humana. No entanto, essa "inteligência" é, na verdade, uma simulação baseada em padrões aprendidos, sem a verdadeira compreensão ou intenção que caracteriza o pensamento humano.

A assimetria entre a comunicação humana e os modelos generativos pode ser compreendida, sobretudo, a partir da complexidade técnica e da capacidade de interpretação.

Quanto à complexidade técnica, a IA generativa opera com base em modelos probabilísticos que analisam vastos conjuntos de dados para gerar respostas ou imagens. Essa complexidade pode criar uma barreira de entendimento para os usuários, que muitas vezes não têm conhecimento técnico sobre como esses sistemas funcionam. Isso leva a uma comunicação

---

XIX e o início do século XX, o marco mais significativo desse processo é reconhecido na Constituição de Weimar, de 1919, que destacou a primazia dos direitos da pessoa ao legitimar a tutela dos interesses econômicos apenas quando vinculados a esses direitos (Doneda, 2006).

assimétrica, em que a IA pode gerar respostas de maneira eficiente, mas os humanos podem ter dificuldade em formular perguntas que maximizem a utilidade dessas respostas.

Já a capacidade de interpretação diz respeito ao entendimento e à geração de conteúdo sintético pela IA generativa com base em padrões aprendidos. No entanto, a interpretação humana é influenciada por emoções, contextos culturais e experiências pessoais que a IA não possui. Isso pode resultar em mal-entendidos ou em respostas que não atendem às expectativas do usuário. Assim, a interação acaba não sendo uma verdadeira comunicação, mas sim uma interpretação do comportamento da máquina.

Por fim, os modelos generativos, como os baseados na arquitetura de transformadores, operam com uma capacidade ativa limitada em razão da estrutura do espaço de representação, sendo este uma abstração dos dados de entrada, no qual as informações são codificadas em um formato que o modelo pode manipular. Nesse sentido, a capacidade ativa dos modelos generativos é limitada em razão da forma como processam e representam os dados.

4.3 Proposta de Beckers e Teubner aos regimes de responsabilidade jurídica atribuídos a agentes autônomos

Conforme demonstrado acima, Beckers e Teubner (2021) prezam pela necessidade de uma abordagem multidisciplinar para compreender e regular a digitalidade e os algoritmos na sociedade. Para tanto, os autores compreendem o comportamento da máquina a partir de uma tipologia que se divide em três categorias: o comportamento individual da máquina, que analisa as características de um único algoritmo; o comportamento híbrido homem-máquina, que resulta da interação entre humanos e máquinas; e o comportamento coletivo da máquina, que emerge da interconectividade de múltiplos agentes.

O comportamento individual refere-se ao funcionamento de um único algoritmo, considerando suas propriedades intrínsecas e a forma como interage com o ambiente. Os riscos associados são determinados pelo código-fonte e design do algoritmo. Já o comportamento híbrido é o resultado da interação entre máquinas e humanos, em que ambos os agentes influenciam o resultado das operações. Esse tipo de comportamento pode ser analisado através da teoria da rede de atores, que atribui um *status* de quase-ator a essas associações. O comportamento coletivo, por sua vez, se refere à dinâmica de sistemas inteiros formados pela interconexão de múltiplos agentes de máquina. Aqui, a análise do comportamento individual se torna menos relevante, pois os riscos emergentes são resultados de interações complexas dentro do sistema.

Essa tipologia visa identificar os riscos e benefícios das tecnologias digitais a partir da identificação da instituição sociodigital e é essencial para discutir a responsabilidade legal relacionada ao comportamento de máquinas, algoritmos e humanos. Isso porque a interrelação

entre tecnologia digital, instituições sociais e regras de responsabilidade é complexa e dinâmica, estando a tecnologia influenciando as instituições e vice-versa.

Nesse sentido, Beckers e Teubner (2021) também entendem pela necessidade de uma tipologia de riscos de responsabilidade, que surge da interação entre os comportamentos digitais e as instituições sociodigitais, alinhando-se à configuração do AI Act europeu (à época da escrita da obra, ainda uma proposta de legislação em construção), que regulamentou a matéria da IA a partir da gravidade dos riscos[58].

Assim, a proposta dos autores também é por uma classificação tripla em relação aos riscos associados ao uso de IA: o risco de autonomia, que envolve a capacidade das máquinas de tomar decisões independentes; o risco de associação, que surge da colaboração entre humanos e máquinas; e o risco de interconectividade, que se relaciona à operação conjunta de algoritmos em redes complexas.

Nessa medida, em relação ao risco de autonomia, associado ao comportamento individual das entidades não humanas, é necessária uma análise das doutrinas de responsabilidade que podem qualificar o *status* jurídico dos algoritmos, considerado como a responsabilidade que pode ser atribuída a ações de máquinas que operam de forma autônoma. No risco de associação, o comportamento híbrido das instituições que operam nesse espaço misto pode levar ao desenvolvimento de novas regras de responsabilidade coletiva, reconhecendo que a responsabilidade não é apenas individual, mas compartilhada entre humanos e máquinas. Por fim, o risco de interconectividade, associado ao comportamento coletivo, está intimamente intrínseco ao conceito de responsabilidade distribuída, em que a responsabilidade legal se torna desindividualizada e se aplica ao sistema como um todo, em vez de a um único agente.

O que pressupõem Beckers e Teubner (2021) é que a política jurídica não deve buscar uma abordagem única e generalizada para a responsabilidade civil que envolva agentes digitais em sua relação com humanos, pois isso resultaria em soluções simplistas e em regras de responsabilidade que não capturam a complexidade das interações. Em vez disso, é necessário um modelo jurídico adaptativo que reconheça as especificidades de cada tipo de comportamento de máquina, evitando arbitrariedades e garantindo um tratamento equitativo das situações.

Partindo-se do entendimento de "assistência digital"[59] de Beckers e Teubner (2021), a

[58] A legislação europeia acerca da IA categoriza os riscos desses sistemas em inaceitável (aqueles que representam uma ameaça clara para a segurança, os direitos fundamentais ou os valores da União Europeia), alto (sistemas que podem afetar significativamente a segurança ou os direitos fundamentais das pessoas, a exemplo dos sistemas baseados em modelos generativos), limitado (que requerem transparência, mas não são considerados de alto risco) e mínimo (que representam um risco mínimo ou inexistente para os direitos ou a segurança das pessoas e, por isso, não estão sujeitos a regulamentações específicas sob o AI Act).

[59] Para Beckers e Teubner (2021, p. 45, traduzido pela autora), o conceito de agência ou "(...) 'assistência digital' tem as suas origens na consagrada instituição social de 'representação humana'. Alguém entra e age no lugar de outra pessoa vis-à-vis a um terceiro. A instituição social de representação constitui, isto é, promulga e produz, um tipo de

autonomia dos algoritmos levanta questões sobre responsabilidade no direito privado, uma vez que tal autonomia é comparada à representação humana, na qual um agente digital (Alter) atua em nome de um humano (Ego). No entanto, assumem que essa nova forma de agência traz riscos, como a dificuldade de identificar o agente responsável, a falta de compreensão entre humanos e algoritmos, e a possibilidade de decisões algorítmicas se desviarem da intenção original. Assim, os autores reconhecem a complexidade da identificação do agente digital em virtude de os algoritmos dependerem de dados externos, complicando a atribuição de responsabilidade.

Em hipótese mais lógica, Beckers e Teubner (2021) propõem que tanto humanos quanto algoritmos possuam uma "*potestas vicaria*", permitindo que os algoritmos atuem autonomamente, mas dentro de um regime de responsabilidade que ligue suas ações aos humanos que representam. Essa abordagem poderia transformar as práticas sociais e jurídicas, reconhecendo as ações algorítmicas como constitutivas da ação humana e criando formas de responsabilidade legal.

O conceito de "risco de autonomia", nesse sentido, torna-se central, referindo-se à delegação de ações sociais para a esfera digital, o que pode resultar em danos devido ao comportamento imprevisível das máquinas. Assim, a assistência digital não deveria ser equiparada a agentes humanos com personalidade jurídica plena, mas sim receber uma personalidade jurídica limitada. Como exemplo, pode-se citar os contratos algorítmicos, cuja

---

atuação chamada ‘agência representativa’”. Frazão (2019), ao tratar do conceito de agência no contexto da responsabilidade civil de administradores de sociedades empresárias, entende que esses não são meros mandatários, mas sim órgãos da pessoa jurídica. Isso significa que não há vontade da pessoa jurídica que não seja manifestada senão por seus órgãos, ou seja, os administradores "presentam" diretamente a vontade da pessoa jurídica, sem propriamente representá-la. Nesse sentido, diferentemente dos sócios, que podem considerar seus interesses pessoais ao exercerem suas prerrogativas (como o direito de voto), os administradores, por serem órgãos, só podem agir em prol do interesse social. Mesmo quando participam de órgãos colegiados, devem votar observando exclusivamente o interesse da pessoa jurídica. A partir desse entendimento, a autora ressalta que a teoria da agência, no sentido tradicional de representação, foi superada no direito societário moderno, na medida em que o administrador não representa a pessoa jurídica como um mandatário, mas sim a "presenta", ou seja, manifesta diretamente sua vontade institucional. No entendimento de Lopes (2021), “Similarmente, nos países adeptos ao sistema de *common law*, define-se agência como a relação jurídica que emerge quando uma pessoa (denominada “principal”) manifesta consentimento para que outra pessoa, seu agente (“*agent*”), atue em seu lugar e sujeita a seu controle, podendo, inclusive, pleitear direitos e criar obrigações em seu nome. Trata-se de um agente em sentido estrito – em oposição à utilização também frequente do termo para denominar, de maneira geral, entes que possuem direitos e deveres próprios, ou seja, pessoas. Distingue-se, portanto, do entendimento filosófico acerca de agência, que se preocupa em determinar ou caracterizar a autonomia de um ator, e não as consequências do seu relacionamento com outros atores. Nesse sentido, possuir agência significa ser o originador de uma ação, ser movido por motivações, propósitos e desejos, relacionando-se, portanto, com a atribuição de intenções, pelo conceito de ação autodirigida ou ‘agir por razões’. (...) Por sua vez, na literatura especializada em ciência da computação, o termo agente representa um conjunto amplo de tecnologias, notadamente aquelas cujos sistemas de processamento de informações são relativamente autônomos. Pode-se aplicá-lo, portanto, a uma variedade de situações em que programas semiautônomos, com diferentes arquiteturas e graus de sofisticação, são utilizados para executar tarefas humanas”.
Transpondo o conceito de agência para a temática da IA, tem-se que “agência” se refere à capacidade de agir intencionalmente, em contraposição à “autonomia” em IA, tida como a capacidade de operar independentemente dos limites definidos. Assim, entendendo que os agentes de IA não possuem autonomia, implicitamente não irão possuir agência, o que ficaria a cargo de pessoas físicas ou jurídicas que manipulam essa tecnologia. Contudo, como será demonstrado ao longo deste estudo, compreende-se que a perspectiva de responsabilidade do dano imputada a uma pessoa física ou jurídica isoladamente é insuficiente.

analogia com a relação de agência é utilizada para a própria formação da relação contratual, aplicando-se, assim, as regras de responsabilidade vicária a essas interações. Isso porque o agente digital pode oferecer soluções inovadoras que o humano não considerou[60].

Os autores, ao tratarem da abordagem histórica de que apenas indivíduos ou seus representantes podiam celebrar contratos, mencionam que existem duas principais abordagens: uma que considera os contratos gerados por algoritmos não vinculativos, e outra que os vê como ferramentas utilizadas por humanos, havendo, inclusive, manifestações de que contratos celebrados por agentes autônomos são inválidos, a menos que haja uma legislação específica que os reconheça, embora vários países já tenham aceitado a validade de documentos eletrônicos (Beckers; Teubner, 2021, pp. 49-54).

Nesse sentido, a solução proposta por Beckers e Teubner (2021) seria a criação de uma lei de agência para agentes eletrônicos, com vistas a equilibrar a distribuição de riscos, responsabilizando um comitente apenas pelas ações de um agente eletrônico que estejam dentro dos limites previamente estabelecidos.

Em síntese, a proposta diz respeito a introduzir uma personalidade parcial para algoritmos e uma nova forma de responsabilidade indireta para humanos que utilizam sistemas autônomos. Assim, a responsabilidade indireta por decisões errôneas de um algoritmo seria aplicável quando um humano delegasse uma tarefa que exigisse liberdade de decisão por parte do agente digital, não sendo previsível e nem explicável pelo programador, devendo a ação executada pelo agente digital necessariamente violar um dever contratual, dando azo à causalidade entre ação e dano (Beckers; Teubner, 2021, p. 139). Como consequência, tem-se o usuário do algoritmo, que delegou a tarefa, como principal responsável.

Com relação à atribuição de responsabilidade por ações que envolvem tanto humanos quanto algoritmos, Beckers e Teubner (2021) sugerem que essas interações sejam entendidas como associações humano-máquina, ou híbridos, que demandam um novo entendimento legal sobre responsabilidade, levando em conta suas propriedades emergentes e a complexidade dessas relações. A discussão se estende a diversos contextos, incluindo a evolução do papel humano na tomada de decisões, em que as máquinas podem dominar em algumas situações, enquanto, em outras, os humanos mantêm o controle.

A título comparativo, os autores fazem uma associação entre os híbridos digitais e a relação de colaboradores dentro de organizações empresariais, suscitando que os humanos "agem para o híbrido da mesma forma que os gerentes de uma empresa não agem em seu

[60] Beckers e Teubner (2021, p. 158) citam o exemplo do caso conhecimento como Robo Advice, que diz respeito a uma discussão processual entre um executivo de Hong Kong e uma corretora de investimentos, acusada de ser responsável por uma perda de US$ 23 milhões, pelo executivo, devido a operações algorítmicas ilícitas de um Robo Advice Computer. Em síntese, esse supercomputador, chamado K1, deveria vasculhar fontes online para prever tanto o perfil do investidor quanto o futuro das ações nos EUA. Ainda que com simulações muito promissoras, o fato é que o computador, devido a uma ordem de *stop-loss*, perdeu a quantia mencionada pertencente ao investidor chinês.

próprio nome, mas como “agentes” em nome do seu “principal”, ou seja, para a empresa” (Beckers; Teubner, 2021, pp. 94-95, traduzido pela autora). Como exemplo de híbridos, os autores citam o uso de robôs por médicos, a tradução de documentos por IA com revisão humana (aqui pode-se entender pelo uso da IA generativa) e o uso, por humanos, de algoritmos de pesquisas de IA, a exemplo dos jornalistas investigativos utilizando desse sistema para identificar práticas ilegais de empresas e indivíduos (caso emblemático retratado na obra dos autores).

Nesse sentido, Beckers e Teubner (2021) propõem a ideia de "agência estendida", sugerindo que a responsabilidade moral deve ser atribuída ao conjunto da interação humano-máquina, em vez de recair sobre partes isoladas, na medida em que as interações entre humanos e máquinas formam associações híbridas que podem coevoluir. A responsabilidade coletiva é enfatizada por eles como um conceito crucial nessas interações, especialmente em sistemas de apoio à decisão baseados em inteligência artificial.

Dessa forma, o entendimento dos autores seria pela introdução do conceito de "pessoas híbridas" nos ordenamentos jurídicos, em que as interações entre humanos e não humanos criam entidades jurídicas, de modo a refletir melhor sobre a ação coletiva desses sistemas.

Assim, para os híbridos digitais, Beckers e Teubner (2021) optam por explorar a responsabilidade empresarial, na qual seria atribuída um regime de responsabilidade digital às legislações, que distribua a restituição de danos entre todos os envolvidos, como operadores, fabricantes e programadores.

Para tanto, a responsabilidade empresarial se aplicaria quando identificada uma situação de cooperação entre humanos e máquinas, na qual tenha havido a tomada de uma decisão errada e as atividades entre ambos estejam densamente entrelaçadas, de modo que a decisão não possa ser atribuída a nenhum dos dois isoladamente e, também, não possa ser estabelecido nexo de causalidade entre as ações individuais e os danos (Beckers; Teubner, 2021, p. 139).

A responsabilidade coletiva, nesse caso atribuída aos produtores, programadores, revendedores e, até mesmo, aos usuários, teria o condão de reconhecer que as ações e decisões do agente seriam resultado de uma rede de interações, e não de sua atuação isolada. Essa abordagem busca evitar a dificuldade de atribuir responsabilidade a um único ator e permite uma distribuição mais justa dos encargos entre os membros da rede.

Já em relação às associações humano-máquinas combinadas à interconectividade entre os múltiplos agentes autônomos, Beckers e Teubner (2021) ressaltam a dificuldade de capturar essa rede interligada do sistema em categorias jurídicas e a necessidade de a lei responder, decretando *pools* de risco para compensar danos e cobrir custos. Nesse sentido, enfatizam que a interconectividade é uma configuração por si só e sua relação com a sociedade resulta em uma instituição sociodigital específica.

Os autores explicam que a complexidade dessa relação híbrida ocorre em dois espaços, sendo o primeiro deles com limitação de comunicação através de interfaces e, o outro, em um vasto espaço de operações algorítmicas internas que permanecem inacessíveis à consciência humana. Isso resulta em uma influência significativa das máquinas interconectadas sobre a sociedade de maneira indireta, criando o que é descrito como "máquinas invisíveis" (Luhmann, 2012/2013 *apud* Beckers; Teubner, 2021, p. 112).

Nesse sentido, o conceito de "acoplamento estrutural com máquinas invisíveis" (entendido como associações humano-máquinas com múltiplos agentes autônomos) é introduzido para descrever a relação não comunicativa entre os sistemas tecnológicos e a sociedade, que opera por meio de processos eletrônicos incompreensíveis (Nassehi, 2019 *apud* Beckers; Teubner, 2021, p. 113).

Para entender as dinâmicas da interconectividade, que resulta em ações coletivas espontâneas, mesmo sem interação direta entre indivíduos, os autores apropriam-se do conceito de "coletividade sem coletivo", utilizado para descrever como multidões agem de forma descoordenada, mas ainda se conectam através de uma infraestrutura. Sendo assim, a interconexão de algoritmos não possui as qualidades necessárias para uma tomada de decisão coletiva, o que gera o risco de interconectividade, dificultando a previsão de cálculos interligados entre máquinas e resultando em comportamentos coletivos inesperados (Beckers; Teubner, 2021, p. 115).

O estudo do comportamento coletivo das máquinas revela que as interações sistêmicas podem gerar novos padrões e recursos (como a comunicação global instantânea), mas também trazem riscos. Para exemplificar o risco da interconectividade, Beckers e Teubner (2021) chamam atenção para o mercado financeiro, no qual os danos causados por negociações algorítmicas são atribuídos ao comportamento coletivo das máquinas, que difere do humano, aumentando a probabilidade de crises de mercado. Isso porque a própria complexidade estrutural do sistema torna difícil identificar responsáveis e prever danos, complicando a atribuição de culpa ou a intenção maliciosa.

Nesse sentido, os autores enfatizam que a responsabilidade não deve recair sobre um único agente, mas sim sobre todos os operadores envolvidos, devido à natureza distribuída e interligada das decisões em sistemas de IA, propondo uma abordagem de responsabilidade conjunta (Beckers; Teubner, 2021, pp. 122-126). Além disso, a dificuldade em provar a cadeia de causalidade e identificar a ação causadora do dano é ressaltada pelos autores, levando à proposta de inversão do ônus da prova, embora isso ainda exija que a vítima prove a existência de uma ação potencialmente causadora (Beckers; Teubner, 2021, pp. 122-126). A proposta final, no entanto, sugere que a vítima prove apenas o dano, criando uma presunção de responsabilidade, que pode gerar desvantagens para os operadores e fabricantes (Beckers;

Teubner, 2021, pp. 122-126).

Assim, Beckers e Teubner (2021, pp. 126-135) propõem uma "responsabilidade por partilha do sistema", em uma abordagem que considera o sistema como um todo[61], com o objetivo de promover uma melhor distribuição dos riscos, considerando, ainda, a possibilidade de soluções obrigatórias de seguros e fundos para compensar vítimas de danos causados por sistemas complexos, priorizando a transparência e a justiça na compensação.

A proposta inclui a formação de "*pools* de risco de IA", em que a responsabilidade é atribuída a partes com vínculos estreitos ao sistema, permitindo um financiamento tanto *ex-ante* quanto *ex-post*. Com isso, os autores pressupõem uma responsabilidade centrada na indústria, especialmente em tecnologias como IA, e sugerem a implementação de taxas para setores específicos, como veículos autônomos (Beckers; Teubner, 2021, pp. 126-133).

Com isso, não sendo aplicáveis nem a responsabilidade indireta, nem a responsabilidade da empresa, caberá compensação à vítima do evento danoso causado pela interconectividade digital, se houver violação de um dever contratual que somente possa ser atribuído a uma série de decisões algorítmicas interconectadas (Beckers; Teubner, 2021, pp. 139-140).

[61] Os autores também usam o termo “responsabilidade da indústria”, no sentido de que uma agência reguladora no ramo relevante da indústria associada ao evento danoso deveria ser a autorizada à administração do fundo (Beckers; Teubner, 2021, p. 139).

5 ANÁLISE DOS MODELOS GENERATIVOS NA ORDEM JURÍDICA BRASILEIRA

No Capítulo 4, o entendimento de Beckers e Teubner (2021) foi analisado do ponto de vista da atribuição de personalidade jurídica a agentes digitais e os regimes de responsabilidade em comparação ao entendimento de Negri (2016 e 2020) e Negri e Lopes (2021).

Como já reforçado na introdução e no capítulo metodológico desta dissertação, Beckers e Teubner (2021) debruçam-se sobre o tema através de uma análise comparativa que vai de encontro não só ao ordenamento civil alemão, como também à apreciação do direito consuetudinário em âmbito mais abrangente.

Em países como os Estados Unidos, em que o ordenamento se baseia no sistema jurídico do *common law*, a própria flexibilidade de adaptação às mudanças sociais faz com que os precedentes judiciais evoluam entre um caso e outro em um curto espaço de tempo, fazendo com que as decisões judiciais em torno do conflito se tornem mais assertivas e contemporâneas. Entretanto, cria-se um alerta maior quanto aos países regidos pelo regramento do *civil law*, em virtude de o próprio ordenamento jurídico ainda não estar preparado legislativamente para lidar com a matéria.

Este Capítulo 5, sendo o último analítico deste estudo, por sua vez, terá por finalidade investigar como a diferenciação técnica entre modelos e sistemas de IA é tratada pelo PL nº 2.338/2023 e analisar as minúcias que afastam a imediata aplicabilidade da teoria de Beckers e Teubner (2021) ao ordenamento jurídico brasileiro, ainda que a abordagem doutrinária dos autores se reconheça como instigante.

5.1 Tratamento de modelos e de sistemas generativos

O legislador, no texto do PL nº 2.338/2023 não traz uma diferenciação precisa entre modelo de IA e sistema de IA sugerindo, a partir de uma análise crítica, que, eventualmente, ambos os conceitos possam ser tratados como sinônimos:

> Art. 4º Para os fins desta Lei, adotam-se as seguintes definições:
> I – Sistema de inteligência artificial (IA): 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, conteúdo, recomendação ou decisão que possa influenciar o ambiente virtual, físico ou real**; (...) (Brasil, 2024, realçado pela autora).

Isso porque, no início da definição de sistema de IA, o legislador explica ser "baseado em máquina". Em Ciência da Computação, o termo máquina é utilizado para referir-se a um

dispositivo ou sistema que executa operações computacionais (Hopcroft *et al*., 2006). Assim, máquina pode ser um fim em si mesma, como uma máquina física (o próprio *hardware*, um computador, um servidor ou um dispositivo móvel) ou uma máquina virtual, isto é, um ambiente de *software* que emula um computador físico (Hopcroft *et al*., 2006).

Ao tratar do sistema de inteligência artificial de propósito geral, por outro lado, o legislador brasileiro direciona o conceito de forma moderadamente mais assertiva ao associar esse sistema a um modelo de IA treinado a partir de uma base de dados ampla, com o objetivo de servir a diferentes finalidades, aproximando-se, assim, da ideia de um modelo fundacional.

> III – sistema de inteligência artificial de propósito geral (SIAPG): sistema de IA baseado em modelo de IA treinado com bases de dados em grande escala, capaz de realizar ampla variedade de tarefas distintas e servir a diferentes finalidades, incluindo aquelas para as quais não foram especificamente desenvolvidos e treinados, (...) (Brasil, 2024, art. 4º, III).

Entretanto, ao final da conceituação, o legislador infere que o sistema de IA de propósito geral (enquanto modelo de IA) possa ser acoplado a diversos outros sistemas ou aplicações:

> (...), podendo ser integrado em diversos sistemas ou aplicações; (...) (Brasil, 2024, art. 4º, III).

É válido retomar, neste ponto, tanto o entendimento de Hopcroft (2006) quanto do AI Act, que reforçam que um sistema é caracterizado pelo aglomerado de componentes que carrega consigo, incluindo o modelo. Assim, não basta que um sistema seja somente baseado em um modelo de IA para ser um sistema. Para um sistema ser integrado a outro, é necessário que haja a coerência técnica entre ambos, ao passo que integrar um modelo a um sistema condiz com a própria natureza do sistema, que se operacionalizará a partir do modelo base utilizado.

O texto do PL nº 2.338/2023 também correlaciona a definição de inteligência artificial generativa ao próprio modelo de IA, o que, por tudo o que exposto até aqui, parece estar distante da definição técnica em Ciência da Computação, que qualifica a IA Generativa como um ramo da inteligência artificial (Stryker; Scapicchio, 2024).

> IV – Inteligência artificial generativa (IA generativa): modelo de IA especificamente destinado a gerar ou modificar significativamente, com diferentes graus de autonomia, texto, imagens, áudio, vídeo ou código de *software*; (...) (Brasil, 2024, art. 4º, IV).

Por fim, no parágrafo primeiro, do artigo 30, do texto aprovado em dezembro de 2024 pelo Senado brasileiro, ainda que se perceba uma tentativa de distinção entre um sistema e um modelo, ainda se verifica certo tratamento equivocado pelo legislador, na medida em que ele entende que um sistema (novamente, modelo mais interface) possa ser fornecido no formato de um modelo autônomo:

> §1º O cumprimento dos requisitos estabelecidos neste artigo independe de o **sistema ser fornecido como modelo autônomo** ou incorporado a outro sistema de IA ou a produto, ou fornecido sob licenças gratuitas e de código aberto, como serviço, assim como por meio de outros canais de distribuição (Brasil, 2024, art. 30, §1º, realçado pela autora).

Ainda que a proposta regulatória do texto do PL nº 2.338/2023 volte-se ao desenvolvimento, fomento e uso ético da IA com base na centralidade da pessoa humana, a questão discutida acima é uma das hipóteses de desdobramentos cabíveis ao ordenamento jurídico brasileiro, na medida em que esta futura legislação trará os conceitos básicos que amoldarão o entendimento e a decisão acerca de conflitos jurídicos relacionados à IA.

Kaufman (2022, pp. 162-163), ao tratar da assimetria informacional entre produtores de tecnologia e legisladores, já predizia a respeito da falta de clareza de reguladores sobre funções específicas das tecnologias. Ademais, para a autora, essa assimetria tende a aumentar na medida em que aumenta a complexidade dos próprios modelos de inteligência artificial[62].

Ainda que a temática possa ser regulada por legislações extravagantes, como a Lei de Software (Lei nº 9.609/1998), é notório que a própria natureza do sistema ou do modelo em análise, alicerçado em inteligência artificial, aproximaria a análise da legislação que virá a regular esse tipo de tecnologia no Brasil. Ainda que assim não fosse, seria preciso, então, que a comunidade jurídica se voltasse à atualização de algumas legislações, como da própria Lei de Software. Pela leitura da citada lei, por exemplo, é evidente elucidar que não há distinção entre modelo e sistema, tratado como programa de computador por essa legislação.

Um modelo de IA de código aberto poderá servir de base para estruturação de um sistema de IA sem que implique em penalidades jurídicas quanto à violação dos direitos de propriedade intelectual do desenvolvedor do modelo, ao passo que utilizar esse mesmo modelo em um sistema construído a partir das mesmas características já existentes em outro sistema poderá motivar o início de uma demanda judicial.

Em virtude disso, vislumbra-se a ausência de clareza da legislação brasileira que vem sendo construída acerca dessa diferenciação técnica, o que, em certa medida, pode levar a eventuais confusões jurídicas e desafios regulatórios quando da sua entrada em vigor.

5.2 Complexidades para a regulação dos modelos generativos em face da personalidade jurídica

No sistema jurídico brasileiro, os valores reconhecidos socialmente como pertencentes

[62] Segundo Kaufman (2022, p. 163): “Se aplicações básicas, como os algoritmos de seleção e classificação de conteúdo utilizado pelas redes sociais, não são plenamente acessíveis, o que dirá dos algoritmos de decodificação do cérebro”.

à personalidade - tidos como direitos atípicos e representados por todos aqueles relacionados ao princípio constitucional da dignidade da pessoa humana - foram integrados à Constituição Federal de 1988. Foram, também, incorporados ao Código Civil brasileiro de 2002, com a previsão dos direitos típicos dos artigos 11 a 21.

A personalidade jurídica, conforme disposição da supracitada dogmática civilista, abrange duas modalidades distintas, sendo a primeira delas a conferência de personalidade a indivíduos por nascimento e, a segunda, à atribuição de personalidade a pessoas jurídicas, que se dá por meio de um registro adequado e da capacidade de exercício, permitindo a realização de atos com efeitos jurídicos.

A autonomia concebida por Beckers e Teubner (2021), como exposto no Capítulo 4, é antes de tudo uma construção social: trata-se de um "atributo relacional" conferido a algoritmos quando a sociedade passa a reconhecê-los como actantes capazes de tomar decisões sob incerteza dentro de um contexto institucional. Essa autonomia não decorre de consciência, livre-arbítrio ou qualquer qualidade ontológica própria, mas de expectativas normativas projetadas sobre o agente digital.

No direito brasileiro, porém, a noção de autonomia está enraizada na autodeterminação da pessoa natural e na capacidade de agir de pessoas jurídicas legitimadas pelo ordenamento, conforme a dogmática civilista e os arts. 1º e 47 do Código Civil. Negri (2020), também referenciado no Capítulo 4, chama a atenção para a impossibilidade de equiparar essa autonomia "forte" — vinculada à dignidade, à consciência e à volição — à autonomia meramente operacional de sistemas de IA. Para ele, a autonomia algorítmica é, no máximo, "fraca" ou aparente, pois inexiste vontade própria que justifique transferência de imputação como se fosse um sujeito de direito.

Desse contraste resulta que, enquanto Beckers e Teubner (2021) concebem algoritmos como agentes vicários passíveis de adquirir uma forma limitada de personalidade quando socialmente reconhecidos, o direito brasileiro, na leitura de Negri (2020), continua ancorado na ideia de que somente seres humanos (ou pessoas jurídicas constituídas por eles) podem exercer autonomia juridicamente relevante.

Na sociedade digital, interconectada à IA, o regime jurídico brasileiro não encontra amparo para absorver a teoria de Beckers e Teubner (2021) automaticamente. No direito civil brasileiro, a personalidade jurídica refere-se à atribuição de direitos e deveres[63] inerentes aos

[63] Negri (2016, pp. 5-6), citando a teoria do *fattispecie*, explica que: "os fenômenos jurídicos seriam compostos por um elemento material e outro formal. Enquanto o primeiro se refere à situação de fato externa, o segundo resulta do complexo de regras que determinam, em termos jurídicos, a qualificação daquele mesmo fato, para a atribuição de efeitos e consequências jurídicas. O processo de qualificação, apoiado na distinção entre relevância e eficácia, assume, assim, o papel de um verdadeiro filtro, capaz de selecionar quais as situações que seriam, realmente, consideradas aptas a ingressar no "sistema dos fenômenos jurídicos". Reduzido à categoria de "pressuposto subjetivo da qualificação", o sujeito, assim codificado, torna-se um elemento formal para a imputação de direitos e deveres."

modelos e sistemas de IA reconhecidos a pessoas singulares ou coletivas, conforme disposição do art. 1º do CC.

Nesse sentido, no direito brasileiro, a "personalidade jurídica é entendida como a suscetibilidade para ser titular de relações jurídicas, isto é, a suscetibilidade para, em abstrato, ser titular de direitos e obrigações" (Mota Pinto, 2005 *apud* Barbosa, 2021, p. 269), não sendo atribuída por lei, mas reconhecida pelo ordenamento. Nessa medida, "a pessoa, só por o ser, é titular de um conjunto de deveres e é, necessariamente, responsável" (Barbosa, 2021, p. 270).

Com isso, tem-se que a personalidade jurídica acaba por conferir certo grau de autonomia às pessoas, o que se difere da capacidade de escolha dos agentes digitais, tida como uma "autonomia tecnológica, fundada nas potencialidades da combinação algorítmica que é fornecida ao *software*" (Barbosa, 2021, p. 270).

O conceito do humanismo digital, discutido no Capítulo 4, teve justamente o intuito de reforçar essa fronteira ontológica. O humanismo digital, assim, funciona como fundamento normativo para recusar qualquer deslocamento da titularidade de direitos e deveres para além de pessoas físicas ou jurídicas, preservando a centralidade da pessoa humana consagrada pelo art. 1º, III, da CRFB.

Ainda que o ordenamento civil brasileiro dê tratamento a uma pessoa coletiva, o que, em primeira análise, meramente conceitual, poderia remeter ao entendimento de Beckers e Teubner (2021) acerca da cadeia de atores envolvidos desde a concepção do modelo de IA até o seu uso efetivo, esse tratamento é essencialmente vinculado ao conceito de pessoa coletiva enquanto pessoa jurídica[64].

Barbosa (2021, pp. 273-277), inclusive, ressalta a ficção legal da personalidade jurídica das pessoas coletivas, que as equipara a pessoas singulares para fins jurídicos. A autora, citando Savigny (1840), alude que, embora as pessoas coletivas não sejam seres humanos, elas são tratadas como tal para a realização de objetivos específicos, abordagem essa que foi criticada por Mota Pinto (2005), que argumenta que a personalidade jurídica é uma criação do direito, não necessitando de uma analogia com pessoas físicas (Barbosa, 2021, pp. 273-274).

A personalidade coletiva é, portanto, vista como um instrumento técnico que serve a interesses humanos coletivos, não uma entidade com vontade própria. Assim, a atribuição de personalidade jurídica de pessoas coletivas a agentes digitais de IA, perante o ordenamento jurídico brasileiro, também não se justificaria, na medida em que não há um interesse humano

[64] Negri e Lopes (2021) ressaltam que a pessoa jurídica foi concebida para permitir a imputação de direitos e deveres a entes coletivos, facilitando a responsabilização por atos praticados em seu nome. No entanto, alertam para os riscos de se adotar um modelo unitário e abstrato de responsabilidade, que desconsidera as particularidades de cada tipo de pessoa jurídica e pode ocultar os verdadeiros responsáveis pelos danos. Observam os autores que, no caso das pessoas jurídicas, a responsabilidade civil é frequentemente tratada de forma generalizante, como se houvesse um "ente único" a ser protegido ou responsabilizado, quando na verdade a pessoa jurídica é composta por uma pluralidade de interesses e agentes (sócios, administradores, beneficiários etc.). Essa abordagem que levaria à chamada "expropriação da subjetividade", tratada no Subcapítulo 4.1 deste estudo.

que a sustente, diferentemente das pessoas coletivas, que têm um substrato social e funcional.

Negri e Lopes (2021) sustentam que o ordenamento jurídico brasileiro, assim como outros sistemas jurídicos tradicionais, reconhece apenas duas categorias de sujeitos de direito: a pessoa natural (física) e a pessoa jurídica. Assim, para os autores, a criação de uma "pessoa eletrônica" — ou seja, a atribuição de personalidade jurídica a agentes autônomos de inteligência artificial ou robôs — não encontra respaldo na estrutura conceitual do direito brasileiro, pois este só admite a existência de pessoas físicas e jurídicas, cada qual com fundamentos, funções e limitações próprias.

Negri e Lopes (2021) argumentam que a personalidade jurídica, no direito brasileiro, é um instrumento técnico que serve para simplificar relações jurídicas complexas, articular patrimônios e permitir a imputação de direitos e deveres a entes coletivos (como sociedades, associações e fundações). A pessoa jurídica, portanto, é uma construção abstrata, um "atalho mental" que facilita a organização e a responsabilização de coletividades humanas, mas que não se confunde com a pessoa natural, nem pode ser estendida indiscriminadamente a qualquer entidade, como robôs ou sistemas de IA, sem que se perca o sentido e a função do instituto (Negri; Lopes, 2021, p. 13).

Contudo, algo na teoria de Beckers e Teubner (2021) que levou o presente estudo a aproximá-la do entendimento da estrutura de construção de sistemas generativos a partir de modelos de mesma natureza é justamente a análise da responsabilidade civil (e, consequentemente, da personalidade) considerando a interferência de diversos agentes.

A personalidade jurídica de agentes digitais de Beckers e Teubner (2021) aplicada à função e ao papel dentro do sistema de IA, acaba por explicitar a cadeia de materialização de um sistema de IA. Assim, ao desenvolvedor e ao fornecedor de um modelo de IA generativa, concentrado em poucas grandes empresas de tecnologia, é aplicado um contexto social e uma forma de comunicação homem-máquina que será totalmente diferente da aplicável ao desenvolvedor e ao fornecedor de sistema de IA que utilize como base o modelo citado, que está em conjunturas sociais bem mais distintas e com a comunicação humano-máquina muito mais amplificada.

A respeito da análise da comunicação humano-máquina, inclusive, vê-se que também é tratada de maneira divergente. Para Beckers e Teubner (2021), existe uma transformação da relação sujeito-objeto em um vínculo Ego-Alter, no qual a máquina assume o papel de interlocutor em redes funcionalmente diferenciadas. A análise luhmanniana de enunciado, informação e compreensão permite falar em “comunicação genuína” sempre que o algoritmo produz respostas que geram expectativas no meio social. O direito brasileiro, entretanto, encara a interface homem-IA como fenômeno essencialmente instrumental: a responsabilidade surge quando a conduta humana (programação, treinamento, utilização) se conecta causalmente ao

dano (arts. 927 e 931 do CC; arts. 12 e 14 do CDC).

Como ressaltado ao longo do estudo, a assimetria informacional, a opacidade algorítmica e o analfabetismo técnico impedem que se alcance a bilateralidade comunicativa pressuposta por Beckers e Teubner (2021). Por isso, a legislação pátria continua exigindo supervisão humana significativa e impõe o ônus da prova ao fornecedor ou inverte-o em favor da vítima (art. 37 do PL nº 2.338/2023) em vez de atribuir relevância jurídica autônoma aos "enunciados" da máquina.

Nesse sentido, para que se possa chegar ao ideal de personalidade legal limitada de Beckers e Teubner (2021), é necessário não só entender que ela será ditada pela congruência entre capacidade de decisão sob incertezas de agentes digitais e responsabilidade humana, mas também pelo entendimento técnico e aprofundado de como se deu a construção desse agente digital. Ao contrário, a lógica civil-constitucional de tutela da dignidade, reforçada pela crítica de Negri (2020) ao unitarismo da pessoa jurídica, impede que se reconheça personalidade limitada a algoritmos sem reformular toda a teoria geral do sujeito de direito.

Desse modo, embora a tipologia de Beckers e Teubner (2021) ofereça insights heurísticos para mapear cadeias de responsabilidade, sua adoção literal colidiria com os princípios e categorias fundamentais do direito brasileiro, exigindo reinterpretação sistêmica em diálogo com a Constituição, o Código Civil e o regime consumerista.

Ainda que demonstrada a impossibilidade de aderência sistêmica do modelo de Beckers e Teubner (2021) à realidade brasileira, na visão deste estudo, ele ampara a cadeia de construção dos sistemas generativos: ora poderá apresentar-se de uma forma e destinado a um público consumidor específico (viés de modelo de IA, por exemplo, como uma "peça" para construção de um sistema de IA direcionado a empresas ou desenvolvedores que se dedicam ao desenvolvimento desses sistemas) e ora com forma e consumidores diferentes (como um sistema de IA, que apresenta robustez após a combinação do modelo de IA a demais interfaces, e direcionado especificamente a pessoas físicas).

À vista disso, pronunciam-se Barroso e Mello (2024, p. 11), fazendo uma analogia entre os atores e fases envolvidos na cadeia de valor da IA e a construção do app Be My Eyes[65]:

> (...) a cadeia de valor da IA, com suas diferentes fases: *design*, desenvolvimento, implementação e manutenção do sistema (HACKER; ENGEL; MAUER, 2023, p. 8-11). Cada uma dessas fases tem seus próprios desafios e riscos, envolvendo diferentes atores, que incluem o desenvolvedor (*developer*)[18] [Seria o caso da OpenAI quanto ao GPT-4 (modelo fundacional)], o implementador (*deployer*)[19] [No caso do ChatGPT (modelo de propósito específico), a OpenAI é desenvolvedora e implementadora. A Be My Eyes, a seu turno, é apenas implementadora do *Virtual Volunteer* (também de propósito específico), adaptado a partir do GPT-4.], o usuário[20] [no caso dos produtos

[65] Aplicativo de IA generativa criado para ajudar pessoas cegas ou com visão limitada, composto por uma comunidade global de pessoas cegas ou com visão limitada, em conjunto com voluntários sem deficiência visual. (Be My Eyes, 2025).

> referidos na nota anterior, usuário é quem gera o texto com o ChatGPT ou consulta o aplicativo de voz, via *Virtual Volunteer*] e o destinatário final (*recipient*)[21] [Quem usa o texto e as orientações de voz. Veja que a notícia produzida pode ou não ser verdadeira. A voz pode ser uma boa orientação ou uma *deep fake*, produzida para iludir o destinatário]. Tais agentes detêm distintas expertises e habilidades, podendo gerar diferentes aportes, danos e responsabilidades (HACKER; ENGEL; MAUER, 2023, p. 8-11). Essa variedade de papéis, como intuitivo, acrescenta alguns graus de dificuldade à regulação da matéria.

Entretanto, também reconhece este estudo a lacuna para a aplicação da teoria de Beckers e Teubner (2021) à realidade brasileira. Mas, entende, ainda, pela insuficiência, em matéria de modelos e sistemas generativos, de imputação de responsabilidade civil a partir da perspectiva de uma pessoa física e/ou jurídica isoladamente.

Negri e Lopes (2021), advertindo sobre a criação de uma "pessoa eletrônica" para robôs ou sistemas de IA repetiria os mesmos problemas do unitarismo da pessoa jurídica, agravados pela ausência de substrato humano. A responsabilidade civil, nesse contexto, poderia ser artificialmente deslocada para o artefato tecnológico, ocultando a responsabilidade real de programadores, fabricantes, operadores e demais agentes humanos envolvidos no desenvolvimento e uso da tecnologia.

Para que fique claro, este trabalho não entende pela criação de uma pessoa eletrônica, mas reconhece que a figura das instituições sociodigitais de Beckers e Teubner (2021) é uma proposta interessante do ponto de vista da aplicação de responsabilidade civil quanto a danos decorrentes do uso de modelos e sistemas generativos. Isso porque, a partir da figura de instituições digitais, é possível que se julgue o cabimento da responsabilidade a partir de uma análise conjunta da multiplicidade de autores que contribuíram para o dano.

A partir da figura de uma instituição digital (ou agentes, ou pessoas físicas e/ou jurídicas em conjunto), o que vislumbra este trabalho é que a aplicação excessiva e/ou desproporcional de responsabilidade a uma pessoa ou outra ou a completa ausência de imputação de responsabilidade[66] poderiam ser minimizadas.

---

[66] Cita-se o acidente com a aeronave da Tam, de 2007, como exemplo de caso em que, diante da investigação de responsabilidade isolada de uma série de indivíduos, o tribunal acabou por entender que não existia concretude material para punir nenhum deles. Abaixo, um breve resumo do caso.

Contexto:

O acidente ocorreu em 17 de julho de 2007, envolvendo o voo TAM 3054, um Airbus A320 da então companhia aérea TAM Linhas Aéreas. A aeronave, procedente de Porto Alegre, ao pousar no Aeroporto de Congonhas, em São Paulo, não conseguiu parar na pista molhada, atravessou a avenida Washington Luís e colidiu com um prédio da própria TAM e um posto de combustíveis. O acidente resultou na morte de 199 pessoas, incluindo passageiros, tripulantes e pessoas em solo, sendo considerado o maior desastre aéreo da história do Brasil até então.

Causas e Discussões sobre Responsabilidade:

Após o acidente, investigações conduzidas pelo Centro de Investigação e Prevenção de Acidentes Aeronáuticos (CENIPA) e pela Polícia Federal apontaram uma série de fatores contribuintes:

• A pista do aeroporto estava escorregadia devido à chuva e não possuía *grooving* (ranhuras para escoamento de água), o que aumentou o risco de aquaplanagem.

• Um dos reversores de empuxo da aeronave estava desativado, o que é permitido pelos manuais, mas exige procedimentos específicos na aterrissagem.

• Houve erro operacional dos pilotos ao configurar os manetes dos motores para o pouso, o que comprometeu a desaceleração da aeronave.

5.3 Aderência aos regimes de responsabilidade de Beckers e Teubner: risco *versus* dano

Em sentido oposto a uma discussão sobre atribuição de personalidade jurídica a agentes digitais face a humanos, o PL nº 2.338/2023, seguindo a regulamentação europeia, tem uma abordagem centrada essencialmente na qualificação do risco decorrente do uso da IA generativa como prerrogativa para a responsabilidade civil.

O AI Act, ainda que estabeleça uma classificação pautada em risco para os sistemas de IA, ao absorver dispositivos para regulamentação da IA de propósito geral em seu texto, dedicou um capítulo específico para tratamento de risco voltado a essa tecnologia, o qual denominou de "risco sistêmico".

O risco sistêmico, nesse sentido, é entendido como aquele que tem um impacto significativo no mercado da UE, devido ao seu alcance ou efeitos negativos reais ou razoavelmente previsíveis na saúde e segurança públicas, nos direitos fundamentais ou na sociedade, e que pode se propagar em larga escala.

Para tanto, o AI Act valeu-se, conceitualmente, de duas categorias para identificação do risco sistêmico em IA de propósito geral, sendo a primeira delas a capacidade de elevado impacto, avaliada com base em ferramentas e metodologias técnicas adequadas, e a segunda à capacidade ou aos impactos equivalentes aos estabelecidos para a capacidade de elevado impacto tendo em vista os critérios estabelecidos no Anexo XIII[67].

Art. 51 Um modelo de IA de uso geral será classificado como um modelo de IA de uso

---

Imputação de Responsabilidade:
Após o acidente, houve intensa discussão sobre a responsabilidade:
• Foram investigados diretores da TAM, funcionários da Agência Nacional de Aviação Civil (ANAC), da Infraero (responsável pela administração do aeroporto) e autoridades do setor aéreo.
• O Ministério Público Federal chegou a denunciar alguns envolvidos por homicídio culposo, incluindo o então diretor-presidente da TAM, funcionários da Infraero e da ANAC.
• A defesa dos acusados alegou que o acidente foi resultado de uma cadeia complexa de fatores, sem dolo ou negligência individual suficiente para caracterizar responsabilidade penal.
Desfecho:
Ao final dos processos judiciais, todos os réus foram absolvidos. A Justiça Federal entendeu que não havia provas suficientes para responsabilizar criminalmente os acusados, reconhecendo a complexidade do acidente e a multiplicidade de fatores envolvidos. Assim, ninguém foi punido criminalmente pelo desastre.

[67] O Anexo XIII do AI Act estabelece critérios para designação de modelos de IA de finalidade geral com risco sistêmico. Em síntese, esses critérios seriam:
a. Número de parâmetros do modelo.
b. Quantidade do conjunto de dados (medida em tokens).
c. Quantidade de cálculo necessária para o treino do modelo, com base em custo, tempo e energia estimados.
d. Modalidades de entrada e saída (texto para texto, texto para imagem etc.).
e. Parâmetros de referência e avaliações da capacidade do modelo: número de tarefas sem treino adicional, adaptabilidade a tarefas novas e distintas, nível de autonomia e escalabilidade e ferramentas a que tem acesso.
f. Elevado impacto no mercado interno devido ao seu alcance.
g. Número de utilizadores finais registrados. (European Parliament, 2024, Anexo XIII, traduzido e resumido pela autora).

> geral com risco sistêmico se atender a qualquer uma das seguintes condições:
>
> (a) tem capacidades de alto impacto avaliadas com base em ferramentas e metodologias técnicas apropriadas, incluindo indicadores e benchmarks;
>
> (b) com base em uma decisão da Comissão, *ex officio* ou após um alerta qualificado do painel científico, tem capacidades ou um impacto equivalente aos definidos no ponto (a) tendo em conta os critérios definidos no Anexo XIII.
>
> 2. Um modelo de IA de uso geral será presumido como tendo capacidades de alto impacto de acordo com o parágrafo 1, ponto (a), quando a quantidade cumulativa de computação usada para seu treinamento medida em operações de ponto flutuante for maior que 10(^25). (European Parliament, 2024, art. 51, traduzido pela autora).

O Projeto de Lei nº 2.338/2023, em 7 de junho de 2024, quando recepcionou as disposições acerca da IA de propósito geral, fez com que, em seu artigo 32 (trechos realçados pela autora), fizesse constar o seguinte:

> Art. 32. O desenvolvedor de um modelo de IA de propósito geral[68] deve, antes de o disponibilizar no mercado ou de o colocar em serviço, garantir que o cumprimento dos seguintes requisitos:
> I - Demonstrar por meio de testes e análises adequados, a identificação, a redução e a mitigação de **riscos razoavelmente previsíveis** para os direitos fundamentais, o meio-ambiente, à integridade da informação e o processo democrático e antes e ao longo de seu desenvolvimento, conforme apropriado;
> II - Documentar dos **riscos não mitigáveis remanescentes** após o desenvolvimento, bem como sobre os impactos ambientais e sociais;
> III - Apenas processar e incorporar conjuntos de dados coletados e tratados em conformidade com as exigências legais, sujeitos a uma adequada governança de dados, em especial de acordo com a Lei nº 13.709, de 14 de agosto de 2018 (Lei Geral de Proteção de Dados Pessoais) e o Capítulo X desta lei;
> IV - Conceber e desenvolver o sistema de modo a permitir que alcance, ao longo do seu ciclo de vida, níveis apropriados de desempenho, previsibilidade, interpretabilidade, corrigibilidade, segurança e a cibersegurança avaliadas por meio de métodos apropriados, tais como a avaliação de modelos, análise documentada e testes extensivos durante a concepção, *design* e desenvolvimento;
> V - Conceber e desenvolver recorrendo às normas aplicáveis para reduzir, considerando o contexto de uso, a utilização de energia, a utilização de recursos e os resíduos, bem como para aumentar a eficiência energética e a eficiência global do sistema;
> VI - Elaborar documentação técnica e instruções de utilização inteligíveis, a fim de permitir que os desenvolvedores posteriores e aplicadores tenham clareza sobre o funcionamento do sistema;

Naquela oportunidade, é notório que o legislador entendeu por bem dar um tratamento diferenciado à IA de propósito geral, no sentido de associá-la a "riscos razoavelmente previsíveis" e a "riscos não mitigáveis remanescentes".

Ato contínuo, em 18 de junho de 2024, o então Senador Carlos Viana propôs a Emenda 58, fazendo com que o caput do artigo 32 fosse alterado para fazer constar "IA de propósito geral de alto risco" no lugar de "IA de propósito geral". Ao que parece, o parlamentar pretendia

[68] Em 11 de junho de 2024, o Senador Alessandro Vieira propôs a Emenda 14, para que fizesse constar, no caput do artigo 32, previsão específica para a IA generativa. A Emenda foi acolhida na sessão da CTIA de 5 de dezembro de 2024, passando o caput do artigo a dispor o seguinte: "Art. 30. O desenvolvedor de sistemas de IA de propósito geral e generativa com risco sistêmico, deve, antes de sua disponibilização ou introdução no mercado para fins comerciais, garantir o cumprimento dos seguintes requisitos: (...)" (Brasil, 2024)

alinhar o tratamento dado a este tipo de IA com os padrões internacionais de regulação que estavam em discussão.

Ocorre que, em breve análise, necessário se faz reforçar que houve um equívoco na emenda proposta, na medida em que o que se pretendia era atribuir uma categoria de risco de uma IA geral a uma IA específica. O alto risco, tanto na legislação brasileira em construção, como também na europeia (entendido como "risco elevado"), hoje em vigor, é atribuído a modelos e sistemas de IA voltados, em síntese, à perfilização humana, administração da justiça, veículos autônomos e segurança e saúde públicas.

De maneira diversa, como já demonstrado acima, o que o legislador europeu fez foi definir um "risco sistêmico" à IA de propósito geral, entendido tanto a partir de um impacto significativo no mercado da UE, como também de um efeito negativo real ou razoavelmente previsível em saúde e segurança públicas, direitos fundamentais e na sociedade. Com razão, a referida emenda foi rejeitada na Sessão de 5 de dezembro de 2024 da CTIA.

Ainda nessa toada, em 28 de novembro de 2024, o Gabinete do Senador Eduardo Gomes, proponente do PL nº 2.338/2023, emitiu um relatório em complementação ao relatório de 18 de junho para fazer constar, então, o "risco sistêmico" no projeto, sendo este um filtro para obrigações específicas e adicionais relativas à IA de propósito geral. Veja-se:

> XXX - risco sistêmico: potenciais efeitos adversos negativos decorrentes de um sistema de IA de propósito geral e generativa com impacto significativo sobre direitos fundamentais individuais e sociais. (Brasil, 2024, art. 4º)

Pela leitura do conceito de "risco sistêmico", no Projeto de Lei nº 2.338/2023, percebe-se que o legislador optou por absorver a conceituação praticada pelo AI Act para a categorização do risco aplicável à IA de propósito geral e à generativa.

Com isso, o texto aprovado no Senado brasileiro, em 10 de dezembro de 2024, no que diz respeito à governança da IA de propósito geral, dentre ela a generativa, seguiu para aprovação pela Câmara dos Deputados com a seguinte redação:

> Seção V Das Medidas de Governança para Sistemas de Inteligência Artificial de Propósito Geral e Generativa
> Art. 29. O desenvolvedor de sistemas de IA de propósito geral e generativa deverá realizar, além da documentação pertinente sobre o desenvolvimento do sistema, sua avaliação preliminar, a fim de identificar seus respectivos níveis de risco esperados, inclusive potencial risco sistêmico.
> Parágrafo único. A avaliação preliminar deverá considerar as finalidades de uso razoavelmente esperadas e os critérios previstos, nos termos da Seção III do Capítulo III desta Lei.
> Art. 30. O desenvolvedor de sistemas de IA de propósito geral e generativa com risco sistêmico, deve, antes de sua disponibilização ou introdução no mercado para fins comerciais, garantir o cumprimento dos seguintes requisitos:
> I – Descrever o modelo de IA de finalidade geral;
> II – Documentar os testes e as análises realizados, a fim de identificar e gerenciar riscos razoavelmente previsíveis, conforme apropriado e tecnicamente viável;
> III – Documentar os riscos não mitigáveis remanescentes após o desenvolvimento;

> IV – Processar e incorporar apenas conjuntos de dados coletados e tratados em conformidade com as exigências legais e sujeitos a uma adequada governança de dados, em especial quando se tratar de dados pessoais, de acordo com a Lei nº 13.709, de 14 de agosto de 2018 (Lei Geral de Proteção de Dados Pessoais) e o Capítulo II desta Lei;
> V – Publicar resumo do conjunto de dados utilizados no treinamento do sistema, nos termos de regulamentação;
> VI – Conceber e desenvolver os sistemas de IA de propósito geral e generativa recorrendo às normas aplicáveis para, considerando o contexto de uso, reduzir a utilização de energia, a utilização de recursos e os resíduos, bem como para aumentar a eficiência energética e a eficiência global do sistema;
> VII – elaborar documentação técnica e instruções de utilização inteligíveis, a fim de permitir que os desenvolvedores, distribuidores e aplicadores tenham clareza sobre o funcionamento do sistema.
> §1º O cumprimento dos requisitos estabelecidos neste artigo independe de o sistema ser fornecido como modelo autônomo ou incorporado a outro sistema de IA ou a produto, ou fornecido sob licenças gratuitas e de código aberto, como serviço, assim como por meio de outros canais de distribuição.
> §2º Os desenvolvedores de sistemas de IA de propósito geral e generativa poderão formular códigos de boas práticas, ou aderir a eles, para demonstrar conformidade às obrigações estipuladas neste artigo.
> Art. 32. Os desenvolvedores de sistemas de IA de propósito geral e generativa disponibilizados como recurso para desenvolvimento de serviços por terceiros, como aqueles fornecidos por meio de API ou outros modelos de integração, devem cooperar, na medida de sua participação, com os demais agentes de IA ao longo do período em que esse serviço é prestado e apoiado, a fim de permitir uma mitigação adequada dos riscos e o cumprimento dos direitos estabelecidos nesta Lei.
> Art. 33. Caberá à autoridade competente, em colaboração com as demais entidades do SIA, definir em quais hipóteses as obrigações previstas nesta Seção serão simplificadas ou dispensadas, de acordo com o risco envolvido e o estado da arte do desenvolvimento tecnológico.
> Parágrafo único. Aplica-se, no que couber, o disposto no Capítulo VI, cabendo à autoridade competente a aprovação de códigos de conduta e de autorregulação de sistemas de IA de propósito geral. (Brasil, 2024, pp. 17-18)

Apesar de entender a importância da definição do risco para determinar a aplicação ou não da regulação ao caso em específico, o presente trabalho, no entanto, debruçou-se em entender o agente responsável a partir do dano[69].

Quanto a isso, o texto do Substitutivo do PL nº 2.338/2023, aprovado no Senado Federal em dezembro de 2024, regula o seguinte a respeito da responsabilidade civil decorrente de danos causados por sistemas de IA:

> CAPÍTULO V DA RESPONSABILIDADE CIVIL
> Art. 35. A responsabilidade civil decorrente de danos causados por sistemas de IA no âmbito das relações de consumo permanece sujeita às regras de responsabilidade previstas na Lei nº 8.078, de 11 de setembro de 1990 (Código de Defesa do Consumidor), e na legislação pertinente, sem prejuízo da aplicação das demais normas desta Lei.
> Art. 36. A responsabilidade civil decorrente de danos causados por sistemas de IA explorados, empregados ou utilizados por agentes de IA permanece sujeita às regras de responsabilidade previstas na Lei nº 10.406, de 10 de janeiro de 2002 (Código Civil), e na legislação especial, sem prejuízo da aplicação das demais normas desta Lei.
> Parágrafo único. A definição, em concreto, do regime de responsabilidade civil aplicável aos danos causados por sistemas de IA deve levar em consideração os

[69] Schuett (2021), opta por concentrar-se na análise do escopo da matéria a ser regulamentada, mas reconhece a importância de se analisar a quem se aplica a matéria regulamentada: “Um desafio enfrentado por todos os formuladores de políticas que trabalham com regulamentação de IA é como definir o escopo de aplicação, que determina se uma regulamentação é ou não aplicável em um caso específico. O escopo de aplicação define o que é regulamentado (escopo material), quem é regulamentado (escopo pessoal), onde o regulamento se aplica (escopo territorial) e quando se aplica (escopo temporal)” (Schuett, 2021, p. 2, traduzido pela autora).

> seguintes critérios, salvo disposição legal em sentido contrário:
> I – O nível de autonomia do sistema de IA e o seu grau de risco, nos termos disciplinados por esta Lei;
> II – A natureza dos agentes envolvidos e a consequente existência de regime de responsabilidade civil próprio na legislação.
> Art. 37. O juiz inverterá o ônus da prova quando a vítima for hipossuficiente ou quando as características de funcionamento do sistema de IA tornarem excessivamente oneroso para a vítima provar os requisitos da responsabilidade civil.
> Art. 38. Os participantes no ambiente de testagem da regulamentação da IA continuam a ser responsáveis, nos termos da legislação aplicável, por quaisquer danos infligidos a terceiros como resultado da experimentação que ocorre no ambiente de testagem.
> Art. 39. As hipóteses de responsabilização previstas por legislação específica permanecem em vigor. (Brasil, 2024)

A última versão do texto do PL nº 2.338/2023 direciona a temática da responsabilidade civil para o direito privado brasileiro, o que não se verificava na versão inicial[70], na qual o legislador pretendia regular as causas e sujeitos responsáveis pelo dano. A própria civilística deve ser capaz de regular temas que naturalmente estão previstos em seu arcabouço, ainda que decorram de situações totalmente inéditas para o ordenamento jurídico. A despeito disso, discorrem Tepedino e Silva (2019).

> A rigor, a enunciação de novo ramo do direito voltado especificamente para as questões da robótica e da inteligência artificial traz consigo o grave risco de tratamento assistemático da matéria. Os fundamentos para a tutela das vítimas de danos injustos não devem ser buscados em novos e esparsos diplomas normativos, mas sim – e sempre – no ordenamento jurídico em sua unidade e complexidade. A disciplina ordinária da responsabilidade civil – tanto em relações paritárias quanto em relações de consumo –, embasada na tábua axiológica constitucional, serve de fundamento suficiente para o equacionamento dos problemas referentes aos danos causados por sistemas autônomos. Advirta-se, por oportuno: o tratamento sistemático ora propugnado deve levar em consideração o ordenamento jurídico em sua unidade e complexidade, sem se cair na armadilha da enunciação de um (mais um chamado micro) sistema próprio de valores da *lex robotica*.
> Nesse contexto, a enunciação de supostos vazios normativos representa problema muito mais grave do que o mero abalo à dogmática consolidada na tradição jurídica. Com efeito, ao afrontar a unidade e a completude do ordenamento, a indicação insistente de lacunas finda por comprometer a própria efetividade da tutela prometida às vítimas de danos injustos, como se das suas necessidades não desse conta o sistema ora vigente. Em vez de buscar – muitas vezes irrefletida – novas soluções e novos diplomas legais, melhores resultados se haverão de alcançar pelo esforço de releitura dos institutos já conhecidos pela civilística (Tepedino; Silva, 2019, pp. 70-71).

Tepedino e Silva (2019) argumentam que, apesar da lacuna existente na disciplina da responsabilidade civil em relação às novas tecnologias, é possível regular as questões relacionadas à inteligência artificial através das disposições do direito privado. Eles ressaltam

---

[70] "CAPÍTULO V DA RESPONSABILIDADE CIVIL
Art. 27. O fornecedor ou operador 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.
§1º Quando se tratar de sistema de inteligência artificial de alto risco ou de risco excessivo, o fornecedor ou operador respondem objetivamente pelos danos causados, na medida de sua participação no dano.
§2º Quando não se tratar de sistema de inteligência artificial de alto risco, a culpa do agente causador do dano será presumida, aplicando-se a inversão do ônus da prova em favor da vítima." Versão inicial do texto do Projeto de Lei 2.338/2023. Disponível em: https://legis.senado.leg.br/sdleg-getter/documento?dm=9347593&ts=1742240889254&disposition=inline. Acesso em 17 mar. 2023.

que, diante da ausência de normas específicas, o intérprete deve buscar soluções fundamentadas nos valores do ordenamento jurídico vigente, reconhecendo que o ineditismo das questões tecnológicas não implica, necessariamente, a necessidade de inovações legislativas.

Nesse sentido, pretende-se analisar os atuais dispositivos que regulam a matéria em âmbito brasileiro *vis-à-vis* a proposta de responsabilização civil por danos causados por sistemas de IA de Beckers e Teubner (2021).

Regula o ordenamento jurídico brasileiro que, para se justificar a responsabilidade civil, é necessário que se verifique a existência de um dano (por ato ilício, segundo art. 927, caput do CC, ou lícito, por disposição do art. 927, parágrafo único do CC) que deverá ser reparado por aquele que lhe deu causa. Tem-se, com isso, o nexo causal, que é justamente o vínculo entre a conduta de um agente e o resultado danoso dessa conduta, sendo necessário provar que o dano sofrido pela vítima é consequência direta da ação ou omissão do agente para que haja a configuração da responsabilidade civil.

Quando o dano é causado por ato ilícito estará configurada a responsabilidade civil subjetiva, que, além da ação ou omissão, do dano e do nexo de causalidade, deverá estar presente, também, a demonstração da culpa ou dolo do agente. Já quando o dano decorrer de ato que não infrinja a ordem jurídica, sem necessidade de configuração de dolo ou culpa, mas que implicar em risco para outrem, restará constituída a responsabilidade civil objetiva, regulada pelos artigos 931 e 932 do CC. O artigo 931, inclusive, segundo Maria Helena Diniz (2020), está embasado no Princípio da Equidade[71].

Em termos de responsabilidade civil objetiva, a Teoria do Risco[72], projetada pela doutrina, baseia-se em cinco vertentes que parecem não ser capazes amoldar-se a um risco em virtude de desenvolvimento ou uso de modelos e de sistemas de IA generativa. Isso em razão não só da indefinição do sujeito a quem o risco pode ser imputado (para configuração do risco criado, a culpa seria exclusivamente imposta ao desenvolvedor ou à IA, considerada como ente autônomo?) e da insegurança jurídica de imputar-se a responsabilidade somente a algum agente[73], como também da abordagem extremamente rígida do nexo causal, entendendo-se

[71] Para a autora, aquele que lucrar com um acontecimento deverá responder pelos riscos e desvantagens dele resultantes. Segundo ela, a equidade permite ao juiz adaptar a aplicação da lei às particularidades de cada caso, buscando uma solução justa e equilibrada. A equidade, nesse sentido, funciona como um complemento ao direito positivo, permitindo que o julgador considere fatores específicos e circunstâncias excepcionais que possam não estar previstos na legislação. Isso é especialmente relevante em situações em que a aplicação estrita da lei poderia resultar em injustiça ou desproporcionalidade.

[72] Segundo a doutrina, a Teoria do Risco está pautada nas seguintes vertentes:
Risco Criado (a responsabilidade é atribuída a quem cria um risco, independentemente de obter ou não vantagem com isso); Risco Proveito (baseia-se no princípio de que quem se beneficia de uma atividade deve também suportar os ônus decorrentes dela); Risco Integral (a responsabilidade é atribuída independentemente de qualquer excludente de responsabilidade, como culpa exclusiva da vítima, caso fortuito ou força maior); Risco Administrativo (o Estado responde pelos danos causados por seus agentes, independentemente de culpa, desde que haja nexo causal entre a atividade administrativa e o dano); e Risco Social (a responsabilidade é atribuída para garantir a reparação de danos que afetam a sociedade como um todo).

[73] Pelo risco proveito, pela própria cadeia produtiva da IA generativa, o proveito econômico poderia ser identificado

como risco tudo o que esteja ligado à cadeia produtiva e de utilização da IA generativa, desincentivando a inovação tecnológica no país.

A rigidez de se vislumbrar risco em qualquer modelo ou sistema de IA generativa aproxima-se da responsabilidade objetiva por risco do empreendimento, regulada pelo art. 927, parágrafo único, do Código Civil, e à responsabilidade solidária do fabricante, fornecedor e desenvolvedor (arts. 12 e 18 do CDC). Nesse sentido, entende a normativa que seria atribuída a responsabilidade civil objetiva ao fabricante do modelo ou do sistema em virtude do risco do desenvolvimento, abarcando não só o funcionamento inadequado, como também a insegurança no uso por outra pessoa que não o fabricante.

Fato é que a responsabilidade civil, em face da IA generativa, apresenta desafios significativos, especialmente no que tange ao nexo causal. Conforme apontam Tepedino e Silva (2019), a complexidade das interações entre sistemas autônomos dificulta a identificação dos agentes responsáveis por danos. À medida que as redes inteligentes se tornam mais interconectadas, a identificação dos responsáveis por ações danosas se torna cada vez mais nebulosa, gerando incertezas não só para vítimas, como também insegurança jurídica para todos os atores envolvidos na cadeia de produção do modelo ou do sistema de IA generativa.

Tratando-se das discussões doutrinárias brasileiras, diversas são as teorias que têm sido propostas para elucidar o nexo causal, destacando-se, segundo Tepedino e Silva (2019), a teoria da equivalência das condições, a teoria da causalidade adequada e a teoria da causa direta e imediata. A teoria da causa direta e imediata, conforme os autores, tem sido amplamente aceita, estabelecendo que o dever de indenizar deve ser atribuído ao agente cuja conduta é a causa direta do dano.

À luz do entendimento de Beckers e Teubner (2021), a seleção de uma das teorias da causa para justificação do nexo causal[74] não se justificaria porque, para eles, a identificação da causa está centrada na identificação primária da instituição sociodigital, e não no evento em si. Analisando-se detidamente as teorias da causa no ordenamento civil brasileiro, estaríamos diante de três hipóteses: ou atribuir-se-ia uma mesma relevância jurídica às condições que contribuíram para a ocorrência do resultado; ou buscar-se-ia a causa mais adequada para o resultado; ou, por fim, concentrar-se-ia na causa mais direta e imediata, desconsiderando a intervenção de outros fatores que levam diretamente ao resultado.

Em sentido contrário, dispõem Beckers e Teubner (2021, p. 154, traduzido e realçado pela autora):

Para distinguir entre assistência digital e hibridismo digital, a relação entre o

em mais de um agente. Pelos riscos integral e social, em contrapartida, somente um agente poderia responder pelo dano causado por toda a cadeia de produção e utilização da tecnologia.

[74] No contexto brasileiro, a discussão sobre identificação da origem da causa, para configuração de nexo causal, diz respeito tão somente à aplicação da responsabilidade civil subjetiva.

> **humano e o algoritmo precisa ser analisada de perto. Um primeiro critério bastante nítido é se o ato ilícito pode ou não ser atribuído com certeza a uma decisão individualizada. Um segundo critério é a densidade da cooperação entre o humano e o algoritmo. Isso pode ser concretizado perguntando se a interação humano-algoritmo ocorre sob a égide de um esforço cooperativo. Um terceiro critério é se existe uma igualdade de cooperação ou uma assimetria de instrução unilateral. Por exemplo, quando humanos e algoritmos cooperam no jornalismo digital, distinguir entre responsabilidade indireta e responsabilidade corporativa exigiria responder às seguintes perguntas:** A ação causadora de danos pode ser atribuída à operação algorítmica ou ao comportamento humano? Em que contexto o algoritmo foi utilizado, ou seja, como ferramenta auxiliar ao jornalismo investigativo humano ou como participante das escolhas jornalísticas? E qual era a relação entre humano e algoritmo em termos de supervisão humana? É certo que, devido à gradualidade envolvida, permanece uma área cinzenta. Como sempre acontece com a gradualidade, ela precisa de uma decisão binária em um ponto da escala móvel. Mas todos esses critérios são igualmente exigidos no mundo quando a lei traça a linha entre uma relação principal-agente e uma parceria.
>
> Ao distinguir entre assistência digital e interconectividade, surgem diferentes questões. Aqui, não é tanto a relação entre humanos e algoritmos que está em jogo, mas a identificação do ato ilícito dentro de toda uma série de operações algorítmicas. A responsabilidade indireta se aplica quando um algoritmo conectado a uma rede algorítmica comete um ato ilícito ou este é um caso de interconectividade? Se o ato ilícito puder ser claramente atribuído ao algoritmo a quem a tarefa foi delegada, aplica-se a responsabilidade indireta. Caso contrário, apenas as soluções de fundos resolverão o problema da responsabilidade. Se não estiver em vigor um regime de fundos, então, lamentavelmente, não haverá indenização para as vítimas. Uma vez que os regimes de fundos não existem com frequência, a procura de falhas algorítmicas identificáveis tornar-se-á ainda mais urgente.

Ou seja, à medida em que a doutrina brasileira busca pela identificação do evento que deu causa ao dano para, assim, atribuí-lo a um responsável, Beckers e Teubner (2021) entendem pelo caminho inverso, identificando a instituição sociodigital para, então, constatar o ato ilícito. Resumidamente, pela teoria de Beckers e Teubner (2021): identifica-se a instituição sociodigital, analisa-se o comportamento, atribui-se o risco e, então, imputa-se a responsabilidade civil. Aliás, outra diferença de entendimento é que, para esses autores, o ato ilícito será identificado a partir da análise de toda a cadeia produtiva e de uso da IA generativa.

Cabe sublinhar, ainda, que a importação da noção de "lei de agência" tal como desenhada por Beckers e Teubner (2021) – isto é, a outorga de uma *potestas vicaria* a agentes algorítmicos para que atuem em nome de um sujeito humano – conflita frontalmente com a dogmática brasileira. No Código Civil de 2002, a agência, no sentido clássico, é tratada por intermédio das figuras do mandato, da representação e dos atos praticados por prepostos (arts. 653-674 e 115-118), todos pressupostos em torno de dois elementos estruturantes: (i) a existência de um representante dotado de capacidade civil e (ii) a manifestação de vontade do representado – pessoa natural ou jurídica – a quem são imputados, em última análise, os efeitos do ato.

O algoritmo, porém, carece de personalidade e de vontade próprias; tampouco se enquadra na qualidade de "órgão" da pessoa jurídica, como o fazem administradores ou prepostos, porque não está submetido à direção hierárquica nem integra o estatuto da empresa

nos moldes do art. 47 do CC. Reconhecer-lhe agência autônoma exigiria, pois, criar um terceiro gênero de sujeito (a "pessoa eletrônica") ou admitir representação sem representante, hipótese incompatível com o sistema brasileiro de imputação, que ancora a responsabilidade na capacidade volitiva do agente.

Por esses motivos, adotar a lei de agência algorítmica significaria romper com a matriz civil-constitucional que exige que toda responsabilização decorra de ato humano ou de pessoa jurídica validamente constituída, implicando alteração legislativa de grande monta.

Outro aspecto relevante, segundo Tepedino e Silva (2019), é a possibilidade de exclusão da responsabilidade do desenvolvedor de inteligência artificial, com base no conceito de risco do desenvolvimento que sugere que, se o desenvolvedor utilizou as tecnologias mais seguras disponíveis no momento de sua criação, ele pode ser isento de responsabilidade por danos resultantes de falhas que não poderiam ser previstas. Assim, temem os autores que em um cenário onde novas soluções mais seguras podem surgir após a implementação de um sistema possa reverberar na exclusão da responsabilidade dos desenvolvedores.

Nas formas de responsabilização propostas por Beckers e Teubner (2021), parece que essa questão não seria aventada. Isso ocorre porque os riscos associados ao comportamento das instituições sociodigitais estão mais bem delineados. Para a instituição assistência digital, o risco será determinado pelo código-fonte e design do algoritmo. Na instituição associação homem-máquina pela teoria da rede de atores, levando-se em consideração a atuação do agente digital em nome do humano. Por fim, na instituição interconectividade digital, o risco seria emergente e resultado de interações complexas entre os múltiplos agentes digitais.

Passando-se à análise do tipo de responsabilidade civil aplicável, Tepedino e Silva (2019) argumentam que, embora a responsabilidade subjetiva possa ser considerada, a responsabilidade objetiva se mostra mais adequada para a proteção das vítimas, justificada pela natureza autônoma dos sistemas de inteligência artificial, que, apesar de não possuírem personalidade jurídica, podem causar danos significativos. Assim, a análise da conduta do desenvolvedor e do usuário se torna essencial para a atribuição de responsabilidades, considerando o grau de intervenção de cada um, sendo o grau de intervenção do usuário um fator também determinante para a responsabilização[75].

Para Beckers e Teubner (2021), a cada instituição sociodigital já estaria atribuída uma responsabilidade. Assim, a responsabilidade indireta (poderia ser uma hipótese de responsabilidade objetiva no ordenamento civil brasileiro?) seria atribuída à instituição assistência digital, a responsabilidade empresarial (poderia ser entendida como a

[75] Apesar dos autores assumirem uma condição de ação autônoma aos agentes digitais, ao falarem de responsabilidade civil, automaticamente direcionam a responsabilidade para o desenvolvedor ou para o usuário. É justamente esse tipo de associação direta, sem análise da multiplicidade de indivíduos e condutas, que o presente estudo refuta.

responsabilidade solidária no ordenamento civil brasileiro?) aos membros da cadeia da IA (homens e máquinas) e a responsabilidade de fundos coletivos aos setores da indústria envolvidos (não se vislumbra tratamento parecido na doutrina brasileira).

Beckers e Teubner (2021), ademais, refutam a priorização de um regime de responsabilização em detrimento de outro. Veja-se:

> Não concordamos com os autores que priorizam um regime de responsabilidade em detrimento de outro, como aqueles que defendem uma prioridade geral para soluções de fundos ao lidar com externalidades algorítmicas. Apontando para a imprevisibilidade do comportamento do computador, alguns autores propõem soluções de seguros e fundos sempre que algoritmos autônomos estão envolvidos. Em alternativa, aqueles que recomendam uma personificação completa dos algoritmos como pessoas eletrônicas enfrentam o problema de saber como concretizar a sua capacidade financeira para pagar indenizações. nós, eles estão sob pressão para introduzir seguros obrigatórios ou fundos de compensação. Por último, todos aqueles que pretendem aplicar a responsabilidade objetiva por riscos industriais em todos os casos de falhas algorítmicas tendem a recomendar simultaneamente soluções obrigatórias de seguros ou de fundos. Todos esses autores mostram uma atitude mais do que generosa em relação à responsabilidade coletiva compulsória sem boas razões (Beckers; Teubner, 2021, p. 155, traduzido pela autora).

Em relação à análise de Tepedino e Silva (2019), entendem os autores que a responsabilidade civil, tal como é concebida atualmente, não é insuficiente para lidar com as evoluções tecnológicas, incluindo a inteligência artificial generativa. Os autores defendem que, apesar das inovações, a estrutura jurídica existente pode ser adaptada para garantir a proteção das vítimas e a responsabilização adequada dos agentes envolvidos.

Entretanto, contrapondo a posição assumida por Tepedino e Silva (2019), quando atribuem a configuração de atividade de risco de sistemas de IA à individualização de imputação de responsabilidade, à proposição de Beckers e Teubner (2021), quanto à figura de instituições sociodigitais, estaremos diante do conflito de se atribuir a responsabilidade civil exclusivamente a uma pessoa física e/ou jurídica, forma da atual conjuntura privada brasileira.

Ademais, a investigação detida de cada atividade de risco, à luz da responsabilidade objetiva, conforme disposto por Tepedino e Silva (2019) e de acordo com o funcionamento do ordenamento jurídico brasileiro, também não vai de encontro ao entendimento de Beckers e Teubner (2021), que entendem pela necessidade de se distinguir situações típicas de responsabilidade de acordo com os diferentes riscos que elas produzem. Observe-se, abaixo, a diferença entre a compreensão dos quatro autores:

> Como se percebe, o reconhecimento da configuração de atividades de risco a partir do emprego generalizado de sistemas de inteligência artificial parece a solução adequada, em linha de princípio, para o equacionamento da questão atinente à individualização do critério de imputação do regime de responsabilidade. O que não parece possível, ao revés, é a invocação indiscriminada e irrefletida da noção de atividade de risco. Deve-se, com efeito, lançar mão dos critérios desenvolvidos pela doutrina para a elucidação do que vem a ser atividade de risco para fins de incidência da correlata cláusula geral de responsabilidade objetiva. Há que se investigar detidamente, em cada atividade, à

> luz das especificidades dos respectivos sistemas e de seu contexto, a possibilidade de caracterização de atividade de risco (Tepedino; Silva, 2019, p. 84).

> Uma regra geral para responsabilidade objetiva pode revelar-se especialmente difícil descrever as particularidades ou o grau necessário de um risco especial e extraordinário de uma forma suficientemente precisa (a fim de evitar, por exemplo, que qualquer utilização de IA em um smartphone possa ser coberta por responsabilidade objetiva) (Beckers; Teubner, 2021, p. 140, traduzido pela autora).

Por fim, reforça-se, ainda, a limitação do próprio CDC em relação a situações conflituosas envolvendo IA generativa. Apesar do que foi disposto, abaixo, por Tepedino e Silva (2019), no sentido de atribuição da responsabilidade objetiva aos fornecedores dos sistemas, com base no artigo 14 do CDC, cabe a análise de situações em que a própria IA não apresente qualquer tipo de falha, tendo o dano advindo do próprio usuário. Nessa hipótese, resta descartada a atribuição ao fornecedor de "reparação dos danos causados aos consumidores por defeitos relativos à prestação dos serviços" (art. 14, caput, CDC). Da mesma forma, também se torna comprometida a imputação de responsabilidade ao fornecedor "por informações insuficientes ou inadequadas sobre sua fruição e risco [da prestação do serviço]" (art. 14, caput, CDC). Desse modo, não só é praticamente impossível fazer com que o conhecimento técnico a respeito da ferramenta chegue ao usuário (consumidor final, nessa hipótese), haja vista que estamos diante de uma cadeia produtiva, como pelo fato de esse conhecimento exigir um nível considerável de estudo na área pelos próprios usuários, o que acabaria por afastar o uso da ferramenta pela maior parte dos consumidores.

> Aduza-se, ainda, à possibilidade de aplicação do regime da responsabilidade pelo fato do produto ou serviço previsto pelo Código de Defesa do Consumidor (CDC). Afinal, a inteligência artificial pode ser utilizada no âmbito de atividades de fornecimento de produtos ou serviços ao mercado de consumo. Caso se configure relação de consumo à luz da disciplina do CDC, torna-se induvidosa a possibilidade de responsabilização de todos os fornecedores integrantes da cadeia de consumo pelos danos decorrentes de fato do produto ou serviço – resguardada, em qualquer caso, a necessidade de aferição dos demais elementos relevantes para a deflagração do dever de indenizar.
>
> A novidade na matéria residiria, segundo parte da doutrina, na possibilidade de imputação do dever de indenizar também aos desenvolvedores de *softwares* ou algoritmos, e não apenas ao elo final da cadeia de fornecedores. Não se trata, contudo, ao menos no direito brasileiro, de conclusão propriamente inédita: conforme estabelecido pelos arts. 12 e 14 do CDC, a regra é a submissão dos variados fornecedores (incluindo o comerciante, guardadas as especificidades de sua responsabilização previstas pelo art. 13 do diploma) ao regime de responsabilidade objetiva pelos danos causados aos consumidores. Uma vez mais, o ineditismo parece estar não na solução jurídica, mas tão somente nas novas manifestações dos avanços tecnológicos sobre o cotidiano das pessoas (Tepedino; Silva, 2019, pp. 84-85).

Um dispositivo legal proposto no Anteprojeto de Lei para Revisão e Atualização do Código Civil de 2002, inclusive, diante do que exposto acima, pode, eventualmente, expor um

vulnerável a uma condição de vulnerabilidade ainda mais extrema[76]. Veja-se:

> Art. . Pessoas naturais que interagirem, por meio de interfaces, com sistemas de inteligência artificial, incorporados ou não em equipamentos, ou que sofrerem danos decorrentes da operação desses sistemas ou equipamentos, têm o direito à informação sobre suas interações com tais sistemas, bem como sobre o modelo geral de funcionamento e critérios para decisão automatizada, quando esta influenciar diretamente no seu acesso ou no exercício de direitos, ou afetar seus interesses econômicos de modo significativo (Anteprojeto de Lei para Revisão e Atualização do Código Civil de 2002, n. p.).

Pela leitura do texto proposto, subentende-se que, na medida em que o usuário tomar conhecimento sobre o funcionamento do sistema de IA (o que, como já reforçado, não possibilitará plena compreensão devido à limitação de conhecimentos técnicos adjacentes) e aceitar utilizá-lo, ao sistema não recairia qualquer responsabilidade.

5.4 Síntese do estudo em relação ao tratamento legislativo dos modelos generativos e da responsabilidade civil

O Projeto de Lei nº 2.338/2023, ao abordar a regulação da inteligência artificial, de forma ampla ou restrita à IA de propósito geral e generativa, carece de uma distinção clara entre os conceitos técnicos de "modelo de IA" e "sistema de IA", uma lacuna que pode ensejar uma interpretação equivocada das diretrizes regulatórias, uma vez que um modelo de IA é frequentemente importado e utilizado como base para desenvolvimento de sistemas de IA brasileiros. A falta de clareza conceitual pode resultar em uma aplicação inadequada das normas, já que a integração de um modelo a um sistema requer uma coerência técnica que não pode ser ignorada.

Assim, a definição precisa e a adequação dos conceitos são essenciais para garantir que as diretrizes regulatórias não apenas sejam aplicáveis, mas também efetivas na promoção de um ambiente ético e não discriminatório, evitando que a ambiguidade técnica comprometa a segurança jurídica em um campo em rápida evolução, como o da inteligência artificial.

A inserção do Brasil nas discussões globais sobre a regulação da IA (e, consequentemente, da IA generativa) é uma necessidade premente, especialmente à luz da crescente interdependência econômica e tecnológica entre nações. A Estratégia Brasileira de Inteligência Artificial, conforme disposto em seu artigo 44, enfatiza a importância de um trabalho conjunto entre os países para o desenvolvimento de diretrizes regulatórias que promovam a inovação e a ética no uso da IA (EBIA, 2021). Essa harmonização legislativa não

[76] A vulnerabilidade do usuário pode ser entendida a partir do conceito de opacidade de Burrell (2016), já mencionado na nota de rodapé 32 deste trabalho.

apenas facilitará a cooperação internacional, mas também permitirá que o Brasil se posicione como um ator relevante no cenário global, contribuindo para a construção de um ambiente regulatório que respeite os direitos humanos e promova a inclusão digital.

Além disso, a adoção de normas que estejam alinhadas com legislações internacionais, como o AI Act, pode proporcionar maior segurança jurídica para empresas e usuários, incentivando investimentos e a adoção responsável de tecnologias emergentes. Portanto, a participação ativa do Brasil nas discussões sobre a regulação da IA é fundamental para garantir que o país não apenas acompanhe as tendências globais, mas também influencie a formação de um marco regulatório que reflita seus valores e necessidades sociais.

Nesse sentido, argumenta Souza (2024, pp. 21-22), acerca da responsabilidade civil por atos danosos de IA e equidade processual, que a desterritorialidade das interações digitais e a falta de legislação adequada tornam a imputabilidade da responsabilidade civil mais complexa, na medida em que empresas fornecedoras desses serviços estão sediadas no exterior, dificultando a proteção dos direitos digitais dos usuários brasileiros. A autora entende, ainda, que a harmonização das legislações entre diferentes nações é vital para garantir a segurança jurídica no uso de tecnologias digitais, sobretudo pelos mais vulneráveis (Souza, 2024, p. 22).

Ademais, entende o presente estudo que a proposta de criação da figura das instituições sociodigitais, conforme delineado por Beckers e Teubner (2021), ainda que consistente, não é capaz de materializar-se no ordenamento jurídico brasileiro no que tange ao contexto da responsabilização civil por danos decorrentes do uso de IA (geral, de propósito geral e, principalmente, a generativa).

No que tange à responsabilidade civil por danos causados por modelos ou sistemas de IA, conforme elucidado por Beckers e Teubner (2021), a proposta seria a atribuição a partir da identificação do(s) responsável(is) (instituições sociodigitais) pelo dano, que se iniciaria a partir da análise do comportamento individual ou coletivo dos autores em potencial. O grande desafio de extensão dessa teoria à realidade brasileira é que não só a responsabilidade civil privilegia a análise a partir do dano[77] em si como também o imputa a uma pessoa física e/ou jurídica.

Com relação à tutela da pessoa humana, o atual texto da normativa parece ser uma proposição eficaz quando analisado pela perspectiva do risco em abstrato. Em síntese, o que

[77] O que se mantém, inclusive, no Anteprojeto de Lei para Revisão e Atualização do Código Civil de 2002, na medida em que, o artigo proposto para tratamento dos direitos da personalidade em face de sistemas de IA, entende pela atribuição da responsabilidade através do princípio da reparação integral dos danos, conforme colacionado:
Art. . O desenvolvimento de sistemas de inteligência artificial deve respeitar os direitos de personalidade previstos neste Código, garantindo a implementação de sistemas seguros e confiáveis, em benefício da pessoa natural ou jurídica e do desenvolvimento científico e tecnológico, devendo ser garantidos:
(...)
IV - A atribuição de responsabilidade civil, pelo princípio da reparação integral dos danos, a uma pessoa natural ou jurídica em ambiente digital.

propõe o PL nº 2.338/2023 é que à autoridade competente[78] caiba a definição de hipóteses de avaliação preliminar (Brasil, 2025, art. 12, §2º), expedição de normativas gerais (Brasil, 2025, art. 16, I) e estabelecimento de critérios gerais e elementos para elaboração de avaliação de impacto algorítmico pelo fornecedor (Brasil, 2025, art. 25, §5º). Especificamente quanto à IA de propósito geral, caberá à autoridade competente a aprovação de códigos de conduta e de autorregulação de sistemas de IA desse tipo (Brasil, 2025, art. 33, § único).

Quanto à tutela de direitos fundamentais, em ampla análise, entende-se que o PL nº 2.338/2023 está alinhado tanto com o diploma legal da LGPD quanto com o Código Civil e seu Anteprojeto. Todos os textos estão harmonizados no que diz respeito à não discriminação, uso de dados visando à proteção de dados, sobretudo os pessoais, sobre os quais pairam os princípios de igualdade e de privacidade, bem como às condições de transparência, auditabilidade, explicabilidade, rastreabilidade, supervisão humana e governança.

Contudo, analisando o texto sob a perspectiva do dano em concreto, parece-nos que, pela própria limitação do tratamento da responsabilidade civil à matéria, os princípios e os direitos norteadores da tutela humana e seus direitos ainda estão distantes de serem eficazes para reparação, ao indivíduo, do dano causado em razão da utilização da IA generativa, o que viola diretamente certos direitos fundamentais.

Quando se fala em IA generativa, e como exaustivamente retratado nesse estudo, o Brasil coloca-se na posição primária de importador de modelos de IA para que possam estruturar seus sistemas de IA generativa. Com isso, ainda que o fornecedor se adeque à proposta legislativa em trâmite, apresentando o relatório de avaliação de impacto algorítmico, ele não será capaz de explicar, por exemplo, qual a base de dados utilizada para treinamento do modelo, nem sequer garantir que não houve mitigação a direitos fundamentais e pretensa discriminação. A isso ainda está conectada a violação do direito à privacidade, considerando que o modelo pode carecer de consentimento para utilização de dados pessoais para treinamento, como também ferir direitos autorais.

Nesse sentido, é imperioso retomar o ponto da harmonização das legislações globais. O GDPR, na União Europeia, oferece mais segurança em relação à utilização de modelos de IA desenvolvidos naqueles países em virtude de estarem subordinados a uma normativa forte de proteção de dados, o que não se verifica nos EUA, por exemplo, ainda mais após Trump ascender à presidência novamente. Fato é que os EUA ainda se mantêm como maior produtor de tecnologia de IA generativa, consolidando as *Big Techs* em seu território. Esses são os modelos generativos que estão servindo de base para a construção de sistemas ao redor do

[78] Com a entrada em vigor do texto normativo, pode ser que a autoridade competente tenha que ter um papel mais atuante e próximo do fornecedor no que tange à avaliação de conformidade. Isso só se verificará na prática. Assim sendo, necessário que o país tenha, ainda que a título piloto, uma legislação vigente sobre a matéria, e que essa possa se modificar à medida em que a própria tecnologia exigir, o que será verificado com bastante frequência.

mundo. Cabe indagar-se se os procedimentos de coleta e armazenamento, processamento, compartilhamento e eliminação dos dados para treinamento desses modelos estão sendo respeitados em solo americano.

A urgência em relação à proteção da pessoa humana é ainda justificada na medida em que as empresas estão mais focadas no desenvolvimento tecnológico do que na garantia de direitos. Dados de uma pesquisa realizada pelo Laboratório de Tecnologia e Direito (TechLab), da Faculdade de Direito da Universidade de São Paulo, revelam que, apesar de 42% das regulações analisadas prezarem pela garantia de direitos, sobretudo em relação à proteção de dados e coibição de discriminação algorítmica, 32% delas ainda entendem pela prevalência do desenvolvimento tecnológico em relação a garantias de direitos.

Novamente, o contorno a essas situações somente será vislumbrado, na prática, com a entrada em vigor de dispositivo legal específico que regule a inteligência artificial generativa em âmbito brasileiro.

6. CONCLUSÃO

A presente pesquisa buscou explorar a complexidade da inteligência artificial generativa e sua interseção com o ordenamento jurídico brasileiro, culminando na análise do Projeto de Lei nº 2.338/2023.

No primeiro capítulo, foram introduzidos os conceitos fundamentais que moldam a discussão sobre a IA generativa, além da digressão histórica que fez com que esse tipo de tecnologia ganhasse ascensão exponencial em torno do mundo. A partir dessa base conceitual, a pesquisa se direcionou para a metodologia adotada, que se revelou adequada para a investigação qualitativa das implicações legais e sociais da IA.

No segundo capítulo, foram delineadas as diretrizes metodológicas que guiaram o desenvolvimento do estudo, destacando a importância de uma abordagem dedutiva para compreender as nuances do Projeto de Lei em questão. A análise qualitativa permitiu uma exploração aprofundada das relações sociais e jurídicas que emergem da implementação da IA generativa, revelando a urgência de um marco regulatório que se alinhe aos valores democráticos e à proteção dos direitos fundamentais.

O terceiro capítulo reforçou a importância de diferenciar conceitualmente os modelos dos sistemas de IA generativa, seja a partir da técnica atribuída pela Ciência da Computação ou dos fatores socioeconômicos que moldam a expansão dessa tecnologia. Essa análise evidenciou a necessidade de uma distinção clara entre "modelo de IA" e "sistema de IA", uma lacuna que pode comprometer a eficácia das diretrizes regulatórias. A falta de clareza conceitual pode gerar interpretações equivocadas e, consequentemente, uma aplicação inadequada das normas, o que reforça a importância de um arcabouço legal robusto e bem definido. Passou-se, ainda, ao exame conceitual da IA de propósito geral, conectando a monopolização de produção dos modelos generativos à hegemonia econômica da IA. Por fim, demonstrou como a expansão da inteligência artificial generativa tem violado direitos fundamentais.

No quarto capítulo, foram apresentadas as bases teóricas que sustentam a proposta de responsabilização dos agentes autônomos, com foco na noção de personalidade digital. A análise das contribuições de Beckers e Teubner (2021) trouxe à tona a relevância de se considerar a violação de direitos fundamentais no contexto da IA generativa, ressaltando a necessidade de um regime de responsabilização que proteja os indivíduos frente a possíveis abusos tecnológicos. A proposta de personalidade digital dos autores foi, inclusive, debatida à luz do entendimento de Negri (2016, 2020 e 2021).

O quinto capítulo explorou a adequação de tudo o que foi discutido nos capítulos anteriores ao ordenamento jurídico brasileiro, entendendo pela necessidade da diferenciação teórica entre modelos e sistemas de IA, pela lacuna que impossibilita a harmonização dos

conceitos de personalidade digital e a responsabilização civil conforme a teoria tríplice de Beckers e Teubner (2021) à ordem jurídica brasileira e, por fim, pela omissão do PL nº 2.338/2023 em regular a violação aos direitos fundamentais a partir do dano causado por IA generativa. .

Por fim, no sexto capítulo, foram apresentadas as considerações finais, que reiteraram a importância de uma legislação que não apenas regule essa tecnologia, mas também promova a centralidade da pessoa humana em sua implementação e seu uso. A pesquisa conclui que, para que o Brasil se insira de maneira efetiva nas discussões globais sobre a regulação da IA, é imperativo que a legislação em construção não apenas aborde os aspectos técnicos de maneira mais enfática, mas também considere as implicações sociais e éticas que emergem dessa nova realidade.

A adequação do ordenamento jurídico para recepção mais harmoniosa da IA generativa é um passo crucial para garantir um futuro em que a tecnologia sirva ao bem comum, respeitando os direitos e a dignidade de todos os indivíduos.

REFERÊNCIAS

AEPD. **Agencia Española Protección Datos. Synthetic data and data protection**. 2023. Disponível em https://www.aepd.es/en/prensa-y-comunicacion/blog/synthetic-dataand-data-protection. Acesso em: 11 jan. 2025.

A&O SHEARMAN. **A&O announces exclusive launch partnership with Harvey**. A&O Shearman website. 2023. Disponível em: https://www.allenovery.com/en-gb/global/news-and-insights/news/ao-announces-exclusive-launch-partnership-with-harvey. Acesso em: 7 jan. 2024.

A&O SHEARMAN. **ContractMatrix**. A&O Shearman website. 2025. Disponível em: https://www.aoshearman.com/en/expertise/markets-innovation-group/contractmatrix. Acesso em: 7 jan. 2024.

AMARAL, Gustavo Rick; XAVIER, Fernando. **A inteligência artificial e o novo patamar da interação humano-máquina**. TECCOGS – Revista Digital de Tecnologias Cognitivas, n. 26, jul./dez. 2022 pp. 6–43. Disponível em: dx.doi.org/10.23925/1984-3585.2022i26p6-43. Acesso em: 15 jan. 2025.

AMERSHI, S., CAKMAK, M., KNOX, W. B., KULESZA, T. **Power to the People: The Role of Humans in Interactive Machine Learning**. AI Magazine, 35(4), 105-120. 2014. https://doi.org/10.1609/aimag.v35i4.2513. Acesso em: 20 fev. 2025.

ANPD. **Radar Tecnológico nº 3: Inteligência Artificial Generativa**. Brasília, 2024. Disponível em: https://www.gov.br/anpd/pt-br/centrais-de-conteudo/documentos-tecnicos-orientativos/radar_tecnologico_ia_generativa_anpd.pdf. Acesso em: 23 jan. 2025.

BARBOSA, Mafalda. **Nas fronteiras de um admirável mundo novo? O problema da personificação de entes dotados de inteligência artificial**. 2021, pp. 251-285. In: BARBOSA, M. M., BRAGA NETTO, F.; SILVA, M. C.; FALEIROS JÚNIOR, J. L. de M. **Direito digital e inteligência artificial: diálogos entre Brasil e Europa**. 1ª Ed. Editora Foco, 2021, ePUB.

BARBOSA, M. M., BRAGA NETTO, F.; SILVA, M. C.; FALEIROS JÚNIOR, J. L. de M. **Direito digital e inteligência artificial: diálogos entre Brasil e Europa**. 1ª Ed. Editora Foco, 2021, ePUB.

BARROSO, Luís Roberto; MELLO, Patrícia Perrone Campos. **Inteligência artificial: promessas, riscos e regulação. Algo de novo debaixo do sol**. Rev. Direito e Práx., Rio de Janeiro, v. 15, n. 4, 2024, pp. 1-45. DOI: 10.1590/2179-8966/2024/84479. ISSN: 2179-8966. Acesso em: 5 fev. 2025.

BECKERS, Anna; TEUBNER, Gunther. **Three Liability Regimes for Artificial Intelligence: Algorithmic Actants, Hybrids, Crowds**. Edição Inglês. E-book Kindle. 2021.

BE MY EYES. **Be My Eyes: Levando a visão para pessoas cegas ou com visão limitada**. 2025. Disponível em: https://www.bemyeyes.com/language/portuguese-brazil. Acesso em: 13 mar. 2025.

BONFIM, Thiago. **Trump derruba regra que impedia discriminação em contratações do governo**. UOL. 2025. Disponível em: https://noticias.uol.com.br/internacional/ultimas-noticias/2025/01/22/trump-derruba-regra-discriminacao-dei-governo-eua.htm. Acesso em: 11 fev. 2025.

BOSTROM, N. **Superinteligência: Caminhos, Perigos, Estratégias**. Oxford, Oxford University Press, 2017.

BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil**. Brasília, 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 5 mar. 2025.

BRASIL, Lei nº 9.609, de 19 de fevereiro de 1998. **Lei de Software**. Brasília, 1998. Disponível em https://www.planalto.gov.br/ccivil_03/leis/l9609.htm. Acesso em: 6 maio 2025.

BRASIL, Lei nº 9.610, de 19 de fevereiro de 1998. **Lei de Direitos Autorais**. Brasília, 1998. Disponível em https://www.planalto.gov.br/ccivil_03/leis/l9610.htm. Acesso em: 6 maio 2025.

BRASIL. Lei n.º 10.406, de 10 de janeiro de 2002. **Código Civil**. Brasília, 2002. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/2002/l10406.htm. Acesso em: 5 mar. 2025.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. **Lei Geral de Proteção de Dados Pessoais**. Brasília, 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm. Acesso em: 5 mar. 2025.

BRASIL. **Projeto de Lei nº 2.338 de 2023**. Brasília, 2023. Disponível em: https://www25.senado.leg.br/web/atividade/materias/-/materia/157233. Acesso em: 5 mar. 2025.

BRASIL. **Minuta do Substitutivo aprovado pelo Senado**. Brasília, 2024. Disponível em: https://legis.senado.leg.br/sdleg-getter/documento?dm=9881643&ts=1738768184212&disposition=inline. Acesso em: 5 mar. 2025.

BROSTOM, S. **A Dynamic Learning Concept in Early Years' Education: A Possible Way to Prevent Schoolification**. International Journal of Early Years Education, 25, 3-15. 2017. Disponível em: https://www.scirp.org/journal/paperinformation?paperid=80795. Acesso em: 17 jan. 2025.

BROWN, S. **Machine Learning, Explained**. MIT Sloan School of Management, Cambridge. 2021. Disponível em: https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained. Acesso em: 17 fev. 2025.

BURRELL, J. **How the machine 'thinks': Understanding opacity in machine learning algorithms**. Big Data & Society. 2016. Disponível em: https://doi.org/10.1177/2053951715622512. Acesso em: 21 jan. 2025.

CARVALHO, M. **Advogado usa ChatGPT para identificar uso de IA em sentença e requer anulação**. JOTA. 2025. Disponível em: https://www.jota.info/justica/advogado-usa-chatgpt-para-identificar-uso-de-ia-em-sentenca-e-requer-anulacao. Acesso em: 16 mar. 2025.

CHOWDHURY, M.; APON, A.; DEY, K. **Data Analytics for Intelligent Transportation Systems**. English Edition. Elsevier, 2017. E-book Kindle.

CITRON, Danielle; SOLOVE, Daniel J. **Privacy Harms**. Boston University Law Review, v. 102, pp. 793-863, 2022.

COLLINS, Hugh. **Introduction to Networks as connected Contracts**. In: TEUBNER, Gunther, Networks as connected Contracts Oxford, Hart, 2011, pp. 1-4. Disponível em: https://www.jura.uni-frankfurt.de/88237515/NetworkConnectedCpntracts2011.pdf. Acesso em:: 14 jan. 2025.

COMETTI, M. T. **Interpretação Constitucional e Direito Ambiental nos EUA: Desafios e Oportunidades**. Legale Educacional. 2024. Disponível em: https://legale.com.br/blog/interpretacao-constitucional-e-direito-ambiental-nos-eua-desafios-e-oportunidades/. Acesso em:: 11 fev. 2025.

COMISSÃO EUROPEIA. **Diretrizes Éticas Para Inteligência Artificial Confiável**. Bruxelas, 2019. Disponível em: https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai. Acesso em: 10 jan. 2025.

CONTE, V. **Disinformazione digitale, fiduciarietà informativa, rimedi contrattuali.** In: Annali dela SISDiC 10/23. Edizioni Scientifische Italiane, 2023, pp. 259-298.

CRAWFORD, Kate. **Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence**. New Haven: Yale University Press, 2021. ISBN 978-0300209570.

DEEPSEEK. **Modelos de IA**. Disponível em: https://deepseek.com.br/topicos/modelos-de-ia/. Acesso em: 29 abr. 2025.

DEGLI-ESPOSTI, Sara. **La ética de la inteligencia artificial**. Madrid: Catarata, 2023.

DINIZ, Maria Helena. **Curso de Direito Civil Brasileiro: responsabilidade civil**. Vol. 7. 34. ed., rev. e atual. Saraiva jus. 2020.

DONEDA, Danilo. **Da Privacidade à Proteção de Dados Pessoais**. Rio de Janeiro: Renovar, 2006.

DREYFUS, Hubert L.; DREYFUS, Stuart E. **Making a Mind Versus Modeling the Brain: Artificial Intelligence Back at a Branchpoint**. Dædalus, v. 1, n. 117, pp. 15-44, 1988. Disponível em: https://www.amacad.org/sites/default/files/daedalus/downloads/Daedalus_Wi98_Artificial-Intelligence.pdf Acesso em:: 17 jan. 2025.

EHRHARDT JÚNIOR, Marcos; MILHAZES NETO, Antonio Luiz. **A questão da autoria e titularidade das obras criadas por inteligência artificial generativa de imagens e suas possibilidades no direito brasileiro**. Revista Brasileira de Direito Civil – RBDCivil, Belo Horizonte, v. 33, n. 3, pp. 301-324, jul./set. 2024. DOI: 10.33242/rbdc.2024.03.012. Acesso em:: 6 maio 2025.

ELOUNDOU, Tyna.; MANNING, Sam.; MISHKIN, Pamela.; ROCK, Daniel. **GPTs are GPTs: an early look at the Labor Market impact potential of Large Language Models**. 2023. Disponível em: https://arxiv.org/abs/2303.10130. Acesso em: 17 jun. 2024.

ESTADOS UNIDOS. **14th Amendment**. 1789. Disponível em: https://constitutioncenter.org/the-constitution/amendments/amendment-xiv. Acesso em: 9 jan. 2025.

EUROPEAN PARLIAMENT. **General-Purpose artificial intelligence**. 2023. Disponível em: https://www.europarl.europa.eu/RegData/etudes/ATAG/2023/745708/EPRS_ATA(2023)745708_EN.pdf. Acesso em: 2 fev. 2025.

EUROPEAN PARLIAMENT. **EU Artificial Intelligence Act.** 2024. Disponível em:

https://artificialintelligenceact.eu/ai-act-explorer/. Acesso em: 2 fev. 2025.

FARAGE da Costa Felipe, B., MULHOLLAND, C. **Filtro Bolha e Big Nudging: A Democracia na Era dos Algoritmos**. Revista Direitos Fundamentais & Democracia, 27(3), 06–18. (2022). Disponível em: https://doi.org/10.25192/issn.1982-0496.rdfd.v27i32275. Acesso em: 14 fev. 2025.

FINANCIAL TIMES. **Allen & Overy rolls out AI contract negotiation tool in challenge to legal industry**. 2023. Disponível em: https://www.ft.com/content/f1aff4d0-b2c5-4266-aa0a-604ef14894bb. Acesso em: 7 jan. 2024.

FINANCIAL TIMES. **OpenAI says it has evidence China's DeepSeek used its model to train competitor**. 2025. Disponível em: https://www.ft.com/content/a0dfedd1-5255-4fa9-8ccc-1fe01de87ea6. Acesso em: 11 fev. 2025.

FINANCIAL TIMES. **Why Nvidia investor are spooked by Chinese AI upstart DeepSeek**. 2025. Disponível em: https://www.ft.com/content/ee83c24c-9099-42a4-85c9-165e7af35105. Acesso em: 11 fev. 2025.

FLORIDI, L., CHIRIATTI, M. **GPT-3: Its Nature, Scope, Limits, and Consequences**. Minds & Machines 30, 681–694. 2020. Disponível em: https://doi.org/10.1007/s11023-020-09548-1. Acesso em: 9 fev. 2025.

FÓRUM ECONÔMICO MUNDIAL. **Comunicados à imprensa: Relatório sobre o Futuro dos Empregos 2025: 78 milhões de novas oportunidades de emprego até 2030, mas é preciso melhorar a qualificação das forças de trabalho urgentemente**. 2025. Disponível em: https://reports.weforum.org/docs/WEF_Future_of_Jobs_2025_Press_Release_PTBR.pdf. Acesso em: 5 maio 2025.

FRAZÃO, Ana. **Discriminação algorítmica – Parte II: compreendendo o que são os julgamentos algorítmicos e o seu alcance na atualidade**. In Portal JOTA. 2021. Disponível em: https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/discriminacao-algoritmica. Acesso em:: 4 maio 2025.

FRAZÃO, Ana. **Marco da Inteligência Artificial em análise: Já não foram mapeados riscos suficientes para justificar uma regulação adequada e com efeitos práticos?** In Portal JOTA. 2021. Disponível em: https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/marco-inteligencia-artificial. Acesso em: 4 maio 2025.

FUCHS, Christian. **Digital Humanism: A Philosophy for 21st Century Digital Society**. Kindle Edition. 2022.

GARTNER. **Gartner Glossary**. Disponível em: https://www.gartner.com/en/informationtechnology/glossary/foundation-models. Acesso em: 2 fev. 2025.

GERHARDT, T. E.; SILVEIRA, D. T. **Métodos de Pesquisa**. Série Educação a Distância. UFRGS Editora. 2009.

GÉRON, Aurélien. **Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow**. 2. ed. Sebastopol: O'Reilly Media, Inc., 2019. ISBN 9781492032649.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2008.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep Learning**. Cambridge, MA: The MIT Press, 2016. 800 pp. ISBN 978-0262035613.

GOODFELLOW, Ian J.; POUGET-ABADIE, Jean; MIRZA, Mehdi; XU, Bing; WARDE-FARLEY, David; OZAIR, Sherjil; COURVILLE, Aaron; BENGIO, Yoshua. **Generative Adversarial Networks**. 2014. DOI: 10.48550/arXiv.1406.2661. Acesso em: 5 fev. 2025.

GOOGLE. **Gemini**. Disponível em: https://deepmind.google/technologies/gemini/#introduction. Acesso em: 2 fev. 2025.

GOOGLE AI FOR DEVELOPERS. **Modelos do Gemini**. Disponível em: https://ai.google.dev/gemini-api/docs/models. Acesso em: 29 abr. 2025.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. New York: Springer, 2009.

HACKER, Philipp; ENGEL, Andreas; MAUER, Marco. **Regulating ChatGPT and other Large Generative AI Models**. FAccT '23, 12-15 jun. 2023, Chicago, IL, USA. Disponível em: https://doi.org/10.48550/arXiv.2302.02337. Acesso em: 17 mar. 2025.

HOPCROFT, J. E.; MOTWANI, R.; ULLMAN, J. D. **Introduction to Automata Theory, Languages, and Computation**. Boston: Pearson, 2006.

HU, Krystal. **ChatGPT sets record for fastest-growing user base – analyst note**. Reuters, 2 fev. 2023. Disponível em: https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/. Acesso em: 14 jan. 2024.

HUQ, Aziz. **Constitutional Rights in the Machine Learning State**. **Cornell Law Review**, v. 105, 2020. Disponível em: https://ssrn.com/abstract=3613282. Acesso em:: 11 fev. 2025.

IEA. **Data Centers and Data Transmission Networks**. 2025. Disponível em https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks. Acesso em: 11 fev. 2025.

JONES, Elliot. **Explainer: What is a foundation model?** Ada Lovelace Institute, 17 jul. 2023. Disponível em: https://www.adalovelaceinstitute.org/resource/foundation-modelsexplainer/#table1. Acesso em: 18 jan. 2025.

JOHNSON, Khari. **OpenAI debuts DALL-E for generating images from text**. VentureBeat. 2021. Disponível em: https://web.archive.org/web/20210105221534/https:/venturebeat.com/2021/01/05/openai-debuts-dall-e-for-generating-images-from-text/. Acesso em: 12 mar. 2025.

JURAFSKY, Daniel; MARTIN, James H. **Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models**. Stanford: Stanford University; 2021.

KAUFMAN, Dora. **Desmistificando a inteligência artificial**. 2022. Autêntica.

KORKMAZ, M. R. D. C. R. **Dados sensíveis na Lei Geral de Proteção de Dados Pessoais: Mecanismos de Tutela para o Livre Desenvolvimento da Personalidade**. Programa de Pós-graduação em Direito e Inovação da Faculdade de Direito da Universidade Federal de Juiz de Fora. 2019. Disponível em: https://repositorio.ufjf.br/jspui/handle/ufjf/11438. Acesso em: 7 jan. 2025.

KORKMAZ, Maria Regina Detoni Cavalcanti Rigolon. **Revisão de decisões automatizadas na**

Lei Geral de Proteção de Dados Pessoais. Tese (Doutorado em Direito Civil) – Faculdade de Direito, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2022. Disponível em: http://www.bdtd.uerj.br/handle/1/18276. Acesso em: 5 maio 2025.

LARA, Patricia. **Wall Street: Nasdaq despenca após entusiasmo com IA chinesa**. CNN Brasil. 2025. Disponível em: https://www.cnnbrasil.com.br/economia/mercado/bolsas-de-ny-fecham-sem-rumo-unico-com-tombo-do-nasdaq-desencadeado-por-deepseek/. Acesso em: 11 fev. 2025.

LENHARO, Mariana. **AI consciousness: scientists say we urgently need answers**. Nature, 21 dez. 2023.

LÉON, L. P. **Meta promete se aliar a Trump contra países que regulam redes sociais. Agência Brasil**. 2025. Disponível em https://agenciabrasil.ebc.com.br/internacional/noticia/2025-01/meta-promete-se-aliar-trump-contra-paises-que-regulam-redes-sociais. Acesso em: 11 fev. 2025.

LIDDY, E.D. **Natural Language Processing. In Encyclopedia of Library and Information Science**. 2001. 2nd Ed. NY. Marcel Decker, Inc. Disponível em: https://surface.syr.edu/istpub/63/. Acesso em: 2 maio 2025.

LOPES, André. **A promessa de Sam Altman que convenceu Trump sobre AGI: "Será lançada no seu mandato"**. Exame. 2025. Disponível em: https://exame.com/inteligencia-artificial/a-promessa-de-sam-altman-que-convenceu-trump-sobre-agi-sera-lancada-no-seu-mandato/. Acesso em: 11 fev. 2025.

LOPES, Giovana F. Peluso. **Inteligência artificial: Considerações sobre personalidade, agência e responsabilidade civil**. Editora Dialética. 2021.

LÓPEZ, Nuria. **EUA divulga novas diretrizes para IA: qual é o impacto disso no consumo?** Daniel Law. 2025. Disponível em: https://www.daniel-ip.com/pt/artigos/eua-divulga-novas-diretrizes-para-ia-qual-e-o-impacto-disso-no-consumo/#:~:text=Em%20resumo%2C%20a%20regulamenta%C3%A7%C3%A3o%20americana,Uni%C3%A3o%20Europeia%20est%C3%A1%20estabelecendo%20regulamenta%C3%A7%C3%B5es. Acesso em: 11 fev. 2025.

LOURENÇO, S. R. S. **A Responsabilidade Civil Extracontratual por Danos Causados por Inteligência Artificial Generativa**. Universidade de Coimbra. 2024. Disponível em: https://estudogeral.uc.pt/handle/10316/85894. Acesso em: 17 mar. 2025.

LUCCIONI, Alexandra Sasha; JERNITE, Yacine; STRUBELL, Emma. **Power Hungry Processing: Watts Driving the Cost of AI Deployment?** In: ACM Conference on Fairness, Accountability, and Transparency (ACM FAccT '24), 3-6 jun. 2024, Rio de Janeiro, Brasil. Disponível em: https://doi.org/10.48550/arXiv.2311.16863. Acesso em: 16 mar. 2025.

LUHMANN, N. **Theory of Society**. Volume 2. Stanford, Stanford University Press, 2012/2013. Ebook ISBN: 9780804787277.

MACHADO, José Ronaldo de Freitas. **Direitos humanos: síntese histórica dos direitos das minorias**. Revista Científica Multidisciplinar Núcleo do Conhecimento. Ano 06, Ed. 04, Vol. 03, pp. 130-147. Abril. ISSN: 2448-0959, DOI: 10.32749/nucleodoconhecimento.com.br/historia/sintese-historica. Acesso em: 12 jan. 2025.

MAYER, H; YEE, L.; CHUI, M.; ROBERTS, R. **Superagency in the workplace: Empowering people to unlock AI's full potential**. McKinsey Digital. 2025. Disponível em: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/superagency-in-the-workplace-empowering-people-to-unlock-ais-full-potential-at-work. Acesso em: 6 fev. 2025.

MELO, J. O. de. **Cortes dos EUA recorrem a multas para conter precedentes falsos gerados por IA**. Consultor Jurídico. 2025. Disponível em: https://www.conjur.com.br/2025-mar-04/cortes-dos-eua-recorrem-a-multas-para-conter-precedentes-falsos-gerados-por-ia/. Acesso em: 16 mar 2025.

META AI. **Models and Libraries**. Disponível em: https:// ai.meta.com/resources/models-and-libraries. Acesso em: 29 abr. 2025.

MCTI. **Estratégia Brasileira de Inteligência Artificial**. Ministério Ciência, Tecnologia e Inovações. Secretaria de Empreendedorismo e Inovação. 2021. Disponível em: https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/arquivosinteligenciaartificial/ebia-documento_referencia_4-979_2021.pdf. Acesso em: 17 mar. 2025.

MCKINSEY. **The economic potential of generative AI: The next productivity frontier**. McKinsey Digital. 2023. Disponível em: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#introduction. Acesso em: 16 jan. 2025.

MICROSOFT. **Serviço OpenAI do Azure**. Microsoft. 2025. Disponível em: https://azure.microsoft.com/pt-br/products/ai-services/openai-service. Acesso em: 11 fev. 2025.

MICROSOFT CORPORATE BLOGS. **Microsoft and OpenAI evolve partnership to drive the next phase of AI**. Microsoft. 2025. Disponível em: https://blogs.microsoft.com/blog/2025/01/21/microsoft-and-openai-evolve-partnership-to-drive-the-next-phase-of-ai/. Acesso em: 11 fev. 2025.

MITCHELL, Tom M. **Machine Learning**. 1. ed. New York: McGraw-Hill Education, 1997. 432 pp. ISBN 978-0070428072.

MOROZOV, Evgeny. **Big Tech: a ascensão dos dados e a morte da política**. São Paulo: Ubu, 2018.

MOTA, Camila Veras. **Pesquisador brasileiro em IA explica por que DeepSeek impressionou: "Fizeram de forma totalmente diferente da maioria das empresas de tecnologia**. G1. Globo.com. 2025. Disponível em: https://g1.globo.com/tecnologia/noticia/2025/02/01/pesquisador-brasileiro-em-ia-explica-por-que-deepseek-impressionou-fizeram-de-forma-totalmente-diferente-da-maioria-das-empresas-de-tecnologia.ghtml. Acesso em: 11 fev. 2025.

MOTA PINTO, C. A. **Teoria Geral do Direito Civil**. 4ª Ed. Coimbra Editora, 2005.

NAEENI, S. K., NOUHI, N. **The Environmental Impacts of AI and Digital Technologies**. AI and Tech in Behavioral and Social Sciences, 2023, 11-18. Disponível em: https://doi.org/10.61838/kman.aitech.1.4.3. Acesso em: 10 fev. 2025.

NASSEHI, Armin. **Muster Theorie der digitalen Gesellschaft**. Munich, 2019. ISBN 978-3-406-74024-4.

NEGRI, Sergio M. C. de A. **As razões da pessoa jurídica e a expropriação da subjetividade**. Civilistica.com. Rio de Janeiro, a. 5, n. 2, 2016. Disponível em: https://civilistica.com/as-razoes-da-pessoa-juridica/. Acesso em: 12 fev. 2025.

NEGRI, Sergio M. C. de A.; LOPES, G. F. P. **Da Personalidade Eletrônica à Classificação de Riscos da Inteligência Artificial (IA).** Teoria Jurídica Contemporânea. Periódico do Programa de Pós-graduação em Direito da Universidade Federal do Rio de Janeiro. Volume 6. 2021. Disponível em: https://revistas.ufrj.br/index.php/rjur/article/view/44818/27367. Acesso em: 11 jan. 2025.

NEGRI, Sergio M. C. de A. **Robôs como pessoas: a personalidade eletrônica na Robótica e na inteligência artificial**. Pensar, v. 25, n. 3, 2020, pp. 1-14. Disponível em: https://ojs.unifor.br/rpen/article/view/10178. Acesso em: 12 fev. 2025.

NEGRI, Sergio; MACHADO, J. S. **Human Rights and Corporate Personhood: A critical aproach to Corporation Constitucional Rights**. In: XXV World Congress on the Philosophy of Law and Social Philosophy, 2017, Lisboa. Peace based on human rights. Lisboa: Faculdade de Direito da Universidade de Lisboa, 2017. pp. 557-558.

O GLOBO. **Trump encerra programas de diversidade e põe funcionários em licença remunerada até que escritórios sejam fechados**. O Globo. 2025. Disponível em: https://oglobo.globo.com/mundo/noticia/2025/01/22/trump-encerra-programas-de-diversidade-e-coloca-funcionarios-em-licenca-remunerada-ate-que-escritorios-sejam-fechados.ghtml. Acesso em: 11 fev. 2025.

OPENAI. **Business terms**. 2023. Disponível em https://openai.com/policies/business-terms/. Acesso em: 21 jan. 2025.

OPENAI. **Compare Models- OpenAI API. 2025**. Disponível em: https://platform.openai.com/docs/models/compare. Acesso em: 29 abr. 2025.

OPENAI. **Harvey Partners with OpenAI to build a custom-trained model for legal professionals**. 2024. Disponível em: https://openai.com/index/harvey. Acesso em: 21 jan. 2025.

OPENAI. **O que é o ChatGPT?** 2024. Disponível em: https://chatgpt.com.br/o-que-e-chatgpt/. Acesso em: 29 abr. 2025.

OPENAI. **Partnering with Axios expands OpenAI's work with the news industry**. 2025. Disponível em: https://openai.com/index/ partering-wtih-axios. Acesso em: 21 jan. 2025.

PAVLIK, George. **O que é IA generativa (GenAI)? Como funciona?** Oracle. 2023. Disponível em: https://www.oracle.com/br/artificial-intelligence/generative-ai/what-is-generative-ai/. Acesso em: 29 abr. 2025.

PEIXOTO, Fabiano Hartmann; BONAT, Debora. **GPTs e Direito: impactos prováveis das IAs generativas nas atividades jurídicas brasileiras**. Sequência Estudos Jurídicos e Políticos, Florianópolis, v. 44, n. 93, pp. 1–31, 2023. DOI: 10.5007/2177-7055.2023.e94238. Disponível em: https://periodicos.ufsc.br/index.php/sequencia/article/view/94238. Acesso em: 17 jan. 2024.

PERLINGIERI, Pietro. **La persona e isuoi diritti**. Napoli: Edizioni Scientifiche Italiane, 2005.

PORTUGAL GOUVÊA, C. et al. **Relatório de Pesquisa Regulação da Inteligência Artificial ao Redor do Mundo**. TechLab – Laboratório de Tecnologia e Direito. Faculdade de Direito da Universidade de São Paulo. 2024. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4803041. Acesso em: 14 mar. 2025.

RAYMOND, Eric S. **The Cathedral & the Bazaar**. Sebastopol: O'Reilly Media, Inc., 2001. ISBN 9780596001087.

REBOLLO DELGADO, Clicerio. **Inteligencia artificial y derechos fundamentales**. Madrid: Dykinson, 2023.

RIBEIRO, Anderson Filipini; SANTOS, Diego Prezzi; CUNHA, Filipe Mello Sampaio. **Impactos da inteligência artificial no constitucionalismo contemporâneo**. Revista do Instituto de Direito Constitucional e Cidadania – IDCC, Londrina, v. 9, n. 1, e103, jan./jun., 2024. DOI: 10.48159/revistadoidcc.v9n1.e103. Acesso em: 10 mar. 2025.

RODOTÀ, Stefano. **A vida na sociedade da vigilância. A privacidade hoje**. Rio de Janeiro: Renovar, 2008. Tradução: Danilo Doneda e Luciana Cabral Doneda.

RODOTÀ, Stefano. **Transformações do corpo**. Revista Trimestral de Direito Civil, v. 19, pp. 91-107, 2004. Disponível em: https://www.lexml.gov.br/urn/urn:lex:br:rede.virtual.bibliotecas:artigo.revista:2004;1000725693. Acesso em: 15 dez. 2024.

ROMANO, S. M. V. **Linguagem church: modelo generativo que pretende unificar as teorias da inteligência artificial**. Revista Processando o Saber, [*s. l.*], v. 3, 119-127, 1 out. 2011. Disponível em: https://www.fatecpg.edu.br/revista/index.php/ps/article/view/108. Acesso em: 16 mar. 2025.

ROSA, Vika. **Brasil é um dos países que mais usa IA generativa, aponta pesquisa do Google; números são surpreendentes**. IGN. 2025. Disponível em: https://br.ign.com/tech/135974/news/brasil-e-um-dos-paises-que-mais-usa-ia-generativa-aponta-pesquisa-do-google-numeros-sao-surpreendent. Acesso em: 11 fev. 2025.

RUEDIGER, M. A. **Desinformação nas Eleições 2018: O Debate sobre Fake News no Brasil**. FGV. 2019. Disponível em: https://repositorio.fgv.br/server/api/core/bitstreams/f7e9f3b0-37d8-4e33-b367-b655ae601eea/content. Acesso em: 11 fev. 2025.

RUSSELL, S. J.; NORVIG, P. **Artificial intelligence: A modern approach**. 2021.4th ed. Pearson.

SANTAELLA, Lucia. **O homem e as máquinas**. In: DOMINGUES, Diana (org.). A arte no século XXI, São Paulo: Editora da UNESP, São Paulo, pp. 33-44, 1997.

SANTAELLA, Lucia. **Inteligência contínua: a sétima revolução cognitiva do Sapiens**. São Paulo: Paulus, 2022.

SAVIGNY, F. von. **System des heutigen römischen Rechts**, II, 1840, 310 s., apud BARBOSA, Mafalda. **Nas fronteiras de um admirável mundo novo? O problema da personificação de entes dotados de inteligência artificial**. 2021, pp. 251-285. In: BARBOSA, M. M., BRAGA NETTO, F.; SILVA, M. C.; FALEIROS JÚNIOR, J. L. de M. Direito digital e inteligência artificial: diálogos entre Brasil e Europa. 1ª Ed. Editora Foco, 2021, ePUB.

SENADO. **Relatório Final dos trabalhos da Comissão de Juristas responsável pela revisão e atualização do Código Civil**. Disponível em: https://www12.senado.leg.br/assessoria-de-imprensa/arquivos/anteprojeto-codigo-civil-comissao-de-juristas-2023_2024.pdf. Acesso em: 17 mar. 2025.

SCHREIBER, Anderson. **Direitos da Personalidade**. 3. Ed. São Paulo: Atlas, 2014.

SCHUETT, Jonas. **Defining the scope of AI regulations**. LawAI Working Paper Series, No. 4. 2021. Forthcoming in Law, Innovation and Technology, v. 15, n. 1. Disponível em: https://law-ai.org. Acesso em: 3 maio 2025.

SILVA, Tarcízio. **Racismo algorítmico: Inteligência artificial e discriminação nas redes digitais**. Edições SESC. 2022.

SKRLJ, Blaz. **From Unimodal to Multimodal Machine Learning: An Overview**. SpringerBriefs in Computer Science. Springer. Disponível em: http://link.springer.com/book/10.1007/978-3-031-57016-2. Acesso em: 6 mar. 2025.

SMITH, John. **Energy Consumption in Data Centers**. Analyst Brief. New York: Tech Insights, 2022.

SOMMERVILLE, Ian. **Software Engineering**. 10. ed. Boston: Pearson, 2015. ISBN 978-0133943030.

SOUZA, L. C. de. **Desafio da Equidade no Processo Judicial nos Conflitos Decorrentes do Uso de Inteligência Artificial**. Revista Brasileira de Direito Internacional. 2024. DOI**:** https://doi.org/10.26668/IndexLawJournals/2526-0219/2024.v10i1.10467. Acesso em: 18 mar. 2025.

STALLMAN, Richard M. **Free Software, Free Society: Selected Essays of Richard M. Stallman**. 2. ed. Joshua Gay (Ed.); Lawrence Lessig (Introd.). North Charleston: CreateSpace Independent Publishing Platform, 2009. 224 pp. ISBN 978-1441436856.

STRUBELL, E., GANESH, A., MCCALLUM, A. (2020). **Energy and Policy Considerations for Modern Deep Learning Research**. Proceedings of the AAAI Conference on Artificial Intelligence. Disponível em: https://doi.org/10.1609/aaai.v34i09.7123. Acesso em: 20 fev. 2025.

STRYKER, Cole. SCAPICCHIO, Mark. **What is generative AI?** IBM. 2024. Disponível em: https://www.ibm.com/think/topics/generative-ai. Acesso em: 11 jan. 2025.

SUTTON, Richard S.; BARTO, Andrew G. **Reinforcement Learning: An Introduction**. 2. ed., in progress. Cambridge, Massachusetts; London, England: The MIT Press, 2018.

TEDAI. **Glossary – Latent Space**. São Francisco. 2024. Disponível em: https://tedai-sanfrancisco.ted.com/glossary/latent-space/. Acesso em: 6 mar. 2025.

TEPEDINO, Gustavo. **A tutela da personalidade no ordenamento civil-constitucional brasileiro**. In: Temas de Direito Civil. Rio de Janeiro: Renovar, 2004, pp. 23-58.

TEPEDINO, Gustavo; SILVA, Rodrigo da Guia. **Desafios da inteligência artificial em matéria de responsabilidade civil**. Revista Brasileira de Direito Civil – RBDCivil, Belo Horizonte, v. 21, pp. 61-86, jul./set. 2019.

THOMSON REUTERS. **New report on ChatGPT & generative AI in law firms shows opportunities abound, even concerns persist**. Thomson Reuters. 2023. Disponível em: https://www.thomsonreuters.com/en-us/posts/technology/chatgpt-generative-ai-law-firms-2023/. Acesso em: 11 jan. 2024.

TEUBNER, Gunther. **Corporate Codes in the Varieties of Capitalism: How Their Enforcement Depends on the Differences Among Production Regimes. Indiana Journal of**

Global Legal Studies. 2017. Disponível em: https://www.repository.law.indiana.edu/ijgls/vol24/iss1/4/. Acesso em: 10 jan. 2025.

TEUBNER, Gunther. Droit et réflexivité: l'auto-référence en droit et dans l'organisation. Trad. Nathalie Boucquey. Bruylant: L.G.D.J., 1996, p. 8.

UNIÃO EUROPEIA. **Relatório Especial 08/2024 - Ambições da UE para a inteligência artificial: Melhor governação, investir mais e com mais orientação: as chaves do futuro**. 2024. Disponível em: https://www.eca.europa.eu/ECAPublications/SR-2024-08/SR-2024-08_PT.pdf. Acesso em: 4 maio 2025.

UNIÃO EUROPEIA. **Resolução do Parlamento Europeu que contém Recomendações à Comissão Sobre Disposições de Direito Civil Sobre Robótica (2015/2103(INL))**. 2017. Disponível em: http://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_PT.html?redirect. Acesso em: 25 jun. 2024.

UNIÃO EUROPEIA. **Resolução do Parlamento Europeu que contém recomendações à Comissão sobre o Regime de Responsabilidade Civil Aplicável à Inteligência Artificial (2020/2014(INL))**. 2021. Disponível em: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_PT.html. Acesso em: 25 jan. 2025.

UNIÃO EUROPEIA. **Carta dos Direitos Fundamentais da União Europeia**. Nice, 2000. Disponível em: https://www.cnpd.pt/bin/legis/internacional/CARTAFUNDAMENTAL.pdf. Acesso em: 9 jun. 2024.

VASWANI, Ashish; SHAZEER, Noam; PARMAR, Niki; USZKOREIT, Jakob; JONES, Llion; GOMEZ, Aidan N.; KAISER, Lukasz; POLOSUHIN, Illia. **Attention Is All You Need.** 2017. DOI: 10.48550/arXiv.1706.03762. Acesso em: 5 fev. 2025.

VIEIRA, Marcelo Gimenes. **Dora Kaufman: diversidade em IA exige esforço de regulamentação.** 2022. Disponível em: https://itforum.com.br/noticias/dora-kaufman-diversidade-em-ia-exige-esforco-de-regulamentacao/. Acesso em: 5 maio 2025.

ZUBOFF, Shoshana. **Surveillance Capitalism or Democracy? The Death Match of Institutional Orders and Politics of Knowledge in Our Information Civilization**. Organization Theory, v. 3, pp. 1-79, 2022.

WINSTON, Patrick Henry. **Artificial intelligence desmystified**. Minuta de 30 set. 2018. (mimeo).