FACULDADE BAIANA DE DIREITO

CURSO DE GRADUAÇÃO EM DIREITO

MILLA DE OLIVEIRA GARDASEVIC

A RESPONSABILIZAÇÃO CIVIL PELO USO NÃO AUTORIZADO DE DADOS PESSOAIS: UMA ANÁLISE CRÍTICA DO TREINAMENTO DE IA GENERATIVA

Salvador
2025

MILLA DE OLIVEIRA GARDASEVIC

A RESPONSABILIZAÇÃO CIVIL PELO USO NÃO AUTORIZADO DE DADOS PESSOAIS: UMA ANÁLISE CRÍTICA DO TREINAMENTO DE IA GENERATIVA

Monografia apresentada ao curso de graduação em Direito da Faculdade Baiana de Direito e Gestão, como requisito parcial para obtenção do grau de Bacharel em Direito.

Orientador: Prof. Dr. Maurício Requião

Salvador
2025

TERMO DE APROVAÇÃO

MILLA DE OLIVEIRA GARDASEVIC

A RESPONSABILIZAÇÃO CIVIL PELO USO NÃO AUTORIZADO DE DADOS PESSOAIS: UMA ANÁLISE CRÍTICA DO TREINAMENTO DE IA GENERATIVA

Monografia aprovada como requisito parcial para obtenção do grau de bacharel em Direito da Faculdade Baiana de Direito e Gestão, pela seguinte banca examinadora:

Nome: _____

Titulação e instituição: _____

Nome: _____

Titulação e instituição: _____

Nome: _____

Titulação e instituição: _____

Salvador,_____/____/ 2025.

AGRADECIMENTOS

Dedico este trabalho à minha mãe, Leila, que sempre foi a luz mais constante no meu caminho. A ela, que acreditou nos meus sonhos antes mesmo que eu pudesse nomeá-los, e que, com sua presença tão inteira, me mostrou o tipo de mulher que eu desejo honrar. Quero levá-la comigo em tudo o que faço, porque a firmeza com que conduz a vida me guia e me dá direção.

Ao meu pai, Aleksandar, pelo amparo firme e pela tranquilidade que sempre me ofereceu. Sua presença me deu a segurança necessária para seguir, mesmo quando o trajeto parecia maior do que a minha força. Guardo nele a certeza de que existe um lugar no mundo onde posso repousar e recomeçar.

A vocês dois, juntos, dedico também este capítulo da minha vida. Tudo o que alcancei nasceu do que construímos como família, na soma das escolhas, dos esforços e do amor que sempre me cercou. Cresci sabendo que podia ir longe porque tinha para onde voltar. Somos nós três, para sempre, e cada passo meu carrega algo de cada um de vocês.

À minha avó Lina, cuja coragem e ternura me sustentaram em tantas etapas. Foi ela quem me ensinou que a bondade também é uma forma de coragem e que o afeto, quando verdadeiro, é capaz de sustentar qualquer caminhada. Tudo que eu procuro ser nasce do exemplo dela, que segue iluminando a vida de todos ao seu redor.

In memoriam, ao meu avô Genaro, que partiu em 2022, mas permanece em quem sou. As lembranças que guardo dele seguem vivas como uma força que me acompanha diariamente. Muito do que me tornei carrega a marca da alegria com que ele seguia meus passos e da inspiração que sua vida deixou. Que este trabalho chegue até onde ele estiver.

Ao meu tio Sérgio, à minha tia Ieda e aos meus primos Lucas e Theo, pela presença firme e afetuosa que sempre fez parte da minha vida. Lembrar que tenho uma família tão maravilhosa conforta meu coração, e guardo vocês nele com muito carinho.

À minha família da Sérvia, que mesmo distante nunca deixou de celebrar cada conquista minha com entusiasmo. A saudade não diminuiu a força do vínculo, apenas o tornou mais bonito. Saber que vibro em dois continentes sempre me deu a sensação de pertencer a um mundo maior.

Às minhas amigas do Módulo, que me acompanham desde o tempo em que a vida parecia grande demais para ser entendida por nós. Vocês são as irmãs que escolhi. Crescemos juntas, e cada etapa vivida ganhou outro sentido porque foi compartilhada. O apoio de vocês sempre chegou no momento certo, do jeito certo, e sou grata por terem acreditado em mim enquanto este trabalho tomava forma. Levo nossa amizade como uma das certezas mais bonitas que já

construí.

Às mães e aos pais dessas amigas, que me acolheram como se eu fosse parte de casa e da família, deixo também meu carinho e eterna gratidão.

Aos meus amigos e amigas da faculdade, que o destino colocou no meu caminho nas horas certas. Obrigada por me acolherem quando tudo ainda era novo demais. Não começamos todos juntos, mas terminamos unidos, e isso diz muito sobre o que construímos. A parceria de vocês tornou a jornada mais alegre, e sou grata por cada ano que dividimos, pelas conversas que me ampararam e pelos momentos que transformaram a faculdade em um lugar que vou lembrar com carinho para sempre.

À Alfa Consultoria e às amizades que nasceram ali. Carrego comigo valores e histórias que sei que ficarão para sempre. Cresci como pessoa e como profissional, e muito disso floresceu porque pude caminhar com vocês.

Às pessoas que a vida me deu fora de todos esses espaços, aquelas que não cabem em categorias e que permanecem comigo como memórias vivas do que fui e do que ainda sou: levo vocês como parte essencial da minha história.

Aos meus chefes, Ana Carolina e Lucas, que estiveram comigo todos os dias e sempre me ofereceram apoio genuíno. Com vocês aprendi que o Direito vai muito além da técnica: ele é feito de sensibilidade e respeito. Agradeço pela confiança, pelo espaço para crescer e pela generosidade com que vocês conduzem tudo no dia a dia. Me espelho muito em vocês.

Aos professores e funcionários da faculdade, que deram vida ao conhecimento e sustentaram a estrutura que possibilita a realização de tantos sonhos.

Ao meu orientador, Maurício Requião, pela leitura atenta, pela disponibilidade e pelas direções que tornaram este trabalho possível. Sou grata pela confiança com que acompanhou este percurso e que me guiou na construção desta pesquisa.

A todos aqueles que cruzaram a minha jornada, também deixo meu agradecimento. Cada presença teve seu lugar na construção de tudo. Agradeço à vida por ter sido generosa comigo e por ter reservado surpresas tão bonitas. Que eu siga reconhecendo essa generosidade e fazendo dela motivo para seguir em frente com esperança e coragem.

"*A inteligência procura, mas quem encontra é o coração.*"

George Sand

RESUMO

O presente trabalho analisa criticamente a responsabilização civil pelo uso não autorizado de dados pessoais no treinamento de sistemas de inteligência artificial (IA) generativa, fenômeno marcado por opacidade técnica e exploração massiva de informações pessoais. Parte-se da constatação de que, na economia digital contemporânea, os dados pessoais foram convertidos em ativo econômico central para plataformas tecnológicas, o que intensifica práticas *web scraping* e reuso informacional em larga escala. Nesse contexto, a IA generativa atua como catalisadora dessa dinâmica, uma vez que seu desempenho depende diretamente do volume e da heterogeneidade dos dados empregados no treinamento, frequentemente incluindo informações identificáveis incorporadas de modo irreversível aos modelos. Diante dessa realidade, o trabalho formula como objetivo geral avaliar os limites e possibilidades do ordenamento jurídico brasileiro para responder ao uso não autorizado de dados pessoais por modelos generativos, investigando especialmente a capacidade da responsabilidade civil de oferecer tutela efetiva. Como objetivos específicos, examina-se o contexto econômico da exploração de dados, os pressupostos clássicos da responsabilidade civil (conduta, dano, nexo causal e imputação), o déficit estrutural da função reparatória *ex post* e a necessidade de reconstrução normativa orientada à proteção da autodeterminação informativa. A metodologia utilizada é qualitativa e se vale do método hipotético-dedutivo, formulando hipóteses sobre a insuficiência do paradigma reparatório tradicional frente aos riscos informacionais difusos e estruturalmente opacos. Além da revisão doutrinária, legislativa e jurisprudencial, o estudo incorpora pontualmente uma análise documental das políticas de privacidade de grandes plataformas de IA, com o objetivo de confrontar os discursos de proteção de dados pessoais com suas práticas efetivas de tratamento e coleta, revelando assim um descompasso que agrava a vulnerabilidade informacional dos titulares. Os resultados apontam que a responsabilidade civil clássica, centrada na recomposição individual e dependente de causalidade e dano delimitáveis, é incapaz de lidar sozinha com prejuízos que se internalizam na arquitetura dos modelos e se reproduzem indefinidamente. O trabalho identifica que a violação decorrente do uso não autorizado de dados pessoais não se limita a danos patrimoniais, mas representa lesão estrutural à autodeterminação informativa e à liberdade do titular, configurando forma contemporânea de dano existencial. A partir desse diagnóstico, defende-se uma reorientação funcional da responsabilidade civil, incorporando também deveres *ex ante* (transparência, governança, rastreabilidade, minimização de dados) e ampliando o uso de medidas não pecuniárias, como auditorias, correções técnicas e exclusões eficazes de dados. Complementarmente, sugere-se a adoção de instrumentos distributivos, como *disgorgement*, multas proporcionais e fundos *cy-près*, para retirar vantagens econômicas do ilícito e recompor assimetrias informacionais. O trabalho contribui para o desenvolvimento de soluções jurídicas que preservem a autodeterminação informativa dos titulares e reafirmem a centralidade da pessoa no contexto da IA generativa.

Palavras-Chave: Proteção de dados pessoais; Responsabilidade civil; Inteligência artificial generativa; Treinamento; Autodeterminação informativa.

ABSTRACT

This study critically analyzes civil liability for the unauthorized use of personal data in the training of generative artificial intelligence (AI) systems, a phenomenon marked by technical opacity and massive exploitation of personal information. It starts from the observation that, in the contemporary digital economy, personal data have been converted into a central economic asset for technological platforms, which intensifies web scraping practices and large-scale informational reuse. In this context, generative AI acts as a catalyst for this dynamic, since its performance depends directly on the volume and heterogeneity of the data used in training, frequently including identifiable information irreversibly incorporated into the models. Given this reality, the general objective of the work is to evaluate the limits and possibilities of the Brazilian legal system in responding to the unauthorized use of personal data by generative models, especially investigating the capacity of civil liability to offer effective protection. As specific objectives, the economic context of data exploitation is examined, as well as the classical elements of civil liability (conduct, damage, causal nexus and imputation), the structural deficit of the ex post reparatory function and the need for a normative reconstruction oriented toward the protection of informational self-determination. The methodology used is qualitative and relies on the hypothetical-deductive method, formulating hypotheses on the insufficiency of the traditional reparatory paradigm in the face of diffuse and structurally opaque informational risks. In addition to doctrinal, legislative and jurisprudential review, the study incorporates a punctual documentary analysis of the privacy policies of major AI platforms, with the objective of confronting personal data protection discourses with their effective treatment and collection practices, thus revealing a mismatch that aggravates the informational vulnerability of data subjects. The results indicate that classical civil liability, centered on individual compensation and dependent on delimitable causality and damage, is unable to deal alone with harm that becomes internalized in the architecture of models and reproduced indefinitely. The work identifies that the violation resulting from the unauthorized use of personal data is not limited to property damage, but represents a structural injury to the informational self-determination and freedom of the data subject, constituting a contemporary form of existential damage. Based on this diagnosis, the work argues for a functional reorientation of civil liability, also incorporating ex ante duties (transparency, governance, traceability, data minimization) and expanding the use of non-pecuniary measures, such as audits, technical corrections and effective data deletion. Complementarily, it suggests the adoption of distributive instruments, such as disgorgement, proportional fines and cy-près funds, to withdraw economic advantages derived from the illicit act and to restore informational asymmetries. The work contributes to the development of legal solutions that preserve the informational self-determination of data subjects and reaffirm the centrality of the person in the context of generative AI.

Keywords: Personal data protection; Civil liability; Generative artificial intelligence; Training; Informational self-determination.

LISTA DE ABREVIAÇÕES E SIGLAS

ADI	Ação Direta de Inconstitucionalidade
Art.	artigo
ANPD	Agência Nacional de Proteção de Dados
CC/02	Código Civil
CDC	Código de Defesa do Consumidor
CPC/2015	Código de Processo Civil
CRFB/88	Constituição Federal de 1988
DPO	*Data Protection Officer*
EC	Emenda Constitucional
EDPB	*European Data Protection Board*
IA	Inteligência Artificial
i.e.	isto é
IoT	*Internet Of Things*
IBGE	Instituto Brasileiro de Geografia e Estatística
IP	Protocolo de Internet
LGPD	Lei Geral de Proteção de Dados
LLMs	*Large Language Models*

MAU	*Monthly Active Users*
nº	Número
p.	Página
PL	Projeto de Lei
STF	Supremo Tribunal Federal
STJ	Superior Tribunal de Justiça
URLs	*Uniform Resource Locator*
v.g	*Verbi Gratia*

LISTA DE TRADUÇÕES

accountability	prestação de contas
big data	grandes volumes de dados
big tech(s)	grande(s) conglomerado(s) tecnológico(s)
cheapest cost avoider	agente capaz de evitar o dano ao menor custo
compliance	conformidade regulatória
corpora	corpos textuais
cy-près	destinados ao fim mais próximo possível
Data Protection Officer	Encarregado de Proteção de Dados
datasets	conjuntos de dados
disgorgement	devolução / restituição
European Data Protection Board	Comitê Europeu de Proteção de Dados
ex ante	prévia / anterior
ex post	posterior
gatekeeper	controlador de acesso
input(s)	dado(s) de entrada
Internet of Things	Internet das Coisas

Monthly Active Users	Usuários Ativos Mensais
nudges	estímulos comportamentais
opt-in	consentimento prévio/expresso
opt-out	manifestação de oposição/exclusão
output(s)	saída(s) / resultado(s)
privacy by default	proteção como configuração padrão
privacy by design	proteção desde a concepção
profiling	perfilamento
prompt	comando
punitive damages	danos punitivos
quantum	montante
ratio decidendi	razão de decidir
ratio legis	propóstio / finalidade da lei
status quo ante	estado anterior das coisas
Uniform Resource Locator	Localizador Uniforme de Recursos
verbi gratia	por exemplo
web scraping	raspagem de dados

SUMÁRIO

1 INTRODUÇÃO

Nas últimas décadas, o avanço tecnológico remodelou profundamente as relações econômicas, inaugurando um cenário em que os dados pessoais passaram a ocupar papel decisivo na geração de valor. O que antes figurava como elemento secundário da atividade econômica, tornou-se recurso estratégico para o desenvolvimento de soluções digitais e a criação de modelos de negócio sustentados pela coleta e análise de informações. Diante disso, os dados pessoais deixam de ser mero subproduto das interações humanas e passam a constituir ativo econômico de primeira grandeza, convertendo os indivíduos em participantes involuntários dessa cadeia de valor e inserindo a tutela dessas informações nas discussões jurídicas contemporâneas.

É nesse cenário que se insere o presente trabalho, cujo tema se delimita de forma precisa à análise da responsabilização civil pelo uso não autorizado de dados pessoais no treinamento de inteligência artificial (IA) generativa. Trata-se, portanto, de um recorte epistemológico que não abrange genericamente o tratamento de dados pessoais ou a responsabilidade civil em sentido amplo, mas especificamente a utilização de dados pessoais como base para o treinamento de modelos generativos de IA sem a autorização dos titulares. Esse recorte concentra a investigação desta prática que é contemporânea, tecnicamente complexa e marcada pela opacidade dos processos de coleta e processamento de dados, características estas que tornam ainda mais difícil delimitar consequências jurídicas.

Antes de adentrar no problema central, faz-se necessário esclarecer alguns pontos fundamentais. Em primeiro lugar, no treinamento de IA generativa os dados utilizados não se limitam a textos, imagens ou códigos desconectados de pessoas, mas frequentemente incluem informações capazes de identificar indivíduos, direta ou indiretamente. Em segundo lugar, ao contrário do que ocorre em operações tradicionais de tratamento de dados, o uso de informações pessoais no treinamento de IA não é facilmente reversível: uma vez incorporada ao modelo, seus efeitos permanecem internalizados, podendo repercutir em *outputs* e inferências algorítmicas futuras. Por fim, há um descompasso evidente entre a evolução tecnológica e a capacidade do ordenamento jurídico de regular práticas que envolvem técnicas sofisticadas de extração e utilização indevida de dados pessoais em larga escala.

A questão, portanto, demonstra ter grande relevância sócio-jurídica. Socialmente,

escancara-se uma realidade em que os usuários das plataformas digitais têm sua experiência cotidiana transformada em matéria-prima informacional, muitas vezes sem plena consciência ou controle da destinação dos seus dados pessoais. O resultado é o enfraquecimento da autodeterminação informativa que passa a ser minada por estruturas tecnológicas capazes de prever e direcionar comportamentos. Juridicamente, a problemática desafia os instrumentos tradicionais de responsabilização civil, estruturados em pressupostos clássicos que precisam se readaptar diante de práticas irreversíveis de uso de dados pessoais. Somado a isso, tem-se o fato que o treinamento de IA generativa inaugura uma dimensão de dano que não é apenas individual, mas estrutural e coletiva, exigindo repensar a própria função da responsabilidade civil.

Diante desse cenário, o trabalho formula como objetivo geral analisar criticamente a responsabilização civil pelo uso não autorizado de dados pessoais no treinamento de sistemas de IA generativa, avaliando os limites e possibilidades do ordenamento jurídico brasileiro de lidar com esse fenômeno.

Por sua vez, como objetivos específicos, busca-se: compreender o contexto econômico e tecnológico que impulsiona a exploração de dados pessoais; examinar os contornos teóricos e práticos dos pressupostos clássicos da responsabilidade civil frente ao ambiente digital; identificar as fragilidades do modelo reparatório individual diante de práticas massivas e opacas; e, por fim, propor caminhos interpretativos capazes de reorientar a responsabilidade civil para uma lógica também preventiva, coletiva e voltada à proteção da autodeterminação informativa.

No que se refere aos recursos metodológicos, o trabalho adota o método hipotético-dedutivo, partindo da formulação de hipóteses decorrentes do problema central identificado: a possível inadequação do modelo tradicional de responsabilidade civil frente ao uso não autorizado de dados pessoais no treinamento de sistemas de inteligência artificial generativa. A pesquisa é predominantemente qualitativa e desenvolve-se por meio de análise bibliográfica, doutrinária, legislativa e jurisprudencial.

Embora não realize pesquisa empírica de campo ou coleta sistemática de dados junto a indivíduos, incorpora, de forma pontual, uma análise documental das políticas de privacidade das principais plataformas de IA generativa (sistematizada em tabela no item 2.3.5) a fim de observar, de maneira concreta, como esses agentes declaram tratar dados pessoais e como tais declarações dialogam, ou não, com a proteção jurídica prevista no ordenamento. Essa etapa

documental funciona como elemento de apoio para testar, em nível prático, as hipóteses formuladas.

Nesse percurso, considera-se que a complexidade técnica e o caráter massivo e opaco do uso de dados pessoais podem revelar que os pressupostos clássicos da responsabilidade civil não são suficientes para assegurar reparação integral, indicando uma possível limitação estrutural do paradigma reparatório *ex post* e levantando a perspectiva de danos que ultrapassam a esfera individual e assumem dimensão coletiva e estrutural. Assim, o método adotado não se limita à descrição abstrata do fenômeno, mas à testagem crítica dessas hipóteses diante das categorias jurídicas existentes e dos elementos documentais analisados, de modo a verificar se o direito vigente consegue responder aos desafios apresentados ou se aponta para a necessidade de reorientação da responsabilidade civil em direção a uma lógica preventiva e sistêmica.

No que tange à sua estrutura, esta monografia contará com cinco capítulos, sendo um capítulo de introdução, três de desenvolvimento e um de conclusão.

O segundo capítulo analisa a lógica predatória de exploração de dados pessoais na economia digital, evidenciando a transformação da informação em ativo econômico e o papel central das *big techs* na coleta massiva e opaca de dados. Em seguida, demonstra como a inteligência artificial generativa intensifica essa dinâmica ao depender de grandes volumes de dados pessoais para treinamento, recorrendo a práticas como *web scraping* e bases públicas que frequentemente incluem informações identificáveis. O capítulo ainda evidencia que o treinamento vai além: os próprios *inputs* fornecidos pelos usuários durante o uso das plataformas continuam sendo aproveitados para refinar modelos, muitas vezes sem plena ciência dos titulares. A análise documental das políticas de privacidade das principais plataformas de IA, sintetizada em tabela própria, revela um descompasso entre o discurso de proteção de dados e as práticas efetivas, reforçando a erosão da autodeterminação informativa e a inserção dos indivíduos em um processo de colonização digital.

O terceiro capítulo enfrenta, em chave dogmática, os limites da responsabilização civil no tratamento indevido de dados pessoais no meio digital, partindo da estrutura clássica dos pressupostos da responsabilidade civil (conduta, dano, nexo causal e nexo de imputação) e testando sua capacidade de resposta diante das especificidades da inteligência artificial generativa. Após reconstruir esses elementos à luz do ecossistema de dados, o capítulo demonstra como a opacidade técnica e a circulação massiva de informações tensionam a

identificação da conduta ilícita, a configuração do dano informacional, a prova do nexo causal e a distribuição de responsabilidade entre múltiplos agentes de tratamento. Em seguida, desenvolve a distinção entre responsabilidade *ex ante* e *ex post*, bem como analisa o ônus probatório em contexto de invisibilidade do dano e assimetria informacional. Por fim, discute a irreversibilidade do uso de dados pessoais e a natureza dos prejuízos indenizáveis, evidenciando que a aplicação dos institutos tradicionais, embora indispensável, mostra-se insuficiente sem uma adaptação funcional orientada à proteção efetiva da autodeterminação informativa.

No quarto capítulo, desenvolve-se uma leitura crítica da responsabilização civil pelo uso não autorizado de dados pessoais no treinamento de sistemas de IA generativa, partindo da constatação de que o dever jurídico de reparar, embora existente, enfrenta limites práticos que comprometem sua concretização efetiva. O capítulo desloca o foco da pergunta "há responsabilidade?" para "como e para que a responsabilidade deve operar?" em um ambiente de risco difuso. Examina-se a insuficiência estrutural do paradigma exclusivamente reparatório *ex post*, sem negar a relevância da compensação pecuniária, mas ressaltando que a lógica compensatória individual, isoladamente, não consegue restaurar a autonomia informativa nem reverter danos incorporados aos modelos de IA. A partir desse diagnóstico, o capítulo defende uma reorientação funcional da responsabilidade civil, sugerindo-lhe dimensões preventivas, coletivas e estruturais, e propondo bases para uma reconstrução normativa ancorada na proteção da autodeterminação informativa e na governança dos fluxos de dados: elementos indispensáveis para que o Direito preserve efetividade regulatória na sociedade da informação.

Assim estruturado, o trabalho busca oferecer uma análise crítica capaz de contribuir para o debate jurídico contemporâneo sobre proteção de dados e responsabilidade civil, abrindo caminho para reflexões que se estendem para além deste estudo, mas que nele encontram um ponto de partida necessário.

2 A LÓGICA PREDATÓRIA DA EXPLORAÇÃO DE DADOS PESSOAIS

No século XX, o desenvolvimento tecnológico acelerado impulsionou o armazenamento e processamento de dados a uma escala sem precedentes (Amaral, 2016, p. 04-12), exigindo uma resposta do mundo jurídico. Nesse contexto, o direito à proteção de proteção de dados pessoais emergiu nas últimas décadas como uma das temáticas mais alarmantes e desafiadoras do direito contemporâneo (Doneda e Zanatta, 2022, p. 14-15).

A preocupação com o tema tomou forma na Convenção nº 108 do Conselho Europeu, de 1981, sendo a primeira convenção jurídica internacional dedicada à tratativa da proteção de dados pessoais. A Convenção reconheceu a necessidade de um sistema de proteção de dados que fosse compatível com a livre circulação de informações e para isso, estabeleceu um conjunto de definções e princípios relativos aos dados pessoais (Santos, 2024, p. 21-22).

Neste hiato, outros esforços normativos internacionais passaram a influenciar o panorama global, impulsionando a necessidade de regulamentação também em países com alto volume de dados, como o Brasil (Loschi, 2025). Embora a Constituição Federal (CRFB/88) já trouxesse previsões em seus artigos (Art.) 5º, X e XII sobre a inviolabilidade da intimidade, da vida privada e do sigilo de dados e comunicações (Brasil, 1988), foi somente 30 anos depois que uma lei geral veio detalhar e regulamentar a proteção desses direitos no ambiente digital brasileiro (Calaza, 2024, p. 08).

Desse modo, promulgada em 2018, a Lei Geral de Proteção de Dados (LGPD) estabeleceu um conjunto de normas e princípios para a proteção de dados pessoais. Tal avanço normativo foi complementado pela jurisprudência do Supremo Tribunal Federal (STF) e pelo próprio texto constitucional. Em sede de julgamento das ADIs nºs 6387, 6388, 6390 e 6393, o STF (2020) reconheceu a proteção de dados pessoais como direito fundamental implícito. Tal compreensão foi ratificada e formalmente consolidada em 2022, com a aprovação da Emenda Constitucional nº 115/2022 que inseriu expressamente a proteção de dados pessoais no Art. 5º, LXXIX, elevando-o à condição de direito fundamental expresso (Scheuermann, 2023, p. 03).

Assim, o reconhecimento constitucional da proteção de dados como um direito fundamental expressou a necessidade de uma tutela mais robusta, pois o Estado não apenas deveria se

abster de intervir indevidamente nos dados dos cidadãos, mas também teria o dever de proteger os indivíduos contra intervenções de terceiros (Alexy, 2024, p. 455-458). Esse "efeito horizontal" dos direitos fundamentais de proteção impõe um dever de cuidado aos agentes de tratamento de dados pessoais, que devem ser aplicados em todos os setores, públicos e privados (Alexy, 2024, p. 523-524).

Em contraste com sólidos avanços legislativos e jurisprudenciais, a realidade da era digital demonstra que a economia está operando sob uma lógica predatória da exploração dos dados pessoais. A tendência mercadológica que se consolidou é a da instrumentalização dos dados pessoais como recursos estratégicos na cadeia de valor informacional (Zuboff, 2021, p.15), indo de encontro com os direitos fundamentais à intimidade, privacidade e à proteção de dados, estabelecidos na CRFB/88 (Brasil, 1988).

As normas protetivas, embora relevantes, têm se revelado insuficientes (Morozov, 2018, p. 134-135) frente aos avanços vorazes nos quais as *big techs*, têm capturado, processado e monetizado fragmentos personalíssimos dos indivíduos: transformando dados pessoais em previsibilidade, e previsibilidade em ativo financeiro (Morozov, 2018, p. 137).

A dinâmica de mineração desses dados inaugura um regime assimétrico, no qual o consentimento do titular de dados, previsto no Art. 7º, I da LGPD, perde densidade jurídica diante da opacidade técnica e contratual das plataformas (Fornasier; Knebel, 2020, p. 15). A suposta transparência informacional, exigida pelo Art. 6º, VI, da LGPD, esbarra na complexidade das operações algorítmicas e na pulverização dos circuitos de tratamento de dados (Brasil, 2018).

Ainda que as *big techs* frequentemente apresentem que a coleta de dados está dentro do seu escopo de liberdade da atuação (Santos; Homci, 2024, p. 06), o que está em curso é a consolidação de uma nova racionalidade econômica, orientada pela apropriação sistemática da experiência humana como insumo de mercado e atribuindo aos titulares dos dados uma função utilitarista (Pinheiro, 2021 p. 282).

Embora o utilitarismo tenha findado suas raízes no século XVIII, tendo como precursores os estudiosos Jeremy Bentham e John Stuart Mill, é possível traçar paralelos com a contemporaneidade: a lógica subjacente à coleta massiva de dados pessoais e à vigilância digital se alinha a uma interpretação distorcida dessa filosofia. Nesse contexto, o "bem maior" buscado pelas plataformas muitas vezes disfarçado de conveniência ou personalização

(Santos; Homci, 2024, p. 8), justifica a dinâmica predatória onde a autonomia do indivíduo é secundarizada em prol da maximização de lucros e do controle mercadológico (Chirita, 2018, p. 160).

É nessa conjuntura que se delineia o fenômeno denominado por Zuboff (2021, p. 15) como capitalismo de vigilância: um sistema econômico que extrai, transforma e comercializa a vivência humana sob a aparência de inovação. A vigilância, antes concebida como instrumento de controle estatal (Forgioni, 2018, p. 104), assume contornos corporativos e se converte em matriz de acumulação privada (Fornasier; Knebel, 2020, p. 11).

Neste novo arranjo, os indivíduos deixam de ser usuários e passam a ocupar a posição de matéria-prima, produzindo, mesmo que inconscientemente, valor informacional enquanto navegam, interagem e se expressam no ambiente virtual, dando origem ao que se convencionou chamar de trabalho digital: uma forma de labor não remunerado e invisibilizado, que alimenta modelos de negócios baseados na publicidade direcionada e no perfilamento comportamental (Fuchs; Sevignani, 2013, p. 01-02).

Assim sendo, o que se observa é o enfraquecimento da autodeterminação informativa diante de um ecossistema orientado não pela garantia de direitos, mas sim pela violação deles (Wanderer, 2023, p. 10-11). A maximização do lucro por meio da certeza é construída com volumes massivos de dados, resultando em uma fórmula com um custo operacional baixíssimo e alta lucratividade pela exploração dos ativos informacionais (Ciuriak, 2018, p. 03).

Para a melhor compreensão da lógica predatória da exploração de dados pessoais na economia digital, é fundamental conceituar o dado isolado e seu fluxo, diferenciando-o do dado pessoal. Diante do exposto, passa-se à análise da Seção sobre o dado pessoal.

2.1 O DADO PESSOAL

Numa primeira análise, o dado em sua essência deve ser entendido como uma representação puramente sinática, uma sequência de símbolos que, por si só, não contém significado. Pode ser um número, uma palavra, uma coordenada geográfica, um clique em uma página da internet ou uma fotografia digitalizada. O dado, enquanto elemento bruto, é neutro e

desprovido de semântica: trata-se de uma entidade quantificada ou quantificável, passível de armazenamento e processamento automático, mas que nada revela sobre a realidade sem um sujeito que o interprete (Setzer, 2015). Sua verdadeira relevância emerge quando correlacionado com outros elementos, pois é nessa articulação que o dado deixa de ser mera forma e se converte em dado pessoal (Amaral, 2016, p. 04-06).

Dessa forma, a LGPD estabelece uma diferenciação crucial. Conforme o Art. 5º, inciso I, dado pessoal é a informação que se vincula diretamente a um indivíduo, tornando-o identificável (Brasil, 2018). A identificação pode ser direta, por meio de um nome ou CPF, ou indireta, quando a informação, combinada a outras, leva à identidade de uma pessoa natural. Em sua essência, um dado só é considerado pessoal quando estabelece uma relação intrínseca e direta com o indivíduo (Machado, 2023, p. 06-07).

Assim, o ponto de inflexão entre dado e dado pessoal encontra-se precisamente na passagem da mera representação simbólica para a informação dotada de contexto e significado. Quando um conjunto de dados, antes neutro, adquire a capacidade de individualizar alguém, seja por referência direta, seja pela possibilidade de associação com outros elementos, deixa de ser apenas um registro estrutural e passa a se inserir no domínio da proteção jurídica. O que caracteriza o dado pessoal não é a sua natureza formal de símbolo, mas o fato de que, ao ser interpretado em determinado contexto, ele se torna veículo de identificação de uma pessoa natural (Machado, 2023, p. 07-08).

Ao delimitar o dado pessoal como categoria jurídica, a LGPD não se limita a fixar um conceito único e homogêneo. Isso porque, embora qualquer dado capaz de identificar um indivíduo mereça tutela, há informações que, pela sua natureza íntima e pelo risco potencial de discriminação, demandam salvaguardas ainda mais rigorosas (Mulholland, 2018, p. 08).

É nesse contexto que surge a noção de dado pessoal sensível, estabelecendo uma distinção qualitativa dentro do próprio universo dos dados pessoais. Assim, a LGPD traz em seu Art. 5º, inciso II a definição de dado pessoal sensível, um conceito que exige uma proteção ainda mais precisa. A categoria engloba dados sobre origem racial ou étnica, convicção religiosa, opinião política, filiação sindical ou a organização de caráter religioso, filosófico ou político, dados referentes à saúde, vida sexual, dados genéticos ou biométricos (Teffé, 2022, p. 15).

Cumpre esclarecer que, uma vez submetido a processos de tratamento, o dado deixa de ser estático e passa a integrar fluxos dinâmicos que envolvem múltiplos agentes e distintas

finalidades. É justamente nesse itinerário que emergem os maiores riscos de violação à privacidade e torna-se necessário compreender o ciclo de vida do dado pessoal (Albers, 2016, p. 13).

A análise do ciclo de vida do dado pessoal permite compreender como a informação, desde o instante em que é coletada, percorre diferentes estágios de tratamento que exigem salvaguardas específicas. Cada uma dessas etapas é juridicamente relevante, pois implica diferentes graus de risco à privacidade e à autodeterminação informativa do titular, exigindo dos agentes de tratamento uma regulação integral do processo (Brasil, 2018).

Nesse percurso, a atuação dos sujeitos responsáveis (controlador, operador e encarregado) revela-se determinante para a concretização da proteção de dados, impondo uma rede de deveres cuja finalidade última é resguardar a dignidade e a autodeterminação informativa da pessoa natural, nos termos dos Arts. 7º; 11; e 37 a 41 da Lei nº 13.709/2018 (Brasil, 2018).

Posto isso, os dados pessoais podem ser produzidos, adquiridos no mercado informacional ou simplesmente coletados por meio de interações digitais e presenciais. A coleta, nesse sentido, constitui etapa inicial do ciclo de tratamento, em que a informação é obtida por formulários, *cookies* ou outras tecnologias, devendo sempre estar vinculada a uma finalidade legítima, específica e informada ao titular, em observância ao princípio da finalidade (Amaral, 2016, p. 17-25).

Superada essa fase, os dados passam a ser processados e armazenados. O processamento compreende qualquer operação que manipule o dado com vistas a um objetivo determinado, ao passo que o armazenamento corresponde à sua guarda em bases de dados. Em ambas as etapas, a legislação impõe a adoção de medidas de segurança técnicas e administrativas aptas a resguardar a informação contra acessos não autorizados, vazamentos e demais incidentes que possam comprometer a privacidade do titular (Amaral, 2016, p. 17-25).

Por fim, a etapa da eliminação ou anonimização dos dados merece ênfase como fase terminal do ciclo de vida. O dado pessoal não pode ser tratado indefinidamente: sua conservação só se legitima enquanto perdurar a finalidade que justificou sua coleta. Uma vez esgotada tal finalidade, a manutenção dos dados se converte em violação ao princípio da necessidade e em potencial afronta ao direito fundamental à proteção de dados (Bioni, 2020, p. 03-05).

Nesse sentido, o esquecimento digital deixa de ser uma faculdade eventual e passa a constituir dever jurídico dos controladores e operadores, sob pena de perpetuar a exploração indevida da

memória informacional do titular. A efetividade do direito ao esquecimento, portanto, não se resume a um imperativo ético ou a uma expectativa social, mas se afirma como comando normativo que obriga os agentes de tratamento a assegurar, no plano concreto, a cessação do uso de informações pessoais quando ausente a finalidade que lhes deu origem (Bioni, 2020, p. 03-05).

Nesse percurso, o primeiro agente a ser considerado é o titular do dado, isto é, a pessoa natural a quem a informação se refere, conforme Art. 5º, V da LGPD (Brasil, 2018). A centralidade da figura do titular não pode ser relativizada: ele não é mero objeto da atividade econômica, mas sim sujeito de direitos fundamentais cuja proteção se sobrepõe a interesses meramente mercadológicos (Pinheiro, 2021, p. 286). A ele cabe, por exemplo, o direito de consentir ou não com o tratamento, de revogar consentimento, de acessar e corrigir informações e de exigir a eliminação de seus dados quando não mais necessários às finalidades inicialmente declaradas, vide Arts. 7º,I e §5º; 8º,§5º; 18,I a VI; e 18, §6º da Lei nº 13.709/2018 (Brasil, 2018). Assim, a figura do titular constitui a medida e o limite de toda atividade de tratamento de dados pessoais.

Ainda que muitas vezes não se apresente de forma nominal, o titular pode ser identificado por meio de elementos técnicos que, isolados ou combinados, permitem sua individualização. Endereços de IP (protocolo de internet), por exemplo, registram o ponto de conexão utilizado e podem ser cruzados com dados de provedores de internet para localizar o usuário (Santos *et al.*, 2016, p. 14-16). *Cookies* vinculados a sessões de navegação permitem reconhecer padrões de acesso e hábitos de consumo, de modo a diferenciar um indivíduo dentro de um grupo de usuários (Lopes *et al.*, 2022, p. 8-9). Geolocalizações captadas por aplicativos ou dispositivos móveis revelam rotinas e trajetos pessoais que, associados a horários e frequências, tornam inequívoca a vinculação a uma pessoa natural (Silva *et al.*, 2022, p. 12).

Do mesmo modo, metadados de arquivos digitais, como a data, o local de criação ou o dispositivo utilizado, funcionam como rastros técnicos que, analisados em conjunto, possibilitam a reconstituição da identidade (Neto; Morais; Bezerra, 2017, p. 16-17). É justamente essa capacidade de reconstrução contextual que justifica a amplitude da noção de dado pessoal adotada pela LGPD, impondo proteção também àquelas informações que, ainda que isoladamente neutras, revelam o indivíduo quando articuladas entre si (Brasil, 2018).

Por outro lado, a LGPD distingue no Art. 7º, incisos VI, VII e IX, os agentes de tratamento do dado entre controlador e operador (Brasil, 2018). O controlador é aquele que toma as

decisões essenciais sobre o tratamento: define finalidades, critérios de coleta, meios de utilização e hipóteses de compartilhamento (Sousa; Vasconcelos; Sousa, 2018, p. 178).

Ainda, o operador, por sua vez, atua sob a determinação do controlador, executando tarefas técnicas, mas sem autonomia decisória sobre o destino da informação. A diferença, longe de ser meramente formal, é de ordem substancial, pois define os contornos da responsabilidade civil e administrativa em caso de violação normativa (Sousa; Vasconcelos; Sousa, 2018, p. 178).

A esses agentes soma-se a figura do encarregado de dados, ou *Data Protection Officer* (DPO), instituído pela LGPD como elo de comunicação entre o controlador, os titulares e a Agência Nacional de Proteção de Dados (ANPD). Trata-se de um papel essencial para a implementação de uma cultura de conformidade, pois é por meio dele que se garante a transparência, a prestação de contas e a efetividade dos direitos dos titulares (Martins e Longhi, 2025, p. 40-41). Dessa forma, o encarregado não é um mero funcionário técnico, mas um instrumento institucional de governança e *accountability* (Pinheiro, 2021, p. 286).

O ciclo de vida dos dados pessoais também deve ser analisado à luz dos princípios gerais do tratamento, previstos no Art. 6º da LGPD (Brasil, 2018). O princípio da finalidade, previsto no inciso I, impõe que os dados sejam coletados para propósitos determinados, explícitos e legítimos, vedada a coleta indiscriminada. O princípio da necessidade exige a limitação do tratamento ao mínimo indispensável, evitando a acumulação de informações em excesso, conforme inciso III do mesmo artigo. Já o princípio da adequação assegura que o tratamento seja compatível com o contexto em que o dado foi fornecido, respeitando o inciso II e impedindo que o titular seja surpreendido por usos desviados da sua expectativa legítima (Teffé; Viola, 2020, p.15).

Outro pilar fundamental é o princípio da transparência, que exige que os titulares sejam informados, de forma clara e acessível, acerca das práticas de tratamento a que seus dados estão submetidos, nos termos do Art. 6º, VI da LGPD (Brasil, 2018). A transparência não se confunde com a simples disponibilização de termos extensos e obscuros, mas demanda uma comunicação efetiva, capaz de permitir que o indivíduo compreenda e exerça seus direitos de maneira consciente. Sem transparência, os direitos de informação, retificação e exclusão tornam-se meramente formais, carecendo de exequibilidade prática (Teffé; Viola, 2020, p. 10).

Além disso, destaca-se o princípio da responsabilização e prestação de contas

(*accountability*), que obriga os agentes de tratamento a demonstrarem, de forma documentada, a observância e a eficácia das medidas adotadas para cumprir a lei. O princípio do Art. 6º, X projeta uma lógica de co-responsabilidade, na qual não basta alegar a adoção de boas práticas: é imprescindível provar sua efetividade, inclusive mediante relatórios de impacto exigidos pelo Art. 38 da LGPD, bem como pela obrigação de manter registros das operações de tratamento e da adoção de medidas de segurança, conforme os Arts. 37, 38, parágrafo único e 48 da mesma lei (Brasil, 2018) . Em um cenário de circulação globalizada de dados, tal dever ganha contornos ainda mais robustos, pois obriga empresas transnacionais a internalizar mecanismos de conformidade regulatória (*compliance*) que transcendem fronteiras (Willemin; Faria; Amante, 2018, p. 102-105).

Contudo, a compreensão do ciclo de vida do dado pessoal não se esgota na análise das obrigações jurídicas. É preciso reconhecer que, por trás desse processo, há uma dimensão econômica que impulsiona e condiciona a própria dinâmica do tratamento de dados. A informação pessoal, ao mesmo tempo em que constitui expressão da identidade individual, converte-se em ativo estratégico para empresas e plataformas digitais, servindo como insumo essencial na geração de valor (Sattler, 2018, p. 43).

Essa dupla natureza, tanto jurídica quanto econômica, explica a necessidade de um regime normativo rigoroso, pois a crescente mercantilização da informação tende a ampliar os riscos de instrumentalização do indivíduo como simples recurso a ser explorado. Assim, a Seção seguinte passa a examinar a economia movida a dados e a mercantilização da informação pessoal, contexto no qual a ambivalência do dado atinge sua máxima expressão.

2.2 A ECONOMIA MOVIDA A DADOS E A MERCANTILIZAÇÃO DA INFORMAÇÃO PESSOAL

À luz de uma análise histórica, constata-se que com o restabelecimento das noções de "concorrência" após longos períodos de uma economia pautada nos monopólios, a Primeira Revolução Industrial (1760) ajudou a redimensionar as bases do comércio europeu. Estabeleceu-se então, a concepção de que, amparados pelos ideais de livre-concorrência, os comerciantes poderiam adotar as práticas que bem escolhessem. Em um ambiente de liberalização econômica, a informação pôde ser percebida como vantagem competitiva,

começando a emergir a importância da informação para a dinâmica mercantil. No entanto, à epoca, o verdadeiro núcleo da geração de valor era o da produção material e da força de trabalho (Forgioni, 2018, p. 63-69).

Sob a ótica brasileira, a própria concepção de liberdade econômica só passa a ter algum respaldo normativo com a Constituição de 1934, especificamente nos Arts. 115 e 117, que delinearam pela primeira vez princípios voltados à livre iniciativa e à intervenção estatal equilibrada na ordem econômica (Forgioni, 2018, p. 104), o que tornava a noção de poder informacional praticamente inexistente, ou ao menos inconcebível em termos econômicos. Mesmo posteriormente, já com ambientes físicos estruturados para conectar ofertantes e demandantes, como feiras e centros comerciais, o papel da informação seguia sendo secundário e acessório (Pinheiro, 2021, p. 33).

Em termos globais, ainda que o papel da informação como vantagem competitiva tenha começado a despontar com a Primeira Revolução Industrial, foi somente com os avanços tecnológicos do século XX que se consolidaram as bases materiais e técnicas para que dados passassem a desempenhar função central na reprodução do capital (Pinheiro, 2021, p. 34). As inovações no campo da computação contribuíram para a popularização da internet na década de 1990, conectando indivíduos e empresas de maneiras inéditas (Pasquale, 2015, p. 107).

A ascensão das redes sociais nos anos 2000 ampliou significativamente a capacidade de interação e troca de informações entre os usuários: em 1986, apenas 1% da informação global encontrava-se em formato digital, na presença das redes, o número que cresceu para 25% em 2000, até alcançar impressionantes 98% em 2013 (Breen, 2005, p. 299 *apud* Zuboff, 2021, p. 232).

Desse modo, com a transição para o meio digital, a relação ofertante-demandante adquiriu uma profundidade inédita, fundamentada na captação contínua de dados comportamentais dos usuários. O salto tecnológico que permitiu a ascensão das plataformas digitais consequentemente abriu espaço para uma exploração facilitada dos dados pessoais dos titulares, visto que a digitalização de processos comerciais trouxe uma capacidade inédita de extrair, armazenar e analisar quantidades massivas de informações (Wanderer, 2023, p. 3-4).

Nesse arranjo, os dados deixam de ser mero subproduto (Zuboff, 2021, p. 90) e viraram o principal ativo da economia digital, na medida em que o valor dos dados decorre de sua constante atualização e do acesso às suas fontes, mais do que de seu conteúdo em si (Ursic,

2018, p. 72-73). Desse modo, é notório que no panorama atual, a economia tem sido veemente movida pelo vigilantismo (Zuboff, 2021, p.22-23), onde não basta apenas o capital produtivo mas sim a datatificação da experiência humana para um melhor direcionamento dos insumos produtivos (Crespo; Santos, 2018, p.179-180).

A digitalização dos processos comerciais permitiu um nível de predição e controle mercadológico impossível de se imaginar em modelos anteriores. Com essas bases estabelecidas, passa-se ao exame do fenômeno que consolida essa nova ordem: o *big data*.

2.2.1 Coleta massiva e opaca dos dados pessoais: o *big data*

A dinâmica contemporânea do fluxo informacional é marcada por um fenômeno inédito na história humana: a extração e acúmulo contínuos, automáticos e em larga escala de dados pessoais. A expressão "*big data*" representa não apenas um conjunto de ferramentas tecnológicas voltadas à coleta e análise de dados, mas principalmente uma nova racionalidade produtiva centrada na coleta massiva, na opacidade dos mecanismos de extração e no uso econômico de dados sensíveis como matéria-prima. Trata-se, portanto, de uma realidade estrutural que escapa ao conhecimento e controle do titular dos dados, promovendo uma assimetria informacional radical entre os indivíduos e os agentes econômicos que operam tais sistemas (Taurion, 2013, p. 32–34).

A lógica do *big data* desloca o sujeito de direito tradicional para a condição de um produtor involuntário e constante de informações. Em cada clique, deslocamento, consumo ou reação em redes sociais, o usuário alimenta bancos de dados que, muitas vezes, nem sequer sabem que existem. O modelo hegemônico de exploração de dados não depende da vontade expressa dos indivíduos: sua arquitetura técnica pressupõe vigilância constante e invisível, promovendo um processo de coleta opaca e quase indetectável (Zuboff, 2021, p. 23).

Nessa linha, mesmo quando há uma aparência de consentimento, como nas chamadas "políticas de privacidade"; "políticas de *cookies*" ou "termos de uso", a lógica contratual vigente é unilateral, mutável e redigida de maneira a dificultar a compreensão do titular médio. Nesse sentido, Zuboff alerta que tais contratos são alteráveis a qualquer tempo, sem consentimento específico, vinculando terceiros e eximindo os fornecedores de responsabilidade, uma lógica regressiva descrita como "sádica" pela professora Nancy Kim

(Kim, 2013, p. 50-69 *apud* Zuboff, 2021, p. 68).

De certo, o que está em curso é a transformação da própria subjetividade humana em ativo mercadológico. As plataformas digitais, ao estimularem e recompensarem a autoexposição contínua, criam um ambiente em que os próprios usuários se tornam cúmplices da vigilância que os atinge. Essa lógica demonstra que os indivíduos são levados à auto devassa voluntária, na esperança de reconhecimento social, pertencimento ou mesmo relevância, ainda que isso signifique renunciar à própria privacidade (Han, 2018, p. 81, 87-88).

Impulsionado pela tecnologia do *big data*, um novo modelo de negócio se consolida, redefinindo o papel do consumidor: este deixa de ser apenas destinatário de bens e serviços para tornar-se ele próprio um trabalhador invisível e não remunerado, criador de valor ao produzir dados que, posteriormente, são monetizados pelas empresas. A interdependência entre consumo e produção de dados cria uma linha tênue entre a atividade econômica e a vivência pessoal, instaurando uma nova forma de exploração econômica das subjetividades (Fálcon, 2024, p. 07).

Esse cenário é agravado pela complexidade e pelo gigantismo dos fluxos de dados. A quantidade de informações geradas diariamente por sensores, dispositivos móveis, redes sociais, assistentes virtuais e câmeras urbanas ultrapassa a capacidade humana de monitoramento consciente. Segundo Taurion, se está diante de uma verdadeira avalanche informacional, cuja análise requer novos métodos e cujas fontes se multiplicam exponencialmente, sobretudo com a Internet das Coisas (IoT), que conecta objetos físicos ao universo digital em tempo real (Taurion, 2013, p. 32–34).

Do ponto de vista jurídico, a coleta massiva de dados pessoais levanta sérias tensões com o princípio da necessidade, princípio basilar do tratamento de dados pessoais. Embora o desenvolvimento de sistemas de inteligência artificial exija grandes volumes de dados para treinamento, a ANPD ressalta que essa exigência técnica não pode se sobrepor à obrigação legal de limitar a coleta ao mínimo necessário para a finalidade específica, devendo-se adotar salvaguardas técnicas e jurídicas rigorosas para compatibilizar desenvolvimento tecnológico e proteção de direitos fundamentais (Agência Nacional De Proteção De Dados (ANPD), 2025, p. 7-8).

Em síntese, a realidade do *big data* projeta um modelo de sociedade em que a transparência é exigida apenas do cidadão, enquanto a arquitetura tecnológica e jurídica das grandes

corporações permanece envolta em opacidade. A coleta massiva de dados, aliada à ausência de controle efetivo e à lógica da autoexposição incentivada, compromete o núcleo da autodeterminação informacional e desafia os alicerces do direito à privacidade em sua dimensão mais profunda e contemporânea. Assim, é exposto na subseção a seguir, como ocorre a transformação dos dados pessoais em informação e como a informação é transformada em ativo econômico.

2.2.2 A transformação dos dados pessoais em informação e da informação em ativo econômico

Durante a redação do presente trabalho, refere-se à "informação" em um sentido amplo, abrangendo seu caráter genérico. Passa-se a discutir, nesse tópico em especial, o conceito em sua acepção estrita, focando na informação pessoal como um elemento dotado de valor específico, estratégico e utilitário. Assim, para propósitos exclusivos desta Subseção, é fundamental distinguir os conceitos de dado pessoal e informação.

Enquanto o dado pessoal pode ser compreendido como um fato identificável sobre um titular, desprovido de benefício imediato, a informação representa um conjunto de dados organizados, processados e contextualizados que adquirem significado e utilidade (Wanderer, 2023, p. 05). A transformação dos dados pessoais em informação é um processo que agrega valor à medida que esses dados são coletados, correlacionados e analisados. O mero registro de um clique, quando este pode ser associado a um titular - seja pelo endereço de IP ou um *cookie* associado a sessão em um site -, é um dado pessoal bruto (Zuboff, 2021, p. 72).

Porém, quando este clique é associado a outros dados pessoais, como um histórico de navegação, localização geográfica e preferências de compra, ele se torna uma informação valiosa sobre o comportamento do usuário (Wanderer, 2023, p. 13). Essa agregação e contextualização enseja na prática do perfilamento (*profiling*), ou seja, a criação de perfis detalhados, que, antes (apesar de existirem) eram inatingíveis em grande escala (Custers; Calders; Schermer, 2013, p. 05).

Uma vez transformados em informação, os dados pessoais adquirem o *status* de ativo econômico. Empresas de tecnologia, especialmente as *big techs*, monetizaram esses conjuntos informacionais ao utilizá-los para direcionar publicidade, personalizar serviços e otimizar

processos (Morozov, 2018, p. 155). Essa nova lógica de valorização não depende da propriedade física de um bem, mas sim da capacidade de processar e utilizar a informação gerada pelos indivíduos em suas interações digitais (Crespo; Santos, 2018, p. 179-180).

Mais do que um simples ativo, a informação pessoal se tornou um insumo estratégico fundamental para a economia digital. Ela alimenta algoritmos de inteligência artificial generativa, que, por sua vez, aprimoram a capacidade de predição e segmentação do mercado. Sem o fluxo contínuo e massivo desses dados, muitos dos serviços e modelos de negócio que hoje sustentam as maiores empresas do mundo não seriam viáveis (Morozov, 2018, p. 155).

Nessa moldura, empresas como Google, Amazon e Facebook são exemplos paradigmáticos dessa lógica de extração de valor. Por meio de *cookies*, histórico de buscas, localização geográfica, preferências de consumo e comportamento *on-line*, constroem-se perfis detalhados dos usuários, que são, então, utilizados para ofertar serviços de publicidade hiper segmentada a anunciantes interessados em alcançar grupos específicos com máxima eficácia (Lundqvist, 2018, p. 195-196).

A partir de informações como essas, algoritmos são constantemente desenvolvidos e refinados com base nos fluxos incessantes de informações coletadas, o que permite a criação de sistemas cada vez mais precisos e autônomos. Nesse cerne, os algoritmos assumem papel central para classificar e recomendar conteúdo ou produtos técnicas avançadas de inteligência artificial para antecipar comportamentos, personalizar experiências e maximizar o tempo de permanência do usuário em seus ecossistemas digitais (Wanderer, 2023, p. 13).

Trata-se, portanto, de um ciclo retroalimentado: quanto mais dados pessoais são coletados, mais informações se consolidam e mais eficazes se tornam os algoritmos; quanto mais eficazes, maior o volume de dados pessoais gerado, consolidando o domínio dessas empresas sobre o mercado da atenção e da informação (Wanderer, 2023, p. 13).

A capacidade de analisar e reagir em tempo real aos padrões de comportamento dos usuários, por meio do uso intensivo de *big data* e algoritmos, permite às empresas obterem uma vantagem competitiva sem precedentes (Pasquale, 2015, p. 19-22). O valor não está no dado pessoal estático, mas em sua constante atualização e no acesso privilegiado às fontes que o geram, que o transformam em informação (Surblytė, 2016, p. 08-09).

Por sua vez, a transformação da informação em ativo econômico se manifesta de diversas formas na economia digital. Primeiramente, ela se torna diretamente comercializável, como

ocorre no mercado de *big data* para fins de marketing e publicidade programática, onde perfis de consumidores são vendidos e comprados entre anunciantes e plataformas. Segundo relatório da IDC Corporate, o mercado global de *big data* movimentou 16,1 bilhões de dólares em 2014, com crescimento estimado para 32,4 bilhões de dólares em 2017, e projeções apontando cifras de até 114 bilhões de dólares em 2018. Tais números evidenciam que a informação, tratada por sistemas cada vez mais complexos, não apenas circula como mercadoria, mas estrutura novos modos de organização econômica, cuja lógica extrativa depende da captura contínua dos dados pessoais e da sua constante transformação em informação (Tsai *et al.*, 2015, p. 02).

Além disso, a informação funciona como material para a inovação, permitindo o desenvolvimento de novos produtos e serviços personalizados, desde recomendações de filmes em plataformas de *streaming* até rotas otimizadas em aplicativos de transporte, todos aprimorados pelo aprendizado contínuo dos padrões de uso. No entanto, esse processo de inovação, longe de ser neutro ou espontâneo, tem sido impulsionado por uma lógica de mercado marcada pela urgência em gerar retornos financeiros imediatos (Zuboff, 2021, p. 319-320).

Sob o manto da disrupção tecnológica, muitas dessas inovações têm operado em regime de "experimentação permanente", frequentemente à revelia de regulamentações e sem a devida consideração pelas consequências sociais e econômicas que produzem. O que se observa, portanto, é uma corrida por dominação de mercado travestida de inovação, em que a consolidação da informação e o uso intensivo de algoritmos ocupam papel importantíssimo na reprodução de assimetrias econômicas (Zuboff, 2021, p. 70).

Acrescente-se que a informação atua como um diferencial competitivo crucial, permitindo que empresas otimizem suas operações e obtenham *insights* estratégicos sobre o mercado. Por exemplo, uma plataforma de comércio eletrônico pode usar dados de compras e navegação para prever tendências de consumo, ajustar estoques e até mesmo influenciar o *design* de novos produtos, reduzindo riscos e aumentando a lucratividade. Esse uso estratégico da informação vai além da simples venda de dados pessoais, consolidando uma vantagem informacional que se traduz em poder de mercado e, por vezes, em monopólios ou oligopólios digitais (Falcón, 2024, p. 07-11).

Em outras palavras, a mercantilização da informação pessoal ocorre mesmo em serviços aparentemente "gratuitos" (Falcón, 2024, p. 07). O usuário, ao utilizar redes sociais,

aplicativos de mensagens ou ferramentas de busca sem custo financeiro direto, na verdade "paga" com seus dados e sua atenção, que são convertidos em valor para as plataformas (Morozov, 2018, p. 133-134). A capacidade de prever comportamentos e influenciar decisões, baseada nesses dados, confere às empresas um poder sem precedentes sobre a vida social e econômica dos indivíduos, caracterizando uma economia de vigilância onde o valor não é mais gerado apenas pela produção material, mas pela datatificação da existência (Crespo; Santos, 2018, p.179-180).

Apesar das ressalvas no início desta subseção, cumpre esclarecer que, nas Subseções, Seções e Capítulos a seguir, os termos dado pessoal e informação pessoal serão novamente empregados como expressões equivalentes. A razão dessa opção repousa justamente na correlação antes traçada: o dado pessoal, enquanto registro isolado e fragmentado, carece de significado autônomo. Somente quando contextualizado e correlacionado adquire a qualidade de informação, e, uma vez que essa informação esteja vinculada a um indivíduo identificável, é juridicamente qualificada como dado pessoal.

Assim, ainda que a teoria distinga dado de informação, no campo normativo a legislação brasileira os utiliza de forma intercambiável (Machado, 2023, p. 02-03), de modo que, a partir deste ponto, ambos os vocábulos serão considerados sinônimos no tratamento da temática.

Posto isso, a lógica de valorização dos dados pessoais e sua transformação em vantagem competitiva encontra concretude nos modelos de negócio adotados pelos grandes conglomerados tecnológicos. São essas corporações que exemplificam, de forma paradigmática, como a informação pessoal se tornou o principal insumo da economia digital. Passa-se a análise da subseção que trata das *big techs*, modelos de negócios baseados em coleta e tratamento de dados.

2.2.3 Modelos de negócio baseados em coleta e tratamento de dados: as *big techs*

Em um cenário marcado pela transformação dos dados pessoais em ativo econômico, torna-se necessário identificar os principais agentes ativos nesse processo: as *big techs*. Também conhecidas como grandes conglomerados tecnológicos, representam um modelo de negócio impulsionado pela coleta e pelo tratamento de dados pessoais (Morozov, 2018, p. 149). No cerne desse ecossistema estão as plataformas digitais, que atuam como intermediárias entre

distintos polos de interesse, conectando consumidores e fornecedores em arranjos típicos de mercados de duas pontas (Wanderer, 2023, p. 04).

O panorama da alta lucratividade do *big data* potencializado pela capacidade das *big techs* de criarem as próprias plataformas digitais, revelou consigo um modelo de negócios altamente assimétrico no que diz respeito à relação titular-controlador dos dados (Zuboff, 2021, p. 130). É evidente que o cenário em questão demanda urgente reavaliação jurídica.

Além da centralidade conferida aos dados pessoais como insumo estratégico, os modelos de negócio alicerçados na coleta e no tratamento de informações operam, também, sob a lógica da chamada economia da atenção (Varoufakis, 2025), na qual o tempo do usuário se torna moeda de troca invisível, mas altamente valorizada (Bentes, 2021, p. 192). A atenção, nesse contexto, constitui o elo essencial entre presença digital e rentabilidade econômica, uma vez que permite a intensificação do monitoramento comportamental, viabilizando a personalização de conteúdos, a segmentação publicitária e a maximização da eficiência algorítmica (Zuboff, 2021, p. 115-116).

O advento dessa nova racionalidade econômica promove uma verdadeira reconfiguração dos critérios de valor: aquilo que antes da era digital era medido por produtos ou serviços concretos, passa a ser dimensionado em métricas de engajamento, tempo de tela e profundidade de interação com os ambientes digitais. O *design* das plataformas, intencionalmente desenhadas para prolongar a permanência do usuário, têm por fim último a intensificação da coleta de dados (Bentes, 2021, p. 192).

Nesse cenário, o papel das plataformas digitais se torna de intermediar as *big techs* e os titulares dos dados: elas não apenas conectam fornecedores e consumidores, mas passam a exercer uma função central de filtragem e ordenação do fluxo informacional, assumindo uma posição de controle que ultrapassa a lógica mercadológica e adentra o campo político e social. É nesse ponto que se consagra a figura do *gatekeeper*, isto é, um ente digital que, por sua posição dominante, consegue decidir o que será visível e relevante dentro do espaço virtual (Wanderer, 2023, p. 3).

Sob a ótica jurídico-consumerista, a atuação das *big techs*, por meio das plataformas digitais, configura relação de consumo nos termos do artigo 2º do Código de Defesa do Consumidor (CDC), que define como consumidor toda pessoa física ou jurídica que adquire ou utiliza produto ou serviço como destinatário final. Ainda que os serviços prestados sejam

formalmente gratuitos, o artigo 3º, §2º, do mesmo diploma legal, ao prever que "serviço é qualquer atividade fornecida no mercado de consumo, mediante remuneração", deve ser interpretado na sua forma ampla, a fim de abarcar modelos de remuneração indireta, como é o caso dos provedores de conteúdo que monetizam a atenção do usuário e os dados pessoais coletados (Brasil, 1990).

O Superior Tribunal de Justiça, através do Recurso Especial n. 1192208 MG 2010/0079120-5, julgado em 2012 pela Relatora da Terceira Turma, Ministra Nancy Andrighi, já consolidou esse entendimento, ao reconhecer que o fato de o serviço prestado ser gratuito não afasta a incidência do CDC, quando houver contrapartida econômica indireta resultante do fornecimento de dados, da exposição à publicidade ou de outros mecanismos de capitalização da atividade digital.

Nesse sentido, as plataformas digitais que operam com base em modelos de negócios fundados na coleta, armazenamento e tratamento de dados pessoais devem ser juridicamente qualificadas como fornecedoras, nos termos do Art. 3º, *caput*, do CDC, pois exercem atividade econômica voltada à oferta de serviços no mercado de consumo. A própria lógica contratual subjacente às interações digitais contemporâneas revela a existência de um vínculo jurídico pautado pela prestação contínua de serviços por parte das plataformas e pela entrega de dados e atenção por parte dos usuários, caracterizando estes como consumidores por equiparação (Art. 29 do CDC), ainda que em uma relação mediada por interfaces tecnológicas. Trata-se, portanto, de uma relação assimétrica e vulnerável, que demanda especial tutela jurídica à luz dos princípios da boa-fé objetiva, da transparência e da proteção contra práticas abusivas, tal como exigido pelo Arts. 4º e 6º do CDC (Brasil, 1990).

Para os consumidores, os impactos vão além da dimensão econômica (Machado e Ruaro, 2017, p. 2). A coleta incessante e o processamento sofisticado de seus dados pessoais comprometem sua autodeterminação informativa e sua liberdade de escolha (Machado; Ruaro, 2017, p. 8).

A personalização algorítmica, ao contrário de ampliar as possibilidades de consumo e informação, tende a criar bolhas de filtro e câmaras de eco, nas quais o sujeito é exposto apenas a conteúdos que reforçam suas crenças preexistentes. Isso limita sua visão de mundo e dificulta o exercício de uma cidadania crítica e bem-informada (Balkin, 2017, p.45-46). O poder das *big techs* não é apenas econômico; ele é, ao mesmo tempo, epistemológico e axiológico, pois define o que pode ser conhecido, acessado e, em última instância, vivido no

ambiente virtual (Bentes, 2021, p. 232-234).

Nessa ótica, acentuam-se as preocupações com a tutela da proteção de dados precisa ser redimensionada para enfrentar um modelo de negócios cujo ativo central é a própria subjetividade do usuário (Morozov, 2018, p. 140).

Entre as chamadas *big techs*, destacam-se especialmente quatro conglomerados: Google (*Alphabet*), Amazon, Apple e Meta (anteriormente *Facebook*). Essas empresas formam o núcleo do poder vetorial contemporâneo e sustentam sua dominação sobre a infraestrutura digital global por meio do controle dos fluxos de dados, das ferramentas de busca e indexação, das plataformas *on-line* e das redes sociais de maior alcance no planeta (Widder; Whittaker; West, 2023, p. 8-9).

Cada uma exerce funções que ultrapassam os limites tradicionais da intermediação digital e invadem domínios como a vigilância comportamental, a definição de padrões tecnológicos e a curadoria algorítmica da realidade. O modelo da Google, por exemplo, se ancora em um duopólio de indexação e publicidade digital; a Amazon estende seu alcance desde a infraestrutura de servidores até o controle logístico sobre cadeias globais de consumo; a Apple regula o ecossistema de seus dispositivos por meio de uma arquitetura fechada e interoperável; e a Meta transforma a sociabilidade em mercadoria, ao rentabilizar interações humanas sob a lógica da atenção e da vigilância. Ainda que suas estratégias comerciais e tecnológicas variem, todas compartilham a característica estrutural de operarem com recursos massivos, acumulando não apenas dados e capital, mas capacidade de moldar normativamente o ambiente informacional contemporâneo (Widder; Whittaker; West, 2023, p. 18).

Visto isso, a economia digital contemporânea não pode mais ser compreendida apenas sob a lógica tradicional do capitalismo industrial. Conforme argumenta Wark (2019, p. 10), o mundo já opera sob o domínio de uma nova classe dominante: os vetoralistas. Diferentemente dos capitalistas clássicos, cujo poder se assentava na posse dos meios físicos de produção, os vetoralistas exercem sua dominação através do controle dos vetores de extração e organização da informação (Wark, 2019, p. 39). O titular dos dados, ao navegar, consumir conteúdo ou interagir com interfaces, realiza uma atividade cognitiva que enriquece a base de dados das empresas vetoralistas, sem que haja qualquer tipo de contraprestação, contrato formal ou mesmo consciência clara da exploração em curso (Zuboff, 2021, p. 282-283).

Esse panorama conduz diretamente ao exame da publicidade comportamental como eixo

central da monetização da atenção.

2.2.4 Publicidade comportamental e monetização da atenção

A valorização do dado como mercadoria não se dá por sua existência, mas pela sua operacionalização nos sistemas de informação, na qual o titular dos dados pessoais, ao entregar suas informações para acessar serviços, tem suas experiências cotidianas transformadas em dados rentáveis (Pasquale, 2015, p. 39-40).

Retomado este raciocínio, torna-se necessário evidenciar a etapa subsequente: a forma pela qual esses dados são instrumentalizados para reconfigurar a lógica da publicidade, convertendo a atenção humana em recurso econômico escasso e, portanto, mercantilizável. Esse processo não apenas redefine práticas mercadológicas, mas também inaugura uma forma de governo da subjetividade, em que o titular é menos agente e mais objeto de manipulação (Carvalho, 2018, p. 71).

A publicidade tradicional, voltada a grandes massas, cede lugar a estratégias hiperpersonalizadas, nas quais cada indivíduo é permanentemente perfilado a partir de seus rastros digitais. Conforme observam Machado e Ruaro (2017, p. 06-08), os arquivos de consumo permitem a transição para uma publicidade contextual capaz de endereçar mensagens a destinatários únicos, ampliando sobremaneira a assimetria informacional. Trata-se de um salto qualitativo: o mercado já não apenas observa preferências, mas condiciona comportamentos, minando o próprio conceito de escolha livre no âmbito do consumo.

Nesse cenário, emerge com força a utilização de mecanismos conhecidos como *nudges*: estímulos comportamentais sutis, mas sem força cogente (Fornasier; Knebel, 2020, p. 22). Embora concebidos como instrumentos de incentivo a escolhas socialmente desejáveis, como doação de órgãos ou combate ao endividamento, sua apropriação pela esfera privada altera substancialmente a finalidade inicial (Fornasier; Knebel; Silva, 2020, p. 10).

O chamado “paternalismo libertário”, que tinha como foco escolhas socialmente desejáveis, converte-se em paternalismo mercadológico, no qual os estímulos invisíveis não buscam o bem comum, mas a maximização de lucros, por meio da alienação progressiva das

necessidades humanas (Thaler; Sunstein, 2008 *apud* Fornasier; Knebel; Silva, 2020, p. 10).

A consequência dessa apropriação é o incentivo a um consumo desenfreado, em que a satisfação é sempre momentânea e rapidamente substituída por novos desejos artificiais. O titular dos dados pessoais assume a posição de consumidor, e acaba por viver em permanente oscilação entre prazer e frustração, submetido a um ciclo incessante de induções publicitárias. Tal dinâmica revela o caráter essencialmente irracional da publicidade comportamental: ao invés de promover autonomia, aprofunda a alienação, transformando a vida cotidiana em espetáculo de consumo contínuo (Fornasier; Knebel; Silva, 2020, p. 10-11).

A publicidade comportamental, contudo, não se limita ao plano puramente econômico. Ela reconfigura relações de poder e afeta diretamente a própria noção de democracia. Como ressalta, o capitalismo de vigilância inaugura uma forma de tirania sem violência explícita, mas fundada na apropriação unilateral da experiência humana como insumo de poder. Ao transformar comportamentos em mercadorias fictícias, esse modelo dissolve a individualidade, impondo um regime de controle privado que substitui as garantias públicas da autodeterminação informativa (Zuboff, 2015, p. 602-606)

Nesse sentido, a captura da atenção não ocorre apenas pela exibição de anúncios, mas pelo desenho arquitetônico das plataformas digitais. As interfaces são planejadas para reter o usuário pelo maior tempo possível, reforçando a monetização por meio do prolongamento da exposição a conteúdos publicitários (Kruger, 2018*)*. Wark denomina esse fenômeno de monopólio da atenção, sustentando que o poder deriva do controle dos vetores de informação. A classe vetorialista explora não apenas a força de trabalho, mas sobretudo a consciência e a subjetividade humanas (Wark, 2019, p. 20-24).

Simultaneamente, a exploração é potencializada pela opacidade estrutural das tecnologias utilizadas: os algoritmos que personalizam buscas e *feeds* constituem verdadeiras "caixas-pretas", imunes à auditoria social. A personalização, que em tese conferiria relevância e eficiência às buscas, acaba por criar bolhas informacionais que reforçam preconceitos e reduzem a diversidade de perspectivas. O indivíduo, acreditando exercer escolhas autônomas, encontra-se de fato aprisionado em um circuito de preferências pré-moldadas pelas corporações (Pasquale, 2015, p. 87-90).

Sob o prisma jurídico, esse contexto desafia diretamente a noção de consentimento informado, pilar da proteção de dados pessoais. Quando escolhas são moldadas por estímulos

imperceptíveis e decisões são guiadas por inferências comportamentais invisíveis ao titular, não se pode falar em autonomia verdadeira. A assimetria informacional é levada ao extremo, exigindo uma reinterpretação crítica do princípio da transparência e do dever de informação previstos no ordenamento brasileiro (Doneda, 2010, p. 61-62).

O que se tem, portanto, é a constituição de uma economia política fundada na expropriação da subjetividade. A publicidade comportamental não apenas explora preferências existentes, mas cria novas necessidades, sustentando um ciclo de consumo infinito e insatisfatório. Esse processo não pode ser compreendido apenas como prática comercial abusiva: ele configura verdadeira colonização da experiência humana pelo capital informacional. Nesse sentido, a publicidade comportamental deve ser analisada como fenômeno que transcende o mercado, impactando dimensões sociais, culturais e políticas da vida contemporânea (Zuboff, 2015, p. 606).

Diante desse quadro, a publicidade comportamental, nesse sentido, antecipa os dilemas que se aprofundarão com a inteligência artificial generativa, pois ambas compartilham a lógica da dependência estrutural de grandes volumes de dados e da manipulação opaca das escolhas individuais. Na próxima Seção, será possível perceber como as técnicas de extração e uso intensivo de dados pessoais, aqui problematizadas, constituem o alicerce do treinamento dos sistemas de inteligência artificial generativa, ampliando exponencialmente os riscos à autodeterminação informativa.

2.3 A INTELIGÊNCIA ARTIFICIAL GENERATIVA COMO CATALISADORA DO USO INTENSIVO DE DADOS PESSOAIS

A inteligência artificial generativa, designa a classe de sistemas capazes de, a partir de entradas (*inputs*) linguísticas ou não linguísticas, inferir autonomamente saídas (*outputs*) que consistem na criação de novos conteúdos textuais, visuais, sonoros ou de código, com graus variáveis de autonomia e adaptação, distinguindo-se por operar não apenas sobre classificações ou previsões, mas sobre a própria síntese de conteúdos culturalmente reconhecíveis e utilizáveis em aplicações gerais ou setoriais (Sengar *et al.*, 2024, p. 4).

A natureza gerativa não é um mero rótulo técnico, ela explica por que tais sistemas funcionam como catalisadores do uso intensivo de dados pessoais. Em primeiro lugar, o desempenho

desses modelos é função direta da escala de dados em que são treinados, sobretudo em arranjos de aprendizado de máquina profundo que exigem volumes maciços e heterogêneos de exemplos para parametrizar redes neurais com milhões ou bilhões de parâmetros, o que acentua a pressão por corpos textuais (*corpora*[1][2]) amplos e variados que, no ambiente da *web*, inevitavelmente contêm dados pessoais em múltiplos contextos e granularidades (Santos, 2025, p. 75).

Em segundo lugar, a própria economia informacional que circunda a IA contemporânea consolidou-se na lógica do *big data* como insumo estruturante do processamento algorítmico, de modo que dados são tratados como matéria-prima indispensável à melhoria incremental de sistemas que aprendem por correções sucessivas e detecção de padrões estatísticos cada vez mais finos (Mantuani, 2025, p. 37).

Por fim, a função catalisadora da IA generativa deve ser situada no tema central do presente trabalho: a responsabilização civil pelo uso não autorizado de dados pessoais no treinamento. Ao transbordar a coleta e o tratamento para dentro da própria materialidade do modelo, a IA generativa cria situações em que o dano potencial não se limita à exposição pontual de um registro, mas pode decorrer da replicação ou inferência indevida em *outputs* gerados a partir de representações paramétricas que contêm informação pessoal em forma não reversível, o que complexifica o nexo causal, a demonstração de violação a deveres de cuidado e a própria reparação específica (Mantuani, 2025, p. 41; Santos, 2025, p. 77).

A Subseção que se segue detalhará, em recortes próprios, a dependência de grandes volumes informacionais da qual se poderá aquilatar, com maior precisão, os contornos da ilicitude observada no uso não autorizado de dados pessoais.

---

[1] DICTIONARY.COM. Corpora. In: Dictionary.com. Disponível em: https://www.dictionary.com/browse/corpus. Acesso em: 30 set. 2025. Definição: plural de corpus. Substantivo. Uma coleção grande ou completa de escritos.

[2] GROTHAUS, Michael. O que é corpus e por que todo mundo no ramo da IA está falando sobre isso. 2023. Disponível em: https://fastcompanybrasil.com/tech/inteligencia-artificial/o-que-e-corpus-e-por-que-todo-mundo-no-ramo-da-ia-esta-falando-sobre-isso/ Acesso em: 30 set. 2025. Explicação: corpora é "uma coleção de dados usados para treinar uma inteligência artificial. É o material que a IA analisa para se tornar inteligente naquilo para o qual foi projetada. Cada uma possui seu próprio corpus, pois são os seres humanos que decidem quais dados serão usados para treiná-la. E a escolha desse corpus depende do que se deseja que a IA seja capaz de realizar."

2.3.1 A dependência de grandes volumes de dados no treinamento dos *softwares* de IA generativa: os *Large Language Models*

Considerando que, a função da IA generativa é criar conteúdo inédito a partir de dados de treinamento, bases de dados extensas são exigidas neste processo (Costa *et al*., 2024, p. 14).

Assim, a arquitetura do funcionamento desta tecnologia reside nos chamados modelos de linguagem de larga escala (*Large Language Models* - LLMs), que são algoritmos capazes de processar volumes massivos de dados para reconhecer padrões estatísticos e replicar estruturas de linguagem ou outros tipos de representação simbólica. Sem essa base quantitativa, os algoritmos não alcançam o refinamento necessário para gerar respostas verossímeis, o que explica a dependência estrutural dessa tecnologia em relação à coleta de dados em escala global (Aaronson, 2023, p. 02-10).

A dependência desses *softwares* de volumes massivos de dados decorre da própria forma de aprendizado adotada. O treinamento a partir da aprendizagem interativa consiste em minimizar o erro de previsão da próxima unidade linguística dada a sequência anterior. Essa tarefa demanda um universo de exemplos suficientemente amplo e heterogêneo para que o modelo internalize não apenas vocabulário e sintaxe, mas também convenções pragmáticas, discursos, jargões e variações sociolinguísticas (Duque-Pereira; Moura, 2023, p. 11-13).

Em termos didáticos, uma IA capaz de redigir petições, redações, contratos e pareceres em português precisa “ter visto” milhares de peças de diferentes ramos, épocas e regiões. Do contrário, tenderá a produzir textos gramaticalmente aceitáveis, porém semanticamente fracos, anacrônicos, enviesados ou descolados da realidade fática (Duque-Pereira; Moura, 2023, p. 18-19).

Esse apetite por cobertura e diversidade temporal, temática e geográfica faz com que dados pessoais acabem, com frequência, integrando o insumo do treinamento, seja porque aparecem de modo indissociável em documentos do mundo real (nomes, endereços, assinaturas, metadados), seja porque são parte do contexto que confere verossimilhança às sequências linguísticas (OECD, 2023, p. 34).

Como assinala a ANPD, existem retornos crescentes de escala na aquisição de dados, de modo que os agentes que coletam maiores quantidades consolidam posição de domínio, perpetuando a concentração de mercado e impondo barreiras a novos entrantes (Costa *et al.*,

2024, p. 20-26). O exemplo do ChatGPT é emblemático: o modelo só foi capaz de se popularizar rapidamente, alcançando milhões de usuários em poucos meses, porque foi treinado em *corpora* que abrange desde livros e artigos científicos até repositórios de código e interações em redes sociais, numa escala inalcançável para atores menores (Aaronson, 2023, p. 02-04).

Logo, modelos que não integram dados recentes degradam ao enfrentar referências normativas, tecnológicas ou culturais dinâmicas, fenômeno particularmente sensível em domínios jurídicos e regulatórios. Em paralelo a própria *corpora*, os *softwares* de IA generativa precisam realizar etapas de alinhamento com *feedback* humano (Duque-Pereira; Moura, 2023, p. 9 e 12).

Enquanto a lei brasileira exige justificativas claras para cada finalidade de uso, os conjuntos massivos de dados que alimentam os LLMs frequentemente reúnem informações heterogêneas, de múltiplas fontes, sem possibilidade de vincular cada fragmento a uma finalidade legítima ou a uma base legal específica prevista no Art. 7º da LGPD (Yan *et al.*, 2025, p. 4; Brasil, 2018).

Nesse ponto, torna-se evidente que a dificuldade não está apenas no volume de dados empregado, mas sobretudo nas condições em que esses dados são obtidos. Se os LLMs dependem de diversidade e constante atualização para manter desempenho competitivo, a pressão por expansão contínua de suas bases de treinamento acaba deslocando o processo de coleta para práticas que operam muito além das fronteiras permitidas no tratamento de dados.

Faz-se necessário então, compreender de onde os dados são extraídos e como passam a integrar os modelos, sendo esta condição indispensável para avaliar o processo de treinamento de IA generativa como um todo. A subseção que se segue detalhará, a origem concreta dos dados que formam os LLMs, com ênfase nas práticas de *web scraping* e na utilização de *datasets* públicos, das quais emergem, com ainda maior nitidez, os contornos da ilicitude no uso não autorizado de dados pessoais.

2.3.2 Origem dos dados de treinamento: *web scraping* e *datasets* públicos

A forma como os dados chegam aos modelos de IA generativa decorre de mecanismos

específicos de coleta e agregação, que se estruturam predominantemente por duas vias principais: (i) a raspagem de informações disponíveis na internet, técnica conhecida como *web scraping*, e (ii) a utilização de *datasets* públicos ou abertos, criados com finalidades diversas, mas posteriormente reaproveitados como insumo para o treinamento (Hong *et al.*, 2025, p. 1-2).

Ambas as estratégias, embora distintas, compartilham um denominador comum: a captação massiva de dados em escala, cuja composição frequentemente abarca dados pessoais, sensíveis ou mesmo informações de menores de idade, gerando implicações jurídicas de relevo à luz da LGPD (Costa *et al.*, 2024, p. 18-20).

No que tange ao *web scraping*, este consiste no uso de programas automatizados, capazes de navegar por páginas eletrônicas, extrair e copiar informações específicas ou gerais, de modo repetitivo e em alta escala. Essa técnica, que pode alcançar muitas páginas em curtos intervalos de tempo, é hoje parte estrutural da economia informacional que sustenta a IA generativa (Gallese, 2024, p. 4-7).

Todavia, a simplicidade de seu conceito contrasta com a complexidade de seus efeitos: ao coletar indiscriminadamente nomes, endereços, opiniões, imagens, vídeos e comentários, a raspagem coloca em xeque os princípios da finalidade e da necessidade, pilares da LGPD, sobretudo quando realizada sem prévio tratamento de anonimização ou sem indicar a base legal - exigida pelo Art. 7º da LGPD - que autoriza tal processamento (Costa *et al.*, 2024, p. 19-20).

Isso indica que, juridicamente, a raspagem levanta um problema sensível: até que ponto dados de acesso público podem ser coletados e reutilizados legitimamente para treinar algoritmos?

A questão ganha contornos ainda mais nebulosos quando se considera que muitos *sites* estabelecem termos de serviço e condições de uso que vedam expressamente a coleta automatizada das suas fontes. Desconsiderar tais cláusulas pode configurar violação contratual e infração de direitos autorais. Um caso paradigma envolvendo o jornal *The New York Times* ilustra que a judicialização da raspagem tem se tornado recorrente, com alegações que vão desde apropriação indevida até quebra de contrato (*Agence France-Presse*, 2023).

Assim, embora a prática do *web crapping* não seja direta e automaticamente ilegal, sua execução em desconformidade com os marcos jurídicos aplicáveis coloca o desenvolvedor em situação de desconformidade jurídica relevante.

A utilização de *datasets* públicos, por sua vez, representa um eixo paralelo de abastecimento informacional para os modelos de IA. Organizações como a *Common Crawl*, por exemplo, disponibilizam gratuitamente vastos repositórios de dados coletados por programas automatizados, permitindo que pesquisadores e empresas acessem volumes massivos de conteúdo textual e multimídia (Costa *et al.*, 2024, p. 18).

Ainda que, em tese, tais bases sejam concebidas como instrumentos de democratização do acesso à informação, sua utilização não elimina riscos de tratamento indevido, uma vez que a abrangência da coleta torna praticamente inevitável a inclusão de dados pessoais e sensíveis. Nesses casos, a conformidade com a LGPD e a obrigação de preservar direitos dos titulares permanecem exigíveis, independentemente da gratuidade ou abertura formal da base (Costa *et al.*, 2024, p. 19).

O reaproveitamento de bancos de dados originalmente criados para outras finalidades levanta, adicionalmente, o debate sobre a expectativa legítima dos titulares. O tratamento de informações extraídas de redes sociais, fóruns ou mesmo repositórios públicos não deve ser confundido com consentimento tácito: o simples fato de uma informação estar acessível não a converte em "manifestamente pública" no sentido jurídico estrito[3], sobretudo quando envolve dados de saúde, sexualidade ou convicções políticas. Nessas hipóteses, a exigência de consentimento explícito, ou de outra base legal protetiva, torna-se incontornável, sob pena de violação a direitos fundamentais (La Diega; Harbinja; Nolan, 2024, p. 02-03).

Os dados demonstram que a dimensão dessas práticas é emblemática: em 2020, o *dataset* da *Common Crawl* utilizado para modelos como o GPT-3 atingiu a marca de 45 *terabytes* de texto cru, o que corresponderia a cerca de 90 milhões de livros diferentes (Brown *et al.*, 2020, p. 08). Apenas cinco anos depois, observa-se um crescimento exponencial na escala de coleta: em agosto de 2025, a *Common Crawl* disponibilizou um único arquivo mensal contendo 424 *terabytes* de conteúdo não compactado, extraído de 2,44 bilhões de páginas *web*. Essa coleta abrangeu 47,5 milhões de *hosts* e 38,5 milhões de domínios registrados, incluindo 675 milhões de URLs inéditos, não presentes em rastreamentos anteriores (Vaughan, 2025).

A discrepância entre os volumes revela a magnitude da operação e a impossibilidade material de rastrear, individualmente, a origem e a base legal de cada fragmento coletado. A fusão

[3] "Art. 7º § 4º É dispensada a exigência do consentimento previsto no *caput* deste artigo para os dados tornados manifestamente públicos pelo titular, resguardados os direitos do titular e os princípios previstos nesta Lei." BRASIL. Lei nº 13.709, de 14 de agosto de 2018

entre raspagem massiva e *datasets* públicos, nesse contexto, não apenas multiplica os riscos de inclusão de dados pessoais sem base legítima, mas também torna ilusória a própria operacionalização de direitos como eliminação e correção, já que as informações são incorporadas de forma difusa e paramétrica ao modelo (Kapilan; Kantor; Kallenbach, 2024).

No caso brasileiro, a utilização de bases públicas, como aquelas mantidas pelo Portal de Dados Abertos (Brasil) também enseja debates relevantes. Embora tais repositórios sejam criados para fomentar transparência e inovação exigidos pela Lei de Acesso a Informação, sua adequação à LGPD exige que os dados disponibilizados passem por etapas de anonimização, a depender de sua natureza. Pesquisas recentes indicam que até mesmo órgãos estatais enfrentam dificuldades para assegurar conformidade plena, razão pela qual estruturas de governança vêm sendo propostas e testadas, visando reduzir riscos de exposição indevida e assegurar que a abertura de dados cumpra sua função pública sem colidir com direitos fundamentais (Marques; Moresi, 2024, p. 19-22).

A análise conjunta do *web scraping* e do reaproveitamento de *datasets* públicos evidencia que a formação das bases informacionais dos modelos de IA generativa está assentada em práticas dificilmente compatíveis com as exigências de finalidade, necessidade, adequação e transparência previstas na LGPD. A própria escala dessas operações torna inviável identificar a procedência específica de cada dado incorporado ao treinamento, o que impede a verificação individualizada de sua base legal ou do atendimento às expectativas legítimas do titular.

No entanto, a captação de dados pelos sistemas de IA generativa não se esgota nas etapas de coleta prévia nem no uso de grandes bases externas. Esses modelos continuam a absorver informações durante o próprio uso, por meio dos *inputs* fornecidos pelos usuários, que funcionam como vetores adicionais de coleta e retroalimentação contínua de dados pessoais.

2.3.3 *Inputs* e *outputs* como instrumentos de captação e retroalimentação de dados pessoais

Nos sistemas de inteligência artificial generativa, o funcionamento técnico é estruturado a partir de dois elementos fundamentais: *inputs* e *outputs*.

Os *inputs* consistem nas informações inseridas pelo usuário que alimentam o modelo e servem como ponto de partida para a geração de respostas. Estas podem assumir forma textual, visual, sonora ou multimodal, abrangendo desde instruções simples até dados pessoais, como nomes, endereços, diagnósticos, imagens e documentos anexados. No entanto, a função dos *inputs* não se limita ao fornecimento imediato de dados para a geração de *outputs*: eles também integram o processo contínuo de aprendizado do sistema, funcionando como uma forma de treinamento dinâmico. A cada nova interação, o modelo ajusta seus parâmetros internos, refina padrões linguísticos e comportamentais e aprimora suas respostas futuras com base nos exemplos fornecidos pelos usuários. Assim, os *inputs* configuram não apenas uma via de comunicação, mas uma etapa adicional de aprendizagem algorítmica, na qual o próprio uso cotidiano atua como fonte de treino e calibração do modelo (Duque-Pereira; Moura, 2023, p. 11).

Por sua vez, os *outputs* são as respostas produzidas pela IA a partir desse material, resultantes de inferências probabilísticas realizadas sobre parâmetros previamente treinados (Li *et al*., 2025, p. 3-4). Assim, enquanto os *inputs* expressam o ato de fornecimento de dados, os *outputs* representam a manifestação gerada pelo sistema, podendo ambos - entrada e saída - envolver tratamento de dados pessoais, seja pela inserção direta de informações identificáveis, seja pela reprodução, inferência ou reconfiguração de dados já existentes.

Nesse contexto, os *inputs* e *outputs* constituem não apenas instrumentos funcionais de operação, mas também vetores autônomos de coleta e retroalimentação de dados. A cada comando (*prompt*) emitido e a cada resposta gerada, estabelece-se um fluxo bidirecional de informações que reconfigura as fronteiras tradicionais da proteção de dados e desafia os limites da autodeterminação informativa (Li *et al*., 2025, p. 3-4).

Como os *inputs* correspondem a todas as informações introduzidas pelo usuário durante a inferência, englobando desde consultas triviais até conteúdos altamente sensíveis, estes acabam tomando natureza híbrida: simultaneamente comunicacional e informacional (Duque-Pereira e de Moura, 2023, p. 11, 18-20). Assim, o ato de inseri-los no sistema de IA generativa configura tratamento de dados nos termos do Art. 5º, X da LGPD, pois envolve coleta, armazenamento e, frequentemente, uso posterior para fins de melhoria de desempenho do modelo (Brasil, 2018).

É nesse ínterim que o precedente internacional reforça a gravidade desse aspecto: em 2023, a autoridade italiana de proteção de dados determinou a suspensão temporária do ChatGPT por

falhas na proteção de dados fornecidos via *prompt*, reconhecendo que o risco não decorre apenas da base de treinamento, mas também da manipulação dos dados inseridos durante o uso (McCallum, 2023). Casos concretos, como o vazamento de códigos confidenciais submetidos por funcionários da Samsung à assistentes de IA, exemplificam como o fluxo interativo de *inputs* pode converter-se em vetor de exposição indevida, comprometendo segredos comerciais e informações pessoais sob sigilo (Li *et al*., 2025, p. 4).

Desse modo, a vulnerabilidade dos *inputs* se agrava pela prática de retenção e catalogação de *prompts*, procedimento adotado por grande parte das plataformas de IA generativa para fins de auditoria, segurança e aperfeiçoamento algorítmico. Tal prática, embora apresentada como medida de governança, implica reprocessamento dos dados fornecidos e amplia o risco de reutilização indevida, inclusive em futuros ciclos de treinamento (Ioscote, 2024, p. 7).

O princípio da necessidade, previsto no Art. 6º, III da LGPD, é tensionado quando o armazenamento de interações não é limitado à finalidade estritamente operacional, mas se estende à formação de novos *datasets* internos, constituindo verdadeiro banco de dados derivado de *inputs* humanos. A engenharia de *prompts*, concebida como técnica para moldar o comportamento dos modelos, também opera como mecanismo de retroalimentação, uma vez que os comandos enviados pelos usuários servem de parâmetro para ajustes e etapas de aperfeiçoamento interno subsequentes, revelando uma forma de coleta ativa e contínua de dados sob o disfarce da customização funcional (Ioscote, 2024, p. 9; Duque-Pereira; Moura, 2023, p. 18-20).

De modo correlato, os *outputs* produzidos pelos sistemas generativos configuram, por si só, manifestações de tratamento de dados pessoais. A geração textual, imagética, sonora ou multimodal não é resultado de simples recombinação neutra de símbolos, mas de inferências estatísticas sobre *corpora* massivos que frequentemente incorporam dados pessoais ou sensíveis, direta ou indiretamente (Wan *et al.*, 2024, p. 2).

Diante do exposto, observa-se que os resultados (*outputs*) das IAs levantam questões jurídicas complexas sobre titularidade da informação produzida, pois a linha divisória entre criação autônoma e reiteração de conteúdo humano torna-se difusa. Essa ambiguidade adquire relevância civil quando o modelo, ao responder a uma consulta, reproduz nomes, endereços ou fragmentos de textos extraídos de bases de treinamento, configurando reutilização não autorizada e possível dano à privacidade do titular. Ainda que o *output* se apresente como “conteúdo gerado”, ele pode carregar, de forma latente, traços de dados pessoais

parametrizados no modelo (Vercelli, 2025, p. 2-3).

No plano técnico, o ciclo de geração e retroalimentação de dados é potencializado pelo fenômeno conhecido como colapso do modelo, no qual modelos sucessivos passam a ser treinados sobre dados produzidos por gerações anteriores, resultando na degradação da distribuição estatística original e na perda progressiva de diversidade informacional (Shumailov *et al*., 2024, p. 2-3).

Sob a ótica jurídica, esse processo representa um círculo vicioso de reprocessamento de dados, inclusive pessoais, em sucessivas camadas de treinamento, sem transparência sobre origem, base legal ou mecanismos de eliminação. A cada interação, os *inputs* anteriores tornam-se *outputs* de novos modelos, estabelecendo um fluxo contínuo e cumulativo de dados que escapa ao controle do titular e dificulta qualquer tentativa de exclusão efetiva (Cao *et al.*, 2023, p. 6-7).

Tal dinâmica agrava o problema da irreversibilidade do uso indevido de dados, uma vez que a informação original, ao ser transmutada em parâmetros numéricos e redistribuída em modelos derivados, perde seu vínculo rastreável, tornando inviável a aplicação prática do direito à eliminação previsto no Art. 18, VI da LGPD (Brasil, 2018).

A retroalimentação algorítmica decorrente dos *outputs* também produz efeitos de segunda ordem: ao servir de insumo para ajuste fino supervisionado ou aprendizado por reforço com *feedback* humano, os resultados gerados passam a reconfigurar o comportamento do próprio modelo. Esse ciclo, descrito pela teoria da interação tecnológica como um processo bidirecional de influência cognitiva entre humano e máquina, revela que a IA generativa aprende não apenas a partir de dados prévios, mas também das respostas e correções que emergem durante o uso (Bai; Wang, 2025, p. 2).

Tal característica desloca a noção de tratamento de dados de um evento pontual para um fluxo permanente, em que cada *input* e *output* constitui, simultaneamente, dado e metadado: respectivamente material de treino e evidência de comportamento (Bai; Wang, 2025, p. 3-4). Assim, a interação cotidiana com assistentes de IA se torna, inadvertidamente, uma forma de coleta contínua e autorreferencial, onde o uso individual retroalimenta o sistema.

Esse ecossistema interativo desafia os fundamentos da responsabilidade civil, pois o tratamento de dados pessoais não se limita à coleta originária - explicitada na subseção 2.3.2 -, mas persiste em fluxos autogerativos que escapam à previsibilidade humana. A circulação

incessante entre *inputs* e *outputs* cria zonas cinzentas de imputação: o dado pessoal pode ser captado em uma interação legítima e ressurgir, transformado, em *outputs* subsequentes, sem que seja possível identificar a relação direta entre o evento inicial e o dano (Foo; Rahmani; Liu, 2025, p. 1-2).

À luz dessas dinâmicas, percebe-se que os próprios fluxos interativos das IAs generativas ampliam e complexificam o tratamento de dados pessoais, projetando-o para além da coleta inicial realizada por *web scraping* ou *datasets* públicos. Ainda assim, esse cenário não exaure as tensões jurídicas envolvidas. Resta examinar como tais práticas, quando confrontadas com a exigência de consentimento informado e de finalidade específica, revelam limites estruturais que desafiam a conformidade dos modelos com a LGPD.

2.3.4 Descompassos entre a legalidade e a prática algorítmica: limites do consentimento informado e da finalidade específica

Já foram apresentados, em momento anterior, os princípios gerais da proteção de dados pessoais, dentre eles o consentimento e a finalidade. Retoma-se agora esses pilares, não em sua formulação abstrata, mas em sua aplicação concreta, especialmente diante das problemáticas decorrentes da dependência estrutural dos LLMs de grandes volumes de dados e do uso de *datasets* públicos extraídas por meio de *web scraping*. O objetivo, portanto, é examinar os descompassos entre legalidade e prática algorítmica, demonstrando como a exigência de consentimento informado e a imposição de uma finalidade específica se fragilizam diante da arquitetura expansiva da inteligência artificial generativa.

Nessa esfera, é importante definir o chamado regime de *opt-out* das IAs generativas. Consiste em um modelo no qual os dados do usuário são presumidamente incluídos no tratamento de dados pela plataforma, cabendo ao titular manifestar-se ativamente caso deseje a exclusão ou restrição desse uso. Em contraste com o *opt-in*, que exige consentimento prévio e expresso, o *opt-out* desloca o ônus da proteção para o indivíduo, pressupondo autorização tácita até que haja oposição formal (Margoni, 2024, p. 11; Malik, 2023). Essa dinâmica é especialmente problemática na seara da inteligência artificial generativa, pois combina alta complexidade técnica com baixa transparência, dificultando que o titular compreenda ou exerça de forma efetiva sua autodeterminação informativa.

Sob esse prisma, para os LLMs, a formação de *corpora* por *web scraping* e o reaproveitamento de *datasets* públicos potencializam precisamente os déficits de consentimento informado e de finalidade específica. A coleta massiva inviabiliza informar, individualmente: quem são os titulares; quais dados foram capturados; para que finalidades serão usados; e por quanto tempo - convertendo o consentimento em mera ficção procedimental e tornando opaca a própria base legal do tratamento no ponto de origem (Tamkin *et al.*, 2021, p. 02).

Mesmo quando plataformas adotam regimes de *opt-in* para treino, mantêm tratamentos paralelos cuja extensão e temporalidade não são plenamente auditáveis pelo usuário, o que tensiona a limitação de finalidade em escala algorítmica (Sousa, 2025, p. 05-09).

Diante da realidade exposta até o presente momento, fica evidente que o tratamento de dados pessoais na era da inteligência artificial generativa tem evidenciado tensões profundas entre os limites formais da legalidade e a realidade prática das operações algorítmicas. Embora o consentimento informado seja edificado como um dos principais instrumentos de legitimação do tratamento, sua concretização enfrenta barreiras estruturais (Tepedino; Teffé, 2020, p. 04).

A complexidade técnica dos sistemas, somada à opacidade dos fluxos de dados, compromete a capacidade de o titular compreender de maneira efetiva as consequências de sua manifestação de vontade. Nesse cenário, a promessa de transparência que acompanha o instituto jurídico acaba muitas vezes reduzida a um ideal normativo, dificilmente realizável em escala algorítmica (Kondrup, 2025).

A exigência de consentimento informado, como forma de garantir a autodeterminação informativa, obriga que o titular tenha acesso a informações claras e acessíveis acerca das finalidades do tratamento (Tepedino; Teffé, 2020, p. 4). Contudo, a não disponibilização integral dessas informações inviabiliza a validade da autorização, já que, excetuados os segredos comerciais e industriais, todos os demais elementos devem ser franqueados ao titular para que este consinta de maneira consciente. A insuficiência de dados fornecidos, portanto, mina a própria essência do instituto, transformando-o em mera formalidade procedimental (Tepedino; Teffé, 2020, p. 14-15; Frazão, 2018).

Nesse ponto, a vinculação entre consentimento e finalidade específica ganha relevância. A legislação brasileira determina que o tratamento só é legítimo quando orientado por propósitos determinados, explícitos e legítimos (Brasil, 2018). Autorizações genéricas ou

amplas são inválidas, justamente por não permitirem que o titular compreenda a extensão do uso de suas informações. A finalidade, portanto, funciona como contrapeso ao alargamento abusivo da coleta, buscando assegurar que os dados circulem apenas dentro do horizonte previamente delineado (Tepedino; Teffé, 2020, p. 16-20).

Na prática, porém, observa-se a tendência de obtenção de consentimentos amplos, por meio de cláusulas de adesão que impõem ao indivíduo a aceitação de termos genéricos de uso, muitas vezes sem que este tenha tempo ou condições de compreender sua extensão. A leitura e aceite de políticas de privacidade, políticas de *cookies* e termos de uso ilustra como o consentimento se transforma em pseudo consentimento, esvaziado de substância reflexiva e reduzido a mero clique. Essa dinâmica compromete a efetividade do instituto e transfere ao usuário a responsabilidade de autoproteger seus dados em contextos nos quais ele não dispõe de meios técnicos para tanto (EDPB, 2020, p. 15-16).

Face ao exposto, o consentimento, para ser juridicamente válido, deve refletir um ato reflexivo do indivíduo, pautado em sua racionalidade e capacidade de autodeterminação (Requião, 2022, p. 28). No entanto, o modelo digital atual impõe custos sociais e cognitivos que tornam essa participação ilusória, deslocando o instituto para um campo meramente simbólico. Nessa linha, a crítica de Zuboff (2021, p. 15) sobre as assimetrias de informação reforça a constatação de que o capitalismo de vigilância opera intencionalmente pela invisibilidade, acumulando conhecimento sobre os indivíduos sem lhes devolver clareza ou poder real de escolha.

Outro ponto que compromete a eficácia do consentimento é a disparidade educacional entre os titulares. Dados recentes do Instituto Brasileiro de Geografia e Estatística (IBGE) revelam que apenas 20,5% da população brasileira possui ensino superior completo (Luz, 2025), o que significa que a maioria carece de preparo técnico para interpretar de forma crítica os termos e condições apresentados pelas plataformas digitais. Essa desigualdade de compreensão evidencia que o consentimento, ainda que formalmente informado, não atinge materialmente seu propósito de garantir escolhas conscientes. Ao invés de instrumento de proteção, converte-se em barreira excludente, beneficiando apenas parcelas mais escolarizadas da sociedade (Siqueira; Moreira, 2023, p. 17-18).

A observância do princípio da finalidade determinada é igualmente tensionada pelas práticas algorítmicas contemporâneas. Não raro, dados coletados sob uma justificativa legítima são posteriormente utilizados para fins secundários, como perfilamento comportamental ou venda

de informações a terceiros, prática frontalmente incompatível com a expectativa do titular. Assim, o instituto da finalidade específica, embora consagrado normativamente, é reiteradamente fragilizado pela lógica expansiva do tratamento algorítmico (*European Commission*, 2013, p. 15-19).

Essas incongruências tornam patente que, embora o ordenamento jurídico estabeleça balizas rigorosas para o consentimento informado e para a finalidade determinada, a materialização prática desses princípios é constantemente abalada pela opacidade e pelo alcance desmedido das práticas de coleta e utilização de dados. O resultado é um descompasso estrutural entre norma e realidade, no qual o titular permanece formalmente protegido, mas substancialmente vulnerável diante da assimetria informacional que caracteriza os modelos de negócio baseados em inteligência artificial generativa (Sousa, 2025, p. 05-09).

Diante dessas fragilidades, torna-se necessário avançar na análise das formas pelas quais as promessas normativas são operacionalizadas pelas plataformas digitais. A leitura crítica das políticas de privacidade revela como o discurso das plataformas se contradiz com a realidade concreta do tratamento de dados, contexto no qual se insere a próxima subseção.

2.3.5 A retórica da proteção de dados *versus* a prática na IA generativa

A fim de evidenciar a distância entre a retórica protetiva e a prática efetiva no ecossistema de IA generativa, passa-se à análise das políticas de privacidade de algumas das plataformas mais utilizadas mundialmente.

Com base no levantamento do Statista (2025), que identificou os sistemas de IA com maior número de usuários ativos mensais (MAU) - cujos dados completos constam no Anexo 1 -, foram escolhidas plataformas que, além da relevância empírica, oferecem material documental apto à análise normativa.

Nessas balizas, foram incluídos no recorte analítico o ChatGPT, a Nova AI, a DeepSeek, o Gemini e o Copilot. Para garantir precisão terminológica e segurança interpretativa no cotejo com a LGPD, foram selecionadas apenas plataformas cujo documento oficial apresenta versão completa em português.

O ChatGPT foi selecionado por ocupar posição de liderança em usuários ativos no

levantamento do Statista (2025) e por constituir o paradigma contemporâneo das LLMs (Tamkin *et al.*, 2021, p. 03), permitindo testar a coerência entre a transparência prometida, as bases legais indicadas e a efetividade do exercício de direitos no tratamento de dados em larga escala.

A Nova AI foi incluída por se situar em terceiro lugar no *ranking* do Anexo 1, enquanto a DeepSeek, posicionada logo na sequência, oferece diversidade regulatória relevante por operar sob jurisdições distintas, o que favorece o exame de transferências internacionais e uso de dados publicamente acessíveis. Já o Gemini, foi considerado por ser o décimo colocado no levantamento.

Por fim, o Copilot foi incorporado ao recorte não pelo levantamento do Anexo 1, mas por relevância dinâmica: seu crescimento acelerado no uso móvel, superando o percentual de crescimento do ChatGPT (Silva, 2025), o torna especialmente útil para a análise documental explorada na presente subseção.

As demais aplicações listadas pelo Statista (DouBao, Remini, Talkie AI, Character AI, ChatOn e Genius) não foram incorporadas à análise por não disponibilizarem, até o momento, documentação oficial integralmente acessível em português. A ausência desse material comprometeria a comparabilidade metodológica e enfraqueceria a aderência conceitual à LGPD, razão pela qual sua exclusão não representa limitação metodológica, mas, ao contrário, reforça a precisão e a consistência do recorte adotado.

O objetivo não é a exaustão do mercado, mas a construção de um recorte suficientemente robusto para evidenciar, com rigor, o descompasso entre o discurso de proteção e as práticas reais verificáveis.

Para conferir sistematicidade ao exame desenvolvido, a Tabela 1 foi construída a partir da análise documental das plataformas selecionadas, constituindo síntese analítica original elaborada com base na interpretação crítica dos respectivos documentos oficiais. Trata-se, portanto, de instrumento produzido especificamente para este trabalho, e não de reprodução ou reorganização de dados provenientes de fontes externas, tendo sido concebido com a finalidade de operacionalizar os critérios discutidos ao longo do capítulo.

Na tabela, as respostas ‘sim’ estão destacadas em verde, os resultados ‘parciais’ aparecem em amarelo e as ocorrências de ‘não’ estão marcadas em rosa.

CRITÉRIO	I. ChatGPT	II. Nova	III. DeepSeek	IV. Gemini	V. Copilot
(1) Identificação clara do controlador?	Sim	Sim	Sim	Sim	Sim
(2) Identificação clara da base legal do tratamento?	Parcial	Parcial	Não	Parcial	Parcial
(3) Indicação da finalidade específica para cada tipo de dado coletado?	Parcial	Sim	Sim	Sim	Parcial
(4) Consentimento como base legal principal ou secundária?	Secundária	Secundária	Secundária	Secundária	Secundária
(5) Mecanismos de *opt-out* para uso em treinamento de modelos?	Sim	Não	Não	Parcial	Parcial
(6) Menciona uso de dados publicamente acessíveis (*web scraping*)?	Sim	Não	Sim	Sim	Sim.
(7) Indica base legal para uso de *datasets* públicos?	Não	Não	Não	Parcial	Parcial
(8) Detalha práticas de perfilamento ou inferência automatizada?	Não	Parcial	Parcial	Parcial	Sim
(9) Indica medidas técnicas e administrativas de segurança?	Sim	Sim	Sim	Parcial	Sim
(10) Lista os direitos do titular conforme a LGPD?	Parcial	Parcial	Parcial	Parcial	Parcial
(11) Indica transferência internacional de dados e países envolvidos?	Parcial	Parcial	Sim	Parcial	Parcial
(12) Nomeia encarregado pelo tratamento e canal de contato?	Não	Não	Não	Não	Não

Tabela 1: Políticas de privacidade das principais IAs generativas - Retórica *versus* Prática

Os itens (1 a 12) que integram a Tabela 1 foram definidos com base nos eixos teóricos e empíricos desenvolvidos ao longo do presente capítulo. Cada critério traduz em indicador verificável um dos elementos centrais da dinâmica predatória analisada no trabalho, permitindo confrontar a retórica protetiva das plataformas com a efetividade material das salvaguardas exigidas pelo regime jurídico da proteção de dados pessoais.

Assim, a inclusão da identificação clara do controlador e do encarregado (itens 1 e 12) decorre da necessidade de aferir o cumprimento dos requisitos mínimos de governança e *accountability*, essenciais para delimitar responsabilidades em ecossistemas caracterizados pela autonomia algorítmica crescente (subseção 2.3.1). A indicação decorre das exigências dos Arts. 9º, inciso III e IV e 41 § 1º da LGPD.

A verificação da base legal utilizada (item 2), da finalidade específica do tratamento (item 3) e do papel atribuído ao consentimento (item 4) relaciona-se diretamente aos limites estruturais do consentimento informado e às críticas à finalidade aberta ou derivada, fenômenos recorrentes na exploração massiva e silenciosa de dados pessoais, como demonstrado em 2.2.1 e 2.3.4. O item (2) atende ao Art 7ºe seus incisos e o item (3) atende ao Art. 9º, inciso I, ambos da LGPD.

A verificação da existência de mecanismos de *opt-out* para uso de dados no treinamento de modelos (item 5) foi incorporada especificamente em razão do debate desenvolvido na seção 2.3.4, que evidencia como a dependência estrutural dos LLMs por grandes volumes de dados, associada à opacidade sobre sua origem e destinação, transforma o *opt-out* em um indicador privilegiado para mensurar a efetividade, ou apenas a aparência, do exercício de direitos. Nesse contexto, a análise do *opt-out* permite avaliar se as plataformas oferecem salvaguardas materiais ou apenas reproduzem um modelo de pseudo consentimento incompatível com a proteção de dados em larga escala.

Os itens relativos ao uso de dados publicamente acessíveis via *web scraping* (item 6) e à indicação da base legal para utilização de *datasets* públicos (item 7) decorrem da discussão apresentada em 2.3.2, que demonstra que a acessibilidade pública não se confunde com licitude e que a compatibilidade contextual constitui elemento essencial para aferir a legalidade do tratamento.

O exame das práticas de perfilamento e inferência automatizada (item 8) foi justificado pela

análise da publicidade comportamental, da monetização da atenção e da opacidade algorítmica (subseções 2.2.4 e 2.3.3), fenômenos que revelam como a extração de dados e a produção de inferências constituem o núcleo econômico da IA generativa e agravam a erosão da autonomia informativa.

A inclusão das medidas técnicas e administrativas de segurança (item 9) responde ao caráter estruturalmente sensível do tratamento realizado por modelos generativos, cujas fases de coleta, inferência, retroalimentação e reuso ampliam significativamente os riscos e tornam indispensável a verificação da substância das salvaguardas, e não apenas sua enunciação formal. O item (9) respeita a exigência do Art. 48, § 1º, inciso III da LGPD.

A menção aos direitos do titular (item 10) exigida pelo Art. 9º, inciso VII da LGPD relaciona-se com o exposto no 2.1 e também com o debate sobre a erosão da autodeterminação informativa, tema que será desenvolvido na seção 2.4, evidenciando a necessidade de verificar se tais direitos são efetivamente operacionalizados e não apenas enunciados de forma abstrata.

Por fim, a análise das transferências internacionais de dados e dos países envolvidos (item 11) decorre do reconhecimento de que os fluxos globais de dados são estruturantes nos ecossistemas de IA generativa, operados majoritariamente por *big techs* (como discutido na subseção 2.2.3), cujas infraestruturas transnacionais tensionam de modo permanente os princípios da necessidade, proporcionalidade e finalidade (seção 2.1). A transparência quanto às jurisdições que regulam cada etapa do tratamento torna-se, assim, indispensável para a adequada compreensão da distribuição de responsabilidades - que serão trabalhadas no 3.1.3 e 3.1.4 - e dos riscos inerentes ao deslocamento internacional da informação pessoal.

Dessa forma, os doze critérios refletem os principais pontos de tensão identificados ao longo do trabalho entre a promessa de proteção e a prática concreta das plataformas, convertendo os fundamentos conceituais discutidos nos capítulos anteriores em indicadores empíricos que permitem mensurar o grau real de transparência e conformidade das políticas de privacidade de IA generativa.

Em suma, a Tabela 1 transforma os princípios em testes verificáveis de transparência material, permitindo aferir o descompasso entre retórica e prática.

Feita essa amarração metodológica, passa-se à análise específica de cada uma das políticas de privacidade de algumas das mais populares IAs generativas.

2.3.5.1. Política de privacidade do ChatGPT

No que diz respeito à política de privacidade do ChatGPT, observa-se um desenho declaratório alinhado às boas práticas formais mas que não supre, em termos de transparência material, o nível exigido para compatibilizar a ampla coleta de dados pessoais.

Em uma primeira análise, a política identifica o controlador como "OpenAI OpCo, LLC (e afiliadas)" o que atende integralmente ao item (1) da Tabela 1 (OpenAI, 2025).

No tocante à base legal, o documento não realiza o mapeamento finalidade; categoria de dado; base jurídica, limitando-se a indicar usos amplos dos dados (prestar, melhorar, segurança, *compliance*), o que deixa (2) e (3) apenas parcialmente satisfeitos à luz das exigências de especificidade e compatibilidade contextual da LGPD (OpenAI, 2025).

O consentimento por sua vez, é mencionado como fundamento "quando utilizamos o consentimento", sem ser a base legal principal e com possibilidade de revogação, sinalizando uso secundário e não regra matriz, o que se coaduna com uma arquitetura de múltiplas bases e cumpre parcialmente o item (4) (OpenAI, 2025).

Em relação ao *opt-out* para treinamento, há previsão expressa para que o titular opte por não permitir o uso do seu conteúdo para treinar modelos, via instruções do Centro de Ajuda, contudo, o texto circunscreve a faculdade ao Conteúdo do Usuário (*prompts*/arquivos) e não explicita o alcance sobre metadados e telemetria, de modo que o item (5) é atendido quanto à existência do mecanismo, mas com escopo material a ser melhor delimitado (OpenAI, 2025).

A política afirma coletar informações publicamente disponíveis na Internet para desenvolver os modelos, o que responde afirmativamente ao item (6), mas não explicita a base legal específica para esse reuso de *datasets* públicos sob a ótica da LGPD, ponto central para o item (7), que permanece descoberto no texto fornecido (OpenAI, 2025).

No que tange a perfilamento e inferências automatizadas, inexiste explicação sobre a lógica decisória ou os critérios utilizados: há apenas uma advertência de que as respostas podem não refletir "precisão factual", o que não supre o dever informativo exigido pelo item (8) (OpenAI, 2025).

As medidas de segurança são descritas em termos de "medidas técnicas, administrativas e

organizacionais comercialmente razoáveis", sem granularidade técnica, o que cumpre o item (9) quanto à existência de salvaguardas (OpenAI, 2025).

Quanto aos direitos do titular, a política elenca acesso, exclusão, atualização, portabilidade, restrição, revogação do consentimento e oposição, além de canal para exercício, o que atende substancialmente ao item (10), ainda que não traga uma lista exaustiva nos exatos termos da LGPD e sem quadro operacional por categoria de dado. Assim, o item (10) é cumprido apenas parcialmente (OpenAI, 2025).

No tema de transferências internacionais, a OpenAI declara tratamento e armazenamento nos Estados Unidos e em "vários territórios", com uso de mecanismos válidos de transferência, mas não lista todos os países de destino, de modo que o item (11) é cumprido parcialmente (OpenAI, 2025).

Por fim, a política não identifica nominalmente um Encarregado/DPO, restringindo-se à oferta de canais de contato genéricos (privacy@openai.com), o que não atende ao item (12) (OpenAI, 2025).

Em síntese, a política apresenta conformidade declaratória com diversos requisitos formais, mas permanece aquém do nível de transparência material esperado. Embora ofereça mecanismos e referências que, em tese, dialogam com as exigências da LGPD, a ausência de granularidade quanto às bases legais, ao escopo da coleta e ao uso de dados revela um modelo incompleto e ainda marcado por opacidades estruturais. Assim, a política cumpre parcialmente seu papel de orientar o titular, mas não elimina assimetrias informacionais relevantes.

2.3.5.2. Política de privacidade da Nova

A política da Nova (ScaleUp) apresenta configuração declaratória que combina elementos de conformidade procedimental com lacunas relevantes quando observada à luz da LGPD e dos vetores específicos de risco identificados para sistemas de IA generativa. A invocação expressa do ordenamento turco como lei aplicável reforça a matriz jurídica adotada pela empresa, com consequências relevantes à compatibilidade documental com o regime brasileiro (Nova, 2023).

A documentação identifica a controladora como “ScaleUp” e fornece endereço e sede na Turquia, o que satisfaz, em termos formais, o critério de identificação do controlador do item (1) (Nova, 2023).

No plano da especificidade das finalidades, a política apresenta listagens detalhadas de categorias de dados e fins administrativos e comerciais (comunicação, vendas, suporte, segurança, publicidade e desenvolvimento de produto) e associa, em diversos trechos, categorias de dados a operações concretas, atendendo ao item (3) (Nova, 2023).

A política invoca bases legais diversas sob a égide da legislação turca: contrato, obrigação legal, interesse legítimo e consentimento para atividades de *marketing*; sem, contudo, mapear essas bases por finalidade e categoria de dados, conforme se exige para avaliação de compatibilidade contextual no ordenamento brasileiro. Assim, o item (2) é parcialmente atendido e o item (4) demonstra que o consentimento é base legal secundária (Nova, 2023).

No que tange ao uso dos dados no contexto de IA generativa, o texto afirma utilização de Interface de Programação de Aplicações de terceiros e descreve práticas de integração com Kits de Desenvolvimento de *Software* e provedores de nuvem, bem como parcerias com provedores de análise e publicidade. Todavia, não há previsão expressa de mecanismos de *opt-out* específicos para o uso de conteúdo de usuário no treinamento de modelos, nem é explicitada a prática de *web scraping* ou a utilização de *datasets* públicos com a respectiva base jurídica (Nova, 2023). A ausência dessas menções impede a aferição sobre a extensão do escopo de treino e sobre a possibilidade de exclusão de dados pessoais do ciclo de aprendizagem, o que faz com que a Nova não cumpra os critérios dos itens (5), (6) e (7).

Quanto ao perfilamento e às inferências automatizadas, a política descreve o uso de ferramentas de análise e publicidade dirigida, assim como a geração de segmentações para fins de mensuração e direcionamento de campanhas (Nova, 2023). Contudo, não explicita a lógica decisória, os critérios algorítmicos ou a possibilidade de contestação automática, ficando o item (8) apenas parcialmente atendido. Essa opacidade informativa é especialmente problemática no contexto da monetização da atenção e da venda de segmentos de audiência, elementos centrais identificados na subseção 2.2.4.

Em contrapartida, a Nova dedica espaço extenso às medidas técnicas e administrativas de segurança, atendendo ao item (9). Tal descrição indica diligência procedimental e um nível de detalhamento superior ao observado em muitas políticas similares. A retenção de dados por

período definido (até seis anos salvo exclusão) e a possibilidade de exclusão de conta com remoção dos históricos apontam para práticas de governança de retenção, mas requerem validação operacional sobre prazos por categoria de dado (Nova, 2023).

No tocante aos direitos do titular, a política menciona direitos de acesso, retificação, exclusão, portabilidade e o direito de solicitar a não divulgação de dados para fins de venda/marketing, bem como procedimentos de exercício de direitos voltados a usuários europeus e californianos (Nova, 2023). Todavia, a operacionalização desses direitos em termos de prazo e rotinas aplicáveis ao contexto brasileiro não é apresentada de forma explícita, pelo que o item (10) merece classificação parcial.

No que se refere a transferências internacionais, a Nova admite uso de provedores e infraestruturas em nuvem, citando provedores como Google e Cloudflare, e indica possibilidade de transferências para o exterior (Nova, 2023), porém não lista países destinatários nem descreve os mecanismos jurídicos adotados para garantir padrões equivalentes de proteção, atendendo apenas parcialmente ao item (11). A lacuna é relevante em razão da operação transnacional típica das *big techs* e dos provedores de nuvem, cujo papel na articulação de fluxos de dados impõe exigências adicionais de transparência sobre jurisdições e salvaguardas.

A política não nomeia um Encarregado/DPO por identificação nominal, limitando-se a canais de contato genéricos (websupport@novaapp.ai) (Nova, 2023), o que compromete a observância do item (12) e reduz a previsibilidade dos canais de governança e responsabilização.

Em síntese, a política da Nova combina práticas robustas de segurança operacional e um rol detalhado de finalidades administrativas com omissões cruciais a respeito do reuso de dados para treinamento de modelos, da base legal aplicada a *datasets* públicos, da transparência sobre critérios de inferência automatizada, da identificação nominal do encarregado e da indicação pormenorizada de transferências internacionais.

Esses pontos configuram, no recorte deste trabalho, os aspectos nodais do descompasso entre retórica de conformidade e transparência material exigida para a proteção efetiva da autodeterminação informativa em ambientes de IA generativa.

2.3.5.3. Política de privacidade da DeepSeek

A política de privacidade da DeepSeek apresenta um desenho declaratório mais robusto do que o observado em algumas plataformas emergentes, sobretudo pela combinação de identificação clara do controlador, suplementos regionais detalhados (especialmente para União Europeia e Reino Unido) e explicitação ampliada das categorias de dados tratadas.

O documento identifica a controladora como 'Hangzhou DeepSeek Artificial Intelligence Co., Ltd.', com sede na China, e fornece canal de contato direto (privacy@deepseek.com), o que satisfaz os critérios de transparência institucional previsto no item (1) (DeepSeek, 2025).

Todavia, o documento não realiza mapeamento finalístico sob a ótica da LGPD, tampouco estabelece correspondência entre categoria de dado, finalidade e base legal conforme exige a compatibilidade contextual brasileira, o que reduz o atendimento do item (2) (DeepSeek, 2025).

A política da DeepSeek atende integralmente ao item (3), pois descreve de maneira clara e abrangente as finalidades associadas a cada categoria de dado coletado, especificando usos como prestação do serviço, melhoria dos modelos, segurança, comunicação e cumprimento de obrigações legais (DeepSeek, 2025). A formulação é suficientemente detalhada para caracterizar indicação de finalidade específica.

Quanto ao item (4), embora o consentimento apareça como hipótese de tratamento, ele não exerce papel central na arquitetura jurídica da plataforma, funcionando apenas como base legal complementar em cenários específicos. O tratamento principal repousa em execução contratual, interesses legítimos e obrigações legais. Assim, o consentimento é secundário (DeepSeek, 2025).

O ponto de maior relevância crítica é a admissão explícita do uso de dados pessoais: "*Podemos obter Dados Pessoais disponíveis publicamente por meio de fontes online para treinar nossos modelos e fornecer Serviços*" atendendo ao item (6). No entanto, sem indicação de fundamento jurídico compatível com a LGPD, sem análise de compatibilidade contextual e sem previsão de *opt-out* específico, os itens (5) e (7) não foram atendidos (DeepSeek, 2025).

A política afirma não realizar criação de perfil para fins decisórios significativos, mas não descreve a lógica de inferência ou o funcionamento do mecanismo gerativo, o que deixa o item (8) apenas parcialmente atendido (DeepSeek, 2025).

Por fim, a DeepSeek descreve um conjunto abrangente de medidas de segurança e informa de modo claro que os dados são processados e armazenados na China, o que permite considerar atendidos os itens (9) e (11) em sua dimensão formal (DeepSeek, 2025).

Não existe identificação nominal de encarregado no modelo da LGPD, o que impede que o item (12) seja contemplado (DeepSeek, 2025)

Em síntese, a política da DeepSeek combina identificação clara do controlador, suplementações regionais detalhadas e descrição abrangente das categorias de dados, projetando maior solidez formal que outras plataformas emergentes. Persistem, porém, lacunas relevantes: não há mapeamento finalístico compatível com a LGPD, nem fundamento jurídico para o uso de dados publicamente acessíveis no treinamento, tampouco previsão de *opt-out*.

A explicação sobre inferências automatizadas é limitada e não alcança a lógica operacional do modelo. Embora a política apresente medidas de segurança e explicite o processamento na China, a ausência de detalhamento sobre mecanismos de transferência mantém restrições à transparência material e evidencia o descompasso entre declaração e prática no ecossistema de IA generativa.

2.3.5.4. Política de privacidade do Gemini

No caso dos aplicativos do Gemini, observa-se uma política extensa, articulada em torno da infraestrutura de privacidade do Google e complementada por avisos específicos para operações de IA generativa (Google, 2025).

Em termos de transparência institucional, o documento identifica as entidades responsáveis pelo tratamento - Google Ireland Limited (EEE/Suíça) e Google LLC (demais países) - e utiliza a central de privacidade da Conta Google como principal canal de gestão, o que atende à identificação do controlador e ao item (1) (Google, 2025).

Quanto às bases legais, o aviso lista fundamentos aplicáveis sob a legislação da União Europeia e Reino Unido (execução contratual, consentimento quando cabível, interesse legítimo e obrigações legais), mas não realiza o mapeamento sistemático entre categoria de dado, finalidade e base jurídica exigido pela LGPD, resultando em atendimento apenas parcial

do item (2) (Google, 2025).

O mesmo ocorre com a indicação de finalidades específicas do item (3): o documento descreve múltiplos propósitos - fornecimento do serviço, personalização, melhoria, segurança, prevenção de abusos e pesquisa. O consentimento aparece como fundamento secundário e residual dentro da arquitetura ampla de bases, o que cumpre parcialmente o item (4) (Google, 2025).

Em relação ao uso de dados para treinamento de modelos, o Gemini admite que conversas, interações e conteúdos compartilhados podem ser utilizados para "melhorar a IA do Google", e oferece mecanismos de controle por meio das configurações de "Atividade nos *apps* do Gemini", exclusão automática e gerenciamento de retenção. Esses mecanismos não constituem um *opt-out* integral e específico para treino, mas representam alternativa parcial de contenção de fluxo, razão pela qual o item (5) é atendido parcialmente (Google, 2025).

O documento também reconhece o uso de dados publicamente acessíveis e de informações provenientes de outros serviços Google para treinar modelos, o que satisfaz o item (6). Todavia, não indica base jurídica específica para reutilização de *datasets* públicos sob a ótica da LGPD, deixando o item (7) apenas parcialmente contemplado (Google, 2025).

Sobre práticas de perfilamento e inferência automatizada, a política descreve personalização, revisão humana e utilização de dados de aplicativos conectados, mas não revela a lógica algorítmica nem o funcionamento interno dos mecanismos gerativos. Assim, o item (8) é parcialmente atendido, pois há descrição funcional, mas não há explicação sobre critérios utilizados para produzir inferências (Google, 2025).

No tocante à segurança da informação, o aviso remete às salvaguardas técnicas e administrativas gerais do Google (controles de acesso, infraestrutura distribuída, prevenção de abusos e ferramentas de detecção), mas sem indicar um quadro próprio e específico dos aplicativos do Gemini. Essa remissão garante cobertura básica, porém não fornece detalhamento técnico suficiente para auditoria independente, de modo que o item (9) recebe atendimento parcial (Google, 2025).

A política oferece ferramentas de exercício de direitos - exportação, exclusão, gerenciar histórico, configurar retenção - integradas à Conta Google, mas sem alinhamento expresso aos prazos, procedimentos e garantias específicas da LGPD, razão pela qual o item (10) também é apenas parcialmente atendido (Google, 2025).

Por fim, as transferências internacionais de dados são amplamente reconhecidas: o processamento pode ocorrer em múltiplas jurisdições e estruturas, com dados integrados a outros serviços do Google. Apesar disso, o aviso não identifica países de destino nem detalha, para fins de LGPD, quais mecanismos jurídicos legitimam esses fluxos, o que limita o atendimento do item (11) a um grau meramente formal (Google, 2025).

Não há, contudo, nomeação expressa de encarregado, razão pela qual o item (12) permanece não atendido (Google, 2025).

Em síntese, o aviso dos aplicativos do Gemini apresenta elevado nível de detalhamento descritivo e se beneficia da infraestrutura consolidada do Google, mas permanece marcado por lacunas relevantes sob o prisma da LGPD: ausência de mapeamento finalístico, *opt-out* incompleto para treinamento, fundamentação jurídica apenas parcial para uso de dados públicos, transparência limitada sobre lógica algorítmica e insuficiência na explicitação dos mecanismos de transferência internacional. Esses elementos reforçam o descompasso entre a retórica protetiva e a prática material que atravessa o ecossistema de IA generativa.

2.3.5.5. Política de privacidade do Copilot

No que se refere à política e à documentação técnica do Microsoft 365 Copilot, observa-se um enquadramento voltado à operação em contexto empresarial, integrado ao ecossistema Microsoft 365 e ancorado em compromissos contratuais e de conformidade vigentes (Microsoft, 2025).

Em uma primeira análise, a documentação identifica de forma clara as entidades responsáveis pelo serviço e a sua articulação com o Microsoft Graph e com os provedores de infraestrutura, permitindo identificar adequadamente o controlador operacional do tratamento, o que satisfaz o requisito de identificação institucional previsto no item (1) (Microsoft, 2025).

Quanto à base legal do tratamento, o texto enfatiza a conformidade com regimes aplicáveis em contextos corporativos europeus, e apoia grande parte do tratamento em fundamentos contratuais, em interesses legítimos e em obrigações de segurança e residência de dados acordadas contratualmente com clientes empresariais. Esse enquadramento evidencia uma base jurídica formal, mas a ausência de mapeamento público exaustivo que correlacione

finalidade, categoria de dado e fundamento jurídico para cada operação torna o atendimento ao item (2) apenas parcial (Microsoft, 2025).

Em similar linha, a indicação de finalidades - fornecimento do serviço, melhoria, segurança, personalização e pesquisa - está presente na documentação, porém sem uma correlação sistemática por categoria de dado que permita aferir compatibilidade contextual no âmbito da LGPD, razão pela qual o item (3) também é parcialmente atendido (Microsoft, 2025).

O papel do consentimento figura de forma secundária no desenho jurídico do Copilot, especialmente porque o produto se destina majoritariamente ao uso corporativo, em que predominam bases contratuais e controles de administrador. O consentimento aparece fortemente ligado a funcionalidades ou cenários específicos e não como fundamento principal, o que justifica a classificação parcial do item (4) (Microsoft, 2025).

No que tange a mecanismos de *opt-out* para uso em treinamento de modelos, a Microsoft afirma que *prompts*, respostas e dados acessados via Microsoft Graph não são usados para treinar os LLMs de base do Copilot e disponibiliza ferramentas administrativas que permitem ao cliente gerenciar o uso e a retenção de interações. Essas garantias e controles administrativos representam uma mitigação relevante, mas não configuram um *opt-out* individual universal aplicável a todas as formas de reuso técnico em todos os níveis do ciclo de tratamento, de modo que o item (5) é atendido de forma parcial (Microsoft, 2025).

A documentação admite o uso de conteúdo público e de consultas à *web* para melhorar respostas em determinados fluxos, bem como a integração de contexto proveniente de serviços externos, o que responde afirmativamente ao item (6) (Microsoft, 2025).

Contudo, a justificativa jurídica específica para o reuso de *datasets* públicos sob a ótica de regimes distintos do europeu, notadamente a LGPD, não é objeto de mapeamento detalhado no material público do produto, de modo que o item (7) restou apenas parcialmente contemplado (Microsoft, 2025).

No plano do perfilamento e das inferências automatizadas, o Copilot descreve mecanismos de personalização, geração de resumos e segmentação contextual, além de mecanismos de revisão e filtros aplicados para mitigar danos e enviesamentos. Assim, existe descrição de uso e salvaguardas, o que atende ao item (8) (Microsoft, 2025).

As medidas técnicas e administrativas de segurança estão amplamente descritas no material: isolamento lógico por locatário, controles de acesso baseados em função, criptografia em

trânsito e em repouso, integrações com Microsoft Purview e certificações de conformidade relevantes, o que satisfaz o item (9) na dimensão descritiva e documental (Microsoft, 2025).

Quanto aos direitos dos titulares, a Microsoft oferece mecanismos para exportação e gestão de histórico, bem como controles administrativos que possibilitam a operacionalização desses direitos no contexto organizacional, mas a documentação orientada ao produto não detalha pormenorizadamente prazos, fluxos e procedimentos específicos adaptados à LGPD para titulares individuais fora do âmbito corporativo, o que determina atendimento parcial do item (10) (Microsoft, 2025).

Por fim, a política explicita compromissos de residência de dados e descreve a possibilidade de processamento em múltiplas jurisdições, bem como ofertas específicas de residência e fronteira de dados para a União Europeia, circunstância que demonstra reconhecimento e gestão de fluxos transnacionais. Não obstante, o material público não traz pormenorizado país a país nem detalhamento exaustivo dos mecanismos jurídicos adotados para cada transferência à luz da LGPD, razão pela qual o item (11) é atendido parcialmente em termos formais (Microsoft, 2025).

Não há indicação pública de nomeação de encarregado por pessoa natural no formato exigido pela LGPD, razão pela qual o item (12) permanece não atendido (Microsoft, 2025)

Em síntese, a documentação do Microsoft 365 Copilot apresenta forte enquadramento contratual e técnico voltado à operação corporativa, com medidas robustas de segurança e controles administrativos que aumentam a previsibilidade operacional. Mantêm-se, porém, lacunas de transparência finalística e jurídico-operacional quando a avaliação é realizada sob a ótica da LGPD, sobretudo no que concerne ao mapeamento detalhado de bases legais por finalidade e categoria de dado, ao *opt-out* individual universal para reuso em treino e à exposição técnica da lógica de inferência.

A análise comparativa das políticas de dados evidencia uma assimetria estrutural: enquanto a retórica empresarial invoca princípios de proteção e conformidade, a prática revela ausência de mecanismos eficazes de controle, fiscalização e responsabilização. Tal discrepância não apenas fragiliza a autodeterminação informativa dos titulares, como também aprofunda desigualdades econômicas e sociais, questão que será detalhada na Seção seguinte.

2.4 A EROSÃO DA AUTODETERMINAÇÃO INFORMATIVA E SEUS IMPACTOS ECONÔMICOS E SOCIAIS: O PAPEL DA IA GENERATIVA NA COLONIZAÇÃO DIGITAL

A autodeterminação informativa, fundamento explícito do Art. 2º, II da LGPD, pressupõe capacidade efetiva do titular de governar o ciclo de vida das informações que o identificam. A matriz dogmática desse direito mostra que a proteção acompanha o dado pessoal onde quer que circule e contra quem quer que o manipule (Oliveira, 2020, p. 4).

No entanto, tal direito se encontra em processo de corrosão diante da arquitetura tecnológica contemporânea, na qual o uso intensivo de dados pessoais por sistemas de inteligência artificial generativa desloca a centralidade do consentimento informado do indivíduo para a lógica de acumulação informacional das plataformas digitais (Tepedino; Teffé, 2020, p. 04).

Cabe, pois, destacar, o precedente paradigmático do Tribunal Constitucional Alemão, que, em 1983, já havia reconhecido que o livre desenvolvimento da personalidade estaria em risco sempre que o indivíduo perdesse a capacidade de controlar o destino de seus dados. Esse entendimento contribuiu para a noção de que a proteção de dados é condição para o exercício efetivo da democracia (Carvalho, 2024, p. 79). Hoje, a massificação da coleta automatizada e a opacidade dos sistemas algorítmicos desafiam essa conquista histórica, reconfigurando a relação entre autonomia individual, economia digital e estruturas de poder.

Na economia de dados, a conversão de informações pessoais em ativo econômico sustenta modelos de negócio fundados no princípio do acesso “gratuito” a serviços digitais, em troca da cessão tácita de dados que passam a compor complexos mecanismos de publicidade e vigilância comportamental. Nesse cenário, o indivíduo não exerce propriamente uma escolha, mas se insere em um ciclo compulsório de troca desigual, no qual sua atenção e seu comportamento são convertidos em mercadoria (Oliveira, 2020, p. 06-07).

Todavia, essa dinâmica não se limita à dimensão mercantil. A extração de dados pessoais e metadados é capaz de produzir perfis psicológicos, prever comportamentos futuros e condicionar preferências políticas, como revelaram os episódios envolvendo a Cambridge Analytica e, de forma mais estrutural, os serviços preditivos desenvolvidos por corporações como o Facebook e a IBM. O “superávit comportamental” é a matéria-prima de uma economia de vigilância que converte traços triviais da vida cotidiana em instrumentos de

controle e lucro (Zuboff, 2021, p. 331-336).

A intensificação dessa lógica por meio da inteligência artificial generativa reforça a assimetria informacional entre usuários e *big techs*. Dotadas de infraestrutura de *big data* e capacidade de processar bilhões de dados por segundo, essas corporações consolidam posições de domínio global, controlando fluxos informacionais e impondo aos cidadãos a condição de sujeitos colonizados digitalmente. Trata-se de uma concentração sem precedentes, na qual a convergência de mídias e serviços digitais gera dependência estrutural e fragiliza a possibilidade de autodeterminação (Carvalho, 2024, p. 126-130).

Há de se considerar que este processo encontra respaldo em mecanismos de opacidade jurídica e técnica. A proteção conferida ao segredo industrial, somada à complexidade inerente dos algoritmos, dificulta a fiscalização pública e a compreensão das decisões automatizadas (Pasquale, 2015, p. 15-19). O resultado é a consolidação de sistemas decisórios herméticos, que operam na interseção entre interesse econômico e poder tecnológico, escapando ao escrutínio social e jurídico (Fama, 2024, p. 19).

A colonização digital, entretanto, não é apenas uma questão de concentração empresarial, mas de reconfiguração das relações sociais. A transformação do indivíduo em marca e a economia da atenção dissolvem fronteiras entre vida pessoal e exposição permanente, criando uma forma de sujeição que Varoufakis (2025) identifica como “tecnofeudalismo”: um regime no qual o capital-nuvem fragmenta a identidade em dados manipuláveis e esvazia os espaços de liberdade individual.

Nesse ambiente, a inteligência artificial generativa assume papel central ao consolidar padrões de classificação e exclusão. Algoritmos de recomendação reproduzem desigualdades estruturais, reforçando estereótipos e marginalizando populações já vulneráveis. O discurso de neutralidade algorítmica oculta as escolhas humanas embutidas na programação, tornando a erosão da autodeterminação um fenômeno não apenas individual, mas coletivo e estrutural (Burrell, 2016, p. 02-03).

No Brasil, a realidade da exclusão digital revela a face mais cruel desse processo. Enquanto setores privilegiados se beneficiam das promessas de eficiência da IA, milhões de pessoas permanecem sem acesso à internet, sobretudo nas regiões menos desenvolvidas e entre grupos vulneráveis, como idosos e pessoas das classes D e E (Montini, 2024). A inteligência artificial, assim, ao invés de reduzir desigualdades, corre o risco de ampliá-las, acentuando a

distância entre incluídos e excluídos no ambiente digital (Carvalho, 2024, p. 61).

Diante desse quadro, a discussão jurídica não pode se restringir à constatação da erosão da autodeterminação informativa, devendo avançar para a análise dos mecanismos de responsabilização aplicáveis à exploração indevida de dados pessoais. Se a colonização digital operada pela IA generativa resulta em impactos econômicos e sociais de grande magnitude, cumpre ao direito indagar quais são os limites da responsabilização civil nesse contexto e em que medida o ordenamento jurídico é capaz de oferecer instrumentos eficazes de tutela. É precisamente essa a reflexão que orientará o próximo capítulo, dedicado aos limites da responsabilização civil no tratamento indevido de dados pessoais no meio digital.

3 LIMITES DA RESPONSABILIZAÇÃO CIVIL NO TRATAMENTO INDEVIDO DE DADOS PESSOAIS NO MEIO DIGITAL

A consolidação da sociedade informacional caracteriza-se pela progressiva desmaterialização dos bens e pela centralidade da lógica predatória da exploração de dados pessoais como recurso econômico e cultural (Zuboff, 2021, p. 15). Esse fenômeno não apenas alterou os padrões de consumo e de interação social, como também impôs novos desafios à responsabilização civil, exigindo uma releitura crítica dos seus institutos tradicionais (Andrade; Faccio, 2019, p. 03-04).

No ambiente digital, a circulação de dados pessoais, antes dispersa em arquivos físicos e burocracias institucionais, passa a constituir verdadeira memória social coletiva. Tal realidade, embora amplie o acesso à informação, também expõe os titulares a riscos inéditos, desde fraudes até ataques anônimos e violações massivas de direitos fundamentais, desafiando a adequação dos pressupostos tradicionais da responsabilidade civil, concebidos anteriormente em um cenário de danos materiais e identificáveis (Rosenvald; Farias; Netto, 2025, p. 923-928).

A sofisticação tecnológica da inteligência artificial generativa, acentua essa problemática. Como demonstrado anteriormente, especialmente na seção 2.3, tais sistemas não apenas processam volumes de dados em escala inédita, como produzem decisões e *outputs* com efeitos práticos sobre direitos individuais e coletivos.

Ocorre que, no campo jurídico, a identificação da autoria do ato lesivo nem sempre é clara. Diferentemente das relações interpessoais clássicas, em que a conduta humana é diretamente vinculada ao dano (Fernandes, 2011, p. 02), a mediação algorítmica pode tornar difusa a imputação, colocando em xeque tanto a teoria da culpa quanto a lógica objetiva de risco da atividade. A ausência de transparência, aliada à imprevisibilidade de comportamentos emergentes dos sistemas, produz uma zona cinzenta que dificulta a fixação da responsabilidade e, em última instância, compromete a efetividade do direito à reparação (Tepedino; Silva, 2019, p. 07-08).

A experiência comparada demonstra que esse dilema não é exclusivo do ordenamento brasileiro. No contexto europeu, a edição da Resolução do Parlamento Europeu de 2017 sobre

robótica, seguida do debate em torno do *AI Act*[4], revela a preocupação com lacunas regulatórias em matéria de responsabilidade civil (Aldmour, 2025, p. 10-13; Tepedino; Silva, 2019, p. 8).

Discute-se se seria necessária a criação de uma "personalidade eletrônica" para sistemas autônomos ou, ao contrário, se bastaria reinterpretar os institutos clássicos à luz das novas tecnologias (Almada, 2019, p. 11). Parte da doutrina sustenta que insistir em reconhecer lacunas normativas pode gerar insegurança e fragmentação, sendo preferível aplicar de forma sistemática os instrumentos já existentes no direito civil e consumerista. Essa tensão entre inovação e continuidade jurídica ilustra o principal limite: a dificuldade em acomodar a dinâmica tecnológica em um sistema jurídico marcado pela estabilidade e pela previsibilidade (Tepedino; Silva, 2019, p. 09-11).

Quando se trata de responsabilidade civil, o ordenamento brasileiro recorre à cláusula geral do Art. 927 do Código Civil (CC/02) como ponto de partida. Esse dispositivo, ao prever a obrigação de reparar o dano decorrente de ato ilícito (Arts. 186 e 187), permite enquadrar condutas violadoras da legislação de proteção de dados como ilícitos civis, ainda que não gerem prejuízo patrimonial imediato. Ademais, o parágrafo único do mesmo artigo viabiliza a responsabilização objetiva quando a atividade desenvolvida for de risco (Brasil, 2002).

Diante desse cenário, é importante reconhecer que o regime de responsabilidade civil - subjetivo ou objetivo - aplicável ao tratamento indevido de dados pessoais por sistemas algorítmicos ainda não apresenta posição consensual na doutrina.

Há, de um lado, autores que defendem a permanência do modelo subjetivo, sustentando que a aferição de culpa ainda seria possível mediante técnicas de auditoria e deveres de transparência (Florence, 2021, p. 9-12). De outro, há quem identifique, na lógica de risco inerente à operação de sistemas de IA generativa, fundamentos para a adoção de um regime objetivo, sobretudo quando a atividade desenvolvida cria perigo especial aos direitos da personalidade (Meo, 2022, p. 164).

O presente trabalho, embora reconheça a coexistência dessas correntes, adotará - por razões que serão detalhadas na seção 4.3.1 - uma leitura orientada à responsabilidade objetiva,

[4] O *AI act* é o Regulamento Europeu da Inteligência Artificial (AI Act), aprovado em 2024 pelo Parlamento Europeu. Disponível em: https://www.europarl.europa.eu/news/pt/press-room/20240308IPR19015/regulamento-inteligencia-artificial-parlamento-aprova-legislacao-historica

especialmente em função da vulnerabilidade informacional do titular e da opacidade estrutural que marca a dinâmica algorítmica. Tal escolha metodológica não pretende encerrar o debate, mas apenas situar o marco teórico a partir do qual se desenvolverá a análise crítica nos capítulos seguintes.

Posto isso, apesar dessa moldura analítica e da possibilidade de enquadrar o tratamento indevido de dados pessoais como ilícito civil nos termos do CC/02, é igualmente necessário reconhecer que as plataformas de inteligência artificial generativa, enquanto fornecedoras[5] de serviços digitais ao público, também se submetem ao regime do CDC (Almada, 2019, p. 5). O CDC, como marco normativo transversal, oferece instrumentos relevantes para a imputação de responsabilidade, especialmente quando a atividade apresenta riscos e quando há falhas na prestação do serviço ou defeitos decorrentes de sua lógica algorítmica (Almada, 2019, p. 6-8).

Contudo, sua aplicação encontra limites diante da complexidade técnica dos sistemas, da dificuldade de caracterizar o "defeito" em ambientes algorítmicos e da diluição da autoria nas cadeias digitais, o que reforça a necessidade de interpretação contextualizada dos institutos clássicos. (Almada, 2019, p. 05-06).

O desafio da responsabilização é ainda mais agudo diante da lógica opaca de funcionamento das redes neurais profundas, cujo caráter de "caixa-preta" compromete a verificabilidade do nexo causal. Se, por um lado, a responsabilidade objetiva aparece como instrumento para neutralizar a prova impossível, por outro, a aplicação indiscriminada desse regime pode gerar custos excessivos e inibir a inovação. Surge, assim, o dilema entre assegurar reparação justa às vítimas e evitar a criação de barreiras desproporcionais ao desenvolvimento tecnológico (Almada, 2019, p. 10).

Em síntese, os limites da responsabilidade civil no tratamento indevido de dados pessoais no meio digital decorrem de múltiplos fatores interligados: a desmaterialização da informação, que desafia a concepção tradicional de dano; a imprevisibilidade dos comportamentos algorítmicos, que compromete a reconstrução causal; a heterogeneidade das soluções normativas, que dificulta a uniformidade interpretativa; e a tensão entre a tutela efetiva dos direitos da personalidade e a preservação de um ambiente propício à inovação tecnológica (Leal; Garbaccio; Mallmann, 2024, p. 12-13).

---

[5] "Art. 3° Fornecedor é toda pessoa física ou jurídica, pública ou privada, nacional ou estrangeira, bem como os entes despersonalizados, que desenvolvem atividade de produção, montagem, criação, construção, transformação, importação, exportação, distribuição ou comercialização de produtos ou prestação de serviços. [...] " BRASIL. Lei nº 8.078, de 11 de setembro de 1990.

Nesse contexto, a aplicação dos institutos clássicos da responsabilidade civil, embora juridicamente necessária, revela-se limitada para abarcar a complexidade das arquiteturas digitais contemporâneas, exigindo uma releitura crítica dos fundamentos dogmáticos e uma adaptação funcional capaz de preservar a efetividade das garantias sem romper com a tradição civilista.

3.1 ESTRUTURA DOGMÁTICA DOS PRESSUPOSTOS CLÁSSICOS DA RESPONSABILIDADE CIVIL

A presente seção delimita a moldura dogmática que dará suporte às análises subsequentes, retomando os pressupostos clássicos da responsabilidade civil (conduta, dano e nexo causal) e a sua articulação com o nexo de imputação.

Parte-se de uma plataforma mínima: a responsabilidade pode assentar-se na culpa (compreendendo dolo e culpa em sentido estrito) ou em bases objetivas vinculadas ao risco da atividade, sem que isso dispense, em qualquer caso, a verificação de um prejuízo juridicamente relevante e do liame causal entre o comportamento e o resultado lesivo. É nessa chave que se compreendem as cláusulas do CC/02 sobre o ato ilícito (Arts. 186 e 187) e a regra geral de reparação (Art. 927), cuja leitura segue a tríade conduta–dano–nexo e admite, conforme o caso, um fundamento objetivo de responsabilização (Fernandes, 2011, p. 2-4, 11-12).

A esse tripé soma-se o nexo de imputação, entendido como o juízo normativo que conecta o fato ao lesante, categoria em que se examinam a censurabilidade da conduta e a própria imputabilidade do agente, de modo que "agir com culpa"[6] significa atuar em termos de merecer reprovação jurídica, a partir de critérios de possibilidade e dever de agir de outro modo (Varela, 1970, p. 582 *apud* Portugal, 2009).

A transposição dessa estrutura para o ambiente informacional impõe um ajuste metodológico sem ruptura do edifício clássico: a circulação massiva e colaborativa de informações pessoais alteram os contextos de risco e ampliam a superfície de danos, exigindo que o intérprete

[6] [...] Agir com culpa significa actuar em termos de a conduta do agente merecer a reprovação ou censura do direito. E a conduta do lesante é reprovável, quando, pela sua capacidade e em face das circunstâncias concretas da situação, se concluir que ele podia e devia ter agido de outro modo." (ANTUNES VARELA, Das Obrigações..., vol. I, citado, p. 582).

privilegie um raciocínio atento ao contexto digital e às novas formas de interação, evitando tanto o conservadorismo que recusa o novo quanto a tentação de soluções avulsas (Rosenvald; Farias; Netto, 2025, p. 925).

Ante a ausência de um microssistema normativo específico voltado à regulação dos danos causados por sistemas de inteligência artificial, a chamada "*lex robotica*", a melhor via neste primeiro momento repousa na releitura sistemática do direito comum de danos, buscando enquadrar os novos desafios dentro da lógica e dos valores do ordenamento jurídico vigente, sem recorrer à criação autônoma de um sistema normativo fragmentado e apartado da tradição civilista (Tepedino; Silva, 2019, p. 10-11).

A partir dessa moldura teórica, torna-se possível examinar cada um dos pressupostos clássicos da responsabilidade civil à luz das especificidades do ambiente digital, com especial atenção às transformações provocadas pela inteligência artificial generativa. Inicia-se, portanto, pela análise da conduta, elemento central na atribuição de responsabilidade, cuja compreensão demanda uma reflexão sobre os limites entre a ação humana e a autonomia algorítmica.

3.1.1 Conduta: entre a ação humana e a autonomia algorítmica

O pressuposto inicial da responsabilidade civil é a conduta, tradicionalmente concebida como ação ou omissão humana apta a produzir efeitos jurídicos. No plano dogmático clássico, ela é condição essencial para a configuração do ato ilícito, por constituir a exteriorização de um comportamento contrário ao ordenamento, cujo exame envolve tanto a antijuridicidade objetiva[7] quanto à imputabilidade subjetiva[8] (Gonçalves, 2024, p. 118; Rosenvald; Farias; Netto, 2025, p. 214-218).

Essa compreensão, porém, encontra limites quando transportada para a realidade dos sistemas de IA, em especial aqueles que operam com autonomia decisória, pois já não se trata apenas

[7] O comportamento antijurídico se instala no momento em que o agente ofende o dever genérico e absoluto de não ofender, sem consentimento, a esfera jurídica alheia. [...] Seja por ação ou por omissão, a contradição do comportamento com o sistema – tido aqui como conjunto de princípios e regras produz a antijuridicidade. (Rosenvald; Farias e Netto, 2025, p. 215)

[8] Enquanto a antijuridicidade é um juízo sobre a conduta, a imputabilidade é um juízo sobre o agente. [...] O imputável é aquela pessoa a quem se pode legitimamente atribuir um comportamento antijurídico. O imputável pode ser censurado e reprovado por suas condutas comissivas ou omissivas contrárias ao direito. [...] Haverá imputabilidade quando o autor do comportamento antijurídico for dotado de maturidade e sanidade. (Rosenvald; Farias e Netto, 2025, p. 216)

de rastrear a vontade humana que os programou, mas de lidar com operações de aprendizado e tomada de decisão não previstas anteriormente (Calo, 2015, p. 21-22; Matthias, 2004, p. 2).

A singularidade da IA reside em sua capacidade de aprendizado contínuo e adaptação, que a diferencia de algoritmos determinísticos[9] dos computadores do passado. Ao operar com aprendizado de máquina, redes neurais artificiais e processamento autônomo de dados pessoais, a IA generativa é capaz de produzir respostas distintas à situações idênticas a partir da acumulação de experiências anteriores, aproximando-se de uma forma de "agir" que transcende a mera execução mecânica de comandos (Cerka; Grigienè; Sirbikytè, 2015, p. 3-6).

O resultado é um cenário em que a fronteira entre a atuação humana e o funcionamento autônomo dos sistemas algorítmicos se torna progressivamente difusa (Calo, 2015, p. 47-50), exigindo uma releitura da compreensão de conduta.

A dificuldade reside no fato de que, diferentemente de um agente humano, a IA é um bem imaterial, carente de consciência ética e de capacidade de ser sujeito de direito, razão pela qual não pode, no atual estado normativo, ser responsabilizada de forma direta. A responsabilidade, por isso, deve recair sobre aqueles que a projetam, controlam, comercializam ou exploram economicamente (Bonnet[10], 2015, p. 14-16).

Isso não elimina a necessidade de reconhecer que reconhecer que decisões tomadas pelo sistema, ainda que não previstas pelo programador, não podem ser vistas como fatos neutros, mas sim como prolongamentos de escolhas humanas que definiram parâmetros de coleta e preparação de dados pessoais (Mulhohand; Kerner, 2020, p. 11-12; Bonnet[11], 2015, p. 14-15). Atribuir à máquina uma autonomia ética equivaleria a esvaziar o papel da dignidade humana como núcleo da juridicidade, transformando o direito em mero instrumento de eficiência técnica (Barbosa, 2017, p. 19-21).

---

[9] Determinismo: significa que, dado um estado inicial e um conjunto de regras, o resultado de um processo será sempre o mesmo. Não há incerteza, e os resultados são previsíveis e replicáveis. Por exemplo, em um algoritmo determinístico, a mesma entrada gera sempre a mesma saída. (Macario, 2024). Disponível em: https://medium.com/@michel.macario/llms-determin%C3%ADsticos-ou-estoc%C3%A1sticos-7658da42e971 . Acesso em: 01 out. 2025.

[10] Para melhor compreensão da referência bibliográfica Bonnet (2015), utilizou-se ferramenta de inteligência artificial para tradução do original em francês para o português, preservando o sentido jurídico do texto. O *prompt* empregado encontra-se disponível em: https://chatgpt.com/share/68df2083-903c-800b-a998-af799e59d174 .

[11] Para melhor compreensão da referência bibliográfica Bonnet (2015), utilizou-se ferramenta de inteligência artificial para tradução do original em francês para o português, preservando o sentido jurídico do texto. O *prompt* empregado encontra-se disponível em: https://chatgpt.com/share/68df2083-903c-800b-a998-af799e59d174 .

Nesse ponto, ganha relevância a definição da conduta específica que fundamenta a responsabilização civil no treinamento da IA generativa: a utilização não autorizada de dados pessoais em seu treinamento.

Trata-se de uma conduta que não se esgota no momento técnico do treinamento, mas que se concretiza por meio de um encadeamento de atos: coleta; extração; agregação; limpeza; rotulagem e retenção. Estes são executados sem base legal adequada ou em desconformidade com os princípios do consentimento, finalidade, adequação, transparência e necessidade. Esse encadeamento de atos integra o próprio ilícito, na medida em que projeta efeitos diretos na esfera jurídica do titular (Mendes; Fonseca, 2020, p. 11-12).

Ao violar deveres normativos de proteção, configura-se o ilícito civil, pois a antijuridicidade resulta diretamente do desrespeito ao dever geral de não lesar (Rosenvald; Farias; Netto, 2025, p. 214) e às obrigações específicas que cercam o tratamento de dados pessoais.

O problema torna-se ainda mais evidente quando os dados utilizados carregam vieses estruturais ou categorias sensíveis, cuja reprodução em sistemas de decisão automatizada aprofunda desigualdades. Nesses casos, a conduta ilícita está também em omitir informações significativas sobre a lógica envolvida e suas finalidades, frustrando o direito à informação e impedindo o exercício de defesa pelo titular (Bigonha, 2018, p. 5-6).

A análise da conduta no contexto da IA generativa, portanto, evidencia que o ilícito não se esgota no plano técnico, mas se projeta sobre efeitos concretos na esfera jurídica do titular. Uma vez delimitado o agir relevante - o uso não autorizado de dados pessoais para treinamento - torna-se necessário examinar como esse comportamento repercute na posição jurídica do indivíduo. Avança-se então, para o exame do segundo elemento: o dano, cuja configuração, em ambientes algorítmicos, apresenta peculiaridades próprias.

3.1.2 Dano decorrente da utilização não autorizada de dado pessoal

A doutrina tradicional civilista identifica no dano o gatilho necessário do sistema reparatório, mesmo quando admite a atenuação ou até a irrelevância de outros filtros, como a culpa ou o nexo causal (Rosenvald; Farias; Netto, 2025, p. 304-305).

Em um cenário de multiplicação de atividades de risco e de proliferação de danos imateriais,

o instituto jurídico deslocou-se do eixo centrado no autor da conduta para um paradigma centrado na vítima e em seus interesses legítimos: sejam estes patrimoniais ou extrapatrimoniais (Rosenvald; Farias; Netto, 2025, p. 305-306).

A utilização não autorizada de dados pessoais deve ser compreendida à luz dessa inflexão, pois a lesão não se esgota na invasão da esfera informacional, mas repercute em múltiplas dimensões da personalidade, tornando necessário compreender o conceito de dano no ambiente digital (Silva; Muniz, 2024, p. 8).

Como explicitado na subseção anterior, sistemas dotados de capacidade de aprendizado autônomo podem produzir efeitos lesivos não apenas imprevisíveis, mas também descolados da intenção original do programador. Exemplos concretos, como a geração de imagens falsas de pessoas reais ou a sugestão de conteúdos inadequados para crianças em plataformas automatizadas, evidenciam que o dano surge da lógica própria da IA, na qual escolhas autônomas se apoiam em dados pessoais coletados e tratados, muitas vezes, sem consentimento ou base legal. Assim, a materialização do dano, no caso da IA generativa, encontra-se no cruzamento entre a exploração indevida do dado e os efeitos que essa exploração projeta na vida dos titulares (Cerka; Grigiene; Sirbikyte, 2015, p. 7-9).

A LGPD fornece parâmetros decisivos para identificar quando a utilização de dados configura dano reparável. O Art. 42 dispõe que o agente de tratamento que, no exercício de atividade de tratamento, causar dano patrimonial, moral, individual ou coletivo em violação à lei é obrigado a repará-lo. Já o Art. 44 explicita que a irregularidade no tratamento é constatada quando não houver observância da legislação ou não for fornecida a segurança que o titular razoavelmente espera, sendo relevantes, para tanto, o modo da realização, os resultados e os riscos esperados (Brasil, 2018).

O legislador, portanto, presume que o dano se conecta diretamente à ausência de licitude, finalidade ou segurança, deslocando a análise para o plano objetivo da violação dos deveres de proteção (Góes; D'Albuquerque, 2022, p. 449-450). Esses parâmetros, embora fundamentais, não esgotam a problemática, especialmente quando se trata de atividades algorítmicas baseadas em coleta massiva de dados.

Em se tratando de treinamento de IA generativa com uso não autorizado de dados pessoais, a definição do instante em que o dano se consuma é uma questão relevante e com consequências dogmáticas e probatórias importantes. Reconhece-se, aqui, que a coleta,

retenção e reutilização indevida de dados podem produzir efeitos jurídicos imediatos. Contudo, a problematização detalhada do instante da formação do dano - e a defesa da possibilidade do seu tratamento como dano presumido (*in re ipsa*) - será desenvolvida de forma aprofundada na subseção 4.3.2. Neste ponto, limita-se a indicação de que o tema merece abordagem específica e criteriosa, que será apresentada adiante para evitar antecipações e manter a estrutura argumentativa do trabalho.

Nesse contexto, o próprio fato de existir um debate sobre o momento em que o dano se consuma permite compreender que ele não se limita às situações em que os efeitos lesivos são evidentes ou imediatamente perceptíveis. Nas práticas que envolvem o uso de dados pessoais por sistemas algorítmicos, as consequências se manifestam de forma gradual, silenciosa e muitas vezes difusa.

Observa-se então, que o ambiente tecnológico contemporâneo introduz padrões de dano quase invisíveis. Um desses padrões é a técnica do *profiling*, que classifica indivíduos em grupos estatísticos a partir de dados coletados, gerando efeitos discriminatórios ou restritivos mesmo quando o titular não é diretamente identificado. Essa dinâmica cria microlesões que, por vezes, não são perceptíveis isoladamente, mas que, acumuladas, fragilizam direitos fundamentais de liberdade e igualdade. O que parece mera classificação estatística pode resultar em discriminação de preços ou em manipulação de preferências políticas, o que evidencia que o dano, nesses casos, se manifesta como prejuízo coletivo e transindividual, ainda que os indivíduos afetados não percebam imediatamente a violação (Fonseca, 2021, p. 28-31; Mulholland; Kerner, 2020, p. 9).

Nessa lógica, ignorar esses efeitos sob o argumento de que não há dano seria perpetuar uma falsa premissa, de que responsabilidade preventiva significaria responsabilidade sem dano. O que ocorre, em verdade, é que o dano se manifesta em modalidades menos tangíveis e mais complexas, exigindo do direito uma adaptação conceitual capaz de abarcar também essas lesões difusas e invisíveis (Fonseca, 2021, p. 29-30).

Assim, o dano decorrente da utilização não autorizada de dados pessoais em sistemas de IA generativa deve ser compreendido como um fenômeno multifacetado, que se consuma desde a própria violação da legalidade no tratamento até a propagação de efeitos patrimoniais, morais e coletivos na esfera dos titulares (Mulholland; Kerner, 2020, p. 16-20) Reconhecer essa pluralidade é essencial para assegurar que a reparação acompanhe a gravidade das lesões informacionais e não se limite a prejuízos econômicos quantificáveis.

Assim, para que a responsabilização civil se efetive, não basta constatar a existência do dano: é necessário estabelecer o liame entre a conduta ilícita e o prejuízo experimentado, especialmente em um ambiente digital marcado por complexidade técnica. É nesse ponto que se insere a análise do nexo causal, elemento que se torna particularmente problemático quando deslocado para a realidade da inteligência artificial generativa.

3.1.3 Nexo causal no contexto digital

O nexo causal, no contexto digital, demanda uma releitura capaz de ampliar, sem romper, os contornos da dogmática clássica da responsabilidade civil. A análise não se limita à conexão fático-natural entre evento e resultado, mas exige também a identificação de uma cadeia informacional que envolve atos de coleta, correlação e inferência de dados pessoais, bem como um juízo normativo de imputação orientado pelos deveres próprios do tratamento desses dados. Essa construção evita a redução do liame causal a um teste simplificado de condição necessária[12], permitindo alinhar o exame às garantias da LGPD e ao regime tradicional da responsabilidade civil, sem instituir uma causalidade excepcional para a tecnologia (Lima Junior; Rodrigues; Moraes, 2025).

Nesse contexto, a causalidade manifesta-se em duas vias que dialogam diretamente com o que já foi construído no presente trabalho: (i) uma dimensão ligada ao tratamento inicial dos dados, na qual a irregularidade do acesso e sua inserção no ciclo de treinamento repercutem diretamente sobre a esfera jurídica do titular, e (ii) uma dimensão subsequente, na qual o *output* do sistema projeta efeitos adicionais e específicos sobre o titular, como *profiling* e memorização de conteúdo.

Em ambas, o nexo não é uma mera abstração: ele resulta de uma sequência operativa - de coleta, preparo, treino e disponibilização - regida por deveres legais cujo descumprimento torna o resultado juridicamente atribuível ao agente de tratamento (Bigonha, 2018, p. 5-6).

A investigação do nexo causal no tratamento indevido de dados pessoais exige reconhecer

[12] "*O teste "but-for" é um teste comumente usado tanto no direito civil quanto no direito penal para determinar a causalidade real . O teste questiona: "Se não fosse pela existência de X, Y teria ocorrido?" . Em direito de responsabilidade civil, a causalidade "se não fosse por" é um pré-requisito para a responsabilização , em conjunto com a causalidade próxima . Na ausência de qualquer uma delas, uma parte não pode ser responsabilizada.*" (Legal Information Institute, 2025). Disponível em: https://www.law.cornell.edu/wex/but-for_test. Traduzido pelo Google. Acesso em: 01 out. 2025.

que, no ecossistema digital, grande parte das interações sociais é intermediada por sistemas informáticos, que transformam comportamentos cotidianos em registros duradouros e continuamente correlacionáveis. O ambiente digital se organiza, assim, como um banco de dados praticamente universal e operado em lapsos ínfimos de tempo, deslocando o problema causal de modelos lineares para cadeias distribuídas e opacas, em que a mesma ocorrência pode se apresentar como condição, ocasião e catalisadora do dano. Esse desenho reforça a necessidade de uma reconstrução minuciosa das trajetórias do dado, sob pena de invisibilizar as contribuições causalmente relevantes e fragilizar a tutela da autodeterminação informativa (Prazeres, 2022, p. 12-17).

Essa complexidade se manifesta também na pluralidade de agentes de tratamento, o que, por si só, não rompe o nexo causal: ele o distribui. Em cadeias técnico-contratuais de treino, a concausalidade é a regra e justifica a solidariedade reparatória prevista nos incisos do § 1º do Art. 42 da LGPD. O critério de imputação, aqui, é o risco-proveito: quem se beneficia da atividade tecnológica que instrumentaliza dados pessoais responde pelos resultados típicos dessa mesma atividade, sobretudo quando decide ou tem melhores condições de prevenir o dano (Almeida, 2023, p. 92-93, 135-136).

Isso reclama uma metodologia que considere não apenas a dimensão fática do evento, mas também a sua relevância normativa, afastando concepções mecanicistas e admitindo a pluralidade causal como elemento inerente às atividades digitais de alta complexidade (Almeida, 2023, p. 129). Em vez de romper o liame quando concorrem fatores externos, o direito deve reconhecer que a eficácia causal se mantém sempre que a conduta do agente de tratamento favoreceu, potencializou ou viabilizou a ocorrência do dano, ainda que em concorrência com outras circunstâncias supervenientes (Almeida, 2023, p. 123).

Outro aspecto relevante é o estatuto probatório da causalidade em cenários de opacidade algorítmica: a dificuldade probatória típica não elimina o liame, mas requer mecanismos que o tornem verificável. A complexidade em rastrear com precisão as etapas técnicas não autoriza concluir pela inexistência de nexo, mas impõe padrões de prova capazes de lidar com incerteza. A doutrina processual aponta que o objetivo não é alcançar a verdade absoluta, mas formar um juízo racional de probabilidade lógica, apoiado em indícios convergentes e máximas de experiência (Carpes, 2013, p. 39-45).

Essa concepção, ao privilegiar a verossimilhança em detrimento da certeza, alinha-se ao contexto da sociedade da informação e reforça que, no campo dos dados pessoais, a

suficiência probatória se constrói a partir de elementos indutivos, e não de demonstrações exaustivas de todo o percurso algorítmico. Destaca-se que a probabilidade, neste contexto, não é um expediente arbitrário, mas uma técnica de imputação fundada na solidariedade social e na proteção da vítima, permitindo que se atribua responsabilidade sempre que a atividade tiver contribuído, de forma significativa, para a probabilidade de ocorrência do dano (Rodrigues Junior, 2013, p. 5).

Esses parâmetros de suficiência probatória alcançam seu ponto mais sensível no contexto da inteligência artificial generativa, em que a opacidade dos modelos e a multiplicidade de bases de treinamento tornam ainda mais árduo o rastreamento do percurso causal. No âmbito do tratamento de dados pessoais para IA generativa, a dificuldade de individualizar qual base de dados deu origem a um *output* discriminatório ou lesivo não pode servir como escudo de irresponsabilidade, sob pena de inviabilizar a tutela da autodeterminação informativa (Calaza, 2024, p. 10).

Superada a dimensão probatória - que será devidamente examinada em momento próprio e posterior na seção 3.3 -, importa observar que a própria análise econômica do direito fornece critérios adicionais para estabilizar o juízo causal em ambientes tecnológicos complexos: o critério do *cheapest cost avoider*. De acordo com essa perspectiva, deve-se atribuir a responsabilidade à parte que tinha melhores condições de evitar o dano a custos menores, o que, no contexto da IA generativa, recai sobre os controladores e operadores que detêm expertise técnica, acesso aos registros de treinamento e poder decisório sobre a arquitetura do modelo (Battesini, 2025, p. 22-25).

Assim, propõe-se uma matriz triádica de aferição da causalidade: (a) Teste de traçabilidade: existem registros técnicos que conectam a base de dados ao modelo e o modelo ao efeito?; (b) Teste de previsibilidade *ex ante*: o resultado integra o espectro de consequências adequadas à atividade, segundo o estado da técnica e a experiência?; e (c) Teste de evitabilidade: o resultado teria sido evitado com cumprimento dos deveres de finalidade, necessidade, segurança e *accountability*? (Battesini, 2025, p. 22-25). A satisfação cumulativa desses testes traduz, no digital, a passagem da causalidade meramente fática para a causalidade juridicamente relevante.

Posto isso, o nexo causal em sistemas digitais de inteligência artificial deve ser compreendido como elemento de articulação entre fato e norma em chave transnacional. A circulação global de dados e a inexistência de fronteiras técnicas para os fluxos de informação tornam ineficaz

qualquer tentativa de enquadrar a causalidade em fronteiras puramente nacionais (Silva; Muniz, 2024, p. 13).

Assim, importa destacar que, mesmo quando o tratamento é realizado por empresas sediadas no exterior, a LGPD não permite a desvinculação causal pela mera alegação de ausência de estabelecimento no Brasil. O Art. 3º da lei estabelece que estão submetidos ao regime jurídico brasileiro todos os agentes que tratem dados no território nacional ou cujo tratamento tenha por objeto pessoas localizadas no Brasil, ainda que a operação seja conduzida fora do país. Nessas hipóteses, o controlador estrangeiro deve manter representante legal constituído no Brasil, nos termos do Art. 1.138 do CC/02, justamente para assegurar a efetividade do nexo causal perante titulares brasileiros (Brasil, 2018; Brasil, 2002).

A causalidade, nesse cenário, não é apenas um elo lógico entre ato e dano, mas o critério decisivo que permite identificar quem deve responder pelos danos aos titulares (Bonnet[13], 2015, p. 10). Compreender essa lógica é essencial para evitar que estruturas técnicas complexas sirvam como álibi para a desresponsabilização de agentes de tratamento que se beneficiam da atividade.

A partir desse quadro, torna-se possível enfrentar, de forma crítica, argumentos que pretendem artificialmente romper o liame causal. É nesse ponto que, antes de falar sobre a imputação, se deve realizar a análise das falácias frequentemente mobilizadas para excluir a responsabilidade civil no domínio informacional, iniciando-se pela alegação de fato exclusivo do titular.

3.1.3.1. A falácia do fato exclusivo do titular: limites à excludente de responsabilidade civil no tratamento de dados pessoais

Superadas as dificuldades ligadas à construção do nexo causal em ambientes digitais, importa agora enfrentar uma objeção recorrente, mas falaciosa, que busca deslocar toda a responsabilidade para o próprio titular dos dados. Atribuir ao titular o fato exclusivo pelo dano informacional somente porque forneceu seus dados pessoais é raciocínio incompatível

[13] Para melhor compreensão da referência bibliográfica Bonnet (2015), utilizou-se ferramenta de inteligência artificial para tradução do original em francês para o português, preservando o sentido jurídico do texto. O *prompt* empregado encontra-se disponível em: https://chatgpt.com/share/68df2083-903c-800b-a998-af799e59d174 .

com a própria dogmática do consentimento na proteção de dados e com a estrutura de vulnerabilidade técnica e econômica que permeia as relações digitais (Capanema, 2020, p. 5).

Para delimitar com precisão por que a falácia do fato exclusivo do titular é particularmente inadequada, é indispensável retomar a mecânica técnica do fenômeno: modelos de IA generativa são treinados com volumes maciços de dados que formam a *corpora* para prever sequências e sintetizar conteúdos. Esses modelos projetam *outputs* textuais, visuais, sonoros ou de código a partir de *inputs* enviados pelos usuários em *prompts*.

Embora os *prompts* não sejam, em regra, concebidos como canais formais de coleta de dados pessoais, na prática os usuários frequentemente inserem, de modo espontâneo e até inadvertido, informações identificáveis sobre si ou sobre terceiros. Esse fenômeno decorre tanto da dinâmica interativa dos sistemas generativos quanto da ilusão de privacidade produzida pela interface conversacional, o que evidencia que a presença de dados pessoais nos *inputs* não é excepcional, mas um reflexo da forma como os modelos são utilizados cotidianamente (Weidinger, *et al.*, 2021, p. 20; Floridi; Chiriatti, 2020, p.4).

Paralelamente, a própria *corpora* utilizada no treinamento - formada por conteúdos coletados via *web scrapping* e *datasets* públicos - agrega dados pessoais colhidos em larga escala, muitas vezes sem finalidade legítima compatível, o que significa que a exposição ou reuso indevido decorre não só da interação do usuário, mas também da apropriação originária de dados pessoais em contextos nos quais jamais houve consentimento informado (Costa *et al*., 2024, p. 19-20).

Em ambos os casos (*inputs* episódicos e treinamento por *corpora*) não há qualquer fundamento para deslocar ao titular a responsabilidade exclusiva. A ausência de consentimento válido e específico, na verdade, revela que o problema reside na própria arquitetura do tratamento e nas escolhas do controlador (Tepedino; Teffé, 2020, p. 15-16).

Nesse sentido, tem-se que o consentimento não é uma cláusula de exoneração automática, mas um processo que demanda informação adequada, inteligível e contínua, em cenário de assimetria informacional acentuada. Tratá-lo como “carta branca” para qualquer reuso ou desvio de finalidade subverte sua função de salvaguarda da autodeterminação informativa e ignora a incidência de vícios e defeitos do negócio jurídico, especialmente quando obtido por meio de políticas de privacidade extensas e herméticas (Dias, 2022, p. 48-49).

Mesmo quando se adota a classificação do consentimento no tratamento de dados como

negócio jurídico, tal enquadramento não o transforma em instrumento de mercantilização da personalidade nem confere ao controlador um poder ilimitado. Ao contrário, enquanto negócio jurídico, o consentimento está submetido aos requisitos de validade e às balizas de proteção desde a concepção (*privacy by design*) e proteção como configuração padrão (*privacy by default*), que exigem configurações protetivas de partida e manifestação ativa do titular para qualquer redução de proteção (Requião, 2022, p. 25-29).

Assim, se o desenho do instrumento impede a livre escolha, compromete a informação ou amplia finalidade de modo inespecífico, incidem as tutelas típicas dos defeitos do negócio jurídico e se frustra a pretensão de invocar o fato exclusivo da vítima. A categorização como negócio jurídico, longe de esvaziar a proteção, oferece um repertório dogmático mais robusto para controlar amplitude, duração, cessão a terceiros e demais cláusulas, inclusive por meio de invalidação, modulação de efeitos e deveres anexos de transparência (Requião, 2022, p. 33).

A compreensão do consentimento como negócio jurídico unilateral e autônomo aprofunda a crítica à falácia do fato exclusivo do titular. Se o consentimento subsiste como ato unilateral distinto do contrato principal que viabiliza o serviço digital, sua revogação não configura inadimplemento contratual, mas cessa a eficácia do tratamento, preservando o núcleo da autodeterminação e impedindo que o usuário seja aprisionado por pactos de adesão que operam como armadilhas. A tentativa de transmutar a revogação ou os limites materiais do consentimento em argumento de autorresponsabilização do titular subverte a estrutura do negócio unilateral e afronta a própria previsão legal de retratabilidade, além de ignorar que o desenho documental usualmente dissocia “termos de uso” e “política de dados”, revelando a cisão estrutural entre a prestação do serviço e o tratamento de dados pessoais (Requião, 2022, p. 29-33).

No plano normativo, o Art. 43, III, da LGPD positivou a hipótese de exclusão do dever de indenizar quando houver fato exclusivo do titular, mas a sua incidência é estrita e depende da demonstração, pelo agente de tratamento, de que o resultado danoso decorreu única e diretamente da conduta do próprio titular, com ruptura completa do nexo em relação à atividade sob seu controle. Cabe ainda ao agente de tratamento comprovar que agiu em conformidade diligente aos princípios de finalidade, necessidade, transparência e segurança, exigidas pelo Art. 6º, incisos I, III, VI e VII do mesmo diploma legal (Brasil, 2018).

Sob o prisma da responsabilidade civil, a figura do fato exclusivo da vítima é hipótese

excepcional de ruptura do nexo causal e não um atalho retórico para desonerar atividades de risco informacional. A distinção entre causalidade e imputação impede que se confunda ausência de responsabilidade do fornecedor com exclusividade causal do titular, e a concorrência de causas, quando presente, reclama repartição equitativa do dano, jamais supressão total da responsabilidade do agente econômico (Rosenvald; Farias; Netto, 2025, p. 526-527).

Transposto ao domínio do tratamento de dados por plataformas e modelos algorítmicos, o fornecimento voluntário de dados pelo usuário não neutraliza a causalidade advinda de desenho de sistema, práticas de coleta opacas, reuso para finalidades novas ou insuficiência de salvaguardas (Capanema, 2020, p. 7). Mesmo na responsabilidade objetiva, a excludente só opera quando o comportamento da vítima é causa direta e imediata do resultado (Rosenvald; Farias; Netto, 2025, p. 527), o que não se verifica quando o dano decorre, ao menos em parte, de escolhas técnicas e organizacionais do controlador.

O chamado paradoxo da privacidade reforça esse raciocínio. Ele demonstra que a divergência entre intenções declaradas e condutas efetivas de divulgação de dados resulta de dinâmicas de recompensas imediatas e assimetria informacional, e não de livre e esclarecida assunção de riscos pelo usuário (Bioni, 2019, p. 211-212; Hsing, 2016, p. 41). Desloca-se o centro de gravidade da responsabilidade para quem organiza o ambiente decisório e captura seus benefícios. A exclusividade causal da vítima, nesses casos, além de empiricamente implausível, contraria o critério de causalidade adequada e a boa-fé objetiva, pois transforma a previsibilidade explorada pelo fornecedor em álibi de exoneração (Hsing, 2016, p. 39-44).

Essa precariedade estrutural do consentimento é agravada por instrumentos contratuais de adesão elebrados por simples clique ou pela navegação em site, expansíveis, mutáveis unilateralmente e redigidos para dissuadir a leitura, que foram normalizados por uma lógica condescendente com renúncias implícitas de direitos. Em tal cenário, imputar ao titular a culpa exclusiva por ter “aceitado” termos intermináveis e atualizáveis sem aviso, equivale a negar a dimensão democrática do contrato e a banalizar a assimetria informacional que impede o discernimento real sobre o tratamento de dados e inviabiliza a invocação séria da excludente (Zuboff, 2021, p. 68-69).

Por fim, evidencia-se que no terreno da IA generativa, o comportamento do titular raramente é causa exclusiva do dano, porque a lesão é, via de regra, resultado de uma constelação de decisões técnicas e organizacionais dos agentes de tratamento, o que afasta o fato excludente e

recoloca no centro o dever de conformidade com o regime jurídico de proteção de dados (Almada, 2019, p. 9-11).

Superada a crítica à tese da culpa exclusiva do titular e esclarecido que o fornecimento de dados, isoladamente, não rompe o nexo causal nem exonera os agentes de tratamento, impõe-se agora enfrentar outra construção igualmente falaciosa: a ideia de que a transformação de dados pessoais em conteúdos sintéticos por modelos generativos teria o condão de neutralizar a ilicitude e afastar a responsabilidade civil.

3.1.3.2. A falácia da neutralização: por que gerar dados sintéticos não rompe o nexo causal pelo uso não autorizado de dados pessoais

A ascensão dos chamados dados sintéticos tem sido celebrada por parte da indústria tecnológica como uma solução apta a conciliar inovação em inteligência artificial e conformidade regulatória em proteção de dados. A ideia de que a geração de conteúdos artificiais, produzidos a partir de padrões estatísticos extraídos de bases reais, afastaria a incidência das normas jurídicas de proteção tem sido apresentada como um argumento de neutralização da ilicitude (Sato, 2023).

Essa construção, porém, incorre em equívoco: o simples fato de a saída gerada pelo modelo não corresponder a registros literais não elimina a materialidade da conduta ilícita inicial, qual seja, a apropriação não autorizada de dados pessoais para treinamento e otimização de modelos de larga escala da IA generativa (Costa *et al*., 2024, p. 21).

Sob uma perspectiva técnica simplificada, os modelos generativos trabalham a partir de dois mecanismos principais. O primeiro é o chamado gerador, responsável por criar novos conteúdos que imitem estatisticamente os dados reais. O segundo é o discriminador, que tenta identificar se a informação recebida é autêntica ou apenas uma simulação. Esse processo de confronto se repete até que os dados artificiais se tornem praticamente indistinguíveis dos dados originais (Costa *et al*., 2024, p. 11-12).

De forma semelhante, modelos baseados em transformadores pré-treinados, como os LLMs, utilizam padrões extraídos de grandes bases textuais para prever e criar novas sequências de palavras que aparentam ser reais. Ainda que sofisticada, essa transformação não rompe a cadeia causal, pois todo o funcionamento do sistema continua dependente do uso inicial de

dados pessoais coletados sem autorização (Costa *et al.*, 2024, p. 14-15). Por isso, a geração de dados artificiais - diga-se, dados sintéticos - não deve ser vista como uma causa que interrompe o nexo de responsabilidade, mas como uma etapa subsequente de um mesmo processo ilícito.

O argumento empresarial de que a artificialidade da saída excluiria a responsabilidade civil esbarra na constatação de que os modelos não podem ser construídos sem acesso prévio a dados reais. A ausência de etapas de pré-tratamento adequadas faz com que dados pessoais, inclusive sensíveis, ingressem nas bases, de modo que a posterior transformação em representações sintéticas não elimina a lesão já consolidada à autodeterminação informativa. Nesse ponto, a geração de conteúdo sintético não apenas mantém o nexo causal, como o reforça, pois projeta para o ambiente externo efeitos potencialmente danosos, como *outputs* inverídicos ou discriminatórios (Costa *et al.*, 2024, p. 18-25).

A tentativa de qualificar dados sintéticos como técnica de anonimização, isentando os agentes de tratamento de obrigações legais (Santos, 2025), constitui uma falácia que ignora o fundamento normativo da proteção de dados pessoais. A utilização de *datasets* com dados reais para construção de conteúdos artificiais configura desvio de finalidade sempre que não houver base legal adequada, pois a exploração econômica da informação se realiza a partir de substratos pessoais concretos (Costa *et al*., 2024, p. 18-25).

Do ponto de vista jurídico, sustentar que *outputs* sintéticos excluem responsabilidade significaria admitir que a adulteração estatística do material bastaria para purgar a ilicitude.

Tal raciocínio compromete a lógica da tutela conferida pela LGPD, que protege não apenas contra a reprodução literal de informações, mas também contra o tratamento indevido em toda a sua extensão, desde a coleta até qualquer forma de reutilização, vide Art. 5º, X da LGPD (Brasil, 2018).

Essa compreensão é corroborada por organismos internacionais de proteção de dados, que advertem para os riscos de se atribuir aos dados sintéticos caráter de anonimização plena. Como assinala o Supervisor Europeu de Proteção de Dados, a sintetização do dado só pode ser considerada válida se for antecedida de uma avaliação robusta de risco e de privacidade, que assegure que o resultado não permita inferências sobre indivíduos reais. Na ausência dessa garantia, os conteúdos artificiais permanecem juridicamente vinculados ao tratamento ilícito de origem, de modo que não rompem o nexo causal e ainda perpetuam seus efeitos

(*European Data Protection Supervisor*).

A esse cenário soma-se a fragilidade estrutural das técnicas de anonimização, historicamente apontada pela literatura especializada. Bruno Bioni (2019, p. 109) entende que a simples aplicação de técnicas de ocultação ou de substituição estatística não impede a reidentificação de titulares, pela existência do chamado "efeito mosaico"[14]. Essa constatação reforça que a confiança na síntese como salvaguarda plena é ilusória, pois inexistem garantias concretas de que conteúdos artificiais sejam imunes à reversão ou a inferências capazes de reconstituir informações pessoais.

Longe de configurar uma excludente, a utilização de dados sintéticos mantém íntegra a responsabilidade civil dos agentes de tratamento, já que o risco de reidentificação subsiste e a violação da autodeterminação informativa se perpetua. Assim, reconhecido que a síntese artificial não rompe o nexo causal nem purga a ilicitude do tratamento, passa-se ao exame do nexo de imputação para definir, no ecossistema de IA generativa, a distribuição concreta de responsabilidade entre os diversos agentes envolvidos.

3.1.4 Nexo de imputação e distribuição de responsabilidade nos ecossistemas de IA generativa

O ponto de partida do nexo de imputação, em ecossistemas de IA generativa, é distinguir o juízo sobre a antijuridicidade da conduta do juízo sobre o agente a quem se atribui o resultado. A IA não é sujeito imputável. Não decide com liberdade ética nem possui capacidade para suportar efeitos jurídicos. A imputação repousa em pessoas físicas ou jurídicas que projetam, integram, treinam, disponibilizam e exploram economicamente o sistema (Albiani, 2019, p. 5-7).

A chave dogmática é a imputabilidade, compreendida como a possibilidade de atribuir legitimamente a alguém um comportamento ofensivo ao Direito e censurá-lo por ele (Rosenvald; Farias; Netto, 2025, p. 215-217). Nesse enquadramento, toda cadeia humana que, com capacidade de autodeterminação e poder de agir de outro modo, põe dados pessoais em

[14] "*Por essa lógica, qualquer dado pessoal anonimizado detém o risco inerente de se transmudar em um dado pessoal. A agregação de diversos "pedaços" de informação (dados) pode revelar (identificar) a imagem (sujeito) do quebra-cabeça, a qual era até então desfigurada (anônimo) – o chamado efeito mosaico"* (Bioni, 2019, p. 109).

fluxo de treinamento sem base legal responde pelo ilícito correspondente.

O que se verifica é a violação objetiva dos deveres de tratamento de dados pessoais, cuja autoria não pode ser atribuída à máquina. A responsabilidade recai sobre o fornecedor ou agente de tratamento que, ao tomar decisões estruturantes inaugura a cadeia causal do ilícito (Albiani, 2019, p. 9). Com isso, consolida-se a premissa delineada na subseção 3.1.1: a conduta não é da IA, mas sim do agente humano ou corporativo que governa o ciclo técnico de dados e modelo. É esse o nexo de imputação aplicável ao uso não autorizado de dados pessoais no treinamento de IA generativa.

A solução também precisa lidar com a opacidade técnica e a pluralidade de contribuições causais presentes nesses sistemas. A responsabilidade subjetiva mostra baixa aderência prática quando os tribunais sequer dispõem de parâmetros concretos para definir o que constituiria diligência adequada no treinamento da IA (Lopes, 2020, p. 121).

Quando a cadeia técnica impede a identificação precisa de qual elo isolado contribuiu para o resultado danoso, a análise deve deslocar-se para modelos de responsabilidade compartilhada, ancoradas em critérios de distribuição racional de riscos. Nessa perspectiva, o fornecedor - i.e., a plataforma de IA generativa - pode ser o destinatário primário da obrigação reparatória perante o titular, sem prejuízo de eventual direito de regresso, caso haja elementos que permitam individualizar falhas posteriores. Nos cenários em que a repartição interna se mostra inviável, a imputação objetiva e conjunta funciona como uma espécie de seguro compulsório em favor do titular, especialmente quando o dano decorre do uso não autorizado de dados e não é possível decompor com segurança epistêmica as falhas individuais. O objetivo não é penalizar a inovação, mas impedir a normalização de um espaço de irresponsabilidade que a opacidade tecnológica poderia legitimar (Lopes, 2020, p. 121-126).

Sob a ótica dogmática, o que antes aparecia como nexo causal clássico se reconstrói como nexo de imputação normativo (Barbosa, 2012, p. 482-484). Quando o ordenamento delineia que certo modo de tratar dados faz recair sobre o agente a assunção de riscos típicos, o liame relevante deixa de ser a cadeia físico-natural e passa a ser a adesão do caso concreto ao programa de proteção da norma. O uso não autorizado de dados pessoais no treinamento de IA é precisamente o tipo de evento que o legislador quis atrair para a órbita de imputação direta, pois desrespeita deveres de finalidade, necessidade, transparência e segurança (Almeida, 2023, p. 155).

Em síntese, o nexo de imputação, nos ecossistemas de IA generativa, deve recair sobre quem estrutura e obtém proveito econômico do processamento de dados pessoais, e não sobre a máquina que apenas executa operações estatísticas. O critério relevante é identificar o agente que, ao tomar decisões sobre o treinamento do modelo, cria ou intensifica o risco típico que a LGPD pretende evitar.

Encerrada a análise dos pressupostos clássicos da responsabilidade civil no contexto do uso não autorizado de dados pessoais no treinamento de sistemas de IA generativa, cumpre agora avançar para a investigação dos modelos de responsabilização que podem ser mobilizados diante dessas violações. É nesse horizonte que se insere a próxima seção.

3.2 RESPONSABILIDADE *EX ANTE* E *EX POST*

A distinção entre responsabilidade (i) *ex ante* e (ii) *ex post* fornece a chave de leitura para um modelo de tutela adequado a riscos informacionais de alto alcance e opacidade técnica: a primeira opera preventivamente, por deveres de prudência, de informação, de estruturação de salvaguardas e por medidas inibitórias e de mitigação; enquanto a segunda reage à materialização do dano, distribuindo seus custos e recompondo, tanto quanto possível, a esfera jurídica atingida. Em termos de dogmática, ambas se articulam sem hierarquia: a tutela preventiva não substitui a reparatória, mas a condiciona, e a reparatória não exonera o devedor de deveres prévios de segurança (Rosenvald, 2025, p. 134-136).

A categoria preventiva encontra, no princípio da precaução, um vetor normativo capaz de deslocar o eixo da responsabilidade civil para a antecipação do perigo, ainda que sob incerteza científica. A experiência francesa registra um "balanço mitigado" quanto à influência indenizatória do princípio, mas reconhece sua força no adensamento de deveres de prudência e, sobretudo, no reforço da função inibitória em situações de risco não dissipado (Boutonnet[15], 2014, p. 11-12).

No plano regulatório europeu, a engenharia de controles *ex ante* do *AI Act* precisa ser pareada por um regime *ex post* apto a enfrentar a autonomia, a opacidade e as cadeias complexas de

[15] Para melhor compreensão da referência bibliográfica Boutonnet (2014), utilizou-se ferramenta de inteligência artificial para tradução do original em francês para o português, preservando o sentido jurídico do texto. O *prompt* empregado encontra-se disponível em: https://chatgpt.com/share/68df268f-d660-800b-808c-0d27d24969f5 .

fornecimento da IA generativa. A crítica que emerge sugere que a efetividade da responsabilização *ex post* depende de soluções mais robustas de imputação (v.g., responsabilidade objetiva), capazes de alinhar incentivos de segurança ao longo da cadeia e de evitar um descompasso entre complexidade normativa e a velocidade/escala de danos potencialmente difusos (La Diega; Bezerra, 2024, p.7-8).

Nessa linha, parte da doutrina contemporânea propõe a expansão funcional da responsabilidade civil para além do paradigma exclusivamente *ex post* reparatório, admitindo "sanções" e providências antes e durante o curso do evento danoso. A matriz teórica distingue respostas anteriores ao dano e concomitantes, sem suprimir a função compensatória, mas integrando-a num sistema de gestão de riscos que legitima a tutela inibitória e providências materiais sub-rogatórias (Carrá, 2016).

A experiência do direito ambiental oferece um roteiro operativo para a dimensão *ex ante* também no domínio informacional: a compensação e as medidas estruturais podem ser condicionantes para atividades potencialmente lesivas, com o propósito de reequilibrar sistemas afetados, impedir a socialização dos custos e internalizar, no agente, os ônus do risco imposto à coletividade. Quando a recomposição específica é inviável, mecanismos substitutivos (*in natura* ou pecuniários) cumprem função distributivo-preventiva, sempre sob a diretriz de que a irreversibilidade do dano não exonera o dever de reparar (Ferreira; Silva, 2007, p. 130-136).

Esse desenho sustenta a transposição, com as devidas adaptações, de obrigações de avaliação de impacto e salvaguardas técnicas prévias às operações de sistemas de IA generativa que envolvam dados pessoais, sem prejuízo das respostas *ex post* quando o evento se consuma (La Diega e Bezerra, 2024, p. 17-18; Santos, 2025, p. 3-4).

O resultado é um sistema em que precaução e compensação se combinam pragmaticamente para induzir condutas diligentes no ecossistema digital, com respostas céleres contra a ilicitude e sem abdicar das garantias reparatórias quando necessário (Rosenvald, 2025, p. 134-136).

Por conseguinte, a responsabilidade civil, aplicada à utilização não autorizada de dados pessoais em treinamentos de IA generativa, deve ser pensada em chave bifronte: estruturar, *ex ante*, deveres e condições de funcionamento capazes de reduzir incerteza e risco, inclusive com base em padrões precautórios e em obrigações proporcionais e economicamente viáveis

(Pfeiffer, 2023, p. 4); e, *ex post*, assegurar instrumentos probatórios e modelos de imputação aptos a superar a opacidade técnica, distribuir custos e restaurar, na medida do possível, a integridade das esferas jurídicas afetadas (La Diega; Bezerra, 2024, p. 3).

A combinação dessas dimensões evita tanto a "morte por incerteza" da tutela preventiva quanto a esterilidade de uma reparação tardia, recolocando a responsabilidade civil no seu papel de indução de virtudes de segurança e de justiça distributiva em face de tecnologias de alta complexidade.

Diante desse cenário, torna-se imprescindível examinar como o ônus probatório se comporta frente à invisibilidade estrutural dos danos decorrentes do tratamento indevido de dados pessoais por sistemas de IA generativa.

3.3 ÔNUS PROBATÓRIO E A INVISIBILIDADE DO DANO NO TRATAMENTO INDEVIDO DE DADO PESSOAL

A discussão sobre o ônus da prova no âmbito da responsabilidade civil por tratamento indevido de dados pessoais encontra seu ponto mais delicado na invisibilidade estrutural que marca tais danos.

Ao contrário dos danos típicos de natureza material, que se exteriorizam de forma perceptível, a violação da autodeterminação informativa opera em registros difusos, muitas vezes silenciosos e imateriais, o que dificulta a demonstração judicial de sua ocorrência. É nesse contexto que a lógica probatória do direito civil deve ser revista para atender às especificidades do microssistema de proteção de dados, evitando que a invisibilidade técnica se converta em impunidade normativa (Moraes; Queiroz, 2019, p. 19-20).

Essa invisibilidade adquire contornos ainda mais severos porque, em grande parte dos casos, o titular sequer tem ciência de que seus dados foram utilizados de forma irregular (Bioni, 2019, p. 49-50), o que inviabiliza não apenas a produção de prova, mas até mesmo a percepção inicial do dano e a ativação dos mecanismos de tutela.

A regra geral do Art. 373 do Código de Processo Civil (CPC/2015) atribui ao autor o ônus de provar os fatos constitutivos de seu direito, e ao réu, os fatos impeditivos, modificativos ou extintivos. Essa lógica, entretanto, mostra-se insuficiente no ambiente digital, em que o titular

dos dados, frequentemente alijado do acesso a registros técnicos e a registros de treinamento, encontra obstáculos insuperáveis para individualizar a origem da lesão. Por essa razão, o legislador brasileiro previu hipóteses específicas de redistribuição dinâmica do ônus da prova, reconhecendo que a exigência de comprovação plena por parte do titular conduziria, na prática, à denegação de justiça (Reichelt, 2023, p. 3-5).

A invisibilidade do dano informacional se conecta, ademais, à própria lógica do tratamento de dados. Muitas vezes, a violação não se consuma em um evento ostensivo como o vazamento público, mas na coleta excessiva e no armazenamento além da finalidade de informações. Tais práticas, embora lesivas, não deixam rastros perceptíveis ao titular, razão pela qual a prova do ilícito exige documentos e registros detidos exclusivamente pelo controlador, cuja adequada organização se torna condição para a própria verificabilidade do tratamento (Santana, 2022, p. 95-102).

Além disso, certos instrumentos previstos na LGPD funcionam como autênticos marcadores probatórios, reforçando a governança de evidências exigida para o controle do tratamento. Os registros das operações de tratamento (Art. 37), os relatórios de impacto à proteção de dados pessoais (Art. 38) e o dever de documentar medidas de segurança (Art. 46) criam trilhas auditáveis que não apenas revelam a conformidade do agente, mas materializam os elementos mínimos necessários para a reconstrução causal da atividade de tratamento (Brasil, 2018). Esses mecanismos ampliam a efetividade da inversão e da redistribuição dinâmica do ônus da prova, pois transferem ao controlador a obrigação de demonstrar, de forma estruturada e documental, a licitude das bases legais utilizadas e das medidas adotadas para mitigar riscos.

A consolidação dessa lógica se insere no movimento de responsabilização proativa delineado pela LGPD. Ao exigir não apenas o cumprimento, mas a demonstração do cumprimento das normas de proteção de dados, a lei transfere ao agente de tratamento o encargo de comprovar a adequação de suas práticas. Trata-se de um sistema que vai além da mera reparação *ex post*, impondo uma obrigação permanente de *accountability* que, na prática, redefine o papel do ônus probatório, deslocando-o do titular vulnerável para o ente que efetivamente detém os meios técnicos e informacionais para demonstrar a regularidade de sua conduta (Moraes, 2019, p. 5).

Não se pode ignorar, contudo, que a efetividade desses mecanismos depende de sua implementação concreta pelos agentes de tratamento. Conforme demonstrado no tópico 2.3.5, a análise documental das políticas de privacidade de plataformas de IA generativa revela que

a maior parte delas cumpre formalmente o item (9) da Tabela 1, indicando medidas técnicas e administrativas de segurança. Esse atendimento, porém, cria uma expectativa legítima de que tais medidas sejam efetivamente observadas no plano prático, o que nem sempre pode ser demonstrado pelo titular diante das limitações informacionais que enfrenta. Esse descompasso aprofunda a assimetria probatória e reforça a necessidade de mecanismos processuais que compensem a fragilidade documental do titular.

Nesse cenário, em que a produção documental é central para a reconstrução do tratamento, é importante destacar que a LGPD adota, em consonância com a tradição consumerista, a técnica da inversão probatória como meio de equalizar a assimetria entre as partes. O Art. 42, §2º, autoriza o juiz a inverter o ônus em favor do titular quando sua hipossuficiência for manifesta ou quando a produção de prova se mostrar excessivamente onerosa (Brasil, 2018). Essa disposição aproxima-se do Art. 6º, VIII, do CDC, que consagra idêntico mecanismo para a tutela da parte vulnerável. Ambos os regimes partem do reconhecimento de que o acesso à prova é decisivo para a concretização do direito material (Brasil, 1990).

Outro aspecto central está no regime das excludentes de responsabilidade. O Art. 43 da LGPD prevê que o agente de tratamento só não será responsabilizado se provar que não realizou o tratamento atribuído, que não houve violação à legislação ou que o dano decorreu de culpa exclusiva do titular ou de terceiro. Em todos esses casos, a prova incumbe ao controlador ou ao operador, invertendo-se, de forma expressa, a lógica probatória tradicional. A redação do dispositivo, próxima à do Art. 12, §3º, e 14, §3º, do CDC, reforça o movimento de atribuir ao fornecedor a incumbência de demonstrar a regularidade da atividade e a inexistência de falha (Almeida, 2023, p. 43-46).

A dificuldade, no entanto, se intensifica quando se tratam de empresas estrangeiras que operam no país. A natureza transnacional das operações, associada à dispersão territorial dos servidores e bancos de dados, coloca obstáculos práticos à coleta probatória pelo titular brasileiro. A situação é agravada pelo fato de muitas dessas empresas não manterem sede no Brasil, em afronta ao dever previsto no Art. 1.138 do CC/02, o que reduz ainda mais a capacidade de fiscalização e de obtenção de documentos essenciais. Ainda que a LGPD tenha abrangência extraterritorial (Art. 3º), sua aplicação depende da cooperação e da transparência de atores frequentemente alheios ao sistema jurídico nacional. (Bastos; Von Ende, 2021, p. 118-124) Nesse cenário, a inversão do ônus da prova mostra-se não apenas conveniente, mas necessária para que a tutela jurisdicional não se torne ilusória.

Por fim, a correlação entre invisibilidade do dano e ônus probatório revela um ponto de convergência entre direito material e direito processual. Se, de um lado, a responsabilidade civil por dados pessoais não pode ignorar o caráter imaterial e difuso da lesão, de outro, a tutela jurisdicional só se efetiva quando os obstáculos probatórios são mitigados. Nesse sentido, a inversão do ônus da prova e a exigência de prestação de contas não são privilégios concedidos ao titular, mas mecanismos de reequilíbrio indispensáveis para que a promessa constitucional de proteção de dados seja mais do que retórica normativa (Capanema, 2020, p. 4).

Como desdobramento dessa tensão entre prova e proteção, impõe-se agora examinar os efeitos jurídicos da irreversibilidade no uso indevido de dados pessoais.

3.4. RESPONSABILIDADE DIANTE DA IRREVERSIBILIDADE DO USO DE DADO PESSOAL

A problemática da irreversibilidade do uso de dados pessoais emerge como uma das mais desafiadoras no âmbito da responsabilidade civil contemporânea. Uma vez que informações individuais são coletadas, processadas e inseridas em sistemas de inteligência artificial, especialmente em larga escala, o retorno ao *status quo ante* mostra-se inviável. O dado, ao ser absorvido por um modelo algorítmico, passa a integrar padrões de correlação e inferência que não podem ser desfeitos sem comprometer toda a estrutura do sistema (Costa *et al*., 2024, p. 18).

Esse caráter irreversível distingue o dano relativo à dados pessoais de outras modalidades de lesão, pois não se trata de recompor fisicamente um bem ou de reparar um patrimônio material, mas de enfrentar a perpetuação de efeitos derivados da exploração indevida da informação pessoal (Ehrhardt Jr. e Modesto, 2022, p. 141).

A irreversibilidade não se reduz à impossibilidade técnica de apagar rastros digitais. Trata-se, sobretudo, de uma consequência sistêmica: uma vez que dados pessoais integram a lógica preditiva dos algoritmos, eles passam a gerar *outputs* que podem projetar efeitos jurídicos, econômicos e sociais sobre os titulares. Assim, ainda que se determine a exclusão formal de determinada informação, seu impacto já está reproduzido em inferências e decisões automatizadas (Oliveira, 2020, p. 7-13).

Diante desse quadro, a dificuldade de recomposição deve ser fator de intensificação da tutela. O direito ambiental fornece precedente valioso: quando há lesões irreversíveis ao meio ambiente, a doutrina e a jurisprudência não afastam o dever de reparar, mas estruturam medidas compensatórias capazes de evitar a socialização do prejuízo e garantir proteção coletiva. O princípio conservacionista, nesse domínio, demonstra que a irreversibilidade não elimina a responsabilidade, mas exige sua adaptação em direção a soluções compensatórias e estruturais (Ferreira e da Silva, 2007, p. 7).

Na seara da proteção de dados, essa mesma lógica impõe a construção de instrumentos compensatórios e preventivos aptos a enfrentar riscos que se perpetuam no tempo. A inserção de dados em sistemas de IA generativa, quando feita sem consentimento ou em desvio de finalidade, configura um risco de longa duração que não se extingue com a simples cessação do tratamento (Atata, 2024, p. 4). A difusão do dado, sua replicação e o caráter expansivo da informação digital tornam inviável o retorno ao estado anterior, o que exige da responsabilidade civil uma postura de longa duração, voltada não apenas para reparar, mas também para prevenir e mitigar os efeitos de danos graves e irreversíveis (Rosenvald; Farias; Netto, 2025, p. 133).

A ANPD (2023) reconhece esse desafio ao estabelecer critérios para identificar operações de tratamento de alto risco, entre eles a larga escala e a potencial afetação significativa de direitos fundamentais. Quando associados a tecnologias emergentes ou inovadoras, como a IA generativa, esses riscos assumem contornos de irreversibilidade, dada a impossibilidade de controlar a circulação da informação uma vez disseminada. Nessas hipóteses, a própria regulação já antecipa a necessidade de instrumentos mais rigorosos de responsabilização, justamente porque a reparação clássica não é suficiente para resguardar os direitos dos titulares.

A complexidade da irreversibilidade reforça ainda a necessidade de reconceitualizar os interesses jurídicos protegidos pela tutela de dados pessoais. Como aponta Marion Albers (2016, p. 12-20), a proteção não recai sobre os dados em si, mas sobre as consequências sociais e jurídicas que sua manipulação projeta sobre os indivíduos. Em um contexto em que os dados são constantemente transformados em conhecimento e decisão, a irreversibilidade compromete não apenas a privacidade, mas também a integridade contextual e a liberdade de escolha dos titulares. A responsabilidade civil, diante disso, deve se voltar para salvaguardar posições jurídicas complexas e multidimensionais, que não se reduzem ao controle individual

sobre os próprios dados.

Nesse sentido, a irreversibilidade do uso de dados pessoais não pode ser interpretada como obstáculo à responsabilização, mas como fator que a intensifica. O reconhecimento de que a reparação integral, nos moldes clássicos, é inalcançável não exime o ofensor de responder, mas demanda soluções jurídicas inovadoras, que combinem compensação econômica, medidas estruturais de governança, reforço do dever de transparência e mecanismos de tutela coletiva. O dano irreversível, longe de neutralizar o dever de reparar, amplia a necessidade de responsabilização, justamente porque sua perpetuação ameaça não apenas o titular do dado, mas a própria estabilidade das relações sociais em uma sociedade informacional (Oliveira, 2020, p. 11).

Diante disso, a irreversibilidade do uso de dados pessoais evidencia que a responsabilidade civil, nesse domínio, não se limita à reparação imediata, mas deve lidar com efeitos persistentes e dificilmente quantificáveis. Essa constatação conduz à necessidade de refletir sobre a própria natureza jurídica do dano e os critérios de mensuração dos prejuízos indenizáveis em violações de dados pessoais, tema que se apresenta como o próximo passo da análise.

3.5 NATUREZA JURÍDICA E MENSURAÇÃO DOS PREJUÍZOS INDENIZÁVEIS EM VIOLAÇÕES DE DADO PESSOAL

A determinação da natureza jurídica dos danos oriundos da violação de dados pessoais representa desafio central na dogmática contemporânea da responsabilidade civil. Diferentemente das hipóteses clássicas em que o prejuízo se materializa em bens tangíveis ou em perdas econômicas quantificáveis, os danos relacionados ao tratamento indevido de dados envolvem a colisão entre direitos da personalidade, projeções econômicas da informação e o risco estrutural inerente ao ecossistema digital. O dado pessoal, enquanto projeção da identidade e da autodeterminação informativa do sujeito, não pode ser reduzido à mera mercadoria sem que se comprometa sua dimensão existencial - que será devidamente explorada na subseção 4.3.3 (Requião; Prazeres, 2025, p. 54-55).

A LGPD (Art.42), ao prever expressamente a reparação de danos morais, patrimoniais, individuais e coletivos (Brasil, 2018), consolidou essa ampliação, permitindo que o direito

alcance as múltiplas formas de lesão geradas por práticas ilícitas no tratamento de dados. A exploração indevida de informações pessoais, portanto, deve ser concebida como conduta que já contém em si o potencial lesivo, cuja concretização não depende da demonstração de prejuízo econômico, mas do reconhecimento de que houve violação a interesses jurídicos fundamentais (Góes; D'Albuquerque, 2022, p. 445-446).

O debate, contudo, não se esgota na dicotomia entre dano material e dano moral. A mensuração dos prejuízos indenizáveis nas violações de dados exige reconhecer que o tratamento indevido pode produzir efeitos econômicos relevantes, ainda que de difícil aferição em razão da opacidade algorítmica. Quando empresas exploram dados pessoais em larga escala para fins de monetização, cria-se um valor agregado não partilhado com o titular, mas internalizado pelo agente de tratamento. Ocorre, portanto, uma assimetria em que o titular suporta a violação de seus direitos fundamentais enquanto o controlador aufere ganhos econômicos muitas vezes incalculáveis (Valadão, 2025, p. 22).

A doutrina e a jurisprudência têm recorrido à presunção do dano moral como forma de evitar a inviabilidade prática da tutela. No entanto, surge a necessidade de parâmetros objetivos de quantificação, sobretudo diante de ilícitos massivos ou de violações coletivas. A experiência do direito privado demonstra que, assim como em outras hipóteses de danos difusos, a função reparatória deve ser articulada à função preventiva, de modo a desestimular práticas lesivas estruturalmente lucrativas. Nesse cenário, cogita-se a possibilidade de indenizações escalonadas ou punitivas em hipóteses de reiterado descumprimento das normas de proteção de dados (Antunes, 2019, p. 10).

A análise histórica do dano moral no ordenamento brasileiro auxilia na compreensão desse movimento. Superada a concepção restrita que o vinculava apenas a sentimentos subjetivos de dor ou vexame, consolidou-se a noção de que ele se refere à lesão de interesses existenciais e da dignidade humana. Essa evolução permitiu não apenas a tutela da esfera individual, mas também a expansão para a proteção de bens metaindividuais, admitindo-se a configuração do dano moral coletivo em hipóteses de lesões maciças a valores difusos da coletividade (Rosenvald; Farias; Netto, 2025, p. 438).

Não obstante, a natureza coletiva desses danos ainda suscita divergências. Parte da doutrina sustenta que o dano moral coletivo não se confunde com uma lesão experimentada por um ente abstrato, mas se justifica como verdadeira pena civil com caráter pedagógico e sancionatório. Assim, mais do que compensar a coletividade, tais condenações visam

reequilibrar as relações jurídicas e induzir os agentes econômicos a internalizar os custos sociais de suas práticas abusivas (Rosenvald; Farias; Netto, 2025, p. 445-446).

A mensuração dos prejuízos indenizáveis em violações de dados pessoais, portanto, deve considerar tanto a dimensão moral quanto a econômica. Se, por um lado, os dados integram o núcleo dos direitos da personalidade e, como tal, sua lesão é imprescritível e inalienável em sua essência, por outro, é inegável que possuem expressão patrimonial, cuja exploração indevida pode ser quantificada em termos de proveito econômico. Negar essa duplicidade seria ignorar a realidade de um mercado estruturado sobre a mercantilização da informação pessoal (Requião; Prazeres, 2025, p. 64-65).

Um ponto particularmente sensível reside na assimetria informacional que impede o titular de aferir o alcance de sua lesão. O valor extraído de grandes bases de dados, sobretudo em treinamentos de inteligência artificial generativa, não se apresenta de forma transparente nem passível de mensuração direta pelo indivíduo lesado. Essa circunstância torna insuficiente a reparação meramente compensatória, exigindo que o Judiciário incorpore critérios de razoabilidade e proporcionalidade ao fixar o *quantum* indenizatório (Garcia; Nunes, 2021, p. 3).

Ademais, a responsabilização deve levar em conta o contexto em que os dados são utilizados como moeda de troca, seja em modelos de monetização direta, seja em arranjos indiretos em que a privacidade é convertida em contraprestação (Valadão, 2025, p. 16, 22 e 29). O STJ (2012), no julgamento do Recurso Especial 1192208/MG pela Ministra Nancy Andrighi, que ocorreu em 02 de agosto de 2012, reconheceu que mesmo nos casos em que o serviço é aparentemente gratuito, subsiste relação de consumo, pois há remuneração indireta pelo uso dos dados, o que reforça a necessidade de aferição econômica dos prejuízos.

Diante desse quadro, a natureza jurídica dos danos indenizáveis por violações de dados pessoais deve ser concebida de forma híbrida: extrapatrimonial, por tutelar a dignidade e a autodeterminação informativa, e patrimonial, por reconhecer a exploração econômica subjacente (Valadão, 2025, p. 35). A mensuração dos prejuízos, por sua vez, não pode restringir-se a uma lógica reparatória minimalista, mas deve abranger critérios que contemplem a gravidade da lesão, o alcance coletivo e a função preventiva da responsabilidade civil, de modo a reequilibrar a assimetria estrutural entre titulares e agentes de tratamento (Cardoso, 2020, p. 11-12).

Os desafios mapeados ao longo deste capítulo revelam que a resposta jurídica não pode se restringir à aplicação linear dos pressupostos clássicos, sob pena de esvaziar a garantia à proteção de dados pessoais prometida pelo sistema normativo brasileiro. É justamente a partir dessa constatação que se abre o campo para a análise crítica a seguir, voltada a investigar como reorientar a responsabilidade civil para que produza efeitos jurídicos reais em um ambiente marcado por risco difuso, captura informacional e estruturas tecnológicas de grande escala.

4 A RESPONSABILIZAÇÃO CIVIL PELO USO NÃO AUTORIZADO DE DADOS PESSOAIS NO TREINAMENTO DE IA GENERATIVA: UMA LEITURA CRÍTICA

Este capítulo parte do quanto consolidado no capítulo anterior - a existência do dever jurídico de responsabilização pelo uso não autorizado de dados pessoais - para enfrentar o problema que nele apenas aflora: a aplicabilidade prática desse dever no contexto da inteligência artificial generativa. Em vez de retomar os pressupostos clássicos, toma-os por conhecidos e desloca o foco para os mecanismos que permitem transformar esse dever em consequências jurídicas verificáveis e efetivas (Silva; Muniz, 2024, p. 7).

Busca-se compreender de que modo o direito pode conservar sua função regulatória e garantista em um cenário no qual a exploração de dados pessoais se converte em condição estrutural da economia digital e do aprendizado de máquinas (Ursic*,* 2018, p. 72-73). Trata-se, portanto, de examinar não apenas se há responsabilidade, mas como e para que ela deve operar em um ambiente de risco difuso e cumulativo.

Nesse percurso, adota-se, de forma instrumental, uma abordagem sistemático-teleológica, apenas como uma ferramenta interpretativa apta a preservar a coerência do sistema jurídico diante das novas materialidades de risco e dano introduzidas pela economia dos dados. O detalhamento dessa técnica interpretativa será desenvolvido de modo pontual na seção 4.3, quando sua função metodológica se torna diretamente relevante para a reorientação do regime de responsabilidade civil.

A partir desse enquadramento, a hermenêutica sistemático-teleológica é mobilizada apenas para evitar leituras estritamente literalistas, incapazes de conter as externalidades tecnológicas: a responsabilidade civil deixa de operar apenas como técnica de recomposição individual para assumir também uma função de coerência sistêmica entre normas jurídicas e práticas tecnológicas (Vieira *et al*., 2025, p. 5-9).

A questão hermenêutica, portanto, torna-se relevante não como fundamento autônomo, mas como método auxiliar para ajustar o texto legal às novas materialidades informacionais. A teoria objetiva da interpretação, ao reconhecer a autonomia da norma, permite adaptar o texto legal a novas realidades e ampliar sua eficácia regulatória (Fagundes, 2025, p. 6-8). Em matéria de danos informacionais, o fato jurídico ultrapassa a coleta do dado pessoal e se estende por todo o ciclo de tratamento e reuso, frequentemente em contextos não previstos

pelo titular. Dessa forma, a interpretação passa a exercer papel constitutivo na definição do ilícito e na delimitação do alcance da responsabilidade, não apenas descrevendo o fenômeno, mas revelando as insuficiências do modelo vigente (Citron; Solove, 2021, p. 26).

Diante desse panorama, a responsabilização civil contemporânea se vê entre dois impasses: de um lado, a insuficiência dogmática para captar a natureza difusa e estrutural do dano informacional (Kreimer, 2016, p. 37); de outro, o anacronismo de categorias pré-digitais de conduta, dano e nexo causal, concebidas originalmente para eventos pontuais e individualizáveis (Fernandes, 2011, p. 2-4).

Na sociedade da informação, em que o dado é simultaneamente insumo, produto e moeda, diluem-se as fronteiras entre agente e instrumento, e as relações passam a ser de fluxo, não de evento. A ilicitude deixa de ser episódica e torna-se sistêmica, produzindo uma crise de efetividade do paradigma reparatório tradicional (Souza; Lopes, 2013, p. 15-16).

Desse modo, a responsabilidade civil clássica, fundada na lógica de reparação individual e *ex post*, revela-se insuficiente para enfrentar o uso não autorizado de dados pessoais no treinamento de IA generativa. É justamente essa insuficiência estrutural do modelo reparatório, diante da exploração indevida e silenciosa de dados pessoais, que orientará a análise da próxima seção, voltada a examinar as limitações internas desse paradigma e os caminhos possíveis para uma responsabilização efetiva.

4.1 A INSUFICIÊNCIA ESTRUTURAL DO PARADIGMA REPARATÓRIO FRENTE À UTILIZAÇÃO INDEVIDA DE DADOS PESSOAIS POR SISTEMAS DE IA

Antes de avançar, cumpre esclarecer que o termo função reparatória é aqui utilizado em sentido amplo, para designar o conjunto de mecanismos de responsabilização *ex post* que buscam reagir a um dano já consumado, englobando as dimensões compensatória e punitiva historicamente atribuídas à reparação civil (Del Mastro, 2015, p. 14). Alguns autores, entretanto, operam uma distinção terminológica entre “reparação”, entendida como retorno ao *status quo ante*, e “compensação”, reservada ao pagamento pecuniário (Rosenvald, 2019, p. 1-7).

Para fins do presente trabalho, as expressões “função reparatória” e “função compensatória”

serão empregadas para indicar a lógica retrospectiva da responsabilização civil, isto é, o conjunto de respostas *ex post* orientadas à recomposição ou sanção do ilícito.

É nesse sentido - e sem prejuízo da distinção mencionada acima - que, a lógica clássica da reparação, voltada a recompor a vítima ao *status quo ante* por equivalência pecuniária, mostra-se estruturalmente inadequada quando confrontada com danos informacionais produzidos por sistemas de IA treinados mediante uso indevido de dados pessoais (Ghani, 2017, p. 7-9).

Nesses contextos, a lesão tem natureza expansiva, duradoura e, muitas vezes, irreversível, de modo que a mera compensação não satisfaz o imperativo de justiça nem reequilibra a posição jurídica do titular: a informação já circulou, foi incorporada ao modelo e reempregada em cadeias técnicas que escapam ao controle do lesado. É precisamente essa irrecuperabilidade que evidencia o esgotamento de um paradigma que opera retrospectivamente e supõe reversibilidade fática, como já advertido pela doutrina ao indagar se a indenização, por si, consegue "fazer justiça" diante de danos que não se desfazem (Ferreira, 2014, p. 34).

A responsabilidade civil contemporânea ampliou seu campo de incidência sem que se redefinisse, de modo sistemático, sua finalidade. Essa expansão revelou a fragilidade de um modelo centrado na reparação *ex post*, incapaz de lidar com práticas massivas e opacas como as do treinamento de IA generativa. Diante da dificuldade de imputar condutas e de mensurar danos em ecossistemas algorítmicos, a insistência na função meramente compensatória perpetua uma incongruência: amplia-se o discurso de tutela, mas não se alcança proteção efetiva diante dos novos danos (Ferreira, 2014, p. 66-71).

Há, ademais, uma assimetria estrutural que agrava o problema: o circuito reparatório tradicional acarreta custos administrativos e processuais elevados, consumindo parcela significativa dos recursos destinados à vítima e desacelerando qualquer resposta útil (Ferreira, 2014, p. 68). Em matéria de dados, esse atraso processual equivale a perder a própria oportunidade de tutela, porque o ciclo tecnológico avança enquanto a demanda se arrasta, sedimentando a obsolescência da recomposição. O resultado é um paradoxo de eficiência: quanto mais sofisticado o aparato técnico para o uso não autorizado de dados para treinamento, maior a distância entre a lesão e o momento de eventual provimento jurisdicional, com evidências empíricas históricas de que o gasto transacional do sistema supera, em larga medida, o benefício entregue ao lesado (Lintvedt, 2022, p. 2, 36-37).

A insuficiência também se manifesta no plano dos incentivos. Ao aceitar que a indenização pecuniária “pague a conta” após a ocorrência do ilícito informacional, o sistema transmite ao agente econômico a mensagem de que a apropriação de dados pode ser internalizada como custo ordinário, sobretudo quando a escala do negócio dilui o impacto de condenações unitárias (Zanini, 2018, p. 6).

Em ambientes de exploração algorítmica reiterada, a função compensatória, sozinha, não corrige o cálculo de proveito ilícito nem enfrenta a racionalidade de portfólio que compara vantagem obtida com eventual condenação. Esse fenômeno abrirá espaço, mais adiante, para o diálogo com mecanismos como os *punitive damages*, cuja análise será desenvolvida oportunamente na subseção 4.4.3.

No domínio específico da proteção de dados, o próprio desenho normativo da LGPD reconheceu que não basta ressarcir “depois”: é preciso deslocar o eixo para a demonstração ativa de conformidade e para a prevenção de danos (Brasil, 2018). O regime de responsabilização civil atrelado ao tratamento de dados não pretende apenas recompor perdas, mas impor um dever de prestação de contas que antecipe riscos e redesenhe processos internos, inclusive com consequências coletivas e informacionais. Tal arranjo explicita que “não descumprir a lei” é insuficiente: é necessário provar medidas eficazes e sua efetividade, sinalizando que a tutela adequada exige deslocamento do foco, sob pena de a reparação *ex post* converter-se em remédio tardio e inócuo frente à dinâmica técnica do processamento em larga escala (De Moraes, 2019, p. 2-5).

A objeção de que o reforço de deveres e a objetivação de riscos “inibiriam a inovação” também revela um falso dilema quando transposta ao cenário dos dados pessoais. A experiência comparada demonstra que modelos de imputação mais exigentes não paralisam o desenvolvimento tecnológico; ao contrário, internalizam custos de segurança e conformidade e realocam o risco para quem melhor pode geri-lo, sem confundir periculosidade inerente com risco da atividade (De Moraes, 2019, p. 3-4). Em mercados de dados, isso significa admitir que a própria arquitetura de treinamento da IA integra a álea do negócio e deve ser governada por deveres de cuidado proporcionais ao poder técnico-econômico dos agentes - algo que o paradigma puramente reparatório não alcança, exatamente por chegar tarde e olhar para trás (Freitas, 2019, p. 36-39).

Por fim, a historicidade da crise reforça o diagnóstico: à medida que a técnica deslocou o centro de gravidade dos acidentes pessoais para danos difusos - primeiro com a mecanização,

depois com a digitalização -, a promessa de retorno pleno ao estado anterior se tornou, em larga medida, ficção jurídica (Rosenvald, 2019, p. 7). No plano informacional, a irreversibilidade da replicação e a impossibilidade prática de "destreinar" modelos em escala ordinária corroem a pretensão de equivalência compensatória (Cooper *et al.,* 2024, p.2).

A função reparatória permanece necessária, mas, sozinha, é insuficiente para responder a ilícitos informacionais sistêmicos que desorganizam autonomias e ambientes informacionais de maneira persistente. É essa insuficiência estrutural que prepara o terreno para examinar, a seguir, o déficit da proteção dos dados pessoais e a captura da autonomia do titular como elementos constitutivos do problema.

4.1.1 O déficit estrutural da proteção de dados pessoais e a captura da autonomia do titular

O diagnóstico sobre a exploração predatória dos dados pessoais, desenvolvido no Capítulo 2 já evidenciou que a economia digital opera sob uma lógica de apropriação da experiência humana como insumo econômico. No presente ponto, o interesse desloca-se do fenômeno econômico em si para os efeitos jurídicos-estruturais dessa lógica sobre o regime de proteção de dados e, em especial, sobre sua capacidade de sustentar uma responsabilização civil efetiva.

O núcleo do problema está na dissonância entre a promessa normativa da autodeterminação informativa, consagrada pela LGPD, e a realidade de um sistema técnico e econômico que a neutraliza desde a origem. As categorias dogmáticas que estruturam a proteção de dados (consentimento, finalidade, necessidade, transparência, adequação e segurança) são formalmente mantidas, mas materialmente esvaziadas pela assimetria entre titulares e agentes de tratamento. O resultado é um regime de proteção que opera, na prática, como gestão privada do risco, e não como tutela de direitos fundamentais (Meireles, 2023, p. 20, 22-25).

O consentimento, pilar do modelo de legitimidade do tratamento de dados, tornou-se o principal vetor dessa captura (Mendes; Fonseca, 2020, p.4). Reconhecer o consentimento como processo e como ato jurídico com eficácia autônoma, perspectiva desenvolvida por Maurício Requião (2022, p. 21, 29-32), orienta a avaliação prática da sua eficácia, privilegiando critérios como modo de obtenção e possibilidade real de revogação, sem reduzir

a análise ao mero registro do aceite.

No entanto, na prática, convertido em mero clique de adesão, o consentimento legitima o fluxo de dados sem assegurar reflexão ou compreensão real do titular. Trata-se de um consentimento performático, que preenche uma exigência formal, mas não exprime vontade livre, pois a sua recusa implica exclusão funcional da vida digital. Nessa moldura, a autonomia não é exercida, mas simulada, servindo como instrumento de legitimação de práticas massivas de extração e predição (Grossi, 2023, p. 164-166; Bioni, 2019, p. 172-173).

Essa deformação tem caráter estrutural: deriva não da omissão individual dos titulares, mas do próprio *design* técnico das plataformas, que tornam inviável a oposição concreta do titular ao tratamento de dados. Como demonstrado na Tabela 1 (item 4) do tópico 2.3.5, a análise das políticas de privacidade das plataformas de IA generativa aqui estudadas confirma que o consentimento figura, na prática, como base legal secundária.

O sistema de proteção, ao centrar-se em deveres de informação e em autorizações pontuais, supõe a possibilidade de decisão racional em contextos onde o conhecimento é distribuído de forma precária. A transparência prometida pela LGPD colide com a opacidade algorítmica das operações de IA generativa, tornando impraticável a aferição da extensão do dano (Leal; Paulo, 2023, p. 13-18).

Nessa perspectiva, a LGPD representa avanço inegável ao incorporar deveres de conformidade e mecanismos preventivos. Contudo, a efetividade prática desses instrumentos permanece limitada, pois o modelo regulatório ainda se apoia em respostas predominantemente reativas e casuísticas, incapazes de acompanhar operações que ultrapassam a escala humana de controle (Grossi, 2023, p. 144, 171-172). Diante da datificação total, o indivíduo deixa de ser sujeito de direitos e passa a integrar o ambiente técnico como variável estatística: uma mutação que fragiliza o próprio conceito de dano reparável.

Assim, o que se denomina “déficit estrutural da proteção de dados pessoais” não corresponde à ausência de mecanismos de responsabilização, mas à ineficácia intrínseca de instrumentos *ex post* diante de infrações coletivas e cumulativas. A proteção de dados, concebida para reagir a violações pontuais, encontra-se diante de infrações contínuas, embutidas na própria infraestrutura das plataformas de IA. Nessa condição, o dano não é um evento, mas sim um processo: não há uma conduta isolada, mas uma cadeia de usos e inferências que produzem

erosão progressiva da autonomia do titular.

Portanto, a captura da autodeterminação informativa não é apenas sintoma da lógica de mercado, mas limite estrutural à efetividade da responsabilidade civil. Sem acesso à informação suficiente para traçar causalidade e quantificar prejuízos, o titular encontra-se juridicamente paralisado (Cohen, 2012, p. 5). A tutela privada, que pressupõe um lesado determinado e um agente identificável, colapsa diante da anonimização do dano e da pulverização dos responsáveis, características inerentes ao treinamento de sistemas de IA generativa (Pereira; Ebling, 2023, p. 13-14).

Em síntese, o déficit estrutural da proteção de dados pessoais revela-se como condição antecedente da crise do paradigma reparatório: a responsabilidade civil permanece dependente de um modelo de causalidade e dano individualizado que já não corresponde à realidade técnica dos fluxos informacionais. É dessa fricção entre promessa de tutela e impotência estrutural que emerge a necessidade de reavaliar o papel da responsabilização *ex post*, tema que será desenvolvido a seguir, na subseção dedicada à ilusão compensatória diante da exploração massiva e opaca de dados no treinamento de IA generativa.

4.1.2 A responsabilização *ex post* e a ilusão da compensação individualizada diante da exploração massiva e opaca de dados no treinamento de IA generativa

A racionalidade reparatória que ancora a responsabilidade civil tradicional apoia-se na crença de que todo dano pode ser revertido ou compensado mediante equivalência monetária. Essa lógica, concebida para eventos delimitáveis no tempo e na autoria (Souza; Lopes, 2013, p. 23-24), torna-se ilusória diante de operações automatizadas que produzem ofensas disseminadas, cumulativas e anônimas.

O uso não autorizado de dados pessoais para o treinamento de sistemas de IA generativa constitui um exemplo paradigmático dessa desconexão: a lesão não se esgota no instante da coleta, mas prolonga-se em um processo contínuo de reemprego e retroalimentação informacional, no qual cada inferência ou reuso reforça o desequilíbrio entre o titular e os controladores de dados (Costa *et al*., 2024, p. 20-26).

Ao contrário do dano clássico, que permite identificação de vítima e agressor, o dano

informacional se manifesta por difusão. O dado pessoal, uma vez integrado ao modelo, não pode ser "desaprendido": torna-se parte da própria arquitetura cognitiva do sistema, reproduzindo-se indefinidamente (Cooper *et al.,* 2024, p.2). A compensação monetária, nesse contexto, converte-se em gesto simbólico que não reverte o prejuízo nem previne sua perpetuação. A lógica *ex post* é assim mantida como ritual de legitimação institucional: afirma-se o dever de indenizar, mas o efeito reparador é meramente declaratório (Antunes, 2018, p. 176).

Quando a lesão decorre de práticas automatizadas que atingem milhares de titulares simultaneamente, a pretensão de recomposição individual perde sentido prático. A própria estrutura do dano - produzido por repetições em larga escala e sustentado por fluxos contínuos de informação - impede que se delimitem agentes e vítimas com precisão. O que era pra ser um conflito entre partes identificáveis torna-se um prejuízo que se difunde no ambiente digital e recai sobre a coletividade de titulares afetados (Faleiros Júnior, 2023).

Na maioria das vezes, o titular sequer tem ciência de que foi lesado, pois o dano informacional ocorre em camadas técnicas invisíveis, sem qualquer sinal perceptível de violação (Soares; Ehrhardt Júnior, 2025, p. 11-13). Como visto na seção 3.3, o ordenamento já aciona técnicas de reequilíbrio probatório, mas tais instrumentos revelam-se insuficientes diante da opacidade e da escala do treinamento de IA generativa. Assim, mesmo quando o titular tem alguma consciência da lesão, fica privado de meios técnicos de rastreio, e vê-se reduzido à impotência: ele conhece a existência do dano, mas não dispõe de elementos para prová-lo ou quantificá-lo. A tutela civil, que originalmente pressupõe visibilidade e individualização, colapsa sob a lógica das massas de dados (Faleiros Júnior, 2023).

Nesse cenário, a consequência é a paralisia do titular: a promessa de acesso à Justiça - diga-se, condição de efetividade dos direitos fundamentais - transforma-se em um exercício retórico incapaz de romper o bloqueio informacional imposto pela própria tecnologia (Kreimer, 2016, p. 20).

Quando o acesso à Justiça depende de prova técnica que o titular não tem condições de produzir, a própria exigência de demonstração individual do dano converte-se em instrumento de negação de tutela (Ruzzi; Marchetto, 2024, p. 16). A jurisprudência norte-americana sobre incidentes de segurança envolvendo dados pessoais ilustra essa disfunção: por anos, tribunais negaram legitimidade ativa a vítimas de coleta ilícita de dados sob o argumento de inexistir

"dano concreto"[16] até a ocorrência de dano econômico tangível, como fraudes financeiras ou prejuízo patrimonial direto (Kreimer, 2016, p. 18-21). Tal raciocínio desconsidera a natureza própria do ilícito informacional, cuja lesão se consuma com a aquisição indevida de dados, independentemente de seu uso ulterior - tema já preparado no tópico 3.3 e que será retomado em 4.3.2 sob a chave do dano *in re ipsa*.

Esse quadro evidencia uma assimetria de legitimidade: o sistema jurídico permanece atrelado a parâmetros de tangibilidade e individualização que não se aplicam à materialidade técnica dos fluxos informacionais. O resultado é o esvaziamento do próprio conceito de "lesado", que se dilui na coletividade de titulares expostos a riscos homogêneos e invisíveis.

Além da dificuldade probatória e da dispersão das vítimas, há uma limitação intrínseca de escala. Mesmo quando admitida a reparação por danos coletivos, o *quantum* econômico da indenização dificilmente traduz a magnitude do desequilíbrio informacional produzido. A conversão do dano em cifras cria a aparência de justiça restaurada, mas não restabelece o controle do titular sobre seus dados nem impede novas violações. A indenização monetária, assim, cumpre função catártica e simbólica, legitimando o sistema mais do que reparando o ofendido. Trata-se de um rito declaratório de compensação: a sentença afirma o direito, mas não o reconstitui (Antunes, 2018, p. 176).

O deslocamento progressivo do problema - do indivíduo para a coletividade, do evento para o processo, do ressarcimento para a recomposição estrutural - evidencia a exaustão da racionalidade reparatória como núcleo do sistema de responsabilidade civil em matéria de dados. A tutela individual *ex post* já não basta para conter violações reiteradas e "invisíveis". A economia dos dados exige mecanismos capazes de operar antes do dano e além da indenização, aptos a enfrentar o caráter difuso e contínuo das infrações informacionais. Essa reconfiguração, que implica a passagem de uma tutela centrada na recomposição individual para uma lógica de proteção coletiva e de governança dos fluxos de dados, constitui o objeto da próxima seção.

[16] Em tradução livre, "dano concreto" corresponde ao termo *injury in fact*, utilizado na jurisprudência norte-americana para designar o requisito de lesão real e comprovável como condição para a legitimidade ativa. No contexto citado, refere-se à exigência mencionada por Kreimer (2016) de um prejuízo tangível (como perdas econômicas) para reconhecimento judicial da lesão decorrente de violações informacionais.

4.2 A FUNÇÃO DA RESPONSABILIDADE CIVIL NA SOCIEDADE DA INFORMAÇÃO: PROTEÇÃO COLETIVA DOS FLUXOS DE DADOS E GOVERNANÇA

A superação da centralidade reparatória - já revelada insuficiente para lidar com danos informacionais massivos e difusos - impõe o deslocamento da responsabilidade civil do plano estritamente corretivo para uma racionalidade de governança dos fluxos informacionais. A questão central deixa de ser apenas como reparar e passa a ser como reconfigurar práticas de tratamento que produzem risco difuso e danos estruturalmente irreversíveis, incorporando a *privacy by design* à própria arquitetura tecnológica (Hildebrandt, 2020, p. 274).

Nesse horizonte, a função da responsabilidade civil deve adquirir contornos regulatórios: mais do que recompor situações pretéritas, deve operar como vetor de conformidade tecnológica ao longo de todo o ciclo de vida do dado, induzindo padrões verificáveis de cuidado, transparência e prestação de contas. A lógica da recomposição bilateral, pensada para eventos individualizados (Hironaka, 2001, p. 5-7), não considera fluxos informacionais que transcendem a causalidade "pessoa-a-pessoa", exigindo remédios coletivos e institucionais que internalizem o risco do tratamento e reorganizem as estruturas em que o ilícito se produz.

Em lugar de perdas pontuais e mensuráveis, há desorganização persistente de ambientes informacionais e erosão de autonomias, cuja recomposição monetária não restaura confiança pública, integridade de redes nem autodeterminação dos titulares. Nessa moldura, o eixo da tutela desloca-se para obrigações positivas de governança e mitigação de risco que incidem sobre coleta, curadoria de *datasets*, treinamento e disponibilização de modelos de treinamento de IA generativa. (Doneda *et al*., 2018, p. 10-11). A transformação do dano em estado coletivo de vulnerabilidade impõe respostas processuais e materiais que não se esgotam na soma de pretensões individuais.

O paradigma é ilustrado pela ação civil pública proposta pelo Ministério Público Federal e pelo Idec em face do *WhatsApp* (Proc. nº 5018090-42.2024.4.03.6100, que tramita no Tribunal Regional Federal da 3ª Região), na qual o ilícito se consuma pela coação consentimental em massa, degradando privacidade e autonomia de milhões de usuários: o pedido de valor expressivo a reverter ao Fundo de Defesa de Direitos Difusos evidencia que a resposta adequada é difusa, não atomizada.

Dessa constatação decorre a centralidade da tutela coletiva como remédio estrutural. A

tipologia dos direitos transindividuais[17] oferece o arranjo para escolhas processuais que respeitam a indivisibilidade dos efeitos e a (in)existência de relação jurídica base entre titulares e agentes de tratamento. Nos usos indevidos de dados por sistemas de IA generativa, a ausência de vínculo jurídico prévio e a natureza unitária do ambiente informacional aproximam a tutela do regime dos direitos difusos, sem excluir hipóteses de agregação por homogeneidade quando houver pretensões divisíveis decorrentes de origem comum (Roque, 2019, p. 4-9).

Nesse contexto, a ação civil pública[18] consolida-se como o principal instrumento processual apto a dar efetividade a essa tutela difusa. A legitimidade ativa, por sua vez, requer arranjo institucional proativo: Ministério Público, Defensoria Pública, associações civis e a própria ANPD devem articular ações estruturais com planos de cumprimento faseados e monitoramento permanente, sob pena de perpetuar a assimetria entre titulares fragmentados e plataformas globais. Trata-se de migrar do litígio atomizado para o contencioso estrutural, em que a sentença deixa de ser um ponto final para converter-se em instrumento de governança judicial e administrativa (Roque, 2019, p. 11-13).

Essa perspectiva processual, contudo, precisa ser acompanhada por uma reordenação substancial dos próprios fundamentos da responsabilidade civil. Se a tutela coletiva indica quem deve agir e como agir diante de violações massivas, cabe agora compreender com base em que parâmetros essa atuação deve se estruturar. Em outras palavras, a efetividade do contencioso estrutural depende de uma reformulação dos deveres de cuidado e de imputação aplicáveis às atividades de tratamento, de modo a internalizar a governança e a precaução como elementos constitutivos do próprio dever jurídico

Tal continuidade histórica explica-se pela própria evolução da responsabilidade civil: assim como a industrialização demandou a objetivação da culpa (Hironaka, 2001, p. 2), a datificação generalizada demanda a objetivação do cuidado informacional mediante critérios estruturais

[17]Art. 81. “A defesa dos interesses e direitos dos consumidores e das vítimas poderá ser exercida em juízo individualmente, ou a título coletivo. Parágrafo único. A defesa coletiva será exercida quando se tratar de: I - interesses ou direitos difusos, assim entendidos, para efeitos deste código, os transindividuais, de natureza indivisível, de que sejam titulares pessoas indeterminadas e ligadas por circunstâncias de fato; II - interesses ou direitos coletivos, assim entendidos, para efeitos deste código, os transindividuais, de natureza indivisível de que seja titular grupo, categoria ou classe de pessoas ligadas entre si ou com a parte contrária por uma relação jurídica base; III - interesses ou direitos individuais homogêneos, assim entendidos os decorrentes de origem comum.”. BRASIL, Lei nº 8.078, de 11 de setembro de 1990

[18] BRASIL. Lei nº 7.347, de 24 de julho de 1985. Disciplina a ação civil pública de responsabilidade por danos causados ao meio ambiente, ao consumidor, a bens e direitos de valor artístico, estético, histórico, turístico e paisagístico e dá outras providências. Art. 1º, incisos II e IV.

de imputação e parâmetros de necessidade - compreendidos não como clausulados retóricos, mas como deveres operacionais integrados ao desenho técnico dos sistemas (Bioni, Rielli e Kitayama, 2021, p. 33-34).

Nessa chave, o princípio da necessidade atua em dupla dimensão: (i) em sentido estrito, vedando coleta e uso além do estritamente indispensável ao fim declarado; e (ii) em sentido *lato*, exigindo medidas de mitigação de impacto sobre direitos e liberdades, sob pena de ilicitude estrutural do fluxo de dados (Bioni; Rielli; Kitayama, 2021, p. 33-34).

A *accountability* converte-se em pilar de remediação e prevenção. A obrigação de registro das atividades de tratamento (com inventário de fluxos, finalidades, bases legais, categorias e prazos de retenção) funciona como "contabilidade de dados" que viabiliza auditoria regulatória e judicial, permitindo a detecção de desvios de origem e uso antes da consolidação do dano (Bioni, 2021, p. 268-271). A diretriz do art. 37 da LGPD incentiva modelos organizacionais que documentam proveniência, segurança e descarte, habilitando correções tempestivas (Brasil, 2018).

Por outro lado, o diagnóstico macroestrutural de colonização informacional reforça que a resposta não pode ficar cativa da lógica privada de recomposição: o capitalismo de vigilância opera pela captura da experiência e pelo deslocamento de arenas decisórias para fora do escrutínio democrático, de modo que remédios puramente individuais permanecem catárticos e insuficientes (Zuboff, 2021, p. 238-240). Assim, a responsabilidade civil deve operar como mecanismo de correção institucional e redistribuição de poder informacional, articulando-se com a atuação regulatória e o controle judicial.

Por fim, a coordenação entre tutela civil e sanções administrativas é condição de efetividade estrutural: medidas civis que reorganizam fluxos operacionais de tratamento de dados pessoais e impõem governança devem dialogar com programas sancionatórios e de monitoramento regulatório, evitando sobreposições ineficientes e maximizando efeitos preventivos (Capanema, 2020, p. 2). A responsabilidade civil não abdica da compensação quando cabível; ela abandona sua hegemonia para integrar um arranjo multirremedial e prospectivo (Farias; Rosenvald; Netto, 2025, p. 136), orientado à estabilidade dos ecossistemas informacionais e à proteção coletiva da autodeterminação dos titulares.

A centralidade da tutela coletiva, a qual demanda remédios estruturais e a redefinição da responsabilidade civil para um vetor de governança e conformidade tecnológica, estabelece o

quadro operacional da nova função do instituto. Contudo, para que o contencioso estrutural e os deveres de *accountability* e precaução sejam plenamente efetivos, é indispensável que a atuação judicial e regulatória se fundamente em critérios objetivos de imputação e em fundamentos normativos consistentes. É nesse horizonte que o próximo tópico analisará, por meio da hermenêutica sistemático-teleológica, os elementos interpretativos necessários à consolidação dessa racionalidade emergente da responsabilidade civil.

4.3 A REORIENTAÇÃO DA RESPONSABILIDADE CIVIL: HERMENÊUTICA SISTEMÁTICO-TEOLÓGICA COMO FERRAMENTA INTERPRETATIVA

A opção pela hermenêutica sistemático-teleológica, neste trabalho, não pretende estabelecer um eixo metodológico central, mas funcionar como ferramenta interpretativa auxiliar, capaz de resguardar a coerência do sistema jurídico frente às transformações impostas pela inteligência artificial e pelo uso indevido de dados pessoais.

O contexto informacional contemporâneo impõe ao Direito um problema que transcende a mera subsunção de fatos à norma: exige a reconstrução de significados à luz das finalidades constitucionais e legais de proteção à pessoa humana, à privacidade e à autodeterminação informativa (Aguiar Júnior, 1989, p. 10-14). Nesse sentido, a hermenêutica aqui adotada opera como suporte argumentativo, permitindo que as conclusões desenvolvidas ao longo do capítulo se ancorem nos fins do ordenamento, sem substituir outros elementos dogmáticos relevantes.

A hermenêutica sistemático-teleológica permite interpretar o texto normativo em diálogo com os valores que o animam, reconstituindo a *ratio legis* a partir do bem jurídico tutelado. No caso da proteção de dados, essa *ratio* não se esgota na proteção patrimonial ou compensatória, mas orienta-se à salvaguarda da liberdade informacional e da igualdade substancial entre titulares e agentes de tratamento. A interpretação teleológica, como lembra Aguiar Júnior, supera a lógica formal para alcançar o fim social da norma e os efeitos concretos da decisão, introduzindo um elemento material e valorativo que impede a neutralização do sentido ético do Direito (Aguiar Júnior, 1989, p. 12-13). Essa abordagem revela-se especialmente adequada quando o dano ultrapassa a esfera individual e assume feição difusa, como ocorre nos fluxos automatizados de dados empregados no treinamento de sistemas de IA generativa.

Na formulação de Mozetic, o intérprete atua como mediador entre a generalidade da norma e a singularidade dos fatos, realizando uma fusão de horizontes entre o texto jurídico e a realidade tecnológica que o desafia (Mozetic, 2016, p. 225-233). Essa fusão é indispensável em um campo em que a norma foi concebida sob paradigmas analógicos, mas os fatos pertencem a um universo digital imprevisível. Assim, compreender o direito à proteção de dados demanda reconhecer a distância entre a abstração legal e a concretude do fenômeno algorítmico, produzindo um sentido que não se reduz à literalidade, mas emerge da relação entre norma, técnica e valores fundamentais.

A hermenêutica sistemático-teleológica também responde ao desafio apontado por Mozetic de criar novos instrumentos analíticos aptos a lidar com as realidades digitais, nas quais as categorias clássicas se tornam insuficientes para apreender fenômenos descentralizados e dinâmicos. A transição do mundo analógico ao digital exige do jurista não apenas adaptação, mas verdadeira reinvenção conceitual, pois o Direito passa a interagir com realidades técnicas que alteram as formas de poder, comunicação e risco. Ao assumir essa postura, o intérprete não abandona a legalidade, mas a compreende como campo de sentido aberto, cuja efetividade depende da contínua atualização de seus propósitos (Mozetic, 2016, p. 129-132).

A aplicação dessa hermenêutica ao tema da responsabilidade civil pelo uso não autorizado de dados pessoais visa restabelecer a integridade do sistema jurídico frente à defasagem entre norma e realidade tecnológica. A literalidade, isolada, não é capaz de proteger o titular em um ambiente governado por algoritmos opacos e decisões automatizadas. Como assinalam Ruzzi e Marchetto (2024, p. 12-17), a ausência de regulamentação efetiva e a assimetria entre grandes conglomerados e indivíduos transformam a interpretação judicial em um fator decisivo para a concretização dos direitos fundamentais à privacidade e à proteção de dados, muitas vezes negados por leituras restritivas.

Ademais, impõe-se reconhecer que a própria legislação protetiva de dados - no Brasil representada pela LGPD, sancionada em 2018 e entrando em vigor em setembro de 2020 - não estava preparada para absorver o posterior e célere "*boom*" da inteligência artificial generativa. O diploma legislativo adveio de uma realidade tecnológica ainda dominada por modelos tradicionais de tratamento de dados[19] e não contemplava a real dimensão dos

[19] No presente contexto, a expressão modelos tradicionais de tratamento de dados foi empregada para designar práticas anteriores ao advento da inteligência artificial generativa, centradas em fluxos lineares e finalidades delimitadas, como o uso de *cookies* para personalização de anúncios, o armazenamento de cadastros em bancos de dados corporativos ou o monitoramento de interações em plataformas digitais convencionais.

algoritmos generativos, dos fluxos automatizados de treino e das interconexões massivas de dados que se consolidaram em meados de 2023 (Chui *et al.*, 2023). Dessa forma, a defasagem entre o escopo normativo da LGPD e o salto qualitativo e quantitativo da IA torna-se patente, exigindo que a hermenêutica funcione como ponte entre norma e realidade, preservando a coerência do sistema.

A reorientação proposta se justifica porque a própria função da responsabilidade civil deixou de ser apenas corretiva para assumir dimensões preventivas e distributivas (Rosenvald; Farias; Netto, 2025, p. 136-137). Essa leitura não é arbitrária: ela resulta da exigência de coerência interna do ordenamento, em que as normas devem ser interpretadas de modo a realizar seus fins e harmonizar-se com os valores superiores da ordem jurídica - guiada pelos princípios da dignidade humana e da proporcionalidade.

A hermenêutica sistemático-teleológica, por fim, é a que melhor preserva o papel humanista do Direito diante da técnica: ela recusa o mecanismo interpretativo que reduz o intérprete a mero executor de comandos (Mozetic, 2016, p. 225-226). Nessa medida, a escolha metodológica feita neste trabalho não se resume a uma opção acadêmica, mas representa um recurso interpretativo que permite manter o Direito responsivo aos desafios éticos da era digital, sem pretensão de substituir a dogmática tradicional.

Em suma, faz-se necessário um horizonte teórico que reconstrua a responsabilidade civil no contexto informacional contemporâneo. Ao privilegiar a finalidade sobre a forma e a coerência sobre o isolamento normativo, esse método permite reinterpretar o dever de reparar como dever de cuidado e de governança, compatibilizando o ordenamento jurídico com os riscos emergentes da inteligência artificial. É sob esse prisma que se compreenderão, nas seções seguintes, as bases de imputação, a autonomia do dano informacional e a transformação da responsabilidade civil em instrumento de preservação da liberdade e da integridade dos fluxos de dados pessoais.

4.3.1 Da culpa à teoria do risco-atividade: a vulnerabilidade informacional como elemento de imputação objetiva

Conforme se afirmou reiteradamente no capítulo 3, a responsabilização civil pelo uso indevido de dados pessoais em sistemas de inteligência artificial generativa se ancora em critérios normativos de imputação objetiva. Todavia, por tratar-se de tema discutível na doutrina - que ainda oscila entre leituras subjetivas e objetivas do regime da LGPD - não é metodologicamente admissível o mero lançamento de conclusões, é preciso expor o percurso argumentativo que conduz a esse resultado.

Trata-se, aqui, não de reafirmar o que já foi demonstrado, mas de demonstrar que tal conclusão é o resultado mais coerente quando se interpretam os dispositivos da CRFB/88, do CC/02, do CDC e da LGPD à luz de seus fins teleológicos de tutela da pessoa e da dignidade informacional. Em suma, a presente subsecção não se trata de uma opção puramente apriorística, mas da conclusão necessária de um itinerário interpretativo. Nesse sentido, a hermenêutica sistemático-teleológica opera aqui como ferramenta interpretativa que apenas reforça conclusões já delineadas dogmaticamente, e não como fundamento exclusivo da tese defendida.

A interpretação das normas de responsabilidade civil em matéria de dados pessoais deve ocorrer em diálogo com os valores constitucionais que lhes conferem unidade e sentido. A CRFB/88, ao consagrar a dignidade da pessoa humana como fundamento da República (Art. 1º, III) e ao reconhecer a proteção de dados pessoais no meio digital como direito fundamental autônomo (Art. 5º, LXXIX), impõe ao intérprete que privilegie a tutela da pessoa e não o formalismo da culpa. O direito à proteção de dados é, antes de tudo, uma garantia de liberdade e de igualdade informacional, cuja violação não se esgota na noção clássica de ilícito culposo, mas representa uma ofensa direta ao núcleo de um direito fundamental (Bioni; Dias, 2020, p. 3-4).

A leitura teleológica da LGPD reforça essa conclusão. Embora a redação dos Arts. 42 a 45 não utilize expressamente a expressão “independentemente de culpa”, a própria estrutura da lei evidencia uma racionalidade objetiva fundada na assunção de risco pela atividade (Meo, 2022, p. 174-176). Conduz-se inevitavelmente à objetivação da responsabilidade, bastando a desconformidade material entre o tratamento realizado e o padrão normativo imposto.

O modelo subjetivo, defendido por parte da doutrina (Florence, 2021, p. 1-9; Côrrea; Cho, 2021), mostra-se incompatível com a realidade da inteligência artificial e com a própria teleologia da legislação protetiva.

Exigir prova de culpa individual em cadeias algorítmicas fragmentadas seria negar a efetividade do direito à reparação e perpetuar a assimetria estrutural entre titulares e grandes agentes de tratamento. A lógica da IA generativa (caracterizada por fluxos massivos e opacidade técnica) elimina a possibilidade de individualizar condutas e aferir dolo ou negligência em termos clássicos. Nesses contextos, a responsabilidade subjetiva converte-se em obstáculo à tutela, pois a "prova impossível" neutraliza o próprio direito (Bioni; Dias, 2020, p. 21; Capanema, 2020, p. 4). Desloca-se o eixo interpretativo do comportamento do agente para a estrutura de risco da atividade.

Sob esse prisma, a teoria do risco-atividade, positivada no Art. 927, parágrafo único, do CC/02, emerge como o instrumento mais coerente com a CRFB/88 e com a LGPD. A responsabilidade não deriva da prova de um erro humano, mas da constatação de que o agente de tratamento (controlador ou operador) escolheu desenvolver atividade potencialmente lesiva aos direitos fundamentais, devendo suportar os riscos correlatos. A decisão de tratar dados pessoais, ainda que lícita em tese, implica um dever de governança e de conformidade permanente, cuja violação enseja imputação direta independentemente de culpa (Meo, 2022, p. 174). Trata-se de um dever objetivo de segurança e licitude, não de um juízo moral sobre o comportamento do agente.

Posto isso, relembrando o já constatado no Capítulo 3, é importante assinalar que as plataformas de IA generativa, ao ofertarem serviços digitais ao público, enquadram-se na categoria de fornecedor e, portanto, submetem-se também ao regime do CDC (Almada, 2019, p. 5; Meo, 2022, p. 73-74; 164).

Assim, a articulação entre o CC/02 e o CDC reforça a necessidade de tal objetivação. O CDC (Arts. 12 e 14) consagra a responsabilidade objetiva do fornecedor por defeitos no produto ou serviço, inclusive quando o defeito decorre de falha na segurança ou na informação (Tepedino; Silva, 2019, p. 25). A violação dos deveres de proteção - como o uso não autorizado de dados em treinamentos algorítmicos - constitui defeito de serviço e enseja responsabilização objetiva de todos os integrantes da cadeia, conforme o modelo solidário de imputação já consolidado no direito do consumo.

Nessa conjuntura, a vulnerabilidade informacional dos titulares - i.e., sua incapacidade de conhecer ou controlar o destino de seus dados (Bonna; Canizo; Calzavara, 2022, p. 12) - constitui, portanto, o elemento de imputação objetiva que justifica a incidência dessa cláusula geral e reclama uma inversão interpretativa semelhante à que ocorreu no direito do consumo[20] e no direito ambiental[21]. Em ambos os casos, a proteção de bens jurídicos coletivos e a assimetria entre as partes fundamentaram a objetivação da responsabilidade.

Por conseguinte, a passagem da culpa à teoria do risco-atividade não representa um rompimento arbitrário com a tradição civilista, mas a evolução coerente de um sistema constitucionalmente orientado à tutela da pessoa frente ao poder tecnológico. A responsabilidade civil, lida à luz da finalidade do ordenamento, transforma-se em instrumento de precaução e redistribuição de riscos, em harmonia com a CRFB/88 e com os microssistemas setoriais de proteção. O modelo subjetivo, ao contrário, pertence a uma racionalidade liberal e individualista que não mais responde à complexidade das infraestruturas digitais contemporâneas.

Desse modo, a hermenêutica sistemático-teleológica não cria a conclusão pela responsabilidade objetiva, mas apenas explicita sua coerência com os princípios estruturantes do sistema, legitimando a adoção da responsabilidade objetiva por risco da atividade como expressão da função social e preventiva da responsabilidade civil na sociedade da informação (Soares, 2009, p. 75).

O uso não autorizado de dados pessoais no treinamento de IA generativa, ao violar o direito fundamental à proteção de dados e ao expor o titular a riscos permanentes e não controláveis, enquadra-se perfeitamente no tipo de atividade que o Art. 927, parágrafo único, e a LGPD buscaram submeter à imputação direta e objetiva. O intérprete, portanto, não cria uma exceção: apenas realiza a finalidade constitucional do sistema jurídico, assegurando que a vulnerabilidade informacional seja reconhecida como elemento central da imputação e como

---

[20] Art. 12. “O fabricante, o produtor, o construtor, nacional ou estrangeiro, e o importador respondem, independentemente da existência de culpa, pela reparação dos danos causados aos consumidores por defeitos decorrentes de projeto, fabricação, construção, montagem, fórmulas, manipulação, apresentação ou acondicionamento de seus produtos, bem como por informações insuficientes ou inadequadas sobre sua utilização e riscos.”; Art. 14. “O fornecedor de serviços responde, independentemente da existência de culpa, pela reparação dos danos causados aos consumidores por defeitos relativos à prestação dos serviços, bem como por informações insuficientes ou inadequadas sobre sua fruição e riscos.” BRASIL. Lei nº 8.078, de 11 de setembro de 1990.

[21] Art. 14, § 1º - “§ 1º - Sem obstar a aplicação das penalidades previstas neste artigo, é o poluidor obrigado, independentemente da existência de culpa, a indenizar ou reparar os danos causados ao meio ambiente e a terceiros, afetados por sua atividade. [...].” BRASIL. Lei nº 6.938, de 31 de agosto de 1981.

fundamento da nova racionalidade da responsabilidade civil.

A consolidação da responsabilidade objetiva pelo risco da atividade conduz inevitavelmente a uma redefinição do próprio conceito de dano. Se a imputação já não depende da culpa, tampouco pode o dano restringir-se à comprovação de prejuízos materiais ou psíquicos concretos. No contexto da proteção de dados, a lesão manifesta-se no próprio ato ilícito de tratamento, que vulnera de imediato a esfera jurídica do titular e perpetua efeitos difusos e contínuos. É nesse ponto que se insere a análise do dano *in re ipsa* e do ilícito estrutural como fundamentos autônomos da responsabilização civil.

4.3.2 O dano *in re ipsa* e o ilícito estrutural: a violação ao direito fundamental à proteção de dados como fato gerador autônomo da responsabilidade

Nos últimos anos, o STJ vinha adotando uma postura restritiva quanto ao reconhecimento do dano moral presumido em casos de violação a dados pessoais. A jurisprudência dominante exigia prova concreta de abalo moral, mesmo diante de vazamentos ou exposições indevidas de informações sensíveis, tratando o dano como evento empírico e não como decorrência direta da ofensa ao direito fundamental à proteção de dados.

Exemplo emblemático dessa orientação é o AREsp n. 2.130.619/SP, julgado em 7/3/2023 pela Segunda Turma, no qual o STJ (2023) afirmou que "o vazamento de dados pessoais, a despeito de se tratar de falha indesejável no tratamento de dados de pessoa natural, não tem o condão, por si só, de gerar dano moral indenizável", exigindo que o titular comprovasse efetivo prejuízo decorrente da exposição.

Tal entendimento refletia uma concepção ainda ancorada no paradigma reparatório tradicional, no qual a violação a direitos da personalidade dependia de demonstração de consequências materiais ou psicológicas individualizadas (Rodrigues Júnior, 2013, p. 3-4).

Cumpre ainda, observar que, até o presente momento, a jurisprudência brasileira apresenta apenas um precedente específico sobre o uso de dados pessoais para treinamento de modelos de IA generativa: o Agravo de Instrumento n.º 1034362-91.2024.8.11.0000, julgado pelo TJ-MT em 2 de maio de 2025. Nesse caso, a corte indeferiu pedido de tutela para suspender a utilização de dados de usuários da plataforma X (antigo *Twitter*) pelo modelo "*Grok*", entendendo lícito o tratamento com base no "legítimo interesse do controlador", por haver

“mecanismo acessível de oposição”. Trata-se, até aqui, da única decisão nacional diretamente atinente ao tema do presente trabalho, o que confere relevância paradigmática à sua análise crítica.

Todavia, a *ratio decidendi* revela o persistente déficit hermenêutico que este estudo busca superar. Ao admitir a utilização de dados sem consentimento específico, bastando a menção genérica em política de privacidade e a possibilidade formal de *opt-out*, o tribunal manteve-se preso à leitura literal do Art. 7º da LGPD, ignorando a natureza estrutural do dano informacional e a impossibilidade prática de oposição efetiva em ecossistemas de coleta massiva.

A decisão reduz a autodeterminação informativa a mera faculdade procedimental, convertendo a dignidade em clique. Tal formalismo colide com a teleologia da LGPD e com o princípio da proteção de dados delineado pela CRFB/88, além de esvaziar a função preventiva da responsabilidade civil em ambientes algorítmicos. Em síntese, o acórdão mato-grossense espelha a crise hermenêutica que perpassa o sistema: aplica categorias analógicas a fenômenos digitais, preservando a aparência de legalidade enquanto perpetua a vulnerabilidade estrutural do titular.

Essa resistência, contudo, começou a ser superada: a recente consolidação jurisprudencial acerca do tratamento indevido de dados pessoais representa um divisor de águas na dogmática da responsabilidade civil brasileira. Em agosto de 2025, no REsp n. 2.201.694/SP, a Terceira Turma do STJ (voto vencedor da Ministra Nancy Andrighi) afirmou que, no âmbito dos bancos de dados regidos pela Lei nº 12.414/2011 (Cadastro Positivo), a disponibilização indevida a terceiros consulentes de informações cadastrais e de adimplemento - cuja circulação a lei restringe ao compartilhamento entre bancos de dados - caracteriza dano moral presumido (*in re ipsa*) e impõe responsabilidade objetiva ao gestor do banco de dados (itens 5–6, 10 e 11 do voto vencedor). A *ratio decidendi* desloca o eixo da responsabilização da prova empírica do dano para a constatação do próprio ilícito, reconhecendo que a violação ao direito fundamental à proteção de dados gera abalo presumido, em virtude da sensação de insegurança e vulnerabilidade imposta ao titular (STJ, 2025).

Essa virada interpretativa não é mera flexibilização probatória. Trata-se de reconhecer a autonomia do bem jurídico “proteção de dados pessoais”, inscrito expressamente como direito fundamental pela EC 115/2022, cuja violação produz, por si, lesão à esfera existencial do indivíduo (Bahia *et al.*, 2024, p. 28–31).

A proteção de dados não se confunde com a privacidade. Enquanto a privacidade tutela a esfera íntima e estática, voltada a limitar intervenções externas e operando como um direito negativo; a proteção de dados assume estrutura diversa: centrada na regulação dos fluxos informacionais que condicionam a autonomia do sujeito, ela impõe limites às formas de tratamento de dados que podem gerar assimetrias de poder e afetar a liberdade decisória (Rodotà, 2008, p. 60 *apud* Bahia *et al.*, 2024, p. 29). Assim, a manipulação ou exposição indevida de dados, ainda sem prova de prejuízo material, viola um direito fundamental de estrutura objetiva, cujo resguardo interessa à coletividade por constituir elemento da cidadania digital.

O reconhecimento da autonomia do ilícito informacional impõe também superar a visão tradicional do dano como evento empiricamente demonstrável. No paradigma informacional, o dano é difuso e cumulativo, resultante de práticas sistêmicas de coleta e reuso de dados pessoais. Não há um fato isolado, mas um ilícito estrutural que se perpetua na arquitetura técnica dos tratamentos, escapando ao controle individual do titular. O desvio de finalidade, a ausência de base legal adequada e o tratamento opaco constituem ilícitos em si, porque corroem a autodeterminação informativa (Hironaka, 2023, p. 16–20).

É exatamente aqui que a hermenêutica sistemático-teleológica auxilia estender a orientação do STJ ao treinamento não autorizado de sistemas de IA generativa. Embora o precedente tenha sido proferido no setor de proteção ao crédito, seus fundamentos teleológicos - tutela do direito fundamental à proteção de dados; redução da assimetria informacional; prevenção da sensação de insegurança; e responsabilização objetiva do agente que disponibiliza/compartilha dados fora dos limites legais - são transponíveis a contextos em que (a) há coleta massiva (v.g. *web scraping* e *datasets* públicos), (b) uso subsequente para fins distintos dos necessários/consentidos, e (c) indisponibilidade prática de controle pelo titular.

Em tais cenários, a lesão é intrínseca ao ato de treinar sem base legal, porque atinge o núcleo da autodeterminação e instala a mesma sensação objetiva de insegurança que o STJ reconheceu como suficiente para o *in re ipsa* no Cadastro Positivo.

A noção de dano *in re ipsa*, historicamente restrita, encontra agora fundamento constitucional e funcional na tutela dos dados pessoais. Como sustenta Hironaka (2023, p. 2–3), a evolução da responsabilidade civil migra de um modelo sancionatório para um modelo de responsabilidade pressuposta, no qual o dever de indenizar decorre da ocorrência do ilícito quando a atividade implica risco ou vulneração de direitos fundamentais, independentemente

de culpa ou de prova do dano. Trata-se de causalidade normativa: presume-se o dano porque o ordenamento reconhece a gravidade do bem tutelado e o desequilíbrio estrutural entre titular e agente de tratamento.

Aplicada à IA generativa, a categoria do ilícito estrutural revela que o tratamento massivo sem base legal viola imediatamente o direito fundamental, ainda que não haja dano particularizado. A irreversibilidade prática dos efeitos do treinamento e a dificuldade de rastrear a contribuição individual exacerbam a sensação objetiva de insegurança, justificando a presunção do dano e a responsabilidade objetiva pelo simples fato de manter condições técnicas e organizacionais que fragilizam a autodeterminação informativa (Li *et al.*, 2025, p. 4-5; Gondim, 2021, p. 9-12).

Nessa perspectiva, a responsabilidade civil assume sua função regulatória - exposta no 4.2 -: não apenas compensa, mas reordena fluxos e impõe governança, alinhando-se à teoria do risco e à função social do instituto (Hironaka, 2023, p. 18–20). Se o ilícito estrutural é menos "um ato" e mais "um modo de operar", a resposta deve transcender a indenização monetária e incidir sobre a estrutura que produz o dano.

Em síntese, o *in re ipsa* na proteção de dados é a materialização da função preventiva e normativa da responsabilidade civil. À luz da EC 115/2022, da LGPD e da orientação do STJ, não há espaço para exigir prova de prejuízo concreto quando a arquitetura da atividade viola o núcleo essencial do direito: o dever de reparar e de adequar decorrem do simples fato de o agente ter produzido ou mantido condições que minam seus direitos fundamentais.

Se o reconhecimento do *in re ipsa* e do ilícito estrutural desloca o foco da prova empírica do abalo para a própria violação ao direito fundamental, o passo seguinte é necessariamente qualificar que tipo de lesão emerge desse modo de operar: em ambientes de tratamento massivo e opaco, a ofensa não se esgota no desconforto subjetivo, mas incide sobre a autodeterminação informativa e, portanto, sobre dimensões existenciais do sujeito, exigindo repensar as categorias clássicas de dano para abarcar a perda da autonomia como resultado imediato do ilícito informacional.

4.3.3. A superação das categorias clássicas de dano: o dano existencial e a perda da autonomia informativa como resultado do ilícito estrutural

A transposição da responsabilidade civil para o domínio informacional desafia as categorias clássicas de dano, concebidas sob paradigmas patrimonialistas e individualistas. A lógica do "preço da dor" e a tipologia tradicional já reconhecida - dano material, moral, social e estético (Aguiar Júnior, 2012, p. 76-77 e 370) - mostram-se insuficientes para capturar ofensas que se irradiam estruturalmente e comprometem dimensões existenciais da pessoa.

O tratamento indevido de dados pessoais, sobretudo em contextos de automação e inteligência artificial generativa, produz lesões que não se reduzem a sentimentos de dor, vergonha ou desconforto: compromete-se o exercício da liberdade, o controle sobre a própria identidade e o projeto de vida informacional do indivíduo (Frota; Bião, 2017, p. 14-16). Nessa perspectiva, emerge a necessidade de reconhecimento de uma nova modalidade de lesão: o dano existencial informacional, que expressa a perda da autonomia informativa como violação da própria dignidade humana.

Embora o dano existencial tenha sido originariamente desenvolvido no âmbito do direito do trabalho (Buarque, 2019, p. 3-4), aqui se demonstra, com apoio na interpretação sistemática e finalística do ordenamento, que seu alcance pode ser ampliado para a responsabilidade civil sempre que a ofensa atingir o núcleo essencial da existência digna.

Como sustenta Buarque, a evolução civilizatória e a ampliação da consciência social geram a necessidade de reconhecer novas formas de lesão à pessoa, mesmo sem previsão legal expressa, a partir da cláusula geral de tutela da dignidade humana. A cláusula aberta dos direitos da personalidade autoriza o reconhecimento de novos direitos e de novas categorias de dano sempre que a integridade existencial da pessoa for afetada (Buarque, 2019, p. 12-16).

Assim, negar a reparabilidade da violação informacional com base na ausência de tipificação específica seria trair a própria teleologia da Constituição, que consagra a dignidade da pessoa humana como fundamento do ordenamento e a liberdade informacional como sua expressão contemporânea (Brasil, 1988).

A interpretação integradora do CC/02 conduz ao mesmo resultado. Dispositivos como os Arts. 949 a 954 já contêm, de modo implícito, o reconhecimento de danos que ultrapassam a esfera física ou moral, ao preverem a indenização por "qualquer outro prejuízo que o ofendido prove haver sofrido". O ordenamento brasileiro, portanto, já acolhe a lógica reparatória do dano existencial, ainda que sem nomeá-lo (Buarque, 2019, p. 16-19).

A violação à autodeterminação informativa produz um resultado análogo: o indivíduo deixa

de ser autor do próprio projeto de vida digital e passa a ser objeto de tratamento, classificação e predição por sistemas automatizados que o reduzem a um perfil. Essa perda da autodeterminação, embora intangível, é concreta e permanente, e caracteriza lesão à autonomia existencial, isto é, uma interferência heterônoma no desenvolvimento da personalidade e na liberdade de conduzir a própria vida (Hatoum; Colombo, 2022, p. 9-12)

A hermenêutica sistemático-teleológica justifica essa ampliação pela via do valor: se a dignidade humana constitui a finalidade última do sistema jurídico, toda lesão que a comprometa deve ser reconhecida como dano, ainda que o ordenamento não o tenha previamente categorizado (Maia; Neme; Calissi, 2024, p. 9).

A leitura conjugada das lições de Fux e de Rodotà demonstra que a autodeterminação informativa é derivação direta da dignidade, e, portanto, sua violação não é apenas um problema de privacidade, mas um atentado à própria condição de sujeito livre (Montenegro, 2021; Rodotà, 2008, p. 83, 96-97). Nas palavras de Fux, "a dignidade reclama autodeterminação; no limite da vida, da sobrevivência, não se tem autodeterminação e não se ostenta dignidade" (Montenegro, 2021). Assim, a perda da autonomia informativa deve ser compreendida como forma de dano existencial: ela impede o indivíduo de exercer escolhas fundamentais sobre si mesmo, sua identidade e suas relações sociais.

Quando o tratamento massivo de dados - realizado sem consentimento e com potencial discriminatório - compromete essa liberdade, há uma violação direta ao núcleo de existência da pessoa, análoga à frustração do projeto de vida (Buarque, 2019, p. 15). A informação, nesse cenário, integra o próprio ser, e sua captura indevida desestrutura o eixo de autorrealização que sustenta a dignidade. A lesão à autonomia informativa não é mero "sentir", mas um "não poder mais fazer": o indivíduo perde a possibilidade de agir segundo suas próprias escolhas e de gerir seus próprios dados pessoais no espaço digital (Maia; Neme e Calissi, 2024, p. 9).

A repersonalização do direito civil impôs ao intérprete a passagem do "ter" ao "ser", substituindo o paradigma patrimonial pela centralidade da pessoa. Essa mudança legitima o reconhecimento de novas categorias de dano voltadas à proteção da existência em sua concretude (Silva; Lelis, 2024, p. 7-8).

Assim, a perda da autodeterminação informativa, ao impedir o sujeito de definir o que revelar, ocultar e como se projetar socialmente, constitui um dano existencial próprio da era digital,

pois compromete o direito ao livre desenvolvimento da personalidade. A sua supressão representa, portanto, a negação da dignidade em sua dimensão mais elementar (Hatoum; Colombo, 2022, p. 12).

A objeção de que o dano existencial não teria previsão expressa fora do campo trabalhista é superada justamente pela leitura sistemática do ordenamento, que evidencia a unidade axiológica dos direitos da personalidade. O sistema jurídico não pode permanecer estático diante de novas formas de violação que decorrem da transformação tecnológica, sob o risco de esvaziar a função protetiva do princípio da dignidade (Maia; Neme; Calissi, 2024, p. 9; Gil; Rodrigues, 2023, p. 3-5).

O uso não autorizado de dados pessoais em processos de treinamento de modelos de IA generativa, não produz apenas perda patrimonial ou sofrimento psíquico, mas um esvaziamento da liberdade do titular: o sujeito deixa de governar o que dele se sabe e o que dele se decide. Assim como no caso das "pílulas de farinha" ou das fecundações heterólogas analisadas por Hatoum e Colombo (2022, p. 12-16), há aqui frustração profunda do direito de escolha, só que agora no plano da identidade informacional.

Se a dignidade humana se manifesta hoje também como dignidade informacional, e se a liberdade se expressa como autodeterminação informativa (Maia; Neme; Calissi, 2024, p. 14-16), então o ilícito informacional não apenas causa desconforto moral: ele fere o próprio fundamento ontológico do titular dos dados pessoais. O dano não está na dor, mas na impossibilidade de conduzir a própria narrativa existencial em um ambiente digital que o transforma em objeto de predição.

A análise empreendida ao longo da seção 4.3 evidencia que a hermenêutica sistemático-teleológica funcionou como ferramenta estruturante para demonstrar a compatibilidade dessas conclusões com o sistema jurídico (Vieira *et al.*, 2025, p. 7). Delineia-se um instituto que já não se esgota em reparar, mas em reordenar: a responsabilidade civil torna-se mecanismo de coerência normativa e de recomposição das condições de liberdade em ecossistemas digitais. Todavia, reconhecer essa mutação conceitual é apenas o primeiro passo. Resta construir, sobre esses fundamentos, um modelo operativo capaz de traduzir essa racionalidade em deveres concretos, medidas preventivas e instrumentos de governança. É a essa tarefa que se dedica a próxima seção, última do desenvolvimento do presente trabalho.

4.4 BASES PARA UMA (RE)CONSTRUÇÃO NORMATIVA DA RESPONSABILIZAÇÃO CIVIL PELO USO NÃO AUTORIZADO DE DADOS PESSOAIS NO TREINAMENTO DE IA GENERATIVA

O ponto de partida da reconstrução aqui proposta passa, necessariamente, pela constatação de que o ordenamento brasileiro ainda não dispõe de disciplina específica sobre responsabilidade civil aplicada aos usos de dados pessoais no treinamento de sistemas de inteligência artificial generativa.

Tal constatação, porém, tem um contorno institucional relevante: o dever de legislar, em verdade, é do Congresso Nacional. Embora o PL 2.338/2023 (já aprovado no Senado e em análise na Câmara mediante Comissão Especial), reconheça a centralidade da pessoa humana e os direitos à privacidade, à proteção de dados e à autodeterminação informativa, ele ainda não possui força normativa nem contém um regime detalhado de imputação civil aplicável ao tema (Brasil, 2023).

Esse cenário evidencia um descompasso institucional: a demanda social por tutela é imediata, mas a resposta legislativa é incerta quanto ao conteúdo e ao tempo. Mesmo o controle concentrado por omissão não se mostra, em regra, caminho suficiente para delinear minudências de imputação civil em ecossistemas técnicos complexos. A técnica de decisão em ADO tende a compelir providências legislativas, sem substituir o legislador no desenho detalhado de deveres e remédios (Dantas, 2018, p. 192; Brasil, 1999). Por isso, ainda que cabível para afirmar o dever de legislar, não é instrumento hábil para suprir, de imediato, assimetria temporal entre risco tecnológico e resposta normativa.

Esse cenário institucional, marcado simultaneamente por reconhecimento da urgência normativa e pela ausência de um regime jurídico completo, evidencia a necessidade de uma movimentação capaz de orientar a responsabilidade civil enquanto não sobrevier legislação específica. É justamente nesse ponto que se insere a (re)construção normativa proposta neste capítulo.

A (re)construção normativa da responsabilidade civil pelo uso não autorizado de dados pessoais no treinamento de sistemas de inteligência artificial generativa parte da premissa de que o Direito contém, em si, elementos para enfrentar os novos dilemas informacionais.

Ainda que os capítulos anteriores tenham evidenciado limites estruturais dos paradigmas tradicionais, marcados por categorias analógicas e lacunas regulatórias, essas insuficiências não autorizam um estado de inércia protetiva. Ao contrário, impõem ao intérprete o dever de mobilizar os instrumentos já existentes para responder às lesões contemporâneas.

Inspirada na teoria do reconhecimento e nas concepções procedimentais de legitimidade, essa opção metodológica busca restituir densidade ética à responsabilidade civil, convertendo princípios constitucionais em critérios operativos para lidar com novas materialidades de risco e dano algorítmico (Garcia, 2025, p. 200-203).

A reconstrução aqui proposta exige também uma atenção ao modo como o Direito formula seus próprios enunciados. Expressões como "autodeterminação informativa", "risco algorítmico", "dano existencial" ou mesmo "tratamento indevido" operam, inevitavelmente, como conceitos de contornos abertos, cuja concretização depende de critérios interpretativos e valorativos (Requião, 2009, p. 30). Essa abertura não é um defeito das categorias jurídicas, mas a condição que permite ao sistema responder a realidades tecnológicas em constante mutação. Por isso, além de reorganizar deveres e remédios, a reconstrução demanda atribuir sentido operativo a esses conceitos amplos, isto é, explicitar seu núcleo mínimo de significado e os fatores que orientam sua aplicação em casos concretos (Requião, 2009, p. 27-28, 37-39). Uma vez estabelecido esse núcleo - cuja formulação foi sendo construída gradualmente ao longo deste trabalho - e delimitadas as balizas de concreção, abre-se o espaço necessário para que a interpretação avance da abstração constitucional para o plano decisório.

A partir dessa definição mínima de conteúdo, torna-se possível conferir densidade linguístico-normativa às decisões em ambientes de intensa mutabilidade tecnológica. A clareza quanto ao núcleo e aos limites de cada conceito impede que sua abertura se converta em arbitrariedade e permite que o aplicador do Direito justifique, de modo público e controlável, por que determinados fluxos de tratamento de dados pessoais devem ser reordenados pela responsabilidade civil. Em vez de significados vagos ou intuitivos, o sistema passa a operar com critérios explícitos, capazes de orientar a análise de casos concretos e de sustentar respostas proporcionais, coerentes com a proteção da pessoa em contextos informacionais complexos (Krell; Paiva, 2017, p. 18-22; Requião, 2009, p. 33-35).

Desse modo, parte-se de uma premissa inafastável: a proteção de dados pessoais como direito fundamental autônomo e expresso pelo Art. 5º, LXXIX, da CRFB/88, tornando-se parâmetro de validade e de interpretação para todo o subsistema de responsabilidade civil, inclusive

quando o ilícito decorre do treinamento de sistemas de IA generativa. Nessa moldura, a responsabilidade deixa de ser apenas técnica reparatória e assume também função institucional de tutela da liberdade e da igualdade informacional, em coerência com a ordem constitucional de 1988 (Brasil, 1988).

Em vez de apenas ajustar categorias clássicas por analogia, toma-se o direito fundamental à proteção de dados como eixo crítico para reordenar deveres, ônus e remédios, à luz de uma teoria da justiça que combina diagnóstico social e justificação normativa. A "reconstrução normativa", nesse sentido, não inventa fins, mas explicita valores já imanentes ao sistema e os converte em critérios operativos para responder a novas materialidades de risco e dano informacional (Alencar, 2021, p. 3-7; Garcia, 2025, p. 200-203). Com essa base teórica consolidada, torna-se possível delinear, de forma sistemática, os elementos que devem orientar a aplicação prática desse dever de proteção.

Dessas premissas derivam critérios operacionais que orientarão, nas subseções seguintes, a tradução do dever de proteção em deveres de cuidado proporcionais ao risco informacional do treinamento de IA generativa (dimensão *ex ante*), em providências não pecuniárias aptas a recompor ambientes informacionais (dimensão estrutural) e em sanções e cálculos que neutralizem vantagens ilícitas e internalizem custos (dimensão distributiva). A coerência do sistema reclama que essas três dimensões sejam articuladas segundo prova de efetividade, sob pena de transformar a responsabilidade civil em rito simbólico sem capacidade de governança.

Por fim, a reconstrução proposta preserva o papel do legislador e qualifica a atuação jurisdicional e regulatória: enquanto o Parlamento não positiva parâmetros específicos, cumpre aos intérpretes, inclusive a ANPD, densificar, com justificativas públicas, os deveres de prevenção, registro e rastreabilidade, e calibrar remédios e sanções conforme a gravidade e a reversibilidade dos efeitos do treinamento. Trata-se de uma resposta institucionalmente modesta e normativamente necessária, que busca realizar, aqui e agora, o mandato constitucional de proteção de dados pessoais, sem usurpar a competência legislativa e sem aguardar uma lei que, embora desejável, permanece incerta em seu conteúdo e tempo.

4.4.1 Para uma responsabilização *ex ante* e orientada ao cuidado: deveres preventivos e precaução estrutural no desenvolvimento de IA generativa

A responsabilização *ex ante* que aqui se propõe parte de um deslocamento do foco reparatório para a conformação prospectiva de condutas, impondo aos agentes de desenvolvimento de IA generativa um padrão objetivo de cuidado capaz de internalizar riscos que, de outro modo, seriam externalizados ao ambiente informacional e aos titulares.

Esse giro exige dois pilares: (i) a adoção de balizas estruturais de responsabilidade civil como incentivos regulatórios para que riscos algorítmicos sejam considerados desde a concepção do produto e (ii) a incorporação, no domínio da proteção de dados, de uma racionalidade de gestão de riscos que substitua o improviso por processos verificáveis de prevenção e mitigação. Em outras palavras, a arquitetura jurídica deve funcionar como sinal de preço para o risco algorítmico ao mesmo tempo em que ancora a precaução na *accountability* e em mecanismos estruturais de governança do tratamento de dados pessoais (Pfeiffer, 2023, p. 7-8; Bioni; Luciano, 2019, p. 8-10).

A responsabilidade preventiva, tal como entendida, não elimina o papel das imputações clássicas, mas agrega uma camada anterior de tutela: exige planejamento e monitoramento e proporcionais ao potencial de dano sistêmico da tecnologia, com a aplicação de sanção quando a criação de risco desaprovado decorre de falhas. Essa orientação resgata o "desvalor da ação" em face de novos perfis de risco que o modelo *ex post* não consegue abarcar adequadamente, e legitima deveres positivos de acompanhamento contínuo dos efeitos da solução lançada ao mercado (Freitas, 2019, p. 195-202).

Transposta ao desenvolvimento de modelos generativos, a precaução se concretiza em deveres técnicos mínimos de "segurança e privacidade desde a concepção", cujo cumprimento é aferível: cifragem de dados e de artefatos, controle granular de acesso, autorizações em nível de modelo, técnicas de não-usabilidade, preservação de *privacy by design*, e arranjos de aprendizagem descentralizada que evitem centralização desnecessária de bases sensíveis (Li *et al.*, 2025, p. 6-7).

A esses se somam obrigações de rastreabilidade e de removibilidade: *watermarking*[22] de

[22] *Watermarking* (ou marca d'água digital) consiste na inserção de assinaturas ocultas, porém robustas, em conjuntos de dados, parâmetros de modelos ou instruções (*prompts*), com o objetivo de possibilitar a identificação da origem e a rastreabilidade do conteúdo. Essas marcas podem ser verificadas posteriormente para

dados/modelos, *fingerprinting*[23] e trilhas imutáveis de estados; além de desenvolver mecanismos que tornem exequível retirar a influência de dados indevidos sem reconstruir o sistema do zero. Esses elementos não são "melhores práticas" opcionais: compõem o conteúdo jurídico do dever de cuidado *ex ante* em ecossistemas de alto impacto informacional (Li *et al.*, 2025, p. 6-7).

Como vetor normativo, a *accountability* desloca competências decisórias para quem opera diretamente no tratamento dados, mas sempre sob abertura e controle públicos, Isso implica em avaliações de impacto em proteção de dados antecedentes, definição transparente de riscos toleráveis e registro das justificativas que orientam escolhas de arquitetura, desenho e seleção de dados no treinamento. Essa gramática não substitui direitos individuais; ao contrário, estabelece as condições de possibilidade para a própria adoção da tecnologia, permitindo avaliar, de modo antecipado, se, como e em que termos modelos generativos podem ser treinados (Bioni; Luciano, 2019, p. 8-10).

No plano organizacional, a orientação ao cuidado impõe deveres de governança que começam no topo e perpassam por toda a cadeia de fornecimento: integração da matriz de riscos de IA à estratégia corporativa e controles de missão crítica sobre vieses, propriedade intelectual, uso de dados e desinformação. Trata-se de reduzir a distância entre entusiasmo tecnológico e responsabilidade jurídica, impondo que a própria gestão de riscos seja tratada como competência essencial e não como apêndice de conformidade, sob pena de agravar riscos reputacionais e legais já identificados por diagnósticos setoriais de governança e compliance em IA (Pinheiro; Ponzoni, 2024, p. 1-3).

No direito interno, a boa-fé objetiva e seus deveres laterais oferecem um padrão normativo para densificar o conteúdo desses deveres *ex ante*: transparência, lealdade, cooperação, cuidado, segurança, prevenção e prestação de contas não são fórmulas protocolares, mas critérios para avaliar se o agente estruturou, antes do dano, salvaguardas proporcionais à escala e à opacidade do tratamento. Sendo a proteção de dados direito fundamental, sua

---

comprovar a integridade ou autoria de um artefato digital, funcionando como mecanismo técnico de governança e auditoria em sistemas de inteligência artificial. Tradução livre de Li *et al.* (2025, p. 6-7): "*For example, digital watermarking adds hidden yet robust signatures to datasets, model parameters, or prompts*".

[23] *Fingerprinting* (ou impressão digital de modelo) refere-se a técnicas destinadas a identificar um modelo de inteligência artificial por meio de consultas específicas (*crafted inputs*), capazes de revelar suas características internas, ou mediante o uso de funções criptográficas de hashing que geram assinaturas únicas. Tais métodos permitem detectar alterações não autorizadas e garantir a rastreabilidade de modelos e dados em ambientes distribuídos. Tradução livre de Li *et al.* (2025, p. 6-7): "*Model fingerprinting probes a model with crafted inputs to reveal its identity; cryptographic hashing produces unique fingerprints that change upon any bit-level alteration*".

efetividade reclama tutela anterior à lesão; logo, a omissão no cumprimento desses deveres configura inadimplemento qualificado, apto a acionar respostas civis estruturantes (Grossi, 2023, p. 171-172).

Dessa moldura resulta um teste normativo de diligência *ex ante* para treinamentos de IA com dados pessoais: (a) justificativa de necessidade e proporcionalidade do uso de dados, com alternativas tecnológicas examinadas; (b) demonstração de salvaguardas de não-usabilidade e preservação de privacidade adequadas ao risco; (c) mecanismos de rastreabilidade dos insumos e de reversibilidade prática mediante *unlearning*; (d) governança e auditoria independentes da cadeia de dados e de modelos. Em ausência desses elementos, há criação de risco juridicamente desaprovado e violação do dever de cuidado, independentemente de prova de dano individualizado imediato (Li *et al*., 2025, p. 6-7; Bioni; Luciano, 2019, p. 8-10).

Por fim, a precaução estrutural em IA generativa opera como critério de imputação de conduta e como guia de desenho institucional: ela informa a exigência de planos e registros que permitam verificar, *a posteriori*, a fidelidade do agente às escolhas preventivas que afirmou ter adotado. Com isso, a responsabilidade civil deixa de ser mero mecanismo de compensação para tornar-se instrumento de organização de deveres de cuidado em ecossistemas informacionais complexos, sem abdicar dos regimes *ex post* quando o dano se concretiza, mas exigindo, antes, que o agente prove ter feito o que o padrão jurídico do setor impõe (Freitas, 2019, p. 195-202; Grossi, 2023, p. 171-172).

A consolidação de uma responsabilidade civil orientada ao cuidado conduz, naturalmente, à necessidade de repensar as formas de reparação compatíveis com a natureza difusa e estrutural dos danos informacionais. Superado o paradigma exclusivamente compensatório, impõe-se reconhecer que a efetividade da tutela de dados pessoais exige respostas não pecuniárias capazes de restaurar condições de autodeterminação informativa. É nesse horizonte que se insere o exame das medidas não monetárias e da recomposição informacional, voltadas a restabelecer a integridade do ecossistema de dados e a prevenir a reincidência de práticas ilícitas em treinamentos de IA generativa.

4.4.2 Medidas não monetárias e recomposição informacional

A superação da lógica exclusivamente patrimonial da responsabilidade civil é imperativo

categórico diante dos danos produzidos em ecossistemas informacionais complexos, especialmente aqueles oriundos do uso indevido de dados pessoais no treinamento de sistemas de inteligência artificial generativa.

A correspondência direta entre dano e valor monetário, pressuposto pelo Art. 944 do CC/02, revela-se insuficiente e até inadequada para recompor lesões que se manifestam como perda da autodeterminação informativa. O dinheiro, nesse contexto, deixa de ser meio de restauração e se converte em sucedâneo simbólico - incapaz de neutralizar efeitos difusos e duradouros sobre bancos de dados e ecossistemas de decisão algorítmica. A resposta civil deve, por isso, adotar medidas não monetárias que operem sobre o próprio ambiente em que o ilícito se consolida, restaurando padrões de governança, de rastreabilidade e de veracidade informacional (Fajngold, 2022, p. 3-7).

A recomposição informacional, como categoria normativa emergente, traduz-se na obrigação de restabelecer condições de integridade, segurança e verificabilidade dos fluxos de dados: não apenas compensando o titular lesado, mas restaurando o equilíbrio do ambiente informacional afetado. Essa forma de reparação *in natura*, prevista implicitamente no Art. 947 do CC/02[24] e reforçada pelo Enunciado n.º 589 da VII Jornada de Direito Civil[25] , desloca o foco do patrimônio para a funcionalidade dos direitos fundamentais, convertendo o dever de indenizar em dever de recompor estruturas. Na seara dos dados pessoais, isso significa eliminar usos indevidos, corrigir vieses e restabelecer a transparência quanto à origem e ao propósito do tratamento. Tal caminho responde à natureza estrutural do dano informacional: difuso, cumulativo e de difícil reversão por equivalência pecuniária (Fajngold, 2022, p. 8-11).

Nesse horizonte, a recomposição informacional funciona como extensão prática da função regulatória da responsabilidade civil delineada anteriormente, impondo obrigações positivas de reorganização dos fluxos e de certificação de conformidade. O cumprimento de auditorias independentes e a exclusão de dados coletados ilicitamente constituem, assim, remédios jurídicos idôneos à recomposição informacional, na medida em que restauram a confiança sistêmica e previnem reincidências (Almada; Maranhão, 2023, p. 22-24).

A adoção de medidas não pecuniárias encontra fundamento adicional na doutrina da despatrimonialização da responsabilidade civil, que reconhece não o abandono das reparações

[24] Art. 947. “Se o devedor não puder cumprir a prestação na espécie ajustada, substituir-se-á pelo seu valor, em moeda corrente.” BRASIL. Lei nº 10.406, de 10 de janeiro de 2002.
[25] Enunciado n.º 589 da VII Jornada de Direito Civil: “A compensação pecuniária não é o único modo de reparar o dano extrapatrimonial, sendo admitida a reparação *in natura*, na forma de retratação pública ou outro meio.”

em dinheiro, mas a insuficiência de respostas exclusivamente pecuniárias diante de danos cuja materialidade é estrutural e informacional. Como observa Fajngold (2022, p. 8-11; 19), o chamado "movimento de despatrimonialização" do dano moral revela que o dinheiro deixou de ser o único instrumento de recomposição adequada, abrindo espaço para soluções que restituam a integridade dos bens jurídicos atingidos.

No campo institucional, a recomposição informacional também se articula com instrumentos de corregulação e certificação. A LGPD, em seu Art. 50, autoriza a elaboração de códigos de boas práticas e de governança, cuja observância pode funcionar simultaneamente como mecanismo preventivo e como forma de reparação estrutural, ao permitir a verificação pública de padrões técnicos e éticos aplicáveis ao tratamento de dados (Brasil, 2018).

Tais códigos, ao mesmo tempo em que reforçam a responsabilização dos agentes privados, conferem concretude ao princípio da cooperação e reduzem o espaço de arbítrio judicial na definição de remédios não monetários (Lima, 2021, p. 12-13). Nessa lógica, a recomposição não se limita à restauração técnica, mas abrange a adoção de práticas organizacionais permanentes que previnam reincidências e promovam cultura de conformidade informacional.

A função transformadora das medidas não pecuniárias ganha relevo, ainda, ao se considerar o exposto na seção 4.2 - que os danos informacionais são frequentemente coletivos ou transindividuais. Quando o uso indevido de dados pessoais atinge massas de titulares e compromete a integridade de bases inteiras, a recomposição informacional deve assumir feição pública e estrutural: publicização dos fluxos de dados afetados, planos de correção acompanhados por autoridades independentes, relatórios de impacto e criação de mecanismos de reversibilidade técnica são formas de restaurar o ambiente informacional lesado. Nesses casos, a sentença ou o termo de ajustamento de conduta deixam de ter caráter meramente punitivo para se converterem em instrumentos de reconstrução institucional - espécie de remédio estrutural informacional, apto a restabelecer a normalidade dos ecossistemas digitais (Capanema, 2020, p. 6).

Em termos dogmáticos, a recomposição informacional atua como extensão do dever de mitigação e da boa-fé objetiva, impondo aos agentes de tratamento condutas ativas de correção, comunicação e cooperação. Assim como o ilícito informacional é estrutural, também deve sê-lo sua reparação: não basta suprimir a vantagem econômica obtida com o uso indevido do dado, é necessário reordenar o ambiente em que o ilícito se tornou possível. A reparação não monetária - sob forma de auditoria compulsória, suspensão de modelos

treinados com dados ilícitos ou obrigação de criar mecanismos de exclusão e rastreabilidade - concretiza a função ética e preventiva da responsabilidade civil, ao passo que reforça o caráter fundamental do direito à proteção de dados (Fajngold, 2022, p. 3-7; Lima, 2021, p. 12-13).

Por fim, a recomposição informacional deve ser compreendida como corolário da justiça informacional: deve-se restaurar a autonomia e a confiança dos titulares em um ecossistema de dados íntegro. A efetividade do sistema dependerá, portanto, da consolidação de medidas estruturais que se estendam para além do patrimônio e incidam sobre as culturas de tratamento de dados (Fajngold, 2022, p. 8-11).

Ao deslocar o eixo da indenização para a recomposição, a responsabilidade civil reencontra sua vocação ética de proteção da pessoa e de reequilíbrio do poder informacional, preparando o terreno para as sanções e instrumentos distributivos que serão examinados na subseção seguinte .

A consolidação das medidas não monetárias e da recomposição informacional revela que a reparação efetiva, em matéria de dados pessoais, depende menos do ressarcimento individual e mais da restauração estrutural do ambiente informacional. É nesse ponto que se insere o exame das sanções e dos mecanismos de cálculo, orientados à redistribuição e à prevenção, que complementam a função ética e regulatória da responsabilidade civil no domínio da inteligência artificial generativa.

4.4.3 Sanções e cálculo: *disgorgement*, multas proporcionais e fundos *cy-près*

Antes de ingressar na etapa sancionatória, importa fechar o arco reconstrutivo aberto no tópico 4.4: definidos os deveres *ex ante* de cuidado e precaução (4.4.1) e estabelecida a recomposição informacional por medidas não pecuniárias (4.4.2), falta completar a arquitetura por sua dimensão distributiva: aquela que retira a racionalidade econômica do ilícito e realimenta, com recursos do próprio infrator, a integridade do ecossistema de dados. Trata-se de converter os fins constitucionais e as ferramentas sistemático-teleológicas já afirmadas em parâmetros de cálculo e destinação de valores que neutralizem vantagens indevidas.

Desse modo, a etapa sancionatória da responsabilidade civil, no uso não autorizado de dados pessoais por sistemas de inteligência artificial generativa, não possui uma racionalidade e

punitiva, e sim corretiva e distributiva: impedir que o ilícito se pague e internalizar no agente o custo integral da degradação informacional que produziu. Nesse contexto, emerge um tripé sancionatório composto pelo *disgorgement* dos lucros ilícitos, pelas multas civis proporcionais e pela destinação social dos valores mediante fundos *cy-près*, instrumentos que operam sobre dimensões distintas mas convergentes no objetivo de restaurar a integridade do sistema e eliminar incentivos econômicos à violação (Rosenvald; Farias; Netto, 2025, p. 121-128).

Embora o sistema brasileiro rejeite sanções civis de caráter punitivo, a etapa sancionatória não pode ignorar sua dimensão dissuasória. Quando o benefício econômico esperado da violação supera o custo provável da reparação, a responsabilização perde efetividade. É nesse ponto que a doutrina dos *punitive damages* opera como referência comparada, não para ser importada, mas para demonstrar que a responsabilidade civil também desempenha papel de orientação de condutas. A lição central é simples: a estrutura sancionatória deve tornar antieconômica a escolha pelo risco ilícito. No Brasil, esse efeito dissuasório não decorre de penas civis, mas de instrumentos compatíveis com nossa matriz constitucional capazes de internalizar externalidades e prevenir a repetição de condutas lesivas (Eick, 2020, p. 43–47).

É nesse ponto que a restituição de ganhos ilícitos (*disgorgement*) representa a resposta mais direta à lógica econômica que estrutura o tratamento indevido de dados. Quando o dado pessoal é explorado fora dos parâmetros legais, o agente extrai vantagem econômica mensurável, sem que, muitas vezes, o titular sofra dano patrimonial perceptível. A técnica compensatória tradicional, centrada nas perdas da vítima, mostra-se, assim, incapaz de alcançar o núcleo do problema: a assimetria de ganhos que alimenta o ciclo de exploração informacional. O *disgorgement* desloca o foco do prejuízo para o benefício indevido, operando uma justiça corretiva de segunda ordem: retira do ofensor o produto do ilícito e o devolve ao corpo social, reafirmando o princípio de que o ilícito não pode compensar (*tort must not pay*) (Rosenvald; Farias; Netto, 2025, p. 121-128).

Esse remédio restitutório não se confunde com enriquecimento sem causa, tampouco com pena civil. Ele parte da constatação de que o ilícito informacional é estruturalmente lucrativo, pois possui baixo risco de detecção, danos dispersos, vítimas apáticas e lucros concentrados (Pasquale, 2015, p. 39-40). O cenário exige respostas que neutralizem o cálculo racional da violação. Quando o agente estima que o lucro obtido com o uso indevido de dados superará qualquer indenização provável, o sistema jurídico, ao tolerar tal lógica, converte-se em

cúmplice da assimetria informacional. O *disgorgement*, ao expropriar o ganho e não apenas compensar o dano, inverte o vetor de incentivo e restabelece o equilíbrio distributivo entre inovação e dignidade informacional (Sanchez *et al*., 2025, p. 11-13).

Do ponto de vista funcional, a restituição dos lucros ilícitos cumpre três objetivos normativos: (a) reafirma a proibição de obtenção de vantagens econômicas a partir da violação de direitos fundamentais, (b) reforça a função preventiva da responsabilidade civil, enviando sinal inequívoco de que o ilícito não gera retorno, e (c) fornece base financeira para medidas estruturais de recomposição ou de educação digital. A lógica econômica é, aqui, revertida em racionalidade jurídica: quem lucra com o risco ilícito deve financiar a mitigação de seus próprios efeitos. A técnica se coaduna com o Art. 944, parágrafo único[26], do CC/02, ao permitir que, diante de conduta dolosa e reiterada, a indenização supere a mera equivalência de dano e alcance a totalidade do proveito econômico obtido. Essa reconfiguração não amplia indevidamente o poder sancionatório, mas atualiza o princípio da reparação integral em chave distributiva e preventiva (Rosenvald, Farias e Netto, 2025, p. 121-128).

É nessa linha que estudos sobre a atuação da *Securities and Exchange Commission* e sua incorporação em regimes de direito ambiental e contratual demonstram que a restituição dos lucros indevidos é capaz de desestimular práticas ilícitas e financiar medidas coletivas de reparação, inclusive em contextos de alta complexidade técnica (Sanchez *et al.*, 2025, p. 11-13).

No contexto da inteligência artificial generativa, o mesmo raciocínio se aplica: a restituição dos ganhos advindos do uso indevido de dados pessoais pode financiar ações coletivas de remoção de *datasets* ilícitos e fortalecimento de auditorias independentes. O *disgorgement* converte-se, assim, em ferramenta de redistribuição informacional e de justiça tecnológica, ao mesmo tempo corretiva e prospectiva.

Paralelamente, as multas civis proporcionais operam como mecanismos complementares de desestímulo e calibragem institucional. Diferem-se das multas administrativas aplicadas pela ANPD, pois incidem na esfera civil, com fundamento no Art. 412 do CC/02 e nos deveres de boa-fé objetiva. Seu cálculo deve observar parâmetros de proporcionalidade e capacidade contributiva, considerando o benefício econômico obtido, a gravidade da infração e a

[26] Art. 944. “A indenização mede-se pela extensão do dano. Parágrafo único. Se houver excessiva desproporção entre a gravidade da culpa e o dano, poderá o juiz reduzir, eqüitativamente, a indenização.” BRASIL. Lei nº 10.406, de 10 de janeiro de 2002

necessidade de financiamento de programas de conformidade e reparação coletiva. A multa civil proporcional, quando orientada por critérios objetivos, não representa duplicidade sancionatória, mas técnica de internalização de custos - assegurando que o agente suporte, no plano civil, a totalidade das externalidades negativas que produziu. Trata-se de uma ferramenta de política jurídica, cuja função é recalibrar incentivos e tornar antieconômica a violação reiterada de deveres de proteção de dados (Rosenvald, Farias e Netto, 2025, p. 121-128).

Essas sanções pecuniárias, para alcançarem finalidade pública, devem articular-se com mecanismos de destinação coletiva que evitem dispersão ou ineficiência na execução. É nesse ponto que se insere a técnica dos fundos *cy-près*, adaptada à realidade brasileira pelas previsões do Art. 13 da Lei 7.347/1985 e do Art. 100, parágrafo único do CDC. O modelo prevê que, na impossibilidade de reparação direta aos titulares ou de liquidação individual, os valores revertam a fundos destinados à recomposição de bens jurídicos difusos, como o Fundo de Defesa dos Direitos Difusos (FDD). Inspirada na "compensação fluida" norte-americana, essa técnica garante que o produto das condenações e das multas seja reinvestido em finalidades correlatas ao dano, promovendo um ciclo virtuoso de prevenção e reparação estrutural (Rudiniki Neto, 2019, p. 1-3).

A destinação *cy-près* não constitui mera substituição simbólica, mas prolongamento da função reparatória em chave coletiva. Quando os efeitos do uso indevido de dados atingem milhões de titulares ou comprometem ecossistemas digitais inteiros, a recomposição individual é inviável. A reversão de valores a fundos públicos permite financiar programas de reeducação digital, desenvolvimento de tecnologias de *privacy by design*, incentivo a pesquisas sobre ética algorítmica e fortalecimento da infraestrutura de governança de dados. Trata-se de resposta institucional à natureza difusa e transindividual do dano informacional, que converte a sanção em instrumento de reconstrução social e de promoção de cultura de conformidade (Homma, 2017, p. 11-15).

No plano prático, a aplicação coordenada do *disgorgement*, das multas proporcionais e dos fundos *cy-près* exige uma metodologia de cálculo que integre elementos econômicos e normativos. O parâmetro não é apenas o *quantum* pecuniário, mas o impacto estrutural da conduta. A quantificação deve considerar (i) a extensão da vantagem obtida, (ii) o número de titulares potencialmente afetados, (iii) a reversibilidade técnica dos efeitos e (iv) a destinação social mais adequada dos recursos. A sentença civil, nesses casos, deve ser construída como

instrumento de política pública, fixando parâmetros de redistribuição que assegurem proporcionalidade e finalidade. O juiz torna-se gestor da eficácia do sistema, e não mero liquidante de danos (Rudiniki Neto, 2019, p. 1-3; Homma, 2017, p. 11-15).

Por essa via, a reconstrução normativa da responsabilidade civil pelo uso indevido de dados pessoais em sistemas de IA generativa atinge sua plenitude distributiva: o ilícito deixa de ser vantajoso, o ganho ind5 conURSICevido é expropriado, os valores arrecadados são redirecionados à coletividade e as sanções civis, aplicadas de forma proporcional e transparente, funcionam como incentivos para o desenvolvimento ético e sustentável da tecnologia. O Direito Civil reencontra, assim, sua vocação republicana: reequilibrar, pela via da restituição e da destinação social dos recursos, o poder informacional concentrado nas mãos dos agentes tecnológicos e devolver à sociedade o fruto do ilícito digital (Rosenvald; Farias; Netto, 2025, p. 121-128; Sanchez *et al.*, 2025, p. 11-13).

O resultado é um modelo de responsabilidade civil que, sem romper com suas bases dogmáticas, recupera sua função ética e regulatória, reposicionando-se como instrumento de proteção da autodeterminação informativa e de governança dos ecossistemas digitais. É sobre esse pano de fundo que a conclusão adiante sintetizará as premissas normativas e institucionais indispensáveis à consolidação desse novo paradigma.

5 CONCLUSÃO

A análise desenvolvida ao longo desta pesquisa evidenciou que o uso não autorizado de dados pessoais no treinamento de sistemas de inteligência artificial generativa constitui uma forma inédita de violação informacional, marcada pela opacidade, irreversibilidade e exploração massiva de informações. A partir desse diagnóstico, foi possível reconstruir criticamente os limites do paradigma clássico da responsabilidade civil e apontar caminhos interpretativos capazes de conferir proteção efetiva à autodeterminação informativa na sociedade de dados.

A economia digital consolidou um ambiente de coleta constante, muitas vezes invisível, no qual dados pessoais são convertidos em insumo econômico e circulam de forma amplificada e opaca. Nesse contexto, a IA generativa intensifica essa lógica ao depender de grandes volumes de informação para sua parametrização, de modo que, mesmo sem intenção explícita, dados pessoais acabam sendo incorporados aos modelos. A análise documental das políticas de privacidade das principais plataformas (subseção 2.3.5) confirmou a existência de um descompasso persistente entre o discurso de proteção e as práticas efetivas de tratamento, evidenciando que o titular é frequentemente inserido num processo de captura informacional que escapa à sua vontade e ao seu controle.

A estrutura tradicional dos pressupostos da responsabilidade civil encontra dificuldades significativas quando aplicada ao treinamento de IA generativa. Verificou-se que a ilicitude frequentemente decorre da própria forma de coleta e utilização dos dados, cuja escala e opacidade inviabilizam a identificação individualizada de vítimas ou de prejuízos concretos. A irreversibilidade técnica do treinamento e a replicação contínua de inferências ampliam a dificuldade de demonstrar causalidade e revelam um dano que ultrapassa o indivíduo e alcança a integridade do ecossistema informacional. Diante disso, a centralidade exclusivamente compensatória do instituto mostra-se insuficiente para enfrentar riscos contínuos e difusos, exigindo sua reorientação para funções preventivas e distributivas.

Em termos dogmáticos e normativos, a investigação demonstrou a pertinência de se adotar uma responsabilidade objetiva nos casos em que a operação tecnológica cria perigo especial aos direitos da personalidade, bem como o reconhecimento do dano *in re ipsa* e das configurações de dano existencial decorrentes da erosão da autonomia informativa. Essa reinterpretação não demanda a criação imediata de um novo microssistema jurídico, mas

exige a aplicação coordenada da CRFB/88, da LGPD, do CC/02 e até do CDC, uma vez que as plataformas de IA atuam como fornecedoras de serviços digitais e, portanto, submetem-se aos deveres e ao regime de responsabilidade previstos na legislação consumerista.

Contudo, para que essa reorientação produza efeitos concretos é necessário conferir à responsabilização dimensão processual e instrumentária: a LGPD já prevê instrumentos essenciais como registros das operações de tratamento e relatórios de impacto, que devem ser efetivamente exigidos e valorizados como trilhas auditáveis. Torna-se, portanto, fundamental reforçar mecanismos que viabilizem a inversão dinâmica do ônus da prova, o acesso judicial a registros de atividade e relatórios técnicos e a possibilidade de requisição de auditorias independentes, de modo que a invisibilidade técnica não se converta em impunidade.

Além disso, a alocação do risco exige cuidados práticos: a aplicação extraterritorial da LGPD e a exigência de representante legal para controladores sem estabelecimento no Brasil devem ser enfatizadas como instrumentos de responsabilização transnacional.

Do ponto de vista econômico e de imputação, recomenda-se adotar critérios orientadores - em especial a matriz triádica proposta (teste de traçabilidade; teste de previsibilidade *ex ante*; teste de evitabilidade) e o critério do *cheapest cost avoider* - para identificar quem, na cadeia técnica e contratual, tinha melhores condições de prevenir o dano a custos inferiores e, portanto, deve suportar o ônus da responsabilização.

Restou claro também, que a mera produção de dados sintéticos não afasta a ilicitude quando estes derivam de bases obtidas ilegalmente ou sem salvaguardas robustas.

No plano da responsabilização, a pesquisa sustenta a necessidade de combinar medidas pecuniárias e não pecuniárias: além de multas proporcionais e *disgorgement*, o ordenamento deve privilegiar remédios estruturais: ajustes técnicos nos modelos, exclusão ou anonimização eficaz, auditorias independentes, obrigações de correção de assimetrias informacionais e fundos *cy-près* que reinjajem recursos na proteção dos fluxos de dados. Essas medidas têm natureza tanto reparatória quanto preventiva, contribuindo para desincentivar modelos econômicos baseados na exploração ilícita de dados.

A eficácia desse conjunto normativo e instrumental exige atuação coordenada de diversos atores institucionais: a ANPD precisa ter capacidade técnica e sancionatória robusta; o Ministério Público e o Judiciário devem dispor de meios processuais para requisitar e avaliar provas técnicas; e organismos de fiscalização e regulação devem integrar esforços

transnacionais para enfrentar cadeias de tratamento distribuídas. Em sintonia com isso, propõe-se o fortalecimento de auditorias independentes e da *accountability* para agentes que operem em larga escala com dados pessoais.

Reconhecem-se, por fim, limitações inerentes ao escopo adotado e necessidades de continuidade da pesquisa. Por se tratar de um estudo cujo objeto se circunscreve à análise dogmática e normativa da responsabilização civil, a investigação não comportou, nem pretendia comportar, avaliações empíricas ou mensurações econômicas detalhadas. A consolidação plena das propostas aqui formuladas, contudo, demanda etapas que extrapolam o alcance deste trabalho, tais como estudos empíricos sobre a mensuração do valor econômico efetivamente apropriado pelas plataformas, análises técnicas sobre a eficácia e os limites de métodos de anonimização e pesquisas aplicadas que testem, em ambiente real, instrumentos como o *disgorgement* e os fundos *cy-près*. Tais caminhos configuram desdobramentos naturais para evolução futura do tema e podem aprofundar a compreensão prática dos instrumentos regulatórios sugeridos.

Em síntese, conclui-se que o ordenamento jurídico brasileiro contém, em seus princípios e institutos, instrumentos aptos a enfrentar o uso não autorizado de dados pessoais no treinamento de IA generativa, desde que reinterpretados e operacionalizados com ênfase preventiva, processual e coletiva. A responsabilidade civil, assim reorientada, passa a exercer função institucional de governança informacional: não apenas reparar prejuízos individuais, mas prevenir riscos, recompor assimetrias e desincentivar práticas exploratórias que ameaçam a dignidade informacional.

Preservar dados é, hoje, preservar a liberdade: em uma sociedade mediada por algoritmos, tutelar o uso indevido de dados no treinamento de IA generativa é preservar os fundamentos da vida digna e da autonomia humana em tempos de hiperconectividade.

REFERÊNCIAS

AARONSON, Susan Ariel. ***Data Dysphoria: The Governance Challenge Posed by Large Learning Models*.** SSRN - Social Science Research Network, George Washington University - Elliott School of International Affairs, p. 1-29, 18 set. 2023 DOI: https://doi.org/http://dx.doi.org/10.2139/ssrn.4554580. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4554580. Acesso em: 14 ago. 2025.

Agence France-Presse. The New York Times processa OpenAI, criadora do ChatGPT, e Microsoft por violação de direitos autorais: Jornal norte-americano exige indenização dos prejuízos que podem chegar na casa dos bilhões de dólares, além de uma ordem para que as empresas deixem de utilizar seu conteúdo e apague os dados já compilados. **G1 Tecnologia.** [S. l.], 27 dez. 2023. Disponível em: https://g1.globo.com/tecnologia/noticia/2023/12/27/the-new-york-times-processa-openai-e-microsoft-por-violacao-de-direitos-autorais.ghtml. Acesso em: 10 nov. 2025.

AGUIAR JÚNIOR, Ruy Rosado de. **Interpretação**. Revista da AJURIS, Porto Alegre, v. 16, n. 45, p. 7-20, mar. 1989. Disponível em: https://share.google/06PCmjgSPd5NI1bVa. Acesso em: 17 out. 2025.

AGUIAR JÚNIOR, Ruy Rosado de (org.). V Jornada de Direito Civil. Brasília: Conselho da Justiça Federal, 2012. 388 p. ISBN 978-85-85572-95-2. Disponível em: https://www.cjf.jus.br/cjf/corregedoria-da-justica-federal/centro-de-estudos-judiciarios-1/publicacoes-1/jornadas-cej/vjornadadireitocivil2012.pdf. Acesso em: 10 out. 2025.

ALBERS, Marion. A complexidade da proteção de dados. **Revista Brasileira de Direitos Fundamentais & Justiça**, [S. l.], v. 10, n. 35, p. 19–45, 2016. DOI: 10.30899/dfj.v10i35.93. Disponível em: https://dfj.emnuvens.com.br/dfj/article/view/93. Acesso em: 03 ago. 2025.

ALBIANI, Christine. **Responsabilidade Civil e Inteligência Artificial: Quem responde pelos danos causados por robôs inteligentes?**. 2019. Disponível em: <https://itsrio.org/wp-content/uploads/2019/03/Christine-Albiani.pdf>. Acesso em: 05 de mar. de 2025.

ALDMOUR, Renad. **Civil Liability For Damages Caused By Artificial Intelligence In Light Of Sustainable Development Goals "Sdg3, 9 & 16" "A Legal Analysis Within The Saudi Civil Transactions System"**. Journal of Lifestyle & SDG´ Review, Florida, EUA: SDGsReview, ed. VOL. 5 | e06197|, ano 2025, p. 01-39, 6 jun. 2025. Disponível em: https://doi.org/10.47172/2965-730X.SDGsReview.v5.n05.pe06197. Acesso em: 6 set. 2025.

ALENCAR, Bruno Araujo. A Reconstrução Normativa Em Axel Honneth: Um Novo Princípio Para A Teoria De Justiça. **Revista de Estudos dos Pós-Graduandos em Filosofia**, [S. l.]: Universidade Estadual Paulista, ed. Vol. 13, ano 2021, n. 35, p. 49-60, 17 nov. 2021. Disponível em: https://doi.org/10.36311/1984-8900.2021.v13n35.p49-60. Acesso em: 17 out. 2025.

ALEXY, Robert. **Teoria Dos Direitos Fundamentais**. 3. ed. Brasil: Malheiros, 2024.

ALMADA, Marco. Responsabilidade civil extracontratual e a inteligência artificial. **Revista Acadêmica Arcadas.** São Paulo: Faculdade de Direito da USP, ed. 2, ano 2019, n. 1, p. 88-100, 9 ago. 2019. Disponível em: https://www.researchgate.net/publication/327756096. Acesso em: 6 set. 2025.

ALMADA, Marco; MARANHÃO, Juliano. Contribuições e Limites da Lei Geral de Proteção de Dados para a Regulação da Inteligência Artificial no Brasil. **Revista de Direito Público**, [S. l.], v. 20, n. 106, 2023. DOI: 10.11117/rdp.v20i106.6957. Disponível em: https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/6957. Acesso em: 23 out. 2025.

ALMEIDA, Bruno Costa de. **Causalidade no incidente de violação da segurança do dado pessoal.** 2023. Tese (Mestrado em Direito) - Faculdade de Direito, Universidade do Estado do Rio de Janeiro - UERJ, Rio de Janeiro. Orientador: Profª Dra. Gisela Sampaio da Cruz Costa Guedes. Disponível em: https://www.bdtd.uerj.br:8443/bitstream/1/19910/3/Disserta%c3%a7%c3%a3o%20-%20Bruno%20Costa%20de%20Almeida%20-%202023%20-%20Completa.pdf. Acesso em: 10 set. 2025.

AMARAL, Fernando. **Introdução à Ciência de Dados: Mineração de dados e Big Data.** 01. ed. Rio de Janeiro, Rio de Janeiro: Alta Books, 2016.

ANDRADE, F. S. de; FACCIO, L. G. Notas Sobre A Responsabilidade Civil Pela Utilização Da Inteligência Artificial. **Revista da AJURIS** - Qualis A2, [S. l.], v. 46, n. 146, p. 153–182, 2019. Disponível em: https://revistadaajuris.ajuris.org.br/index.php/REVAJURIS/article/view/983. Acesso em: 06 set. 2025.

ANTUNES, Henrique Sousa. Responsabilidade civil do produtor: os danos ressarcíveis na era digital. **Revista de Direito da Responsabilidade**, ano 1, p. 1476–1485, 2019. Disponível em: https://share.google/B49txVjIybNRCCOyR. Acesso em: 19 set. 2025.

ANTUNES, Henrique Sousa. Das Funções Reconstitutiva E Punitiva Da Responsabilidade Civil Extracontratual. *In.:* **Novos Olhares sobre a Responsabilidade Civil.** Lisboa: Centro Jurídico de Estudos Judiciários, 2018, ISBN: 978-989-8908-36-0. Disponível em: https://share.google/vwxaN5LdyHzAakpB4. Acesso em: 9 out. 2025.

ATATA, Bashirat Bukola. **Artificial Intelligence and the Right to be Forgotten.** International Journal of Research Publication and Reviews, [S. l.]: IJRPR, ed. Vol. 5, ano 2024, n. 8, p. 4300-4310, Disponível em: DOI:10.55248/gengpi.5.0824.2310. Acesso em: 19 set. 2025.

AGÊNCIA NACIONAL DE PROTEÇÃO DE DADOS (ANPD). **Relatório de Impacto à Proteção de Dados Pessoais (RIPD).** Brasília, DF, 05 abr. 2023. Disponível em: https://www.gov.br/anpd/pt-br/canais_atendimento/agente-de-tratamento/relatorio-de-impacto-a-protecao-de-dados-pessoais-ripd. Acesso em: 10 ago. 2025.

AGÊNCIA NACIONAL DE PROTEÇÃO DE DADOS (ANPD). **Nota Técnica nº 12/2025/CON1/CGN/ANPD**. Brasília, 03 abr. 2025. Disponível em:

https://www.gov.br/anpd/pt-br/acesso-a-informacao/participacao-social/outras-acoes/documentos/ts-06-2024-nt-12-2025-consolidacao-das-contribuicoes.pdf/view. Acesso em: 15 mai. 2025.

BAHIA, Claudio José Amaral; SARTORI, Ellen Carina Mattias; MARCANDELI, Raíssa Amarins; PEREIRA, Vanessa Nunes. Direito À Proteção De Dados: Trajetória E Delineações Do Reconhecimento Como Um Direito Fundamental Autônomo No Brasil. **Revista da Faculdade de Direito da UERJ - RFD**, *[S. l.]*, n. 43, 2024. DOI: 10.12957/rfd.2024.67486. Disponível em: https://www.e-publicacoes.uerj.br/rfduerj/article/view/67486. Acesso: 08 out. 2025.

BAI, Yun; WANG, Shaofeng. **Impact of generative AI interaction and output quality on university students' learning outcomes: a technology-mediated and motivation-driven approach.** Scientific Reports, v. 15, n. 24054, 2025. DOI: 10.1038/s41598-025-08697-6. Acesso em: 8 nov. 2025.

BALKIN, Jack M. **Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation.** Social Science Research Network (SSRN). Disponível em: https://ssrn.com/abstract=3038939. Acesso em: 16 mai. 2025.

BARBOSA, Ana Mafalda Castanheira Neves de Miranda. **Do nexo de causalidade ao nexo de imputação: contributo para a compreensão da natureza binária e personalística do requisito causal ao nível da responsabilidade civil extracontratual.** 2012. Tese. (Doutorado em Direito) - Faculdade de Direito, Universidade de Coimbra, Coimbra. Orientador: Dr. Doutor Pinto Monteiro. Disponível em: https://hdl.handle.net/10316/23223. Acesso em: 5 set. 2025.

BARBOSA, Mafalda Miranda. Inteligência Artificial, E-Persons E Direito: Desafios E Perspetivas. **Revista Jurídica Luso-Brasileira**, Lisboa, Portugal: Faculdade de Direito da Universidade de Lisboa, ed. Vol. 3 , ano 2017, n. 6, p. 1475-1503, Disponível em: https://share.google/jXahSuUim4aX6EWvj. Acesso em: 8 set. 2025.

BASTOS, Bruna; VON ENDE, Luiza Berger. A Lei Geral De Proteção De Dados Pessoais E O Tratamento Conferido Às Empresas Estrangeiras. *In:* PIAIA, Thami Covatti; PATZ, Stéfani Reimann; HARTMANN, Gabriel Henrique (Orgs.). **ANAIS do II Seminário Sibre Inteligência Artificial, Proteção de Dados e Cidadania. Santo Ângelo: EdiURI - Editora da URI, 2021, p. 37-61.** Disponível em: https://share.google/akno0W9X8c2YvvQ0a. Acesso em: 5 set. 2025.

BATTESINI, Eugênio. Análise econômica da responsabilidade civil: limites e possibilidades de aplicação do 'incremental learned hand standard' e do 'cheapest cost avoider criterion', com ênfase na responsabilidade civil do Estado em Portugal e no Brasil. **Revista IBERC**. Belo Horizonte, v. 8, n. 2, p. 37–61, 2025. DOI: 10.37963/iberc.v8i2.359. Disponível em: https://revista.iberc.org.br/iberc/article/view/359. Acesso em: 18 set. 2025.

BENTES, Anna. **Quase Um Tique: Economia Da Atenção, Vigilância E Espetáculo Em Uma Rede Social.** Rio de Janeiro: Editora UFRJ, 2021, ISBN: 978-85-7108-480-3. Disponível em: https://pantheon.ufrj.br/bitstream/11422/16510/1/quase-um-tique_2020.pdf. Acesso em: 11 mai. 2025.

BIGONHA, Carolina. **Inteligência artificial em perspectiva**. Panorama Setorial da Internet. Inteligência Artificial e ética. Ano 10, n. 2, out. 2018. Disponível em: https://share.google/LHxjLs3kXh3dpzNfb. Acesso em: 8 set. 2025.

BIONI, Bruno Ricardo. **Compreendendo o conceito de anonimização e dado anonimizado.** Cadernos Jurídicos, São Paulo: Escola Paulista da Magistratura, ed. 53, ano 2020, p. 191-201, Semestral. Disponível em: https://www.tjsp.jus.br/download/EPM/Publicacoes/CadernosJuridicos/ii_9_anonimiza%C3%A7%C3%A3o_e_dado.pdf. Acesso em: 6 ago. 2025.

BIONI, Bruno Ricardo. A Obrigação De Registro Das Atividades De Tratamento De Dados. *In:* BIONI, Bruno Ricardo. (Org.). **Proteção de dados: contexto, narrativas e elementos fundantes.** São Paulo: Appris Editora, 2021. p. 267-271, ISBN: 978-65-995360-0-7. Acesso em: 8 de out. 2025.

BIONI, Bruno Ricardo. **Proteção de dados pessoais: a função e os limites do consentimento.** Rio de Janeiro: Forense, 2019.
ISBN 978-85-309-8328-4. Disponível em: https://www.kufunda.net/publicdocs/Prote%C3%A7%C3%A3o%20de%20dados%20pessoais%20a%20fun%C3%A7%C3%A3o%20e%20os%20limites%20do%20consentimento%20(Bruno%20Ricardo%20Bioni).pdf.Acesso em: 6 nov. 2025.

BIONI, Bruno; DIAS, Daniel. Responsabilidade civil na proteção de dados pessoais: construindo pontes entre a Lei Geral de Proteção de Dados Pessoais e o Código de Defesa do Consumidor. **Civilistica.com - Revista Eletrônica de Direito Civil**. Rio de Janeiro, v. 9, n. 3, p. 1–23, 2020. Disponível em: https://civilistica.emnuvens.com.br/redc/article/view/662. Acesso em: 16 out. 2025.

BIONI, Bruno; KITAYAMA, Marina; RIELLI, Mariana. **O Legítimo Interesse na LGPD: quadro geral e exemplos de aplicação.** São Paulo: Associação Data Privacy Brasil de Pesquisa, 2021. Disponível em: https://share.google/rMc2B6mv2C0vWrGs6 . Acesso em: 16 out. 2025.

BIONI, Bruno Ricardo; LUCIANO, Maria. **O princípio da precaução na regulação de inteligência artificial: seriam as leis de proteção de dados o seu portal de entrada?**. 2019. Disponível em: https://brunobioni.com.br/home/wp-content/uploads/2019/09/Bioni-Luciano_O-PRINCI%CC%81PIO-DA-PRECAUC%CC%A7A%CC%83O-PARA-REGULAC%CC%A7A%CC%83O-DE-INTELIGE%CC%82NCIA-ARTIFICIAL-1.pdf. Acesso em: 23 out. 2025.

BUARQUE, Elaine. O dano existencial como uma nova modalidade de dano não patrimonial: a necessidade da ampliação do princípio da função social da responsabilidade civil e a busca da reparação integral do dano à pessoa. **Revista IBERC**. Belo Horizonte, v. 2, n. 2, 2019. DOI: 10.37963/iberc.v2i2.57. Disponível em: https://revistaiberc.responsabilidadecivil.org/iberc/article/view/57. Acesso em: 16 out. 2025.

BURRELL, J. (2016). **How the machine ‘thinks’: Understanding opacity in machine learning algorithms.** Big Data & Society, 3(1). Disponível em: https://doi.org/10.1177/2053951715622512 . Acesso em: 23 ago. 2025.

BONNA, Alexandre Pereira; CAÑIZO, Amanda de Moura; CALZAVARA, Giovana Ferreira. Consentimento E Lgpd: Desafios Diante Da Hipervulnerabilidade Do Consumidor. **Revista de Direito e Atualidades**, [S. l.], v. 1, n. 3, 2022. Disponível em: https://www.portaldeperiodicos.idp.edu.br/rda/article/view/6231. Acesso em: 9 out. 2025.

BONNET, Adrien. ***La Responsabilité du fait de l’intelligence artificielle. Réflexion sur l’émergence d’un nouvel agent générateur de dommages***. 2015. Tese. (Mestrado em Direito) - Faculdade de Direito, Universidade Pantheon-Assas, Paris, França. Orientador: Nicolas Molfessis. Disponível em: https://share.google/JHQq7WW6NyfGz53DP. Acesso em: 8 set. 2025.

BOUTONNET, Mathilde. **L'influence du principe de précaution sur la responsabilité civile en droit français: un bilan en demi-teinte,** 2014 10-1. Revue de droit du développement durável da Universidade McGill 8, 2014 CanLIIDocs 113. Disponível em: https://canlii.ca/t/2m3g. Acesso em: 18 set. 2025.

BRASIL. Câmara dos Deputados. **Projeto de Lei n.º 2.338**, de 2023. Dispõe sobre o desenvolvimento, o fomento e o uso ético e responsável da inteligência artificial com base na centralidade da pessoa humana. Brasília, DF: Câmara dos Deputados, 2025. Disponível em: https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra?codteor=2868197&filename=PL%202338/2023. Acesso em: 10 ago. 2025.

BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil**. Brasília, DF, 05 out. 1988. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 16 abr. 2025.

BRASIL. Lei nº 6.938, de 31 de agosto de 1981. **Dispõe sobre a Política Nacional do Meio Ambiente, seus fins e mecanismos de formulação e aplicação, e dá outras providências.** Brasília, DF, 31 ago. 1971. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l6938.htm. Acesso em: 20 out. 2025.

BRASIL. Lei nº 7.347, de 24 de julho de 1985. **Disciplina a ação civil pública de responsabilidade por danos causados ao meio-ambiente, ao consumidor, a bens e direitos de valor artístico, estético, histórico, turístico e paisagístico (VETADO) e dá outras providências.** Brasília, DF, 24 jul. 1985. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l7347orig.htm. Acesso em: 18 out. 2025.

BRASIL. Lei nº 8.078, de 11 de setembro de 1990. **Dispõe sobre a proteção do consumidor e dá outras providências.** Brasília, DF, 11 set. 1990. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l8078compilado.htm. Acesso em: 20 mai. 2025.

BRASIL. Lei nº 9.868, de 10 de novembro de 1999. **Dispõe sobre o processo e julgamento da ação direta de inconstitucionalidade e da ação declaratória de constitucionalidade perante o Supremo Tribunal Federal**. Brasília, DF, 10 set. 1999. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l9868.htm. Acesso em: 23 out. 2025.

BRASIL. Lei nº 10.406, de 10 de janeiro de 2002. **Código Civil.** Diário Oficial da União: seção 1, Brasília, DF, ano 139, n. 8, p. 1-74, 11 jan. 2002. Acesso em: 20 ago. 2025

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. **Lei Geral de Proteção de Dados Pessoais (LGPD)**. Brasília, DF, 14 ago. 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm . Acesso em: 16 abr. 2025.

BRASIL. Superior Tribunal De Justiça. Recurso Especial n. 1192208/MG - Proc. 2010/0079120-5. Recorrente: Google Brasil Internet Ltda. Recorrido: Roberto Santos Barbieri. Relatora: Min. Nancy Andrighi. Brasília, DJe 02 ago. 2012. Disponível em: https://www.jusbrasil.com.br/jurisprudencia/stj/22209374. Acesso em: 19 mai. 2025.

BRASIL. Superior Tribunal de Justiça. Recurso Especial n° 2201694/SP - Proc. 2025/0081134-2. Recorrente: Jesse da Silva. Recorrido: Boa Vista Serviços S.A. Relator: Min. Ricardo Villas Bôas Cueva. Brasília, DJe 15 ago. 2025. Disponível em: <https://share.google/dzoS9eNmNUTOks3oy>. Acesso em: 12 set. 2025

BRASIL. Superior Tribunal de Justiça. Agravo em Recurso Especial n° 2.130.619/SP - Proc. 2022/0152262-2. Agravante: Eletropaulo Metropolitana Eletricidade De Sao Paulo S.A. Agravado: Maria Edite De Souza. Relator: Min. Francisco Falcão. Brasília, DJe 10 mar. 2023. Disponível em: https://www.jusbrasil.com.br/jurisprudencia/stj/1780119718 . Acesso em: 17 out. 2025.

BRASIL. Supremo Tribunal Federal. Medida Cautelar De Urgência Na Ação Direta De Inconstitucionalidade ADI 6.387. Requerente: Conselho Federal Da Ordem Dos Advogados Do Brasil. Relatora: Min. Rosa Weber. Brasília, DJe 12 nov. 2020. Disponível em: https://www.stf.jus.br/arquivo/cms/noticianoticiastf/anexo/adi6387mc.pdf. Acesso em: 02 ago. 2025.

BRASIL. Portal de Dados Abertos. Brasília, *[s.d]*. Disponível em: https://dados.gov.br/home. Acesso em: 01 nov. 2025.

BROWN, Tom B.; MANN, Benjamin; RYDER, Nick; SUBBIAH, Melanie; KAPLAN, Jared; DHARIWAL, Prafulla; NEELAKANTAN, Arvind; SHYAM, Pranav; SASTRY, Girish; ASKELL, Amanda; AGARWAL, Sandhini; HERBERT-VOSS, Ariel; KRUEGER, Gretchen; HENIGHAN, Tom; CHILD, Rewon; RAMESH, Aditya; ZIEGLER, Daniel M.; WU, Jeffrey; WINTER, Clemens; HESSE, Christopher; CHEN, Mark; SIGLER, Eric; LITWIN, Mateusz; GRAY, Scott; CHESS, Benjamin; CLARK, Jack; BERNER, Christopher; MCCANDLISH, Sam; RADFORD, Alec; SUTSKEVER, Ilya; AMODEI, Dario. ***Language models are few-shot learners.*** *Advances in neural information processing systems*, v. 33, p. 1877-1901, 2020. Disponível em: https://arxiv.org/abs/2005.14165v4 . Acesso em: 16 ago. 2025.

CALAZA, Tales. Evolução E Regulação Da Privacidade E Proteção De Dados No Contexto Da Internet Das Coisas No Cenário Brasileiro. **Revista do Centro Acadêmico Afonso Pena**. Belo Horizonte, ed. Vol. 29, N. 1, ano 2024, p. 1-18, 1 jul. 2024. Semestral. Disponível em: https://periodicos.ufmg.br/index.php/caap/article/download/52353/44983/202015 . Acesso em: 11 mai. 2025.

CALO, Ryan. *Robotics and the Lessons of Cyberlaw.* ***California Law Review***. Disponível em: https://digitalcommons.law.uw.edu/cgi/viewcontent.cgi?article=1022&context=faculty-articles. Acesso em: 7 nov. 2025.

CAO, Yihan; LI, Siyu; LIU, Yixin; YAN, Zhiling; DAI, Yutong; YU, Philip S.; SUN, Lichao. ***A comprehensive survey of AI-generated content (AIGC): a history of generative AI from GAN to ChatGPT.*** Cornell University. Disponível em: https://arxiv.org/abs/2303.04226v1. Acesso em: 9 nov. 2025.

CAPANEMA, Walter Aranha. A responsabilidade civil na Lei Geral de Proteção de Dados. **Cadernos Jurídicos - Direito Digital e proteção de dados pessoais**. São Paulo: Escola Paulista da Magistratura, ed. 53, 2020. Disponível em: https://share.google/T938kDc4ibvHFj3Q8. Acesso em: 12 set. 2025.

CARDOSO, João Victor Gontijo. **A responsabilidade civil objetiva e o dano moral presumido em decorrência do tratamento indevido de dados.** Percurso Acadêmico, Belo Horizonte, v. 10, n. 20, p. 132–154, jul./dez. 2020. Disponível em: https://share.google/iv79kEKOX5ZW1iNPV. Acesso em: 19 set. 2025.

CARPES, Artur Thompsen. **A Prova Do Nexo De Causalidade Na Responsabilidade Civil.** 2013. Tese. (Doutorado em Direito) - Faculdade de Direito, Universidade Federal do Rio Grande do Sul - UFRGS, Porto Alegre, Rio Grande do Sul. Orientador: Prof. Dr. Danilo Knijnik. Disponível em: https://lume.ufrgs.br/handle/10183/207209. Acesso em: 10 set. 2025.

CARRÁ, Bruno Leonardo Câmara. **É possível uma responsabilidade civil sem dano?**. Consultor Jurídico. Disponível em: https://www.conjur.com.br/2016-abr-25/direito-civil-atual-possivel-responsabilidade-civil-dano-ii/. Acesso em: 10 abr. 2025.

CARVALHO, Cesar Augusto Rodrigues de. **Autodeterminação informativa e sociedade de controle: um estudo sobre as relações entre liberdade e poder na era da informação.** Universidade de São Paulo. Faculdade de Direito, 2024. DOI: https://doi.org/10.11606/9786599682681 Disponível em: www.livrosabertos.abcd.usp.br/portaldelivrosUSP/catalog/book/1509 .

CARVALHO, Victor Miguel Barros de. **O Direito fundamental à privacidade ante a monetização de dados pessoais na internet: apontamentos legais para uma perspectiva regulatória.** 2018. Tese. (Mestrado em Direito) - Faculdade de Direito, Universidade Fedral do Rio Grande do Norte - UFRGN, Natal, Rio Grande do Norte. 145f. Dissertação (Mestrado em Direito) - Centro de Ciências Sociais Aplicadas, Universidade Federal do Rio Grande do Norte, Natal, 2018. Disponível em: https://repositorio.ufrn.br/jspui/handle/123456789/26851 . Acesso em: 13 ago. 2025.

CERKA, Paulius; GRIGIENÈ, Jurgita; SIRBIKYTÈ, Gintare. **Liability for damages caused by artificial intelligence.** Computer Law & Security Review, Kaunas, Lithuania: Elsevier, ed. Vol. 31, p. 376-389, Disponível em: http://dx.doi.org/10.1016/j.clsr.2015.03.008. Acesso em: 8 set. 2025.

CITRON, Danielle Keats; SOLOVE, Daniel J. Privacy Harms. **Boston University Law Review**, Boston, Massachusetts, v. 102:793, n° 3, 2022. Disponível em: https://share.google/utnUZuwQyF0UpKOB4. Acesso em: 05 out. 2025.

CIURIAK, Dan. **On economic strategy in the data-driven economy.** Social Science Research Network (SSRN). Disponível em: https://ssrn.com/abstract=3243962 . Acesso em: 11 maio 2025.

CHIRITA, Anca D. *The Rise of Big Data and the Loss of Privacy. In*: BAKHOUM, Mor; GALLEGO, Beatriz Conde; MACKENRODT, Mark-Oliver; SURBLYTĖ-NAMAVIČIENĖ, Gintaré (Eds.). ***Personal Data in Competition, Consumer Protection and Intellectual Property Law: Towards a Holistic Approach?.*** Berlim, Alemanha: Springer, 2018, ISBN: 978-3-662-57646-5. Disponível em:https://www.google.com/url?q=https://drive.google.com/file/d/1v0JS4TxI0YQGqMqtDlSwtPgufhJlMjkL/view?usp%3Dsharing&sa=D&source=docs&ust=1764768535327049&usg=AOvVaw1rtbBc9MbQn7ZjjK3BJcR2. Acesso em: 10 mai. 2025.

CHUI, Michael; HALL, Bryce; SINGLA, Alex; SUKHAREVSKY, Alexander; YEE, Lareina. **O estado da inteligência artificial em 2023: o ano do crescimento explosivo da IA Generativa.** McKinsey & Company. Disponível em: https://www.mckinsey.com.br/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year/pt-BR. Acesso em: 13 out. 2025.

COHEN, Julie E. ***Configuring the Networked Self: Law, Code, and the Play of Everyday Practice.*** New Haven; London: Yale University Press, 2012, EISBN 978-0-300-17793-0. E-book.

COOPER, A. Feder; CHOQUETTE-CHOO, Christopher A.; BOGEN, Miranda; KLYMAN, Kevin; JAGIELSKI, Matthew; FILIPPOVA, Katja; LIU, Ken; CHOULDECHOVA, Alexandra; HAYES, Jamie; HUANG, Yangsibo; TRIANTAFILLOU, Eleni; KAIROUZ, Peter; MITCHELL, Nicole Elyse; MIRESHGHALLAH, Niloofar; JACOBS, Abigail Z.; GRIMMELMANN, James; SHMATIKOV, Vitaly; DE SA, Christopher; SHUMAILOV, Ilia; TERZIS, Andreas; BAROCAS, Solon; VAUGHAN, Jennifer Wortman; BOYD, Danah; CHOI, Yejin; KOYEJO, Sanmi; DELGADO, Fernando; LIANG, Percy; HO, Daniel E.; SAMUELSON, Pamela; BRUNDAGE, Miles; BAU, David; NEEL, Seth; WALLACH, Hanna; CYPHERT, Amy B.; LEMLEY, Mark A.; PAPERNOT, Nicolas; LEE, Katherine. ***Machine Unlearning Doesn't Do What You Think: Lessons for Generative AI Policy, Research, and Practice.*** *Cornell University.* Disponível em: https://arxiv.org/abs/2412.06966. Acesso em: 15 out. 2025.

CORRÊA, Leonardo; CHO, Tae. **Responsabilidade civil na LGPD é subjetiva**. ln: Consultor Jurídico. OPINIÃO. [S.l.], 29 jan. 2021. Disponível em: https://www.conjur.com.br/2021-jan-29/correa-cho-responsabilidade-civil-lgpd-subjetiva/. Acesso em: 14 out. 2025.

COSTA, Albert França Josuá; CEBRIAN, Fabiana Faraco; ANJOS, Lucas Costa dos; GUEDES, Marcelo Santiago; MORAES, Thiago Guimarães. Inteligência Artificial Generativa. *In.:* **Agência Nacional de Proteção de Dados < radar tecnológico >**, n° 3.

Brasília, DF, 2024. Disponível em: https://share.google/bGLH2rmbfTSk5vdTe. Acesso em: 1 ago. 2025.

CRESPO, Marcelo Xavier de Freitas; SANTOS, Coriolano Aurélio de Almeida Camargo. Como será o futuro dos negócios com a vigência do Regulamento Geral de Proteção de Dados Europeu?. In: PINHEIRO, Patrícia Peck (coord.). **Direito digital aplicado 3.0**. São Paulo: Revista dos Tribunais, 2018. p. 179-182.

CUSTERS, Bart; CALDERS, Toon; SCHERMER, Bart. ***Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases.*** Research Gate. Disponível em: https://www.researchgate.net/publication/321596991_Discrimination_and_Privacy_in_the_Information_Society_Data_Mining_and_Profiling_in_Large_Databases. Acesso em: 12 mai. 2025.

DANTAS, Paulo Roberto de Figueiredo. **Direito Processual Constitucional.** 8. ed. São Paulo: Saraiva Educação, 2018. ISBN 9788547231064. Disponível em: https://doceru.com/doc/ss85c0x . Acesso em: 15 nov. 2025.

DEEPSEEK. DeepSeek. Política de Privacidade do DeepSeek. [S.l.]. DeepSeek, 2025. Disponível em: https://cdn.deepseek.com/policies/en-US/deepseek-privacy-policy.html. Acesso em: 3 set. 2025.

DEL MASTRO, André Menezes. A Função Punitivo-Preventiva Da Responsabilidade Civil. **Revista da Faculdade de Direito**, Universidade de São Paulo, São Paulo: USP, ed. 110, ano 2015, p. 765-817, Anual. Disponível em: https://share.google/nHlbTZscFLsor38kS. Acesso em: 6 out. 2025.

DE MORAES, Maria Celina Bodin. LGPD: um novo regime de responsabilização civil dito proativo. **Civilistica.com -** Revista Eletrônica de Direito Civil, Rio de Janeiro, v. 8, n. 3, p. 1–6, 2019. Disponível em: https://civilistica.emnuvens.com.br/redc/article/view/448. Acesso em: 4 out. 2025.

DIAS, Fernanda Rêgo Oliveira. Limites à utilização do consentimento como base legal adequada para o tratamento de dados pessoais. In: REQUIÃO, Maurício. (Org.). **Proteção de dados pessoais: novas perspectivas.** Salvador, Bahia: EDUFBA, 2022. p. 34-58. Disponível em: https://repositorio.ufba.br/handle/ri/35799. Acesso em: 08 set. 2025.

DONEDA, Danilo; ZANATTA, Rafael A. F. *Personality Rights in Brazilian Data Protection Law: A Historical Perspective. In*: ALBERS, Marion; SARLET, Info Wolfgang (Eds.). ***Personality and Data Protection Rights on the Internet. Ius Gentium: Comparative Perspectives on Law and Justice***. 01. ed. Suíça: Springer Charm, 2022. E-book.

DONEDA, Danilo. **A proteção de dados pessoais nas relações de consumo: para além da informação credíticia**. Brasília: Escola Nacional de Defesa do Consumidor, v. 2, 2010. 124 p. Disponível em: https://share.google/Ip2jZIz1qS5T2xxBy. Acesso em: 11 ago. 2025.

DONEDA, Danilo Cesar Maganhoto; MENDES, Laura Schertel; SOUZA, Carlos Affonso Pereira de; GOMES DE AN, Norberto Nuno Martin. Considerações iniciais sobre inteligência

artificial, ética e autonomia pessoal. **Pensar – Revista de Ciências Jurídicas**, Fortaleza: Conselho Superior de Editoração – UNIFOR, ed. 23, ano 2018, n. 4, p. 1-17, Anual. Disponível em: https://doi.org/10.5020/2317-2150.2018.8257. Acesso em: 9 out. 2025.

DUQUE-PEREIRA, Ives da Silva; MOURA, Sergio Arruda de. **Compreendendo A Inteligência Artificial Generativa Na Perspectiva Da Língua.** SciELO Preprints, [s. l.], 6 out. 2023 DOI: https://doi.org/https://doi.org/10.1590/SciELOPreprints.7077. Disponível em: https://preprints.scielo.org/index.php/scielo/preprint/view/7077. Acesso em: 13 ago. 2025.

EICK, Luciana Gemelli. **A função dissuasória da responsabilidade civil: fundamentos e proposta de aplicação no direito brasileiro**. 2020. Tese. (Doutorado em Direito) - Faculdade de Direito, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, Rio Grande do Su. Orientador: Prof. Dr. Adalberto de Souza Pasqualotto. Disponível em: https://tede2.pucrs.br/tede2/handle/tede/9212. Acesso em: 7 nov. 2025.

European Data Protection Board (EDPB). ***Guidelines 05/2020 on consent under Regulation 2016/679: Version 1.1***. EDPB. Disponível em: https://share.google/x0enzgnQIrJ0xwPK1. Acesso em: 19 ago. 2025.

*European Commission. **Opinion 03/2013 on purpose limitation.** European Comission*. Disponível em: https://share.google/uqA6b8yMZCnlfDY0s. Acesso em: 1 set. 2025.

European Data Protection Supervisor. **Synthetic Data**. Disponível em: https://www.edps.europa.eu/press-publications/publications/techsonar/synthetic-data_en. Acesso em: 15 set. 2025.

EHRHARDT JR, Marcos; MODESTO, Jéssica Andrade. Limites à utilização do consentimento como base legal adequada para o tratamento de dados pessoais. In: REQUIÃO, Maurício. (Org.). **Proteção de dados pessoais: novas perspectivas.** Salvador, Bahia: EDUFBA, 2022. p. 123-164. Disponível em: https://repositorio.ufba.br/handle/ri/35799. Acesso em: 08 set. 2025.

FAGUNDES, Marcos Geromini. Hermenêutica e Interpretação Jurídica: Distinções, Necessidades e Critérios Metodológicos. **Revista Jurídica da Amazônia**, Porto Velho, Brasil, v. 2, n. 1, p. 118–132, 2025. DOI: 10.63043/4ra9qe14. Disponível em: https://revista.mpro.mp.br/amazonia/article/view/102. Acesso em: 5 out. 2025.

FAJNGOLD, Leonardo. A reparação não pecuniária do dano moral coletivo e a tutela coletiva das vítimas de dano moral individual. **Civilistica.com** - Revista Eletrônica de Direito Civil, Rio de Janeiro, v. 11, n. 3, p. 1–23, 2022. Disponível em: https://civilistica.emnuvens.com.br/redc/article/view/831. Acesso em: 22 out. 2025.

FALCÓN, Maria Lucia de Oliveira. Desenvolvimento econômico e a transformação digital: processos e desafios. **Revista Eptic**, Aracaju, Sergipe, ed. Vol. 26, ano 2024, n. 3, p. 86-101, 1 nov. 2024. Disponível em: https://periodicos.ufs.br/eptic/article/download/21391/16395/70799 . Acesso em: 12 mai. 2025.

FALEIROS JÚNIOR, José Luiz de Moura. **Incidentes de segurança multitudinários e a reparação fluida (fluid recovery) na LGPD**. ln: Migalhas. Migalhas de Responsabilidade Civil. [S. l.], 16 mai. 2023. Disponível em: https://www.migalhas.com.br/coluna/migalhas-de-responsabilidade-civil/386481/incidentes-de-seguranca-multitudinarios-e-a-reparacao-fluida-na-lgpd. Acesso em: 12 out. 2025.

FAMA, Josué Sá. Inteligência artificial e privacidade: implicações legais e éticas na era digital. **Revista Científica Multidisciplinar Núcleo do Conhecimento**. Ano. 10, Ed. 01, Vol. 01, pp. 15-39. Janeiro de 2024. ISSN: 2448-0959, Link de acesso: https://www.nucleodoconhecimento.com.br/tecnologia/implicacoes-legais, DOI: 10.32749/nucleodoconhecimento.com.br/tecnologia/implicacoes-legais . Acesso em: 25 ago. 2025.

FERNANDES, Mariane Santos. Elementos Da Responsabilidade Civil. **Revista Hórus**, [S. l.], v. 6, n. 01, p. 9–15, 2022. Disponível em: https://estacio.periodicoscientificos.com.br/index.php/revistahorus/article/view/983. Acesso em: 06 set. 2025.

FERREIRA, Gabriel Luis Bonora Vidrih; SILVA, Solange Teles da. Análise dos fundamentos da compensação ambiental: a responsabilidade civil *ex ante* no direito brasileiro. **Revista de Informação Legislativa**, Brasília, v. 44, n. 175, p. 125-137, jul./set. 2007. Disponível em: <https://bdjur.stj.jus.br/handle/2011/174045>. Acesso em: 10 abr. 2025.

FERREIRA, Keila Pacheco. **Responsabilidade Civil Preventiva: Função, Pressupostos E Aplicabilidade**. 2014. (Tese). (Doutorado em Direito) - Faculdade de Direito, Universidade de São Paulo - USP, São Paulo. Orientador: Dra. Teresa Ancona Lopez. Disponível em: https://www.teses.usp.br/teses/disponiveis/2/2131/tde-27102016-092601/pt-br.php. Acesso em: 4 ago. 2025.

FLORENCE, Tatiana Magalhães. Apontamentos sobre a responsabilidade civil no tratamento de dados. **Revista Brasileira de Direito Civil – RBDCivil**, Belo Horizonte, v. 30, p. 223-235, out./dez. 2021. DOI: 10.33242/rbdc.2021.04.010. Disponível em: https://rbdcivil.ibdcivil.org.br/rbdc/article/view/527 . Acesso em: 6 out. 2025.

FLORIDI, Luciano; CHIRIATTI, Massimo. ***GPT-3: Its Nature, Scope, Limits, and Consequences.*** *Research Gate*. Disponível em: https://www.researchgate.net/publication/346490743_GPT-3_Its_Nature_Scope_Limits_and_Consequences. Acesso em: 15 nov. 2025.

FONSECA, Aline Klayse. Delineamentos jurídico-dogmáticos da inteligência artificial e seus impactos no instituto da responsabilidade civil. **Civilistica.com - Revista Eletrônica de Direito Civil,** Rio de Janeiro, v. 10, n. 2, p. 1–36, 2021. Disponível em: https://civilistica.emnuvens.com.br/redc/article/view/671. Acesso em: 06 set. 2025.

FOO, Lin Geng; RAHMANI, Hossein; LIU, Jun. **Ai-Generated Content (Aigc) For Various Data Modalities: A Survey.** Cornell University. Disponível em: https://arxiv.org/abs/2308.14177. Acesso em: 10 nov. 2025.

FORNASIER, Mateus de Oliveira; KNEBEL, Norberto Milton Paiva; SILVA, Fernanda Viero da. Mineração De Dados E Publicidade Comportamental: Impasses para a regulação do spam e dos nudges na sociedade burocrática do consumo dirigido. **Rei - Revista Estudos Institucionais**, [S. l.], v. 6, n. 3, p. 1536–1559, 2020. DOI: 10.21783/rei.v6i3.506. Disponível em: https://www.estudosinstitucionais.com/REI/article/view/506. Acesso em: 8 ago. 2025.

FORNASIER, Mateus de Oliveira; KNEBEL, Norberto Milton Paiva. **O titular de dados como sujeito de direito no capitalismo de vigilância e mercantilização dos dados na Lei Geral de Proteção de Dados.** Scielo Brasil - Revista Direito e Práxis. Disponível em: https://doi.org/10.1590/2179-8966/2020/46944 . Acesso em: 12 mai. 2025.

FORGIONI, Paula A.. OS FUNDAMENTOS DO ANTITRUSTE. 10. ed. São Paulo: Revista dos Tribunais, 2018.

FRAZÃO, Ana. **Nova LGPD: a importância do consentimento para o tratamento dos dados pessoais: A terceira parte de uma série sobre as repercussões para a atividade empresarial.** Disponível em: https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/nova-lgpd-a-importancia-do-consentimento-para-o-tratamento-dos-dados-pessoais. Acesso em: 19 ago. 2025.

FREITAS, Mário Gamaliel Guazzeli de. **Função preventiva da responsabilidade: Limites para a discussão de uma responsabilidade sem dano**. Tese. 2019. (Mestrado em Direito) - Faculdade de Direito, Universidade de São Paulo - USP, São Paulo. Orientador: Dr. Francisco Paulo De Crescenzo Marino. Disponível em: https://www.teses.usp.br/teses/disponiveis/2/2131/tde-20082020-174435/pt-br.php. acesso em: 6 out. 2025.

FROTA, Hidemberg Alves da; BIÃO, Fernanda Leite. A dimensão existencial da pessoa humana, o dano existencial e o dano ao projeto de vida: reflexões à luz do direito comparado. **Cadernos da Escola de Direito**, v. 2, n. 13, 19 jun. 2017.Disponível em: https://portaldeperiodicos.unibrasil.com.br/index.php/cadernosdireito/article/view/2688 Acesso em: 5 out. 2025.

FUCHS, Christian; SEVIGNANI, Sebastian. What is Digital Labour? What is Digital Work? What's their Difference? And why do these Questions Matter for Under- standing Social Media?. tripleC: Communication, Capitalism & Critique., Austria, ed. Vol. 11, ano 2013, n. 2, p. 237-293, Disponível em: https://doi.org/10.31269/triplec.v11i2.461 . Acesso em: 13 mai. 2025.

GALLESE, Chiara. *Web scraping and Generative Models training in the Directive 790/19.* **Rivista di Scienze Giuridiche, Scienze Cognitive ed Intelligenza Artificiale,** v. 16, n. 2. Itália, 2023. Disponível em: https://doi.org/10.6092/issn.1825-1927/18871. Acesso em: 9 nov. 2025.

GARCIA, Maria Carolina Brunharotto; NUNES, Paula Freire Santos Andrade. Responsabilidade civil, dano moral e tratamento de dados pessoais: estudo prático de jurisprudência sobre como se dará o dever de indenizar. 3 ago. 2021. Disponível em: https://share.google/MXkAYcnQCQsFBxb8G. Acesso em: 19 set. 2025.

GARCIA, Ranieri Rodrigues. **Reconstrução E Procedimento: Habermas E A Crítica Do Direito Moderno.** Tese. 2025. (Doutorado em Filosofia) - Faculdade de Filosofia, Universidade Federal do Rio Grande do Sul - UFRGS, Rio Grande do Sul. Orientador: Prof. Dr. Felipe Gonçalves Silva. Disponível em: https://lume.ufrgs.br/handle/10183/294343. Acesso em: 20 out. 2025.

GHANI, Najua Samir Asad. **A Crise da Responsabilidade Civil e a Responsabilidade Civil sem dano.** Tese. 2017. (Pós-Graduação em Direito) - Instituto Brasiliense de Direito Público - IBDP, Brasília. Orientador: Prof. MSc. Adisson Taveira Rocha Leal. Disponível em: https://repositorio.idp.edu.br/handle/123456789/2286. acesso em: 6 out. 2025.

GIL, Arilson Garcia; RODRIGUES, Maurício Andreiuolo. A revolução da tecnologia e seu impacto na Hermenêutica Constitucional: Um Estudo de Caso Stf V. Telegram, Liberdade e Verdade. **HomaPublica- Revista Internacional de Derechos Humanos y Empresas**. p. 2526-0774. v.7, no.1 e:0111, 2023. Disponível em: https://periodicos.ufjf.br/plugins/generic/pdfJsViewer/pdf.js/web/viewer.html?file=https%3A%2F%2Fperiodicos.ufjf.br%2Findex.php%2FHOMA%2Farticle%2Fdownload%2F41490%2F26806%2F185683 . Acesso em 20 de out. 2025.

GÓES, Bárbara Veiga; D'ALBUQUERQUE, Teila Rocha Lins. Responsabilidade civil no descumprimento da nova Lei Geral de Proteção de Dados Pessoais (Lei no 13.709/2018). In: REQUIÃO, Maurício. (Org.). **Proteção de dados pessoais: novas perspectivas**. Salvador, Bahia: EDUFBA, 2022. p. 427-458. Disponível em: https://repositorio.ufba.br/handle/ri/35799. Acesso em: 08 set. 2025.

GONÇALVES, Carlos Roberto. **Responsabilidade Civil.** 23. ed. São Paulo: SaraivaJur, 2024. E-book.

GONDIM, G. G. A responsabilidade civil no uso indevido dos dados pessoais. **Revista IBERC**, Belo Horizonte, v. 4, n. 1, p. 19–34, 2021. DOI: 10.37963/iberc.v4i1.140. Disponível em: https://revistaiberc.responsabilidadecivil.org/iberc/article/view/140. Acesso em: 06 set. 2025.

GOOGLE. **Central de Privacidade dos apps do Gemini.** Google. Disponível em: https://support.google.com/gemini/answer/13594961?hl=pt-BR#zippy=%2Cquais-s%C3%A3o-as-bases-jur%C3%ADdicas-do-google-para-o-processamento-de-d.... Acesso em: 3 set. 2025.

GROSSI, Bernardo Menicucci. **O desafio da regulação dos dados pessoais: entre a autonomia e a heteronomia.** 2022. Tese. (Doutorado em Direito) - Faculdade de Direito, Pontifícia Universidade Católica de Minas Gerais, Belo Horizonte. Orientador: Prof. Dr. Leonardo Macedo Poli. Disponível em: https://www.researchgate.net/publication/367075319_O_desafio_da_regulacao_de_dados_pessoais_entre_autonomia_e_heteronomia. Acesso em: 11 out. 2025.

GROTHAUS, Michael. **O que é "corpus" e por que todo mundo no ramo da IA está falando sobre isso: Bill Gates, o CEO do Reddit e outros líderes de tecnologia têm falado cada vez mais sobre "corpus". Mas o que significa esse termo?**. ln: Fast Company Brasil.

IA. [S.l.]. 5 jul. 2023. Disponível em: https://fastcompanybrasil.com/tech/inteligencia-artificial/o-que-e-corpus-e-por-que-todo-mundo-no-ramo-da-ia-esta-falando-sobre-isso/. Acesso em: 12 ago. 2025.

HAN, Byung-Chul. **No enxame: Perspectivas do digital.** Tradução: Lucas Machado. Petrópolis: Editora Vozes, 2018. 135 p. (2ª reimpressão, 2019.). Título original: Im Schwarm - Ansichten des Digitalen. ISBN: 978-85-326-5851-7.

HATOUM, Nida Saleh; COLOMBO, Maici Barboza dos Santos. Da necessidade de identificação do dano existencial na responsabilidade civil. **Civilistica.com - Revista Eletrônica de Direito Civil**, Rio de Janeiro, v. 11, n. 3, p. 1–19, 2022. Disponível em: https://civilistica.emnuvens.com.br/redc/article/view/701. Acesso em: 17 out. 2025.

HILDEBRANDT, Mireille. **Law for Computer Scientists and Other Folk.** Nova York: Oxford University Press, 2020. 341 p. ISBN: 978–0–19–886088–4. Disponível em: https://share.google/YEuQkqEWGkfDdmtTm. Acesso em: 17 set. 2025.

HIRONAKA, Giselda Maria Fernandes Novaes. Tendências atuais de responsabilidade civil: marcos teóricos para o direito do século XXI. **Revista Brasileira de Direito Comparado**, n. 19, p. 189-206, 2001. Disponível em: http://www.idclb.com.br/revistas/19/revista19%20(16).pdf. Acesso em: 6 out. 2025.

HIRONAKA, Giselda Maria Fernandes Novaes. Responsabilidade pressuposta: evolução de fundamentos e de paradigmas da responsabilidade civil na contemporaneidade. **Revista IBERC**, Belo Horizonte, v. 6, n. 1, p. 139–161, 2023. DOI: 10.37963/iberc.v6i1.253. Disponível em: https://revistaiberc.responsabilidadecivil.org/iberc/article/view/253.

HONG, Rachel; HUTSON, Jevan; AGNEW, William; HUDA, Imaad; KOHNO, Tadayoshi; MORGENSTERN, Jamie. **A Common Pool of Privacy Problems: Legal and Technical Lessons from a Large-Scale Web-Scraped Machine Learning Dataset**. Cornell University. Disponível em: https://arxiv.org/abs/2506.17185v1. Acesso em: 15 out. 2025.

HOMMA, Fernanda Lissa Fujiwara . Execuções Judiciais Pecuniárias De Processos Coletivos: Entre A Fluid Recovery, A Cy Pres E Os Fundos. **Revista Eletrônica de Direito Processual**, Rio de Janeiro, v. 18, n. 2, 2017. DOI: 10.12957/redp.2017.28306. Disponível em: https://www.e-publicacoes.uerj.br/redp/article/view/28306.Acesso em: 23 out. 2025.

HSING, Chen Wen. **Coleta de dados pessoais e paradoxo da privacidade: Um estudo entre usuários de aplicativos móveis.** 2016. Tese. (Doutorado em Economia) - Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo - USP, São Paulo. Orientador: Prof. Dr. Cesar Alexandre de Souza. Disponível em: https://teses.usp.br/teses/disponiveis/12/12139/tde-03062016-124303/publico/CorrigidaChen.pdf. Acesso em: 10 set. 2025.

IOSCOTE, Fabia. Ia Generativa: Potencial E Limitações Dos Large Language Models E Prompts Na Produção De Notícias. **Revista UNINTER de Comunicação,** [S. l.], v. 12, n. 20, p. 59–80, 2025. DOI: 10.21882/ruc.v12i20.985. Disponível em: https://www.revistasuninter.com/revistacomunicacao/index.php/revista/article/view/985. Acesso em: 8 nov. 2025.

KAPILAN, Prasanth; KANTOR, Susan; KALLENBACH, Paul. **AI and scraped data: Data protection implications: We explore the data protection implications that arise when data scraping is conducted for the purpose of training artificial intelligence tools.** ln: MinterEllison. 12 mar. 2024. Disponível em: https://www.minterellison.com/articles/ai-and-scraped-data-data-protection-implications. Acesso em: 27 ago. 2025.

KONDRUP, Emma. **Informed Consent, Redefined: How AI and Big Data Are Changing the Rules.** ln: The Petrie-Flom Center - Health Law Policy, Biotechnology, and Bioethics at Harvard Law School. Artificial Intelligence. Harvard Law School, 23 Everett Street, Cambridge, MA 02138, 11 abr. 2025. Disponível em: https://petrieflom.law.harvard.edu/2025/04/11/informed-consent-redefined-how-ai-and-big-data-are-changing-the-rules/. Acesso em: 19 ago. 2025.

KREIMER, Seth F. **'Spooky Action at a Distance': Intangible Injury in Fact in the Information Age**. University of Pennsylvania Journal of Constitutional Law, Philadelphia, Pennsylvania: University of Pennsylvania Law School, ed. 18, ano 2016, n. 3, p. 745-792, Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2661018. Acesso em: 4 out. 2025.

KRELL, Andreas J.; PAIVA, Raíi Moraes Sampaio de. Hermenêutica jurídica e uso deficiente de métodos no contexto da aplicação do direito no Brasil. **Revista Brasileira de Direitos Fundamentais & Justiça**, [S. l.], v. 11, n. 37, p. 185–218, 2017. DOI: 10.30899/dfj.v11i37.128. Disponível em: https://dfj.emnuvens.com.br/dfj/article/view/128. Acesso em: 20 out. 2025.

KRUGER, Daniel. **Social media copies gambling methods 'to create psychological cravings'.** ln: Institute for Healthcare Policy & Innovation - University of Michigan. News & Briefs . [S.l.]. 8 mai. 2018. Disponível em: https://ihpi.umich.edu/news/social-media-copies-gambling-methods-create-psychological-cravings. Acesso em: 13 ago. 2025.

LA DIEGA, Guido Noto; BEZERRA, Leonardo C. T. Can there be responsible AI without AI liability? Incentivizing generative AI safety through ex-post tort liability under the EU AI liability directive. **International Journal of Law and Information Technology**, [s.l.], v. 32, eaae021, 2024. Disponível em: https://doi.org/10.1093/ijlit/eaae021. Acesso em: 10 abr. 2025.

LA DIEGA, Guido Noto; HARBINJA, Edina; NOLAN, Katherine. **BILETA's Response to ICO's Generative AI First Call for Evidence: The Lawful Basis for Web Scraping to Train Generative AI Models.** SSRN - Social Science Research Network, [s. l.], p. 1-5, 20 fev. 2024 DOI: https://doi.org/http://dx.doi.org/10.2139/ssrn.4814018. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4814018. Acesso em: 15 ago. 2025.

LEAL, Mônia Clarissa Hennig; PAULO, Lucas Moreschi. A Lei Geral de Proteção de Dados, a vulnerabilidade dos usuários da internet e a tutela dos direitos: linhas introdutórias à dinâmica dos dados, do Big Data, da economia de dados e da discriminação algorítmica. **Civilistica.com - Revista Eletrônica de Direito Civil**, Rio de Janeiro, v. 12, n. 3, p. 1–30,

2023. Disponível em: https://civilistica.emnuvens.com.br/redc/article/view/911. Acesso em: 12 out. 2025.

LEAL , Rodrigo de Lima; GARBACCIO, Grace Ladeira; MALLMANN, Jean A responsabilidade civil no contexto da inteligência artificial: perspectivas comparadas entre Brasil e Portugal (2023-2024). **Revista Da Agu**, [S. l.], v. 23, n. 4, 2024. DOI: 10.25109/2525-328X.v.23.n.4.2024.3551. Disponível em: https://revistaagu.agu.gov.br/index.php/AGU/article/view/3551. Acesso em: 07 set. 2025.

Legal Information Institute. But-for test. Cornell Law School. Disponível em: https://www.law.cornell.edu/wex/but-for_test. Traduzido pelo Google. Acesso em: 01 out. 2025.

LI, Yiming; SHAO, Shuo; HE, Yu; GUO, Junfeng; ZHANG, Tianwei; QIN, Zhan; CHEN, Pin-Yu; BACKES, Michael; TORR, Philip; TAO, Dacheng; REN, Kui. **Rethinking Data Protection in the (Generative) Artificial Intelligence Era.** The State Key Laboratory of Blockchain and Data Security, Zhejiang University; Nanyang Technological University; University of Maryland; IBM Research; CISPA Helmholtz Center for Information Security; University of Oxford, 3 set. 2025. Disponível em: https://arxiv.org/pdf/2507.03034. Acesso em: 20 out. 2025.

LIMA, Cintia Rosa Pereira de. A Proteção De Dados Pessoais No Contexto Da Economia Informacional: Desafios Regulatórios Do Marketing Comportamental. **Revista Eletrônica do Curso de Direito da UFSM**, [S. l.], v. 16, n. 2, p. e64767, 2021. DOI: 10.5902/1981369464767. Disponível em: https://periodicos.ufsm.br/revistadireito/article/view/e64767. Acesso em: 22 out. 2025.

LIMA JÚNIOR, Francisco de Assis Ferreira; RODRIGUES, Ian de Almeida Bispo; MORAES, Yan Gomes. Responsabilidade civil em casos de violação da LGPD por algoritmos de inteligência artificial. **Revistaft RA - Ciências Sociais Aplicadas**, v. 29, n. 146, maio 2025. Orientação: Ingryd Stéphanye Monteiro de Souza. Disponível em: https://doi.org/10.69849/revistaft/ch10202505161613.. Acesso em: 24 set. 2025.

LINTVEDT, Mona Naomi. **Putting a price on data protection infringement.** International Data Privacy Law, Oxford, v. 12, n. 1, p. 1–15, fev. 2022. Disponível em: https://ssrn.com/abstract=4283877. DOI: https://doi.org/10.1093/idpl/ipab024. Acesso em: 4 out. 2025.

LOPES, Alexandra Krastins; VARGAS, Andressa Girotto; OLIVEIRA, Davi Téofilo Nunes de; MAIOLINO, Isabela; BARBOSA, Jeferson Dias; CARVALHO, Lucas Borges de; GUEDES, Marcelo Santiago; MORAES, Thiago. Guia Orientativo: Cookies e proteção de dados pessoais. *In.:* **Agência Nacional de Proteção de Dados Pessoais (ANPD).** Brasília, DF, 2022. Disponível em: https://www.gov.br/anpd/pt-br/centrais-de-conteudo/materiais-educativos-e-publicacoes/guia-orientativo-cookies-e-protecao-de-dados-pessoais.pdf. Acesso em: 7 ago. 2025.

LOPES, Giovana Figueiredo Peluso. INTELIGÊNCIA ARTIFICIAL (IA): Considerações sobre personalidade, imputação e responsabilidade. Orientador: Dr. Brunello Souza Stancioli. 2020. 148 f. TCC (Especialização) - Curso de Mestrado em Direito, UNIVERSIDADE

FEDERAL DE MINAS GERAIS, Belo Horizonte, 2020. Disponível em: https://hdl.handle.net/1843/34056. acesso em: 11 set. 2025.

LOSCHI, Marília. **Internet chega a 74,9 milhões de domicílios do país em 2024.** Agência IBGE Notícias - PNAD Contínua. Disponível em: https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/44031-internet-chega-a-74-9-milhoes-de-domicilios-do-pais-em-2024. Acesso em: 8 out. 2025.

LUNDQVIST, Björn. *Big Data, Open Data, Privacy Regulations, Intellectual Property and Competition Law in an Internet-of-Things World: The Issue of Accessing Data. In*: BAKHOUM, Mor; GALLEGO, Beatriz Conde; MACKENRODT, Mark-Oliver; SURBLYTĖ-NAMAVIČIENĖ, Gintaré (Eds.). ***Personal Data in Competition, Consumer Protection and Intellectual Property Law: Towards a Holistic Approach?.*** Berlim, Alemanha: Springer, 2018, ISBN: 978-3-662-57646-5. Disponível em: https://doi.org/10.1007/978-3-662-57646-5. Acesso em: 10 mai. 2025.

LUZ, Solimar. Cresce para 20,5% a população com ensino superior no Brasil: A maioria dos graduados frequentou ensino médio em escola pública. *ln:* Agência Brasil - Empresa Brasil de Comunicação (EBC). **Educação.** Brasília, 13 jun. 2025. Disponível em: https://agenciabrasil.ebc.com.br/radioagencia-nacional/educacao/audio/2025-06/cresce-populacao-com-ensino-superior-no-brasil. Acesso em: 11 nov. 2025.

MACARIO, Michel. **LLMs: determinísticos ou estocásticos?**. ln: Medium. -. [S. l.], 25 out. 2024. Disponível em: https://medium.com/@michel.macario/llms-determin%C3%ADsticos-ou-estoc%C3%A1sticos-7658da42e971. Acesso em: 1 out. 2025.

MACHADO, Diego. Considerações iniciais sobre o conceito de dado pessoal no ordenamento jurídico brasileiro. **Civilistica.com - Revista Eletrônica de Direito Civil.** Rio de Janeiro, a. 12, n. 1, 2023. Disponível em: <http://civilistica.com/consideracoes-iniciais-sobre-o-conceito/>. Acesso em 04 ago. 2025.

MACHADO, Fernando Inglez de Souza; RUARO, Regina Linden. Publicidade Comportamental, Proteção De Dados Pessoais E O Direito Do Consumidor. Conpendi Law Review, Braga, v. 3. n. 2, p. 421-440, jul/dez 2017. Disponível em: https://repositorio.pucrs.br/dspace/bitstream/10923/11550/2/PUBLICIDADE_COMPORTAMENTAL_PROTECAO_DE_DADOS_PESSOAIS_E_O_DIREITO_DO_CONSUMIDOR.pdf. Acesso em: 10 mai. 2025.

MAIA, Bruno Alberto; NEME, Eliana Franco; CALISSI, Jamile. A proteção de dados e o princípio da dignidade humana: uma compreensão acerca da autodeterminação informativa. **Revista do Direito Público**, Londrina, v. 19, n. 1, p. 201, abr. 2024. DOI: https://ojs.uel.br/revistas/uel/index.php/direitopub/article/view/46952 . Acesso em: 16 out. 2025.

MALIK, Omer Imran. **Consentimento *Opt In* vs. Consentimento *Opt Out*: Qual é a diferença?**. ln: Securiti. Data Consent Automation. [S.l.]. 13 ago. 2023. Disponível em: https://securiti.ai/pt-br/blog/opt-in-vs-opt-out/. Acesso em: 27 ago. 2025.

MANTUANI, Matheus. Inteligência Artificial e Proteção de Dados Pessoais: reflexões sobre a base legal adequada para fundamentar o treinamento via web scraping. In: BRANCO, Sérgio (Coord.). **Inteligência Aritificial e Sociedade Conectada**. Rio de Janeiro: ITS - Instituto de Tecnologia E Sociedade, 2025, p. 35-55. Disponível em: https://share.google/PQfYQHt02HNurGrIu. Acesso em: 6 mai. 2025.

MARGONI, Thomas. ***TDM and generative AI: lawful access and opt-outs.*** Social Science Research Network (SSRN). Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5036164. Acesso em: 8 nov. 2025.

MARQUES, Glauco Lauria; MORESI, Eduardo Amadeu Dutra. Framework de adequação de bancos de dados legados à Lei Geral de Proteção de Dados Pessoais (LGPD): um estudo para órgãos públicos brasileiros. **Revista do Serviço Público**, [S. l.], v. 75, n. 4, p. 830-856, 2024. DOI: 10.21874/rsp.v75i4.10941. Disponível em: https://revista.enap.gov.br/index.php/RSP/article/view/10941. Acesso em: 8 ago. 2025.

MARTINS, Guilherme Magalhães; LONGHI, João Victor Rozatti. Nota introdutória: Os desafios contemporâneos do direito privado na transformação digital. *In:* MARTINS, Guilherme Magalhães; LONGHI, João Victor Rozatti (Coords.). **Direito digital: direito privado e internet.** São Paulo: Editora Foco, 2025, p. 1.015. Disponível em: https://konektacommerce.nyc3.cdn.digitaloceanspaces.com/TEXT_SAMPLE_CONTENT/direito-digital-direito-privado-e-internet-6a-ed-2025-270617-1.pdf . Acesso em 6 ago. 2025.

MATTHIAS, Andreas. *The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata.* ***Ethics and Information Technology***, v. 6, n. 3, p. 175-183, 2004. Disponível em: https://eclass.uoa.gr/modules/document/file.php/PHILOSOPHY875/%CE%9C%CE%A0%CE%9F%CE%A5%CE%A4%CE%9B%CE%91%CE%A3%20%CE%9F%CE%A0%CE%9B%CE%91%20%CE%9C%CE%91%CE%96%CE%99%CE%9A%CE%97%CE%A3%20%CE%9A%CE%91%CE%A4%CE%91%CE%A3%CE%A4%CE%A1%CE%9F%CE%A6%CE%97%CE%A3%20DRONES/The%20responsibility%20gap%20Ascribing%20responsibility%20for%20the%20actions%20of%20learning%20matthias2004.pdf. Acesso em: 7 nov. 2025.

MATO GROSSO. Tribunal de Justiça do Estado de Mato Grosso. Agravo de Instrumento n° 1034362-91.2024.8.11.0000. Quarta Câmara de Direito Privado. Relator: Des. Serly Marcondes Alves. Quarta Câmara de Direito Privado. Mato Grosso, Dje: 02 maio 2025. Disponível em: https://www.jusbrasil.com.br/jurisprudencia/tj-mt/4822351402. Acesso em: 01 nov. 2025.

MCCALLUM, Shiona. **ChatGPT foi banido na Itália devido a preocupações com a privacidade.** ln: BBC. Technology. [S. l.], 1 abr. 2023. Disponível em: https://www.bbc.com/news/technology-65139406. Acesso em: 28 out. 2025.

MEIRELES, Adriana Veloso. Privacidade no século 21: proteção de dados, democracia e modelos regulatórios. **Revista Brasileira de Ciência Política**, n. 41, p. 1-35, 2023. DOI: 10.1590/0103-3352.2023.41.265909. Aprovado em: 04 maio 2023. Disponível em: https://www.scielo.br/j/rbcpol/a/my3M8sH3tfpm4WmXhrNcMjK/?format=html&lang=pt . Acesso em: 10 out. 2025.

MENDES, Laura Schertel; FONSECA, Gabriel C. Soares da. Proteção de dados pessoais e regulação: uma análise institucional. **Revista Estudos Institucionais**, v. 6, n. 2, p. 507-533, maio/ago. 2020. DOI: 10.21783/rei.v6i2.521. Disponível em: https://www.estudosinstitucionais.com/REI/article/view/521/510. Acesso em: 11 out. 2025.

MÉO, Rodrigo Amaral Paula de. **Inteligência Artificial: Reflexos Na Responsabilidade Civil.** 2022. Tese. (Doutorado em Direito) - Faculdade de Direito, Universidade de São Paulo - USP. Orientadora: Profª. Dra. Silmara Juny de Abreu Chinellato. Disponível em: https://www.teses.usp.br/teses/disponiveis/2/2131/tde-28052024-130641/pt-br.php. Acesso em: 10 set. 2025.

MICROSOFT. Learn.microsoft. **Dados, privacidade e segurança para o Microsoft 365 Copilot.** [S.l.]. © Microsoft 2025, 2025. Disponível em: https://learn.microsoft.com/pt-br/copilot/microsoft-365/microsoft-365-copilot-privacy. Acesso em: 3 set. 2025.

MONTENEGRO, Manuel Carlos. **Dignidade humana está na origem da autodeterminação da LGPD, afirma Fux**. ln: Portal CNJ - Conselho Nacional de Justiça. Agência CNJ de Notícias. [S.l.], 22 nov. 2021. Disponível em: https://www.cnj.jus.br/dignidade-humana-esta-na-origem-da-autodeterminacao-informativa-da-lgpd-afirma-fux/. Acesso em: 15 out. 2025.

MONTINI, Alessandra. **Desigualdade digital: o papel da IA na redução ou ampliação das diferenças sociais: Segundo levantamento, pessoas com 60 anos ou mais e pessoas negras são os maiores percentuais sem acesso à internet no país.** ln: FebranTech. Educação. [S.l.]. 12 ago. 2024. Disponível em: https://febrabantech.febraban.org.br/especialista/alessandra-montini/desigualdade-digital-o-papel-da-ia-na-reducao-ou-ampliacao-das-diferencas-sociais. Acesso em: 2 set. 2025.

MORAES, Maria Celina Bodin de. LGPD: um novo regime de responsabilização civil dito "proativo". **Civilistica.com.** Rio de Janeiro: a. 8, n. 3, 2019. Disponível em: <http://civilistica.com/lgpd-um-novo-regime/>. Acesso em: 15 set. 2025.

MORAES, Maria Celina Bodin de; QUEIROZ, João Quinelato de. Autodeterminação informativa e responsabilização proativa: novos instrumentos de tutela da pessoa humana na LGDP. In: ADENAUER, Cadernos. (Org.). PROTEÇÃO DE DADOS PESSOAIS: PRIVACIDADE VERSUS AVANÇO TECNOLÓGICO. 3 ed. Rio de Janeiro, Rio de Janeiro: Anja Czymmeck, 2019. p. 113-135, Disponível em: https://share.google/t8XBmZDx5wnwAAI9z . Acesso em: 15 set. 2025.

MOROZOV, Evgeny. **Big Tech: a ascensão dos dados e a morte da política.** Trad. Claudio Marcondes. São Paulo: Ubu Editora, 2018. E-book.

MOZETIC, Vinícius. **A Hermenêutica Jurídica (Crítica) Da Tecnologia Pós-Moderna Como Resposta Para O Problema Da Compreensão, Interpretação E Aplicação Do Direito.** 2016. Tese. (Doutorado em Direito) - Faculdade de Direito, Universidade do Vale do Rio dos Sinos – UNISINOS, São Leopoldo. Orientador: Prof. Dr. Lênio Luiz Streck. Disponível em: https://repositorio.jesuita.org.br/handle/UNISINOS/6009. Acesso em: 6 out. 2025.

MULHOLLAND, Caitlin. Dados pessoais sensíveis e a tutela de direitos fundamentais: uma análise à luz da lei geral de proteção de dados (Lei 13.709/18). **Revista de Direitos e Garantias Fundamentais**, [S. l.], v. 19, n. 3, p. 159–180, 2018. DOI: 10.18759/rdgf.v19i3.1603. Disponível em: https://sisbib.emnuvens.com.br/direitosegarantias/article/view/1603. Acesso em: 7 ago. 2025.

MULHOHAND, Caitlin; KERNER, Bianca. Responsabilidade civil por danos causados pela violação do princípio da igualdade no tratamento de dados pessoais. In: TEPEDINO, Gustavo; SILVA, Rodrigo da Guia. (Coords.). **O Direito Civil Na Era Da Inteligência Artificial**. 1 ed. São Paulo: Revista dos Tribunais, 2020. p. 565-584. Disponível em: https://www.academia.edu/44242325/Responsabilidade_Civil_por_danos_causados_pela_violação%C3%A7%C3%A3o_do_princ%C3%ADpio_da_igualdade_no_tratamento_de_dados_pessoais. Acesso em: 17 set. 2025.

NETO, Elias Jacob de Menezes; MORAIS, José Luis Bolzan de; BEZERRA, Tiago José de Souza Lima. O projeto de lei de proteção de dados pessoais (PL 5276/2016) no mundo do big data: o fenômeno da dataveillance em relação à utilização de metadados e seu impacto nos direitos humanos. **Revista Brasileira de Políticas Públicas**, v. 7, n. 3, p. 184-198, 2017. Disponível em: https://www.rel.uniceub.br/RBPP/article/viewFile/4840/3636. DOI: 10.5102/rbpp.v7i3.4840. Acesso em: 8 set. 2025.

NOVA. **Política de Privacidade**. Disponível em: https://novaapp.ai/privacy. Acesso em: 3 set. 2025.

OECD. **AI Language Models: Technological, Socio-Economic and Policy Considerations.** OECD Digital Economy Papers, n. 352, Abr. 2023. Disponível em: https://www.oecd.org/content/dam/oecd/en/publications/reports/2023/04/ai-language-models_46d9d9b4/13d38f92-en.pdf. Acesso em: 1 out. 2025.

OLIVEIRA, Bruna Pinotti Garcia. Inteligência Artificial E Proteção De Dados: Sobre A Audodeterminação Informativa E A Manipulação Informacional Por Machine Learning. **Revista Multidisciplinar Humanidades e Tecnologias (FINOM)**, Minas Gerais: Faculdade do Noroeste de Minas, ed. Vol. 26, ano 2020, p. 162-186, 15 jul. 2020. Disponível em: https://revistas.icesp.br/index.php/FINOM_Humanidade_Tecnologia/article/view/1356/1013. Acesso em: 31 ago. 2025.

OPENAI. **Política de privacidade.** Disponível em: https://openai.com/pt-BR/policies/row-privacy-policy/. Acesso em: 3 set. 2025.

PASQUALE, Frank. ***The Black Box Society: The Secret Algorithms That Control Money and Information.*** ISBN 978- 0- 674- 36827- 9. E-book.

PEREIRA, Lucca Ramos Alves; EBLING, Maurício. Responsabilidade civil na utilização da inteligência artificial. **TECNOLOGIAS EM PROJEÇÃO**, [S. l.], v. 14, n. 2, p. 72–87, 2024. Disponível em: https://projecaociencia.com.br/index.php/Projecao4/article/view/2404. Acesso em: 12 out. 2025.

PFEIFFER, Marc H. ***First, Do No Harm: Algorithms, AI, and Digital Product Liability – Managing Algorithmic Harms Through Liability Law and Market Incentives.*** Rutgers Edward J. Bloustein School of Planning and Public Policy; Center for Urban Policy Research | Bloustein Local, set. 2023. Disponível em: https://doi.org/10.48550/arXiv.2311.10861. Acesso em: 1 out. 2025.

PINHEIRO, Denise; PONZONI, Luiz. O uso responsável da IA generativa como chave para a inovação ética. **Revista DOM, Fundação Dom Cabral**, *[s.i.]*, 2024, p. 48–51. Disponível em: https://ci.fdc.org.br/AcervoDigital/Artigos%20FDC/Artigos%20DOM%20Contexto%20abr.%202024/O%20uso%20respons%C3%A1vel%20da%20IA%20generativa%20como%20chave%20para%20inova%C3%A7%C3%A3o%20%C3%A9tica.pdf. Acesso em: 24 out. 2025.

PINHEIRO, Patrícia Peck. **#DIREITODIGITAL**. São Paulo: Saraiva Jur, 2021, rev. amp. atual. E-book.

PORTUGAL. Supremo Tribunal de Justiça. Acórdão no processo 3413/03.2TBVCT.S1. Relator: Paulo Sá. Lisboa, 26 maio 2009. Disponível em: https://www.dgsi.pt/jstj.nsf/954f0ce6ad9dd8b980256b5f003fa814/dda829b22a91e69a802575ca002f61bf?OpenDocument. Acesso em: 20 set. 2025.

PRAZERES, Gustavo Cunha. Autodeterminação informacional vs. regulação do risco: uma abordagem sistêmica da regulamentação digital. **Revista Direito e Praxis**, Rio de Janeiro, v. 13, n. 2, 2022, p. 808–829. Disponível em: https://orcid.org/0000-0002-2389-5228. Acesso em: 13 set. 2025.

REICHELT, Luis Alberto. A inversão do ônus da prova em sede de responsabilidade civil decorrente do tratamento de dados pessoais na Lei Geral de Proteção de Dados sob a ótica dos direitos fundamentais processuais. **Revista Direito, Inovação e Regulações - Centro Universitário de Cascavel (UNIVEL)**, Cascavel, Jan. 2023; V. 2 (4): 100-109. ISSN-e: 2965-0860. Disponível em: https://periodicos.univel.br/ojs/index.php/redir/article/view/389 . Acesso em: 15 set. 2025.

REQUIÃO, Maurício. A natureza jurídica do consentimento para tratamento de dados pessoais. *In*: REQUIÃO, Maurício. (Org.). **Proteção de dados pessoais: novas perspectivas.** Salvador, Bahia: EDUFBA, 2022. p. 16-33. Disponível em: https://repositorio.ufba.br/handle/ri/35799. Acesso em: 08 set. 2025.

REQUIÃO, Maurício. **Normas de textura aberta e interpretação: uma análise no adimplemento das obrigações.** 2009. Tese. (Mestrado em Direito) - Faculdade de Direito, Universidade Federal da Bahia - UFBA, Salvador. Orientadora: Profª. Dra. Roxana Cardoso Brasileiro Borges. Disponível em: https://repositorio.ufba.br/handle/ri/10792. acesso em: 11 nov. 2025.

REQUIÃO, Maurício; PRAZERES, Gustavo Cunha. Natureza Jurídica Dos Dados Pessoais: Entre Projeções Existenciais E Os Direitos Patrimoniais. *In*: EHRHARDT JUNIOR, Marcos; CATALAN, Marcos. (Coords.). **Dados Pessoais E A Proteção Dos Direitos Da Personalidade Na Era Da Inteligência Artificial.** Belo Horizonte: Fórum, 2025. p. 47-68,

RODOTÀ, Stefano. **A vida na sociedade da vigilância: A privacidade hoje.** Tradução: Danilo Doneda; Luciana Cabral Doneda. Rio de Janeiro. São Paulo. Recife: Renovar, 2008. ISBN: 9788571476882.

RODRIGUES JUNIOR, Otavio Luiz. Nexo Causal Probabilístico: Elementos para a crítica de um Conceito. **Revista de Direito Civil Contemporâneo**, [S. l.], v. 8, p. 115–137, 2017. Disponível em: https://ojs.direitocivilcontemporaneo.com/index.php/rdcc/article/view/46. Acesso em: 24 set. 2025.

ROQUE, Andre. A Tutela Coletiva Dos Dados Pessoais Na Lei Geral De Proteção De Dados Pessoais (Lgpd). **Revista Eletrônica de Direito Processual**, Rio de Janeiro, v. 20, n. 2, 2019. DOI: 10.12957/redp.2019.42138. Disponível em: https://www.e-publicacoes.uerj.br/redp/article/view/42138. Acesso em: 16 out. 2025.

ROSENVALD, Nelson; FARIAS, Cristiano Chaves de; NETTO, Felipe Braga. **Curso de Direito Civil: Responsabilidade Civil**. 12. ed. São Paulo: JusPodivm, v.3, 2025.

ROSENVALD, Nelson. Prevenção e responsabilidade. Função preventiva da responsabilidade civil no Anteprojeto de Reforma do Código Civil. In: PACHECO, Rodrigo (Org.). **A Reforma do Código Civil: Artigos sobre a atualização da Lei no 10.406/2002.** Brasília, DF, Senado Federal, 2025, p. 131-146. Disponível em: https://share.google/o5F1kJOPlNuF6St98. Acesso em: 18 set. 2025.

ROSENVALD, Nelson. Responsabilidade civil: compensar, punir e restituir. **Revista IBERC**, Belo Horizonte, v. 2, n. 2, 2019. Disponível em: https://revista.iberc.org.br/iberc/article/view/48. Acesso em: 7 nov. 2025.

RUDINIKI NETO, Rogério. Fundamentos E Diretrizes Para A Destinação Alternativa De Verbas Pecuniárias (Cy-Près) Em Termos De Ajustamento De Conduta Celebrados Pelo Ministério Público. **PIÁ – Paraná Inteligência Artificial: Teses 2019**. Disponível em: https://site.mppr.mp.br/sites/hotsites/arquivos_restritos/files/migrados/File/MP_Academia/Teses_2019/Rogerio_Rudiniki_Neto_-_Fundamentos_e_diretrizes_para_a_destinacao_alternativa_de_verbas.pdf. Acesso em: 22 out. 2025.

RUZZI, Mariana; MARCHETTO, Patrícia Borba. Obstáculos à efetividade do direito à privacidade e à proteção de dados na era do big data e da inteligência artificial. **Revista Internacional Consinter de Direito**, Paraná, Brasil, v. 10, n. 19, p. 193–213, 2024. DOI: 10.19135/revista.consinter.00019.07. Disponível em: https://revistaconsinter.com/index.php/ojs/article/view/657. Acesso em: 16 out. 2025.

SANCHEZ, José Guilherme Machado; DIXO, Leonardo José Carvalho; BERNARDES, Luis Guilherme Maquiné; AGUIAR, Valdir Hilgenberg; QUEIROZ, Paulo. Disgorgement: indenização pelo ilícito lucrativo: Disgorgement: compensation for the lucrative illicit act. **RCMOS - Revista Científica Multidisciplinar O Saber**, Brasil, v. 1, n. 1, 2025. DOI: 10.51473/rcmos.v1i1.2025.928. Disponível em: https://submissoesrevistarcmos.com.br/rcmos/article/view/928. Acesso em: 22 out. 2025.

SANTANA, Rafael da Silva. A necessidade como elemento modulador da validade dos atos

de tratamento de dados pessoais. *In*: REQUIÃO, Maurício. (Org.). **Proteção de dados pessoais: novas perspectivas.** Salvador, Bahia: EDUFBA, 2022. p. 84-105. Disponível em: https://repositorio.ufba.br/handle/ri/35799. Acesso em: 08 set. 2025.

SANTOS, Bruno P.; SILVA, Lucas A. M.; CELES, Clayson S. F. S.; BORGES NETO, João B.; PERES, Bruna S.; VIEIRA, Marcos Augusto M.; VIEIRA, Luiz Filipe M., GOUSSEVSKAIA, Olga N.; LOUREIRO, Antonio A. F. Loureiro. Internet das coisas: da teoria à prática. **Minicursos SBRC – Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos**, v. 31, p. 1-16, 2016. Disponível em: https://homepages.dcc.ufmg.br/~mmvieira/cc/papers/internet-das-coisas.pdf. Acesso em: 8 set. 2025.

SANTOS, Helena Beatriz da Silva. **O Tratamento De Dados Pessoais De Saúde À Luz Do Regulamento Geral Da Proteção De Dados.** 2024. Tese. (Mestrado em Direito) - Faculdade de Direito, Coimbra Business School ISCAC, Coimbra. Disponível em: http://hdl.handle.net/10400.26/53898. Acesso em: 2 ago. 2025.

SANTOS, Lucas Cunha Imbiriba dos; HOMCI, Janaina Vieira. A análise do utilitarismo e do hedonismo na renúncia ao direito à privacidade na internet ante a caracterização da vulnerabilidade algorítmica do consumidor. **Revista Jurídica Cesumar - Mestrado**, [S. l.], v. 24, n. 1, p. 253–264, 2024. DOI: 10.17765/2176-9184.2024v24n1.e12310. Disponível em: https://periodicos.unicesumar.edu.br/index.php/revjuridica/article/view/12310. Acesso em: 13 mai. 2025.

SANTOS, Romualdo Baptista dos. Função preventiva da responsabilidade civil e tutela inibitória substantiva na reforma do Código Civil brasileiro. **Revista de Direito da Responsabilidade**, ano 7, p. 324-339, 2025. Disponível em: <https://www.academia.edu/129078626/FUN%C3%87%C3%83O_PREVENTIVA_DA_RESPONSABILIDADE_CIVIL_E_TUTELA_INIBIT%C3%93RIA_SUBSTANTIVA_NA_REFORMA_DO_C%C3%93DIGO_CIVIL_BRASILEIRO>. Acesso em: 18 set. 2025.

SANTOS, Rodrigo Santana dos. Coleta de Dados Pessoais para o Treinamento de Inteligência Artificial Generativa: Um Desafio para a Proteção de Dados e a Privacidade. *In*: BRANCO, Sérgio (Coord.). **Inteligência Artificial e Sociedade Conectada**. Rio de Janeiro: Instituto de Tecnologia e Sociedade - ITS, ed. 1, ano 2025, p. 72-93, Mensal. Disponível em: https://share.google/PQfYQHt02HNurGrIu. Acesso em: 6 mai. 2025.

SANTOS, Tibério. **Synthetic Data: Como Dados Falsos Criados por IA Resolvem LGPD E Treinam Algoritmos Sem Violar Privacidade (R$ 4,2 Bi em Economia).** Eunerd. Disponível em: https://encontreumnerd.com.br/blog/synthetic-data-como-dados-falsos-criados-por-ia-resolvem-lgpd-e-treinam-algoritmos-sem-violar-privacidade-r-4-2-bi-em-economia?srsltid=AfmBOorRNSeTeymUrEL4pEy5OHCRNU-FOQakIgLk8zhvZXNOmGGnDBlw. Acesso em: 14 set. 2025.

SÃO PAULO. Tribunal Regional Federal - 3ª Região (TRF-3). Ação Civil Pública Cívell nº 5018090-42.2024.4.03.6100. 2ª Vara Cível Federal de São Paulo. Juiz Federal: Luís Gustavo Bregalda Neves. Julgado em 14 ago. 2024. Disponível em:

https://www.trf3.jus.br/documentos/acom/anexos-noticias/2024/decisao_Whatsapp_-_5018090-42.2024.4.03.6100-1723652607148-14052.pdf. Acesso em: 16 abr. 2025.

SATO, Luiza. **Dados sintéticos: o brigadeiro de *whey protein* da proteção de dados.** Disponível em: https://tozzinifreire.com.br/artigos/dados-sinteticos-o-brigadeiro-de-whey-protein-da-protecao-de-dados. Acesso em: 11 set. 2025.

SATTLER, Andreas. *From Personality to Property?*. *In*: BAKHOUM, Mor; GALLEGO, Beatriz Conde; MACKENRODT, Mark-Oliver; SURBLYTĖ-NAMAVIČIENĖ, Gintaré (Eds.). ***Personal Data in Competition, Consumer Protection and Intellectual Property Law: Towards a Holistic Approach?.*** Berlim, Alemanha: Springer, 2018, ISBN: 978-3-662-57646-5. Disponível em: https://doi.org/10.1007/978-3-662-57646-5. Acesso em: 10 mai. 2025.

SCHEUERMANN, Gabriela Felden. Dados pessoais como um direito fundamental autônomo a partir da Emenda Constitucional nº 115/2022. **Revista da Defensoria Pública do Estado do Rio Grande do Sul**, Porto Alegre, v. 2, n. 33, p. 253–274, 2023. Disponível em: https://revista.defensoria.rs.def.br/defensoria/article/view/600. Acesso em: 29 set. 2025.

SENGAR, Sandeep Singh; HASAN, Affan Bin; KUMAR, Sanjay; CARROLL, Fiona. ***Generative Artificial Intelligence: A Systematic Review and Applications***. Cornell University. Disponível em: https://arxiv.org/abs/2405.11029. Acesso em: 10 set. 2025.

SETZER, Valdemar W.. Depto. de Ciência da Computação, Universidade de São Paulo. Dado, Informação, Conhecimento e Competência. São Paulo: USP, 2015. Disponível em: https://www.ime.usp.br/~vwsetzer/dado-info.html. Acesso em: 17 ago. 2025.

SHUMAILOV, Ilia; SHUMAYLOV, Zakhar; ZHAO,Yiren; PAPERNOT, Nicolas; ANDERSON, Ross; GAL, Yarin. AI models collapse when trained on recursively generated data. Nature, v. 631, p. 755–759, 25 jul. 2024. DOI: 10.1038/s41586-024-07566-y. Disponível em: https://www.nature.com/articles/s41586-024-07566-y . Acesso em: 8 nov. 2025.

SILVA, Amanda Moreira Mota da; LELIS, Mariana Nascimento Santana. O dano existencial a partir da concepção de dignidade humana em Giovanni Pico Della Mirandola. **Revista de Doutrina Jurídica**, Brasília, DF, v. 115, n. 00, p. e024001, 2024. DOI: 10.22477/rdj.v115i00.909. Disponível em: https://revistajuridica.tjdft.jus.br/index.php/rdj/article/view/909. Acesso em: 21 out. 2025.

SILVA, Ana Marília Dutra Ferreira da; SILVA, Carlos Eduardo da; SIQUEIRA, Mariana de; MARQUES, KayoVictor Santos. Proteção de dados pessoais e direito à privacidade no contexto da pandemia de covid-19: uma análise das aplicações de contact tracing à luz da proporcionalidade. **RevistaDireito GV**, São Paulo, v. 18, n. 3, set./dez. 2022, e2232. https://doi.org/10.1590/2317-6172202232. Disponível em: https://www.scielo.br/j/rdgv/a/fmmZDmbxS9tGzyWB3NTR3fF/?format=html&lang=pt. Acesso em: 8 set. 2025.

SILVA, Hellen Eduarda Rodrigues; MUNIZ, Aline de Assis Rodrigues do Amaral. Responsabilidade Civil Na Era Digital: Desafios e perspectivas. **Revista Acadêmica Online**,

[S. l.], v. 10, n. 50, p. 1–18, 2024. DOI: 10.36238/2359-5787.2024.v10n50.26. Disponível em: https://revistaacademicaonline.com/index.php/rao/article/view/26. Acesso em: 9 set. 2025.

SILVA, Rogério. Microsoft Copilot lidera crescimento em usuários móveis nos EUA, supera ChatGPT. ln: MPI SOLUTIONS. Tecnologia. [S.l.]. 3 set. 2025. Disponível em: https://www.mpisolutions.com.br/noticias/microsoft-copilot-lidera-crescimento-em-usuarios-moveis-nos-eua-supera-chatgpt/#:~:text=Microsoft%20Copilot%20lidera%20crescimento%20em%20usu%C3%A1rios%20m%C3%B3veis%20nos%20EUA%2C%20supera%20ChatGPT,-56%20segundos%20atr%C3%A1s&text=Mais%20de%2085%25%20dos%20usu%C3%A1rios,tornando%20o%20uso%20cotidiano%20habitual.. Acesso em: 4 set. 2025.

SIQUEIRA, Dirceu Pereira; MOREIRA, Mayume Caires. O acesso as tecnologias de informação e comunicação no Brasil: os reflexos da exclusão e da desigualdade digital nos direitos da personalidade. **Revista Brasileira de Direito**. Passo Fundo, RS, Brasil, v. 19, n. 1, p. e4836, 2023. Disponível em: https://seer.atitus.edu.br/index.php/revistadedireito/article/view/4836. Acesso em: 30 out. 2025.

SOARES, Eduardo Jonas. **Função Social Da Responsabilidade Civil.** 2009. (Pós-Graduação em Direito) - Faculdade de Ciências Sociais de Florianópolis, Complexo De Ensino Superior De Santa Catarina - UNICESUSC, Florianópolis. Orientador: Profª. Msc. Leilane Zavarizi Da Rosa. Disponível em: https://biblioteca.sophia.com.br/terminalri/2764/VisualizadorPdf?codigoArquivo=9502&tipoMidia=0. Acesso em: 6 out. 2025.

SOARES, Rafael Oliveira; EHRHARDT JÚNIOR, Marcos. Os dados pessoais como bens de valor econômico e a despersonalização das pessoas naturais: a comoditização do indivíduo e sua incompatibilidade com a ordem constitucional brasileira. **Civilistica.com - Revista Eletrônica de Direito Civil.** Rio de Janeiro, v. 14, n. 1, p. 1–20, 2025. Disponível em: https://civilistica.emnuvens.com.br/redc/article/view/1052. Acesso em: 10 out. 2025.

SOUSA, Rosilene Paiva Marinho de; VASCONSELOS, Fernando Antônio; SOUZA, Marckson Roberto Ferreira de. Informação E Privacidade Na Lei Geral De Proteção De Dados. *In*: **Anais do Workshop de Informação, Dados e Tecnologia - WIDaT,** [S. l.], v. 2, p. 4–11, 2018. DOI: 10.22477/ii.widat.129. Disponível em: https://labcotec.ibict.br/widat/index.php/widat2023/article/view/129. ISBN: 978-85-237-1381-2. Acesso em: 13 ago. 2025.

SOUSA, Nuno Miguel Teixeira. Modelos de Governação Informacional em Sistemas de IA Generativa: análise crítica comparativa entre ChatGPT e AMÁLIA. ***Brazilian Journal of Information Science: research trends***, vol. 19, 2025, e025023. DOI: https://doi.org/10.36311/1981-1640.2025.v19.e025023. Disponível em:https://revistas.marilia.unesp.br/index.php/bjis/article/view/17282. Acesso em: 15 ago. 2025.

SOUZA, Maristela Denise Marques de; LOPES, Adriana. Crise Dos Pressupostos Tradicionais Da Responsabilidade Civil. **Revista da AJURIS**, Porto Alegre: Associação dos Juízes do Rio Grande do Sul, ed. 40, ano 2013, n. 129, p. 107-152, Disponível em: https://share.google/hmwhkQeKh9UL1JCm0. Acesso em: 5 out. 2025.

SOUSA, Rosilene Paiva Marinho de; VASCONCELOS, Fernando Antônio de; SOUSA, Marckson Roberto Ferreira de. Informação e privacidade na Lei Geral de Proteção de Dados. *In*: DIAS, Guilherme Ataide; DUTRA, Moisés Lima (Orgs.). **II Workshop De Informação, Dados E Tecnologia**, [S. l.], v. 2, p. 4–11. Paraíba, 2018. DOI: 10.22477/ii.widat.129. Disponível em: https://labcotec.ibict.br/widat/index.php/widat2023/article/view/129. Acesso em: 13 ago. 2025.

Statista Research Department. **Most popular artificial intelligence (AI) applications worldwide in February 2025, by monthly active users.** Statista. Disponível em: https://www.statista.com/statistics/1609163/top-ai-applications-mau-worldwide/. Acesso em: 2 set. 2025.

SURBLYTÉ, Gintare. **Data as a digital resource.** SSRN - Social Science Research Network. Disponível em: https://ssrn.com/abstract=2849303 . Acesso em: 11 maio 2025.

TAMKIN, Alex; BRUNDAGE, Miles; CLARK, Jack; GANGULI, Deep. ***Understanding the capabilities, limitations, and societal impact of large language models*.** Disponível em: https://doi.org/10.48550/arXiv.2102.02503. Acesso em 17 ago. 2025.

TAURION, Cezar. **Big Data.** Rio de Janeiro: Brasport, 2013. E-book.

TEFFÉ, Chiara Antonia Spadaccini de. **Dados pessoais sensíveis: uma análise funcional da categoria e das hipóteses de tratamento.** 2022. Tese. (Doutorado em Direito) - Faculdade de Direito, Universidade do Estado do Rio de Janeiro - UERJ, Rio de Janeiro. Orientador: Prof. Dr. Gustavo Tepedino. Disponível em: https://www.bdtd.uerj.br:8443/handle/1/18291. Acesso em: 12 set. 2025.

TEFFÉ, Chiara Antonia Spadaccini de; VIOLA, Mario. Tratamento de dados pessoais na LGPD: estudo sobre as bases legais. **Civilistica.com - Revista Eletrônica de Direito Civil.** Rio de Janeiro, v. 9, n. 1, p. 1–38, 2020. Disponível em: https://civilistica.emnuvens.com.br/redc/article/view/510. Acesso em: 8 ago. 2025.

TEPEDINO, Gustavo; SILVA, Rodrigo da Guia. Desafios Da Inteligência Artificial Em Matéria De Responsabilidade Civil. **Revista Brasileira de Direito Civil**, Belo Horizonte: RBDCivil, ed. Vol. 21, ano 2019, p. 61-86, 20 set. 2019. Disponível em: https://share.google/ZOhsn5AYuRix1ZZn4. Acesso em: 6 set. 2025.

TEPEDINO, Gustavo; TEFFÉ, Chiara Spadaccini de. O Consentimento Na Circulação De Dados Pessoais. **Revista Brasileira de Direito Civil – RBDCivil**, Belo Horizonte, ano 2020, p. 83-116, Semestral. Disponível em: https://share.google/MP74CyZ1RlK7blchS. Acesso em: 14 ago. 2025.

TERRACAP. **O que é privacy by design e privacy by default?**. Disponível em: https://www.terracap.df.gov.br/index.php/listagem-faq/78-lgpd-lei-geral-de-protecao-de-dados-pessoais/196-53-o-que-e-privacy-by-design-e-privacy-by-default. Acesso em: 01 out. 2025.

TSAI, Chun‐Wei; LAI, Chin-Feng; CHAO, Han-Chieh; VASILAKOS, Athanasios V. **Big data analytics: a survey.** Springer Nature Link. Disponível em: https://doi.org/10.1186/s40537-015-0030-3. Acesso em: 4 mai. 2025.

URSIC, Helena. *The Failure of Control Rights in the Big Data Era: Does a Holistic Approach Offer a Solution?*. *In*: BAKHOUM, Mor; GALLEGO, Beatriz Conde; MACKENRODT, Mark-Oliver; SURBLYTĖ-NAMAVIČIENĖ, Gintaré (Eds.). ***Personal Data in Competition, Consumer Protection and Intellectual Property Law: Towards a Holistic Approach?.*** Berlim, Alemanha: Springer, 2018, ISBN: 978-3-662-57646-5. Disponível em: https://doi.org/10.1007/978-3-662-57646-5. Acesso em: 10 mai. 2025.

VALADÃO, Rodrigo Borges. **Monetização dos dados pessoais: a nova fronteira da privacidade**. Academia. Disponível em: https://www.academia.edu/129225001/Monetiza%C3%A7%C3%A3o_dos_Dados_Pessoais_A_Nova_Fronteira_da_Privacidade. Acesso em: 20 set. 2025.

VAROUFAKIS,Yanis. **Technofeudalism: what killed capitalism.** *[s.i]*. Editora Melville House Publishing, 2025. E-book.

VAUGHAN, Thom. **Common Crawl**. Disponível em: https://commoncrawl.org/blog/august-2025-crawl-archive-now-available. Acesso em: 05 set. 2025.

VERCELLI, Ariel. *Inteligencias artificiales, términos de uso y gestión de contenidos (inputs / outputs)*. **JAIIO, Jornadas Argentinas de Informática**, [S. l.], v. 11, n. 12, p. 1–10, 2025. Disponível em: https://revistas.unlp.edu.ar/JAIIO/article/view/19542/19556. Acesso em: 8 nov. 2025.

VIEIRA, Allan Sarmento; DA SILVA, Wallach Pereira; DANTAS, Mírian Ionnara Abrantes Viana; DE ALBUQUERQUE, Pedro Gustavo Lopes; MAIA, Kyev Moura. A Hermenêutica No Contexto Cibernético: Perspectivas E Desafios Das Normas Jurídicas Frente Às Inovações Tecnológicas. **ARACÊ**, *[S. l.]*, v. 7, n. 5, p. 21679–21692, 2025. DOI: 10.56238/arev7n5-040. Disponível em: https://periodicos.newsciencepubl.com/arace/article/view/4829. Acesso em: 4 out. 2025.

WAN, Zhipeng; CHENG, Anda; WANG, Yinggui; WANG, Lei. **Information Leakage from Embedding in Large Language Models**. Cornell University. Disponível em: https://arxiv.org/abs/2405.11916. Acesso em: 6 nov. 2025.

WANDERER, Bertrand. Economia movida a dados e o papel das plataformas digitais. **Journal of Law and Regulation**, [S. l.], v. 9, n. 2, p. 22–43, 2023. Disponível em: https://periodicos.unb.br/index.php/rdsr/article/view/43231. Acesso em: 17 abr. 2025.

WARK, McKenzie. **Capital is Dead.** United Kingdom: CPI Mackays. E-book.

WEIDINGER, Laura; MELLOR, John; RAUH, Maribeth; GRIFFIN, Conor; UESATO, Jonathan; HUANG, Po-Sen; CHENG, Myra; GLAESE, Mia; BALLE, Borja; KASIRZADEH, Atoosa; KENTON, Zac; BROWN, Sasha; HAWKINS, Will; STEPLETON, Tom; BILES, Courtney; BIRHANE, Abeba; HAAS, Julia; RIMELL, Laura; HENDRICKS,

Lisa Anne; ISAAC, William; LEGASSICK, Sean; IRVING, Geoffrey; GABRIEL, Iason. **Ethical and social risks of harm from Language Models**. Cornell University. Disponível em: https://arxiv.org/abs/2112.04359
. DOI: https://doi.org/10.48550/arXiv.2112.04359. Acesso em: 15 nov. 2025.

WIDDER, David Gray; WHITTAKER, Meredith; WEST, Sarah Myers. **Open (for business): big tech, concentrated power, and the political economy of open AI.** Social Science Research Network (SSRN). Disponível em: https://ssrn.com/abstract=4543807 . Acesso em: 4 maio 2025.

WILLEMIN, Andrea Carmo Name; FARIA, Geralda Magella de; AMANTE, Cláudio José. Informação E Privacidade Na Lei Geral De Proteção De Dados. *In*: DIAS, Guilherme Ataide; DUTRA, Moisés Lima (Orgs.). **II Workshop De Informação, Dados E Tecnologia**, [S. l.], v. 2, p. 4–11. Paraíba, 2018. DOI: 10.22477/ii.widat.129. Disponível em: https://labcotec.ibict.br/widat/index.php/widat2023/article/view/129 Acesso em: 13 ago. 2025.

YAN, Biwei; LI, Kun; XU, Minghui; DONG, Yueyan; ZHANG, Yue; REN, Zhaochun; CHENG, Xiuzhen. **On protecting the data privacy of Large Language Models (LLMs) and LLM agents: A literature review.** Science Direct. Disponível em: https://www.sciencedirect.com/science/article/pii/S2667295225000042. Acesso em: 8 out. 2025.

ZANINI, Vilma Tomaz Lourenço Ferreira. Responsabilidade civil punitiva no direito brasileiro. In: GUERRA, Alexandre Dartanhan de Mello (Coord.). **Estudos em homenagem a Clóvis Beviláqua por ocasião do centenário do Direito Civil codificado no Brasil**. São Paulo: Escola Paulista da Magistratura, 463-481. Disponível em: https://www.tjsp.jus.br/download/EPM/Publicacoes/ObrasJuridicas/cc24.pdf?d=63680. Acesso em: 5 out. 2025.

ZUBOFF, Shoshana. **A Era do Capitalismo de Vigilância. A luta por um futuro humano na nova fronteira de poder**. 1 ed. digital. Rio de Janeiro: Intrínseca, 2021. E-book.

ANEXO 01 - Aplicações de IA com maior base de usuários ativos mensais.