3° PRÊMIO
DANILO
DONEDA
DE ARTIGOS CIENTÍFICOS

Artigos científicos vencedores

ANPD

III Prêmio Danilo Doneda de Artigos Científicos da Agência Nacional de Proteção de Dados

— Prêmio Danilo Doneda 2025 —

Artigos científicos vencedores

Supervisão Editorial
Angela Halen Claro Franco

ANPD
Brasília, DF
2025

Agência Nacional de Proteção de Dados

Diretor-Presidente
Waldemar Gonçalves Ortunho Junior

Diretores
Iagê Zendron Miola
Lorena Giuberti Courtinho
Miriam Wimmer

Comissão julgadora
Diego Carvalho Machado
Gabriela Natacha Bechara
Iagê Zendron Miola
Lucas Borges de Carvalho
Lucas Costa dos Anjos
Mariana Almeida de Sousa Talouki
Miriam Wimmer

Comissão organizadora
Adi Balbinot Junior
Albert França Josuá Costa
Gustavo Andrade Bruzzeguez
Jayme Marrone Júnior
Marcus Vinicius Rossi da Rocha
Roseane Salvio
Uberson Rossa

Revisão e organização textual
Angela Halen Claro Franco
Roseane Salvio
Uberson Rossa
Jayme Marrone Júnior
Albert França Josuá Costa

Projeto gráfico e capa
André Scofano Maia Porto

Editoração eletrônica
Rodrigo Lucas Mendes

1ª edição
Publicação digital – PDF (dezembro / 2025)

ANPD · Agência Nacional de Proteção de Dados
SCN, Qd. 6, Conj. A
Ed. Venâncio 3000, Bl. A, 9º andar
Brasília, DF · Brasil · 70716-900
www.gov.br/anpd

Sumário

Notas do editor

Os artigos foram recebidos e avaliados de acordo com os prazos estabelecidos no Edital Nº 1, de 22 de abril de 2025. Durante o processo editorial desta coletânea, todas as menções à Autoridade Nacional de Proteção de Dados (ANPD) foram substituídas nos artigos por Agência Nacional de Proteção de Dados (ANPD), a fim de refletir a transformação da entidade estabelecida pela Medida Provisória nº 1.317, de 17 de setembro de 2025.

Prefácio

É com imensa satisfação que apresentamos esta coletânea, fruto da Terceira Edição do Prêmio Danilo Doneda de artigos científicos, uma iniciativa promovida pela Agência Nacional de Proteção de Dados. O prêmio não apenas homenageia a memória e o legado do jurista Danilo Doneda, pioneiro nos estudos da proteção de dados e colaborador ativo na criação da Lei Geral de Proteção de Dados Pessoais no Brasil, mas também solidifica o compromisso da ANPD com o fomento à pesquisa, à inovação e à disseminação do conhecimento no campo da privacidade e proteção de dados pessoais.

A missão da ANPD, enquanto órgão central de regulamentação da LGPD e fiscalização de sua observância, transcende a mera aplicação de sanções. Seu papel é fundamental na construção de uma cultura de respeito à privacidade, na orientação de boas práticas e na garantia de que os direitos dos titulares de dados sejam efetivamente protegidos. Para tanto, é imprescindível estimular o debate qualificado e a produção acadêmica que desvendem os complexos desafios que emergem diariamente.

Nesse contexto, o Prêmio Danilo Doneda assume um protagonismo singular. Ao reconhecer e premiar trabalhos de destaque, ele impulsiona a reflexão crítica sobre temas cruciais que impactam a sociedade brasileira no âmbito da privacidade e proteção de dados pessoais. Esta publicação condensa algumas dessas contribuições valiosas, que exemplificam a profundidade e a atualidade das discussões no campo da proteção de dados, abordando com rigor acadêmico e relevância prática, tópicos que estão no epicentro das discussões contemporâneas e que fazem parte da Agenda Regulatória e do Mapa de Temas Prioritários da ANPD.

O trabalho que atingiu a primeira colocação nesta edição do concurso, intitulado "Consentimento como base legal em ambientes digitais: '*Dark Patterns*' e a ilusão de escolha", faz uma análise sob a perspectiva regulatória da validade do consentimento obtido nos serviços digitais que adotam padrões enganosos de interface para induzir as ações do usuário, visando a maximização de lucros. Esses serviços utilizam tecnologias de personalização e estratégias de manipulação de interfaces cada vez mais sofisticadas,

conhecidas como padrões obscuros (*dark patterns*, em inglês), que exploram vieses cognitivos e minam direitos previstos na LGPD, como a transparência e o exercício da autodeterminação informativa. O artigo faz uma síntese das sete principais categorias de *dark patterns*, apresenta um panorama regulatório da União Europeia, Estados Unidos da América e Brasil, e mostra os principais desafios práticos e jurídicos enfrentados nesse processo. Faz também recomendações de produção de diretrizes normativas precisas e a implementação de ações educativas em larga escala.

Na segunda colocação, o artigo "Segredo de negócios e decisões automatizadas discriminatórias: impasses e soluções regulatórias" explora o conflito entre a proteção jurídica do segredo de negócios e os princípios e normas da proteção de dados pessoais, especialmente no que tange à proteção contra decisões automatizadas discriminatórias. O texto discute os fundamentos jurídicos do segredo de negócios e como a proteção ao segredo pode limitar a publicidade das informações sobre operações de tratamento de dados pessoais, gerando opacidade e em especial a opacidade algorítmica em sistemas de tomada de decisão mediados por inteligência artificial. O trabalho faz ainda uma análise sobre a discriminação algorítmica que pode ser gerada por esses sistemas a partir de vieses de dados e variáveis introduzidas na definição e treinamento dos modelos. Compara-se como outras jurisdições têm lidado com essa questão e sugere a regulação responsiva no contexto brasileiro, ancorada no artigo 20, § 2º da LGPD, além da instituição de mecanismos complementares que possam ampliar a publicização de informações relevantes e favorecer um controle social e institucional difuso de tais sistemas.

E em terceira colocação, o trabalho intitulado “A identificação da criança como pressuposto jurídico-operacional para a aplicação proporcional da LGPD: limites técnicos e diretrizes regulatórias” aborda a questão da crescente presença de crianças e adolescentes nos ambientes digitais e a ausência de mecanismos eficazes de identificação etária, e como essa ausência compromete a aplicação efetiva do princípio do melhor interesse de crianças e adolescentes disposto na LGPD, resultando em uma permanente insegurança jurídica e alto risco operacional. Discorre ainda sobre as soluções de

verificação etária disponíveis, sua eficácia e os dados coletados por cada uma delas para atingir o objetivo da identificação da idade do usuário. Por meio do estudo de caso de uma plataforma digital de ampla popularidade global, analisa a convergência regulatória no contexto brasileiro e internacional. O estudo sugere a elaboração de um *framework* baseado em risco para as soluções de verificação de idade, e a adoção de uma postura proativa pelas empresas fornecedoras de produtos e serviços digitais no que se refere à identificação e mitigação de riscos. Além disso, o estudo propõe que os agentes de tratamento adotem, desde a concepção de seus produtos e serviços, o design apropriado à idade. O trabalho destaca a necessidade de incorporação dos princípios de *safety-by-design*, *privacy by design e security by design* nos produtos, e de abstenção da exploração econômica de crianças e adolescentes nos ambientes digitais.

Cada um desses textos ilumina facetas essenciais da implementação da LGPD e dos direitos fundamentais que ela visa salvaguardar. Eles são um testemunho da riqueza intelectual e da dedicação dos pesquisadores em colaborar para a construção de um ecossistema digital mais ético, justo e transparente.

Esta obra é, portanto, um convite à comunidade jurídica, tecnológica, regulatória e à sociedade civil para uma reflexão aprofundada sobre esses temas, reforçando a importância da academia e da pesquisa para o avanço da cultura da privacidade e proteção de dados pessoais em nosso país.

Waldemar Gonçalves Ortunho Junior
Diretor-Presidente da Agência Nacional de Proteção de Dados – ANPD

Iagê Zendron Miola
Diretor da Agência Nacional de Proteção de Dados – ANPD

Lorena Giuberti Courtinho
Diretora da Agência Nacional de Proteção de Dados – ANPD

Miriam Wimmer
Diretora da Agência Nacional de Proteção de Dados – ANPD

Sobre os premiados

1º lugar

Giovanna Diniz

Graduanda em Direito pela Universidade de São Paulo (USP). Pesquisadora nas áreas de Proteção de Dados Pessoais, Inteligência Artificial aplicada ao Direito, Direito Digital, Filosofia e História do Direito.

2º lugar

Isadora Valadares Assunção

Estudante de Direito na Universidade de São Paulo. Realizou iniciação científica sobre responsabilidade civil e discriminação algorítmica. Apresentou e publicou artigos em fóruns nacionais e internacionais, com destaque para temas relacionados à regulação de inteligência artificial e proteção de dados. Atualmente, coordena o grupo de extensão Techlab, que estuda a intersecção entre o direito e as novas tecnologias.

3º lugar

Giuseppe Grando

Bacharel em Direito pela UNISC - Universidade de Santa Cruz do Sul. Analista de Privacidade na El Canary Privacy and Ethics. Certificado pela Data Privacy Brasil. Entusiasta e pesquisador na intersecção de Privacidade e Proteção de Dados com tecnologias emergentes e direito público.

Artigos científicos vencedores

1º lugar

Consentimento como base legal em ambientes digitais: "dark patterns" e a ilusão de escolha

Giovanna Diniz

Resumo

A transição para uma sociedade digitalizada redefiniu os parâmetros de consentimento e reforçou sua centralidade como instrumento de legitimação no tratamento de dados pessoais. Paralelamente, multiplicaram-se mecanismos de indução comportamental conhecidos como padrões obscuros (*dark patterns*), caracterizados por distorcerem a arquitetura da escolha com o objetivo de orientar decisões de forma não consciente. Ao contrário dos nudges éticos, que buscam ampliar a racionalidade das decisões sem comprometer a autonomia, os padrões obscuros atuam com base na exploração de vieses cognitivos e têm sua eficácia ampliada por tecnologias de personalização que operam sobre grandes volumes de dados. A regulação, até o momento, apresenta respostas desiguais. A União Europeia dispõe de marcos avançados, com destaque para o GDPR e a DSA. Os Estados Unidos mantêm uma abordagem segmentada, baseada em ações estatais e iniciativas setoriais. No Brasil, observa-se um movimento inicial, centrado na LGPD e na atuação da ANPD, mas ainda desprovido de normas específicas voltadas à regulação dessas práticas. Entre os desafios abordados, destacam-se a fragmentação jurídica, a natureza in-

tangível dos danos e a escalabilidade das técnicas de manipulação por meio de inteligência artificial. A realidade observada exige da ANPD o desenvolvimento de parâmetros normativos claros para o *design* de interfaces, aliados a programas de conscientização que consolidem práticas informacionais mais alinhadas ao princípio da autodeterminação do usuário.

Palavras-chave: *dark patterns; consentimento; autodeterminação informacional; privacidade; LGPD.*

1. Introdução

O reconhecimento do consentimento como base legal para o tratamento de dados, embora central na Lei Geral de Proteção de Dados Pessoais (LGPD), não garante sua validade quando descolado das condições que o tornam legítimo à luz da autodeterminação informativa. Estratégias de manipulação de interface, como os dark patterns, minam esse princípio ao influenciar silenciosamente a vontade do titular. E justamente nesse horizonte se insere a resposta institucional. Ao incluir o consentimento na Agenda Regulatória 2025-2026, a Agência Nacional de Proteção de Dados (ANPD) sinaliza que o foco regulatório deve ir além do enquadramento formal, exigindo uma reflexão das estruturas que condicionam o ato de consentir. O centro da preocupação desloca-se, portanto, para a qualidade do ambiente decisório em que a escolha é formulada, sobretudo quando se verifica a presença de elementos que, embora não coercitivos em sentido estrito, operam por meio de vieses cognitivos e escolhas de design que reduzem a transparência e a autonomia decisória do titular.

Em conformidade com isso, este artigo parte de uma perspectiva jurídico-comparada para avaliar se o consentimento obtido sob tais condições pode ser considerado válido, tendo como principal base teórica e empírica o estudo conduzido por Baumeister *et al.* (2024) para o Data Standards Chair, que aprofunda o entendimento dos efeitos comportamentais e jurídicos dos padrões enganosos em contextos digitais.

2. Definição e conceituação de "dark patterns"

Nem sempre escolhemos. Às vezes, apenas seguimos o caminho que nos foi sutilmente traçado. A arquitetura da escolha (*choice architecture*[1]) é fascinante nesse sentido, mostrando como nossas decisões são moldadas por terceiros, muitas vezes sem que percebamos. Thaler e Sunstein (2021 *apud* Baumeister *et al.*, 2024, p. 17) cunharam o termo *nudge* para designar essas influênciasCompreendido em português como "empurrãozinho" ou "incentivo", o termo abrange iniciativas que favorecem decisões mais acertadas para o indivíduo e a sociedade, sem impor restrições à sua autonomia (Tipp, 2024). A intenção, nesse contexto, não é coagir, mas oferecer orientação.

Ainda que os nudges se fundamentem em princípios éticos de influência sutil, essa mesma lógica pode ser distorcida e usada para fins menos nobres. É nesse desvio que surgem os chamados padrões obscuros (*dark patterns*, também denominados *deceptive patterns*). Embora ambos se baseiem no *design* intencional da experiência do usuário, diferem profundamente no propósito: o que era pensado como orientação leve torna-se manipulação deliberada, explorando vieses cognitivos e conduzindo a decisões que não seriam tomadas sob plena consciência (Sharma, 2024; Finegold, 2025). De acordo com Baumeister *et al.* (2024) e Posson (2025), boa parte das práticas manipulativas que hoje dominam o digital são transposições de táticas desenvolvidas no comércio físico. Lojas de departamento, por exemplo, constroem seus espaços para reter a atenção do consumidor, direcionar seu percurso e promover contato visual com produtos específicos. Recorrem a estímulos sensoriais (como cores, aromas, fontes e organização espacial) como ferramentas de persuasão. Tudo isso compõe um ambiente projetado

1 Expressão criada por Thaler e Sunstein (2021 *apud* Baumeister *et al.*, 2024, p. 16-17) para designar a prática de influenciar escolhas a partir da forma como o ambiente decisório é estruturado. A complexidade dessa tomada de decisões, segundo Kahneman (2003 *apud* Calonga *et al.*, 2022), reside na interação de dois sistemas: o sistema 1, uma "força invisível e reativa" que utiliza atalhos mentais, muitas vezes resultando em padrões de pensamento enviesados; em contrapartida, o sistema 2, uma faculdade consciente e ponderada, que se baseia na experiência e na habilidade de construir cadeias lógicas de pensamento.

para induzir o consumo. Mas no ambiente digital, a coleta massiva de dados, a personalização em tempo real e a escalabilidade das interfaces tornam os *dark patterns* muito mais eficazes e difíceis de detectar. Basta pensar que, ao contrário de uma loja física, onde há limitações, as plataformas digitais permitem mudanças instantâneas e testes contínuos baseados no nosso comportamento (Baumeister *et al.*, 2024, p. 17). Para Oliveira (2023, p. 60) e *Zac et al.* (2025), esses padrões podem "afetar materialmente a tomada de decisões do usuário", atingindo inclusive indivíduos fora dos grupos tradicionalmente considerados vulneráveis, como idosos ou crianças.

O que se nota é que a onipresença dos *dark patterns* em serviços digitais variados evidencia um problema sistêmico de mercado, e não apenas a prática isolada de alguns agentes. De acordo com Baumeister *et al.* (2024, p. 78), esses padrões se consolidam por sua capacidade de manipular o comportamento do consumidor, ampliar a coleta de dados e elevar lucros. Relatórios recentes mostram sua presença em 11% dos sites populares de comércio eletrônico, em 95% dos aplicativos Android gratuitos e em todos os principais serviços de mídia social (Baumeister *et al.*, 2024, p. 9). Uma análise de 642 sites e aplicativos de assinatura revelou que quase 76% empregam pelo menos um padrão obscuro, enquanto cerca de 67% utilizam mais de um (ICPEN, 2024).

Vale salientar que essa alta incidência não é um acidente, mas sim uma estratégia deliberada impulsionada pelo desejo de "maximizar o lucro por meio de uma melhor compreensão de seus potenciais clientes" (Baumeister *et al.*, 2024, p. 12). A dimensão do problema indica que, na ausência de regulação, o mercado tende a incorporar práticas manipulativas como padrão, o que reforça a necessidade de intervenção externa para conter essas distorções.

3. Tipologias e mecanismos de manipulação

Os padrões obscuros manifestam-se de diversas formas, explorando diferentes vieses cognitivos para influenciar as decisões dos usuários. A literatura especializada e as autoridades de proteção de dados identificam várias categorias, que se tornaram mais

numerosas e sofisticadas ao longo do tempo. A partir de *Zac et al.* (2025), das orientações da European Data Protection Board (2023) e da tipologia proposta por Gray *et al.* (2021), apresentamos uma síntese das sete principais categorias de *dark patterns*:

Categoria	Definição	Padrões incluídos	Exemplos
Obstrução *(Obstruction/"Roach Motel"[2])*	Torna difícil ou impossível ações como o cancelamento de um serviço ou a recusa de consentimento.	"*Dead end*" (ausência de caminho funcional); "*longer than necessary*" (fluxo desnecessariamente longo) e "*misleading action*" (botões que não fazem o que sugerem).	Botão "excluir conta" que leva a dezenas de confirmações; opção "cancelar assinatura" que redireciona para a página inicial sem executar a ação.
Skipping	Interface leva o usuário a esquecer ou ignorar aspectos importantes de proteção de dados.	"*Deceptive snugness*" (aparência de segurança onde não há controle real) e "*look over there*" (elementos visuais desviam atenção dos controles importantes).	Tela de boas-vindas dizendo "tudo pronto!" sem apresentar opções de controle; botão grande com "continuar" e *link* discreto para configurações de privacidade.

2 O nome tem origem em uma marca norte-americana que comercializava armadilha para insetos durante a década de 1970, cujo slogan afirmava: "*Roaches check in, but they don't check out*!" (Baratas fazem *check-in*, mas não fazem *check-out*!). A expressão foi apropriada para criticar práticas digitais em que se inscrever é simples, mas sair exige esforço desproporcional. (Cf. Mazumdar, S., Blue, S. Responsible Design Part 9 of 14: Roach Motel. Think Design – A Havas Company, 2022).

Categoria	Definição	Padrões incluídos	Exemplos
Stirring	Manipula as decisões do usuário com apelos emocionais ou estímulos visuais sutis.	"*Emotional steering*" (apelos emocionais para influenciar) e "*hidden in plain sight*" (informação crítica camuflada visualmente).	Mensagens que fazem o usuário se sentir mal por não aceitar uma oferta ou opção (e.g., "tem certeza de que quer perder esta oportunidade?").
Apresentação instável (*Fickle*)	Interface propositalmente inconsistente e confusa, dificultando a navegação e a compreensão sobre controle de dados.	"*Lacking hierarchy*" (falta de organização visual); "*decontextualising*" (informações fora do contexto); "inconsistent interface" (mudanças visuais confusas) e "*language discontinuity*" (mudança de termos sem explicação).	Menu com ícones não padronizados e nomes diferentes para a mesma função ("privacidade" em um lugar, "segurança" em outro); partes do site em outro idioma sem justificativa

Categoria	Definição	Padrões incluídos	Exemplos
Forçamento (*forced action*) ou muro de *cookie* (*tracking wall*)	Vincular o acesso a funcionalidades essenciais à aceitação de cookies ou termos desnecessários.	Também conhecido como "*take it or leave it*" (pegue ou largue).	Mensagem "para continuar navegando, aceite todos os *cookies*" sem botão de recusa, ou com "recusar" oculto ou desabilitado.
Sobrecarga (*overloading*)	Usuário confrontado com excesso de solicitações, informações ou opções, induzido a compartilhar mais dados ou permitir o tratamento de dados pessoais involuntariamente.	"*Continuous prompting*" (solicitações repetidas e insistentes para consentimento), "*privacy maze*" (dificultar o acesso a controles de privacidade através de múltiplas páginas) e "*too many options*" (excesso de escolhas confusas ou redundantes).	*Pop-ups* constantes com "aceite agora para continuar".

Categoria	Definição	Padrões incluídos	Exemplos
Deixar no escuro (*Left in the dark*)	Esconde informações ou ferramentas de controle, deixando o usuário desorientado sobre seus dados e direitos.	"*Conflicting information*" (explicações contraditórias) e "*ambiguous wording or information*" (linguagem vaga que gera dúvida)	Política de privacidade com termos como "podemos compartilhar seus dados com parceiros confiáveis" sem especificar quem são.

Ao enfraquecerem a integridade do consentimento e desrespeitarem princípios como a autodeterminação informativa e a boa-fé, essas práticas revelam um quadro mais amplo de manipulação. A diversidade crescente de tipologias, que vão da taxonomia de doze classificações de Brignull (2018) às sessenta e oito classificações identificadas por Li *et al.* (2024), reflete uma sofisticação manipulativa que continua a se expandir. Dada essa plasticidade, regulamentações centradas em aspectos visuais específicos, como cores ou disposição de botões, tendem a ser ineficazes. Por isso, uma abordagem baseada em princípios, que avalie a intenção de manipular e os impactos sobre a autonomia do usuário, mostra-se mais promissora no longo prazo (Posson, 2025).

4. Panorama regulatório e casos de aplicação em diversas jurisdições

A crescente inquietação global com a proliferação dos padrões obscuros tem impulsionado diversas jurisdições ao redor do mundo a desenvolver e implementar arcabouços regulatórios específicos. Essas iniciativas, embora distintas em suas abordagens

e níveis de maturidade, convergem na necessidade de conter as práticas manipulativas que se tornaram onipresentes no ambiente digital e que, cada vez mais, ameaçam a autonomia dos usuários. Vamos, então, discutir esses desenvolvimentos.

4.1 União europeia: GDPR, DSA e a atuação de autoridades

A União Europeia tem se destacado como uma das regiões mais rigorosas no enfrentamento aos *dark patterns*. Apesar da ausência de menção explícita no General Data Protection Regulation (GDPR), princípios como a transparência[3] e o consentimento livre e informado[4] já ofereciam base para sua regulação. O EDPB, em suas Diretrizes 03/2022, foi categórico ao considerar inválido o consentimento obtido por meio de mecanismos manipulativos[5]. A Lei de Serviços Digitais (Digital Services Act – *DSA*), por meio de seu artigo 25, consolida essa tendência com uma proibição direta de arranjos de interface que influenciem negativamente o comportamento do usuário[6]. A Commission Nationale de l'Informatique et des Libertés (CNIL) tem se destacado pela atuação rigorosa. Conforme noticiado por Cheminat (2024), a agência aplicou penalidades severas ao Yahoo em 2024 por dificultar a recusa de *cookies* pelos usuários.

3 Art. 5: "*Personal data shall be processed lawfully, fairly and in a transparent manner in relation to the data subject* ('lawfulness, fairness and transparency')" (União Europeia. Regulamento Geral sobre a Proteção de Dados [GDPR]. Bruxelas: Parlamento Europeu e Conselho, 2016).

4 Art. 7 (2): "1. *If the data subject's consent is given in the context of a written declaration which also concerns other matters, the request for consent shall be presented in a manner which is clearly distinguishable from the other matters, in an intelligible and easily accessible form, using clear and plain language. 2. Any part of such a declaration which constitutes an infringement of this Regulation shall not be binding.*" (União Europeia. op. cit., 2016).

5 Cf. Guidelines 03/2022 on Deceptive design patterns in social media platform interfaces: how to recognise and avoid them. Version 2.0, adopted on 14 February 2023

6 Art. 25: "*Providers of online platforms shall not design, organise or operate their online interfaces in a way that deceives or manipulates the recipients of their service or in a way that otherwise materially distorts or impairs the ability of the recipients of their service to make free and informed decisions*" (União Europeia. Digital Services Act [DSA]. Bruxelas: Parlamento Europeu e Conselho, 2022.)

Vale citar que na União Europeia, práticas associadas a padrões obscuros podem ser enquadradas tanto pelo GDPR quanto pela Lei de Serviços Digitais ou pela Lei de Inteligência Artificial, com multas que chegam a 4% ou 6% do faturamento global. A atuação do Garante per la protezione dei dati personali (Garante), na Itália, segue na mesma linha. Corti (2023) relata que a Ediscom S.p.A. foi multada em trezentos mil euros (cerca de R$1.890.180,00) pelo uso de interfaces enganosas, como *pop-ups* reincidentes e botões de recusa ocultos. A autoridade reforçou que todas as empresas, inclusive as de pequeno porte, devem seguir as mesmas normas éticas.

4.2 Eua: FTC, leis estaduais e precedentes de *enforcement*

Nos Estados Unidos, a regulação dos padrões obscuros apoia-se majoritariamente em leis gerais de proteção ao consumidor, embora novas legislações estaduais venham abordando essas práticas de maneira mais direta. Ao analisar a atuação recente da Federal Trade Commission (FTC), Finegold (2025) destaca o uso da Seção 5 do *FTC Act* como ferramenta contra condutas enganosas ou injustas, citando os casos contra a Publishers Clearing House[7] e a DirecTV[8]. Em 2024, uma análise conjunta da FTC, *I*nternational Consumer Protection Enforcement Network (ICPEN) e Global Privacy Experts Network (GPEN) revelou a ampla disseminação de padrões obscuros em plataformas de assinatura, conforme noticiado por Henderson (2024). O Consumer Financial Protection Bureau (CFPB)

7 A FTC processou a PCH por utilizar *dark patterns* para induzir consumidores a acreditarem que a compra de produtos aumentaria suas chances de ganhar prêmios, o que não era verdade. (Cf. Publishers Clearing House, LLC (PCH), v. FTC. Washington, D.C, 2025)

8 A FTC alegou que a DirecTV falhou em informar adequadamente que seus preços promocionais eram limitados aos primeiros 12 meses, apesar de os contratos exigirem compromisso de 24 meses. Além disso, acusou a empresa de adotar práticas de opção negativa ao oferecer canais "gratuitamente" por um trimestre, iniciando posteriormente a cobrança automática sem aviso suficiente. A FTC levou o caso ao tribunal federal da Califórnia, argumentando que tais condutas violavam a ROSCA, legislação que obriga a divulgação clara de termos de assinatura e requer consentimento expresso para cobranças adicionais. A ação, no entanto, foi prejudicada em 2018, quando o tribunal rejeitou boa parte das alegações, levando a FTC a desistir do caso (Cf. DirecTV, FTC v. Washington, D.C., 2015).

também tem atuado contra práticas digitais manipulativas, como demonstrado na sanção aplicada à TransUnion[9]. No âmbito estadual, normas como a California Consumer Privacy Act (CCPA), a California Consumer Privacy Act of 2018 (CPRA) e a Colorado Privacy Act oferecem diretrizes claras sobre padrões obscuros, como destaca Oliveira (2023). A CCPA, por exemplo, estipula que "o acordo obtido através do uso de padrões obscuros não constitui consentimento válido" e exige "simetria na escolha" para as opções de privacidade[10]. De forma geral, a abordagem norte-americana é fragmentada e combina normas federais genéricas com regras estaduais emergentes, criando um cenário complexo. Embora a Lei DETOUR (Dark Patterns Elimination and Oversight Reform Act) tenha sido proposta no Congresso para combater padrões obscuros em grandes plataformas digitais, a medida não prosperou[11]. Diferentemente do modelo coeso da UE, os EUA operam sob um verdadeiro “mosaico” regulatório, onde a identificação de padrões obscuros está condicionada ao contexto em que se manifestam, o que resulta em incerteza jurídica e aplicação desigual entre os estados.

4.3 Brasil: LGPD e a atuação da agência nacional de proteção de dados (ANPD)

A promulgação da LGPD redefiniu o cenário da proteção de dados no Brasil ao estabelecer que o consentimento não pode ser presumido nem imposto, mas deve ser uma manifestação clara da vontade do titular. A consolidação dessa transformação normativa se materializa no artigo 5º, inciso XII, ao caracterizar o consentimento como “manifestação livre, informada e inequívoca”. Vale dizer que essa liberdade não pode ser meramente simbólica, sendo necessário que o titular tenha reais condições de escolha, sem so-

9 A CFPB processou a TransUnion por enganar consumidores com o uso de padrões obscuros, levando-os a se inscrever em assinaturas recorrentes sem consentimento explícito. Os usuários eram levados a clicar em botões que aparentavam conceder relatórios de crédito gratuitos, mas que ativavam, de forma oculta, cobranças mensais (Cf. CFPB vs. TransUnion. Washington, D.C., 2023)

10 Cf. Enforcement Advisory No. 2024-02

11 Cf. *S.3330 - DETOUR Act*

frer qualquer tipo de coação ou manipulação ao exercê-la. E é justamente essa falta de liberdade efetiva que está na raiz das críticas aos padrões obscuros. Não é à toa que Arvigo e Carboni (2022) enfatizam que o descumprimento de quaisquer desses critérios pode resultar na invalidação e nulidade do tratamento de dados.

No tocante à informação, o art. 9º enumera os elementos que devem ser apresentados ao titular, entre eles, a finalidade específica do tratamento, sua duração, a identificação do controlador, o compartilhamento de dados e os direitos do titular. Em muitos contextos digitais, no entanto, a apresentação dessas informações de forma fragmentada, excessivamente técnica ou pouco acessível compromete a compreensão e, por consequência, a validade do consentimento (Rodrigues *et al.*, 2024).

Além disso, a inequivocidade exige que a concordância do titular seja expressa por meio de uma ação positiva. O simples uso de um serviço ou a ausência de oposição não configuram consentimento válido. Trata-se de uma garantia de que o titular exerceu controle sobre sua decisão de forma consciente. A ANPD, embora ainda em processo de consolidação regulatória sobre padrões obscuros, adotou medidas proativas, como demonstra a publicação do Guia Orientativo de Cookies e Proteção de Dados Pessoais[12]. Com o consentimento agora formalmente incluído na Agenda Regulatória 2025-2026[13], o Brasil tem a oportunidade de consolidar uma regulação moderna sobre design ético e controle automatizado. A literatura nacional, com trabalhos de Guerra (2024), Oliveira *et al.* (2024) e Santos (2023), já vem alertando sobre os impactos desses padrões na privacidade e na autonomia do usuário. A punição aplicada à Meta Platforms no caso "Meta IA"[14] (motivada por funda-

12 Cf. Guia Orientativo de Cookies e Proteção de Dados Pessoais

13 Cf. Resolução nº. 23, 09 dez. 2024.

14 Em 2 de julho de 2024, a ANPD determinou a suspensão cautelar da nova política de privacidade da Meta, que permitia o uso de dados pessoais publicados em suas plataformas (Facebook, Instagram e Messenger) para treinar modelos de inteligência artificial generativa. A decisão foi tomada após a Autoridade identificar indícios de violações à LGPD, incluindo o uso de hipótese legal inadequada para o tratamento de dados pessoais, falta de transparência na divulgação das mudanças na política de privacidade, limitação excessiva dos direitos dos titulares e riscos para crianças e adolescentes sem salvaguardas adequadas (Cf. Voto nº. 11/2024/DIR-MW/CD; Processo nº 00261.004509/2024-36).

mentos como base legal inadequada, falta de clareza e restrições aos direitos dos titulares) evidencia o compromisso da agência com a aplicação efetiva da LGPD. Pereira (2024) observa que o grupo de Mark Zuckerberg recorreu à prática de *obstruction/roach motel*, exigindo quase dez passos para que o usuário revogasse o consentimento referente ao uso de seus dados para inteligência artificial, dificultando de forma intencional o exercício desse direito. Miriam Wimmer, em seu voto, endossa essa leitura, ao afirmar que "a opção de opt-out fornecida aos usuários, que permitiria aos titulares se opor ao tratamento de seus dados pessoais, não é disposta de maneira evidente, e a complexidade para exercício dessa opção assemelha-se a um padrão obscuro de mascaramento de informações" e que "o número elevado de ações que o usuário precisa realizar para expressar a sua oposição em relação ao tratamento de seus dados pode levá-lo a tomar decisões que seriam contrárias à sua vontade" (ANPD, 2024, p. 3). A atuação do órgão, ainda que sem referência nominal aos padrões obscuros, já demonstra preocupação com práticas que violam os princípios da LGPD, especialmente no que tange à autodeterminação informativa e ao livre desenvolvimento da personalidade. Cabe observar que, mesmo na ausência de regulação específica, tais práticas já encontram enfrentamento eficaz no ordenamento brasileiro por meio do Código de Defesa do Consumidor. E como bem recorda Frazão (2024), é possível também a aplicação cumulativa da LGPD nos casos que envolvam uso indevido de dados pessoais e da legislação concorrencial quando caracterizado o abuso de posição dominante.

5. Problematizações e desafios na regulação e combate aos dark patterns

Apesar das diversas iniciativas legislativas e das crescentes sanções aplicadas por autoridades de proteção de dados e defesa do consumidor, os padrões obscuros permanecem resistentes ao controle regulatório. A seguir, exploraremos os principais desafios práticos e jurídicos enfrentados nesse processo.

5.1 Fragmentação legal e inconsistência na aplicação

O regime jurídico aplicável aos padrões obscuros é frequentemente descrito como fragmentado, em razão da ausência de uma definição legal uniforme e da sobreposição de esferas regulatórias. A coexistência de normas de proteção ao consumidor e de proteção de dados, aliada às diferenças entre sistemas jurídicos nacionais, contribui para um ambiente normativo labiríntico, que dificulta a aplicação sistemática e previsível (Posson, 2025). A depender do enquadramento jurídico adotado, a aplicação pode recair sob a alçada da proteção ao consumidor (tratada por tribunais nacionais) ou da proteção de dados (supervisionada por autoridades de proteção de dados), o que potencializa o risco de interpretações divergentes e compromete a segurança regulatória. No cenário europeu, Polona (2025) bem observa que, embora a DSA proíba o uso de padrões obscuros por plataformas online, ela exclui da sua aplicação práticas já cobertas pela Unfair Commercial Practices Directive (UCPD) e pelo GDPR. Ou seja, se uma conduta for considerada um dark pattern sob a DSA, mas também constituir infração ao GDPR, prevalece a análise conforme este último, que, por sua vez, não trata diretamente do conceito. Em consequência, a efetividade da DSA é comprometida, dada a predominância de uma avaliação casuística com base na UCPD. Posson (2025) complementa que a prioridade deve ser esclarecer a interação entre os marcos existentes, e não necessariamente criar outros. Castro, Dascoli e Diebold (2022) observam que essa necessidade se torna ainda mais evidente quando voltamos os olhos aos Estados Unidos, onde a convivência de legislações estaduais e federais impõe altos custos de conformidade (*compliance costs*) às empresas e oferece proteção desigual aos consumidores, que varia conforme o estado em que residem. Por isso é crescente o apelo por maior coordenação internacional e por diretrizes harmonizadas que tornem o ambiente regulatório mais coerente e eficiente. O debate atual gira justamente em torno de se novas leis são imprescindíveis ou se uma aplicação mais clara e uniforme das normas já existentes seria suficiente (Posson, 2025; Moore; Culligan, 2024).

5.2 Dificuldade de identificação e prova do dano

O prejuízo à privacidade, em especial, tende a ser invisível quando ocorre e, muitas vezes, sequer é percebido pelo titular dos dados. O desequilíbrio entre o benefício imediato de utilizar um serviço e o custo, mais difuso, da perda de privacidade futura é difícil de mensurar. Como podemos ver, isso se deve à própria natureza, pois operam de forma sutil e se apoiam em mecanismos psicológicos que agem fora da consciência dos usuários (*Zac et al.*, 2025). No Brasil, por exemplo, não há registro de litígios que mencionem expressamente os termos "padrão obscuro", "*dark pattern*" ou "*deceptive pattern*"[15]. Em regra, a materialização do prejuízo se dá apenas a posteriori, em decorrência de vazamentos de dados[16]. Conforme relatado por Lorenzo (2025), o país liderou o ranking global de vazamento de cookies, com mais de sete bilhões de informações pessoais disponibilizadas em mercados ilegais. Entre os dados expostos estavam credenciais de *login* capazes de permitir o roubo de sessões sem o uso de senha, além de nomes de usuário, endereços de *e-mail*, país, cidade, gênero, data de nascimento e, em alguns casos, o endereço residencial dos titulares.

Convém lembrar que grande parte das ferramentas automatizadas existentes não abrangem a totalidade das tipologias de padrões obscuros, o que reforça a importância de avaliações humanas, sejam elas técnicas ou jurídicas. Esse foi o ponto que *Li et al.* (2024) buscaram comprovar ao propor uma taxonomia com ses-

15 A pesquisa foi conduzida na plataforma "Inspira AI" em 17 de junho de 2025, com o uso de filtros que incluíram os principais órgãos jurisdicionais e administrativos brasileiros. Foram consultados os Tribunais Superiores (STF, STJ e TST), os Tribunais Regionais Federais (TRF1 a TRF5), os Tribunais de Justiça de todas as unidades federativas, diversas entidades administrativas tributárias (como CARF, DRJ, SEFAZ estaduais e o TIT-SP), além de entidades regulatórias (BACEN, CADE, CVM e TCU) e os 24 Tribunais Regionais do Trabalho (TRT1 a TRT24).

16 Segundo o entendimento consolidado pelo STJ, a exposição indevida de dados pessoais comuns (i.e., não sensíveis), por si só, não enseja o reconhecimento de dano moral presumido (*in re ipsa*). Nesses casos, cabe ao titular demonstrar que sofreu prejuízo concreto, seja de ordem emocional, reputacional ou econômica, para que haja possibilidade de reparação. Cf.: Efing, A., Andretta, J. Vazamento de dados pessoais e o direito à indenização por dano moral in re ipsa. Civilistica, Rio de Janeiro, v. 14, n. 1, p. 1–16, 2025.

senta e oito categorias distintas. Dessas, apenas trinta e uma foram detectáveis pelas principais soluções automatizadas, demonstrando uma taxa de cobertura inferior a 50% e revelando grandes lacunas na infraestrutura de monitoramento disponível. Não admira, pois, que o combate aos padrões obscuros se mostre limitado, lento e à mercê de ações reativas. Fato é que essa conjuntura abre caminho para a perpetuação de condutas manipulativas com pouca ou nenhuma responsabilização.

5.3 Implicações da inteligência artificial e personalização

A convergência entre inteligência artificial e padrões obscuros sinaliza uma nova fase na evolução dos riscos digitais associados à manipulação do comportamento do usuário. O uso de aprendizado de máquina (*machine learning*) combinado com coleta massiva de dados de comportamento e personalização em tempo real torna esses mecanismos mais eficazes e difíceis de detectar (Oliveira, 2023, p. 61). Ao contrário dos *designs* estáticos, os padrões baseados em IA ajustam-se automaticamente às reações dos usuários, comprometendo assim a capacidade de fiscalização posterior. Rosala (2023) e Potel-Saville (2024) destacam o papel de técnicas como testes A/B[17] e modelos algorítmicos de experimentação contínua na criação de estratégias específicas, moldadas para influenciar decisões individuais com alta precisão. A trajetória observada conduz inevitavelmente a um cenário digital sustentado por mecanismos contínuos de vigilância e indução comportamental automatizada, em que o desafio regulatório passa a ser a capacidade de reagir com a mesma velocidade com que a tecnologia amplia o alcance dessas práticas.

17 O teste A/B constitui um procedimento experimental desenhado para cotejar duas distintas versões de um componente (como uma página eletrônica, uma publicidade ou uma mensagem de correio), com o intuito de discernir qual delas gera o melhor rendimento. Sua funcionalidade baseia-se na criação de uma versão A (servindo como base) e uma versão B (incorporando uma alteração). Os usuários são designados de forma aleatória a uma das versões, e suas interações são documentadas e escrutinadas para se identificar aquela com desempenho superior. Cf.: Gallo, A. A Refresher on A/B Testing. Harvard Business Review, 28 jun. 2017.

5.4 Recomendações para a ANPD

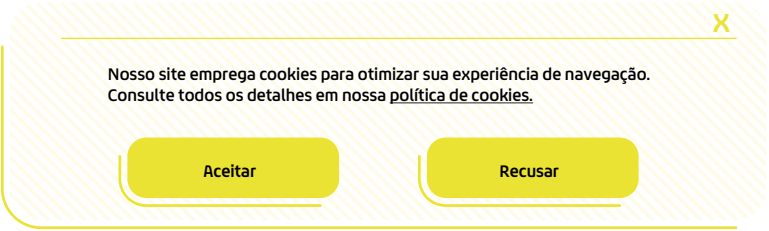
As competências previstas no art. 55-J da LGPD[18] atribuem à ANPD funções estratégicas de regulação e fiscalização. A partir desse marco, delineamos um conjunto de recomendações voltado à consolidação de uma agenda regulatória mais proativa. Entre as frentes disponíveis, duas se destacam pela capacidade de produzir efeitos estruturantes: a produção de diretrizes normativas precisas e a implementação de ações educativas em larga escala. Nos tópicos seguintes, abordamos essas direções com maior profundidade.

5.5 Diretrizes para interfaces transparentes e não manipulativas

Se o *design* pode violar normas por si só, como apontam Arvigo e Carboni (2022), é necessário observância aos aspectos legais e éticos desde o início do processo criativo (*privacy by design*), e não os tratar posteriormente. Moore e Culligan (2024) complementam ao enfatizar a centralidade da transparência: toda informação relevante deve ser exposta com clareza, sem filtros técnicos ou barreiras cognitivas. A simetria entre as opções oferecidas aos usuários é igualmente fundamental, sendo inaceitável que escolhas menos protetivas à privacidade recebam destaque visual ou sejam operacionalmente mais simples do que aquelas que asseguram maior resguardo aos dados pessoais.

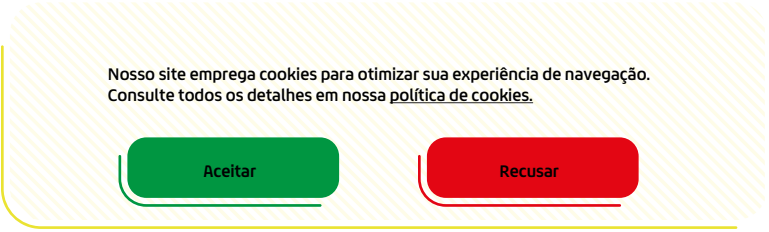
18 Art. 55-J. Compete à ANPD: "[...] VI - promover na população o conhecimento das normas e das políticas públicas sobre proteção de dados pessoais e das medidas de segurança; VII - promover e elaborar estudos sobre as práticas nacionais e internacionais de proteção de dados pessoais e privacidade; VIII - estimular a adoção de padrões para serviços e produtos que facilitem o exercício de controle dos titulares sobre seus dados pessoais, os quais deverão levar em consideração as especificidades das atividades e o porte dos responsáveis" (Cf. Lei nº. 13.709, de 14 de agosto de 2018).

Figura 1 - Interface representando uma boa prática de privacy by design ao oferecer escolhas de consentimento para cookies equilibradas visualmente e um botão de fechamento que garante o acesso ao conteúdo, promovendo a transparência e a autonomia do usuário sem coerção



Fonte: Autora

Figura 2 - IInterface representando uma má prática de design que se manifesta pela assimetria visual, destacando o botão "aceitar" (em verde) em detrimento do "recusar" (em vermelho), caracterizando um padrão obscuro de "stirring". Também se nota a ausência de uma opção para fechar a janela, privando o usuário de acessar o conteúdo sem tomar uma decisão.



Fonte: Autora

Arvigo e Carboni (2022) argumentam que o caminho para uma proteção de dados mais efetiva passa por uma aproximação entre design e direito, de modo a conceber interfaces que informem de maneira clara e possibilitem escolhas conscientes, mas que ao mesmo tempo evitem a saturação de informações que comprometa a tomada de decisão dos titulares. A eliminação de padrões manipulativos, como a pré-seleção de opções ou o carregamento

automático de cookies antes do consentimento[19], é outra diretriz importante. Quanto à retirada do consentimento, espera-se que o processo seja tão simples e direto quanto a sua concessão, sem etapas excessivas ou estratégias dissuasivas. Adicionalmente, há uma ênfase crescente na proteção de grupos vulneráveis, sobretudo crianças e adolescentes, os quais demandam camadas adicionais de proteção. Rodrigues *et al.* (2024, p. 26) relembram que, embora o Brasil não possua uma legislação federal específica sobre design, a garantia e efetivação dos direitos da criança e do adolescente no ambiente digital é pautada pelo princípio da "garantia dos direitos das crianças e adolescentes por design dos produtos e serviços em ambientes digitais", conforme o Art. 3, X, da Resolução n. 245/2024[20]. Por fim, o design voltado ao consentimento informado deve incorporar recursos visuais, interativos e acessíveis, assegurando que o usuário compreenda as implicações de suas escolhas (Rodrigues *et al.*, 2024, p. 32). Rosala (2023, tradução nossa) sugere uma lista com perguntas que devem guiar o processo desde sua concepção:

1. Os usuários poderiam gastar mais ou fornecer mais dados do que pretendiam ou precisavam?
2. Quando os usuários consentem com algo em troca de uma capacidade, produto ou experiência, a troca é justa e apropriada?
3. As informações apresentadas sobre cada escolha estão factualmente corretas?
4. Dada a forma como as informações ou opções são apresentadas, os usuários podem facilmente interpretar mal as escolhas (ou disponibilidade de escolha)?

19 Gralha (2022, p. 28) chama atenção para o fato de que, em muitos casos, não é a compra que ativa a coleta de dados, mas o próprio ato de visitar um ambiente digital comercial. Antes mesmo de qualquer interação relevante, o usuário é instado a consentir com cookies que prometem adaptar a experiência. Embora muitas vezes a navegação não avance e o interesse seja rapidamente abandonado, os dados capturados alimentam sistemas de publicidade que se valem daquela breve interação para moldar futuras abordagens com ainda mais precisão. Cf.: Gralha, P. Os limites materiais do consentimento dos titulares de dados no comércio eletrônico: uma análise crítica das políticas de privacidade dos 30 (trinta) maiores varejistas/marketplaces do brasil em 2022. Dissertação (mestrado). Orientador: Luca Belli. Escola de Direito do Rio de Janeiro da Fundação Getulio Vargas, 2022.

20 Cf. *Resolução nº. 245, de 5 de abril de 2024*

5. Os usuários podem perder outra opção na interface (por exemplo, porque ela está obscurecida ou em um local que o usuário pode não esperar)?
6. Os usuários podem perder uma informação essencial que os ajudaria a fazer uma escolha?
7. Os usuários podem acessar rapidamente todas as informações sobre cada opção disponível?
8. Os usuários podem implementar rapidamente uma escolha que desejam fazer (ou muitas etapas desnecessárias os bloqueiam)?
9. Os usuários são apressados em tomar uma decisão?
10. Os usuários são injustamente pressionados ou manipulados emocionalmente ao fazer uma escolha?
11. Os usuários podem se sentir envergonhados, nervosos ou culpados ao recusar uma escolha?

A ANPD já iniciou sua atuação no enfrentamento das distorções provocadas por práticas de *design* manipulativo, o que é muito bem apreciado. O Guia Orientativo sobre o Uso de Cookies inclui uma seção específica sobre banners (ANPD, 2022, p. 30-36), na qual se delineiam orientações claras sobre o que constitui um consentimento válido e como aplicar os princípios da LGPD ao uso de *cookies*. Mas, apesar de seu mérito, trata-se de um documento técnico em formato estático (.pdf), com alcance reduzido fora dos setores especializados. Considerando o peso que decisões de interface têm sobre a liberdade do titular, a comunicação regulatória exige maior acessibilidade e dinamismo. Nesse sentido, a experiência francesa, com o portal Data & Design by LINC[21], oferece um exemplo de como apresentar diretrizes por meio de uma linguagem visual e interativa, adequada tanto ao público técnico quanto a desenvolvedores em contextos diversos.

Esse tipo de ação ainda faz falta no Brasil. A criação e difusão, pela ANPD, de uma ferramenta com esse perfil teria potencial para democratizar a compreensão regulatória, ampliar o alcance das normas existentes e consolidar a cultura do *privacy by design* no desenvolvimento digital nacional.

21 Cf. https://design.cnil.fr/en/

5.6 Educação e conscientização dos usuários

A capacidade de reconhecer padrões obscuros não elimina a influência que eles exercem sobre o comportamento do usuário. Oliveira (2023, p. 60) já chamava atenção para a fragilidade estrutural do sujeito conectado, percepção que Zac *et al.* (2025) buscaram testar empiricamente no estudo intitulado "*Dark Patterns and Consumer Vulnerability*". Com base em um experimento realizado com mais de mil e quinhentas pessoas em uma plataforma simulada de investimentos, os autores demonstraram que, mesmo quando percebidos conscientemente, esses padrões continuam a afetar escolhas e atitudes, uma vez que se valem de heurísticas e processos inconscientes de tomada de decisão. Para tornar o raciocínio mais palpável, imagine o leitor diante de uma plataforma de reservas que anuncia, com temporizador visível, a última vaga restante em um hotel. Mesmo que desconfie da veracidade do alerta, o impulso gerado pelo medo da escassez pode mobilizar uma resposta rápida, reduzindo a margem para ponderação.

Já em outra abordagem, Calonga *et al.* (2022) exploraram a visão de jovens brasileiros sobre os *dark patterns* por meio de entrevistas semiestruturadas e grupos focais. Houve um consenso de que esses padrões são insidiosamente implementados, priorizando o benefício das empresas em detrimento do usuário. A pesquisa também revelou a crença de que é "impossível evitar os padrões sombrios", impulsionando os usuários a aceitar que terão de "aprender a conviver com eles" devido à sua dependência de certos serviços e empresas (Calonga *et al.*, 2022, p. 16). O que podemos inferir é que a abordagem educacional eficaz não pode se limitar à denúncia das técnicas de manipulação ou à explicitação de elementos de design enganosos. O essencial é compreender os mecanismos mentais que sustentam essas escolhas induzidas, inclusive quando o indivíduo reconhece o padrão. Retomando Calonga *et al.* (2022), observa-se que, embora a conscientização não seja garantia de imunidade, ela continua relevante na mitigação dos efeitos dessas práticas., considerando que a permissividade frente ao uso indevido de dados e à indução de decisões potencialmente sensíveis (como contratações ou compras) representa uma falha ética e regulatória que precisa ser enfrentada com a devida seriedade.

Em uma linha convergente, Baumeister *et al.* (2024, p. 79) sugerem uma perspectiva suplementar: à moda das iniciativas de combate a fraudes bancárias, convém intensificar a sensibilização da coletividade sobre os padrões obscuros. É claro que a prudência gerada pode não ser suficiente para impedir todos os abusos, mas já representa um avanço ao estabelecer um ambiente de vigilância coletiva e de pressão civil sobre os entes reguladores. Somado a isso, inserir tal matéria nos currículos escolares desde a educação básica configura uma chance estratégica de munir os cidadãos contra um universo digital repleto de engodos visuais e emocionais. Compreendendo as bases dessas táticas, o usuário estará em melhor condição para se contrapor a elas e decidir com maior discernimento.

6. Conclusão

Ao longo das páginas deste texto, tornou-se claro que os dark patterns transcenderam a condição de meras anomalias para se firmarem como um componente intrínseco e frequentemente invisível da experiência digital contemporânea. Longe de serem exceções isoladas, essas estratégias, embora mantenham a roupagem de escolhas legítimas e até mesmo convenientes, operam como reproduções sofisticadas de estruturas meticulosamente desenhadas para manipular o comportamento do usuário. Sua gênese, que remonta a táticas de persuasão empregadas no varejo físico, encontrou no ambiente digital um terreno fértil para sua intensificação e expansão, impulsionadas pela vasta coleta de dados e pela capacidade de personalização em tempo real. Essa evolução conferiu-lhes uma sofisticação e amplitude sem precedentes, com um propósito singular: otimizar lucros, muitas vezes à custa da autonomia informacional dos indivíduos. Para Guerrini (2024), essa ubiquidade das táticas de manipulação e engano presentes no *design* digital hodierno faz com que superá-las pareça uma façanha quase impossível. Ainda assim, a confluência de medidas regulatórias, o aumento da lucidez dos consumidores e o advento de novas empresas empenhadas em resolver a questão oferece um alento para a modificação do quadro. Concordamos com ele.

Referências

ANPD. **Guia Orientativo: Cookies e Proteção de Dados Pessoais**. Brasília, DF: ANPD, 2022. Brasília, 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: 24 nov. 2025.

ANPD. **Voto nº 11/2024/DIR-MW/CD no âmbito do processo decisório SEI nº 00261.004509/2024-36**. Conselho Diretor. Diretora Relatora: Miriam Wimmer. Brasília: ANPD, 2024. Disponível em: https://www.gov.br/anpd/pt-br/assuntos/noticias/anpd-determina-suspensao-cautelar-do-tratamento-de-dados-pessoais-para-treinamento-da-ia-da-meta/SEI_0130047_Voto_11.pdf. Acesso em: 27 nov. 2025..

BAUMEISTER, J.; PARK, J.; CUNNINGHAM, A.; VON ITZSTEIN, S.; GWILT, I.; DAVIS, A.; WALSH, J. **Patterns in the Dark: Deceptive Practices in Online Interactions**. Relatório para o Data Standards Chair. University of South Australia, 2024. Disponível em: https://dsb.gov.au/sites/dsb.gov.au/files/2024-11/report-patterns-in-the-dark.pdf?utm_source=chatgpt.com. Acesso em: 24 nov. 2025.

BRASIL. Presidência Da República. Lei nº 13.709 de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais - LGPD. **Diário Oficial da União**: seção 1, Brasília, DF, ed. 157-A, p. 1, 15 agosto 2018.

BRIGNULL, H. **Dark Patterns**, 2018. Disponível em: https://darkpatterns.org/. Acesso em: 10 jun. 2025.

CALIFORNIA PRIVACY PROTECTION AGENCY. **Enforcement Advisory** No. 2024-02: Avoiding Dark Patterns. Califórnia, 2024.

CALONGA, L. O. L.; SOARES, C. D. M.; MELO, T. C.; MACHADO, L. M. Pensa que me engana, eu finjo que acredito: padrões obscuros sob a perspectiva do usuário. In: ENCONTRO DA ASSOCIAÇÃO NACIONAL DE PÓS-GRADUAÇÃO E PESQUISA EM ADMINISTRAÇÃO, 46., 2022, [*s. l.*]. **Anais** [...]. Maringá: ANPAD, 2022. Disponível em: https://anpad.com.br/uploads/articles/120/approved/8a88d5f412f2ad376f8597d28cbd3720.pdf. Acesso em: 24 nov. 2025..

CASTRO, D., DASCOLI, L., DIEBOLD, G. The Looming Cost of a Patchwork of State Privacy Laws. **ITIF – Information Technology & Innovation Foundation**, 24 jan. 2022.

CHEMINAT, J. **Cookies: la Cnil inflige 10 M€ d'amende à Yahoo**. Le Monde Informatique, 18 jan. 2024.

CONSUMER FINANCIAL PROTECTION BUREAU. **CFPB vs. TransUnion**. Washington, D.C., 2023.

CORTI, F. Legal design e strategie di condizionamento del consenso. Cosa ci insegna il caso Ediscom?. **Labor Project**, 05 jul. 2023.

DAVIS, A.; WALSH, J. **Patterns in the Dark: Deceptive Practices in Online Interactions**. Relatório para o Data Standards Chair. University of South Australia, 2024.

EFING, A.; ANDRETTA, J. Vazamento de dados pessoais e o direito à indenização por dano moral in re ipsa. **Civilistica**, Rio de Janeiro, v. 14, n. 1, p. 1–16, 2025.

EUROPEAN DATA PROTECTION BOARD. **Guidelines 03/2022 on Deceptive Design Patterns in Social Media Platform Interfaces: How to Recognise and Avoid Them**. Version 2.0. Bruxelas, 2023.

FEDERAL TRADE COMMISSION. **DirecTV, FTC v. Washington**, D.C.: FTC, 2015.

FEDERAL TRADE COMMISSION. **Publishers Clearing House, LLC (PCH), FTC v. Washington, D.C.:** FTC, 2025.

FINEGOLD, R. Forthcoming Litigation for Companies That Employ Dark Patterns. **The University of Chicago Business Law Review**, 2025.

FRAZÃO, A. Como conter as dark patterns? **Jota**, 12 set. 2024.

GALLO, A. A Refresher on A/B Testing. **Harvard Business Review**, 28 jun. 2017.

GRALHA, P. **Os limites materiais do consentimento dos titulares de dados no comércio eletrônico: uma análise crítica das políticas de privacidade dos 30 (trinta) maiores varejistas/marketplaces do brasil em 2022**. Dissertação (Mestrado em Direito). Escola de Direito do Rio de Janeiro da Fundação Getulio Vargas, FGV, 2022.

GRAY, C.; SANTOS, C.; BIELOVA, N.; TOTH, M.; CLIFFORD, D. Dark Patterns and the Legal Requirements of Consent Banners: An Interaction Criticism Perspective. In: PROCEEDINGS OF THE 2021 CHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS (CHI '21). **Association for Computing Machinery**, New York, NY, USA, A. 172, 1–18.

GUERRA, S.; WIENSKOSKI, L. Dark patterns: uma nova agenda regulatória para o Brasil?. **Revista de Direito Administrativo & Constitucional**. Belo Horizonte, v. 24, n. 98, p. 137–160, 2024. DOI: 10.21056/aec.v24i98.2011.

GUERRINI, F. AI-Driven Dark Patterns: How Artificial Intelligence Is Supercharging Digital Manipulation. **Forbes**, 17 nov. 2024.

HENDERSON, J. **FTC, ICPEN, GPEN Announce Results of Review of Use of Dark Patterns Affecting Subscription Services**. Federal Trade Comission, 10 jul. 2024.

INTERNATIONAL CONSUMER PROTECTION ENFORCEMENT NETWORK . **Dark Patterns in Subscription Services Sweep**: Public Report. 02 jul. 2024.

LI, M.; WANG, X.; NIE, L.; LI, C.; LIU, Y.; ZHAO, Y.; LEI, X.; SAID, K. S. **A Comprehensive Study on Dark Patterns**. arXiv preprint arXiv:2412.09147, 2024.

LORENZO, A. Brasil lidera vazamento de 'cookies' na internet, diz pesquisa. **Olhar Digital**, 13 jun. 2025.

MAZUMDAR, S., BLUE, S. **Responsible Design Part 9 of 14: Roach Motel**. Think Design – A Havas Company, 2022.

MINISTÉRIO DA JUSTIÇA E SEGURANÇA PÚBLICA. Conselho Diretor da Autoridade Nacional de Proteção de Dados – ANPD. Resolução nº 23, de 09 de dezembro de 2024. Aprova a Agenda Regulatória para o biênio 2025-2026. **Diário Oficial da União**: seção 1, Brasília, DF, ed. 238, p. 115, 11 dezembro 2024.

MINISTÉRIO DOS DIREITOS HUMANOS E DA CIDADANIA. Conselho Nacional dos Direitos da Criança e do Adolescente - CONANDA Resolução nº 245, de 5 de abril de 2024. Dispõe sobre os direitos das crianças e adolescentes em ambiente digital. **Diário Oficial da União**: seção 1, Brasília, DF. ed. 68, p. 42, 5 abril 2024.

MOORE, L.; CULLIGAN, R. Dark Patterns: Not a new concept but will now be heavily regulated. **Lexology**, 09 fev. 2024.

OLIVEIRA, T. **A regulação da inteligência artificial e os dark patterns nas redes sociais: uma análise sobre a proteção dos direitos fundamentais**. Dissertação (Mestrado em Direito e Ciência Jurídica). Faculdade de Direito da Universidade de Lisboa, Lisboa, 2024.

OLIVEIRA, T. C.; COLETI, T. A.; MORANDINI, M.; BALANCIERI, R.; OLIVEIRA, A. L. Dark Patterns nos marketplaces: uma Investigação com Base nas Reclamações dos Consumidores. In: WORKSHOP INVESTIGAÇÕES EM INTERAÇÃO HUMANO-DADOS (WIDE), 3, 2024, Brasília/DF. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2024. p. 39-45.

PEREIRA, M. C. M. **ANPD e Meta**: Como a Autoridade pode transformar e orientar o trabalho dos UX Designers? **Blog** Instituto de Pesquisa em Direito e Tecnologia do Recife, Disponível em: https://ip.rec.br/blog/anpd-e-meta-como-a-autoridade-pode-transformar-e-orientar-o-trabalho-dos-ux-designers/. Acesso em: 24 nov. 2025.

POLONA, C. **Regulating dark patterns in the EU: Towards digital fairness**. Think Tank – European Parliament, 13 jan. 2025.

POSSON, V. **Dark Patterns: Why More Laws Won't Help**. European Tech Alliance, 16 abr. 2025.

POTEL-SAVILLE, M. AI-Driven Dark Patterns: How Artificial Intelligence Is Supercharging Digital Manipulation. Entrevistador Federico Guerrini. **Forbes**, 17 nov. 2024.

RODRIGUES, C.; MENDONÇA, E.; MENDONÇA, J.; ZANATTA, R. Contribuições da Data Privacy Brasil para a Tomada de Subsídios de Tratamento de Dados Pessoais de Crianças e Adolescentes da Autoridade Nacional de Proteção de Dados Pessoais. São Paulo: **Associação Data Privacy Brasil de Pesquisa**, 2024. Disponível em: https://bibliotecadigital.stf.jus.br/xmlui/bitstream/handle/123456789/8158/RODRIGUES_MENDONC%CC%A7A_ZANATTA_Contribuic%CC%A7o%CC%83esdodataprivacybrasil.pdf?sequence=1&isAllowed=y. Acesso em: 24 nov. 2025.

ROSALA, M. **Deceptive Patterns in UX: How to Recognize and Avoid Them**. Nielsen Norman Group, 01 dez. 2023.

SANTOS, I. **Violação de (con)sentimentos**: Uma análise psicojurídica da vulnerabilidade de titulares de dados diante de técnicas manipulatórias comportamentais. Trabalho de Conclusão de Curso (Bacharelado em Direito). Universidade de Brasília, Brasília. 2023.

SHARMA, D. The Guidelines on Dark Patterns – Effective or Incomplete? **Law School Policy Review**, 25 mar. 2024.

TIPP, A. **Dark patterns versus behavioural nudges in UX**. UX Planet, 08 mai. 2024.

UNIÃO EUROPEIA. **Digital Services Act [DSA]**. Bruxelas: Parlamento Europeu e Conselho, 2022.

UNIÃO EUROPEIA. **Regulamento Geral sobre a Proteção de Dados [GDPR]**. Bruxelas: Parlamento Europeu e Conselho, 2016.

UNITED STATES CONGRESS. S.3330 - **Deceptive Experiences To Online Users Reduction Act**. Washington, D.C., 2024.

ZAC, A.; HUANG, Y-C.; VON MOLTKE, A.; DECKER, C.; EZRACHI, A. Dark patterns and consumer vulnerability. **Behavioural Public Policy**, [*S. l.*], p. 1-50, 2025. DOI: 10.1017/bpp.2024.49. Disponível em: https://www.researchgate.net/publication/388650236_Dark_patterns_and_consumer_vulnerability Acesso em: 27 out. 2024.

2º lugar

Segredo de negócios e decisões automatizadas discriminatórias: impasses e soluções regulatórias

Isadora Valadares Assunção

Resumo

O artigo aborda os conflitos entre a proteção ao segredo de negócios e a tutela contra decisões automatizadas discriminatórias no contexto da proteção de dados pessoais. Inicialmente, explora-se o fundamento jurídico e axiológico do segredo de negócios, destacando sua importância econômica e suas implicações na inovação tecnológica. Em seguida, analisa-se como a proteção ao segredo contribui para a opacidade algorítmica em decisões automatizadas, identificando diversos tipos de vieses e fontes potenciais de discriminação. São avaliadas experiências regulatórias da União Europeia e dos Estados Unidos para enfrentar a tensão entre segredo comercial e transparência nas decisões automatizadas, além de se explorar a solução regulatória brasileira de auditoria de aspectos regulatórios pela Agência Nacional de Proteção de Dados. Conclui-se que o modelo de regulação responsiva, complementado por mecanismos adicionais de transparência e controle social difuso, pode compatibilizar esses interesses em potencial conflito, mas se demandam mecanismos complementares de controle social.

Palavras-chave: *segredo de negócios; decisões automatizadas; proteção de dados pessoais; LGPD.*

1. Introdução

Os ramos do Direito se intersectam, o que não é diferente com a proteção de dados pessoais. A transversalidade temática de diversos dos casos com os quais as autoridades nacionais de proteção de dados se deparam fica clara. Para exemplificar, cite-se a Medida de 27 de novembro, 2024, do Garante Privacy (2024). Na ocasião, tratava-se de um acordo entre o GEDI Gruppo Editoriale S.p.A e a OpenAI para licenciar o conteúdo cujo titular dos direitos autorais era o grupo editorial para fins de treinamento dos modelos de inteligência artificial (IA) dessa empresa. O licenciamento é, afinal, uma das soluções encontradas pela indústria de IA para continuar o desenvolvimento de seus modelos face aos questionamentos judiciais sobre a legalidade de se utilizar material protegido por direitos autorais sem a autorização dos titulares (Lemley, 2024).

O Garante Privacy iniciou um procedimento de fiscalização *ex officio* e emitiu um aviso em que se notava a possibilidade de violação de diversos artigos da General Data Privacy Regulation (GDPR) caso o compartilhamento do material se concretizasse. Dentre as violações, a ausência de base legal e o infringimento do princípio da finalidade, já que os indivíduos mencionados nos artigos não tinham expectativas razoáveis de que os artigos fossem compartilhados nem haviam consentido (Garante Privacy, 2024). Percebe-se, portanto, que uma solução conformada na seara da propriedade intelectual mostra-se frontalmente conflitante com o sistema de proteção de dados pessoais, ao menos em potencial.

Outra circunstância em que isso ocorre é na proteção ao segredo de negócios, que, apesar de representada como ideal para o fomento da competitividade e da inovação tecnológica, sendo tema central dos ramos do direito concorrencial e da propriedade intelectual, pode conflitar com os princípios e as normas específicas da proteção de dados pessoais no direito brasileiro, como a proteção contra decisões automatizadas discriminatórias (Frazão; Oliva; Tepedino, 2019; Fujimoto, 2023; Bueno, 2023). É o que se passa a analisar.

Para isso, adotar-se-á o método hipotético-dedutivo para responder às perguntas "A proteção do segredo de negócios conflita

com a proteção contra decisões automatizadas discriminatórias? Se sim, há formas de compatibilizá-las?", utilizando-se também do direito comparado para comparar impasses e soluções encontradas em outros ordenamentos, particularmente os Estados Unidos e a Europa. Para isso, apresentar-se-á, em primeiro lugar, os fundamentos da tutela jurídica ao segredo de negócios no plano internacional e nacional. Em seguida, tratar-se-á da relação do segredo de negócios com a opacidade algorítmica e com a proteção de dados pessoais. Especificamente, abordar-se-á dois impasses: como a opacidade algorítmica advinda do segredo de negócios dificulta ou impossibilita o escrutínio sobre decisões automatizadas discriminatórias e, face a esse contexto, qual o papel das autoridades regulatórias, se é que há algum. Finalmente, defende-se que o modelo de regulação responsiva, se complementado por instrumentos que viabilizem o escrutínio difuso, pode compatibilizar o segredo de negócios com a proteção contra discriminações algorítmicas.

2. A proteção jurídica ao segredo de negócios

Preliminarmente, pode-se estabelecer que o segredo de negócios protege informações comercialmente valiosas que não sejam conhecidas do público geral em um determinado setor, e que tenham sido objeto de medidas razoáveis do titular da informação para mantê-la em tal estado de desconhecimento (Organização Mundial da Propriedade Intelectual, 2024). Nesse sentido, é possível perceber um delineamento específico da proteção de tal direito. Apesar da adoção de práticas de ocultação de informações do público em geral, ou de competidores diretos, remeter à Antiguidade (Schiller, 1930), passando pelas reservadas corporações de ofício da Idade Média (Varela-Pezzano, 2010), foi modernamente que se uniu o requisito de valor comercial, confidencialidade e medidas razoáveis como fundamentos do segredo de negócios, subsequentemente, protegendo-o de apropriações indevidas (Organização Mundial da Propriedade Intelectual, 2024).

Antes de remeter aos diplomas internacionais e domésticos que sedimentam tal proteção, e explorar em quais termos o fazem, é necessário explorar qual o fundamento para a tutela jurídica

dos segredos de negócio que, à primeira vista, podem se contrapor ao interesse geral na ampla disseminação de conhecimento e informações.

O debate sobre por que se protegeriam segredos de negócio - e a qual título - ocupou a doutrina, em especial norte-americana, desde o século XIX, do qual datam os casos que primeiro reconheceram uma tutela a informações confidenciais de caráter comercial (Lemley, 2008). Em síntese, discutia-se se o segredo de negócios tinha caráter contratual, delitual (*tort*) ou real (Risch, 2007).

Para os defensores da primeira linha, tutelam-se os interesses das partes na condução contratual, com especial aplicação para os casos de transações de negócios envolvendo cláusulas de confidencialidade e de contratos de trabalho das quais estas também derivariam. Não se explica, entretanto, os casos em que não havia qualquer relação contratual entre a parte que se apropriou da informação e o titular, como nos casos de espionagem corporativa (Lemley, 2008).

Já para os que qualificavam o interesse de negócios como delitual, protegiam-se os interesses do titular contra lesões causadas direta ou indiretamente pela apropriação indevida, reconduzindo-se a *torts* específicos, já que falta ao direito norte-americano uma cláusula geral de responsabilidade civil como a brasileira. Ocorre que isso deixa indefinido o limite entre a apropriação indevida e o trato normal do comércio, que inclui tentativas de identificar a vantagem competitiva do competidor (Lemley, 2008).

Face à insuficiência de tais linhas argumentativas, chegou-se a defender o caráter de propriedade intelectual do segredo de negócios. Mais do que as especificidades do contexto norte-americano, destaca-se que parte de tal defesa foi embasada por questões axiológicas, por se entender que a tutela jurídica ao segredo de negócios se remeteria ao mesmo fundamento dos demais ramos da propriedade intelectual, isto é, fomentar investimentos no desenvolvimento científico-tecnológico e na difusão das informações obtidas por meio desse (Lemley, 2008).

Isto porque, na medida em que o titular das informações tem garantido um direito a controlar o acesso a determinadas informa-

ções de valor comercial, protegendo-as contra apropriações indevidas por terceiros, (I) não têm que investir o mesmo nível de recursos em métodos fáticos de manter o segredo, já que, caso ocorrer a apropriação, terá remédios jurídicos à sua disposição e (II) pode compartilhá-las com terceiros, estabelecendo limites de confidencialidade, pois sabe que estes serão eventualmente aplicados judicialmente. Assim, o segredo de negócios teria como fundamento, a *contrario sensu*, estabelecer um enquadramento no qual um nível menor de segredo fosse comercialmente viável. Em outras palavras, o segredo de negócios fomentaria a publicidade, tal como as patentes, mesmo que essa o fizesse diretamente, enquanto aquele a fizesse indiretamente (Lemley, 2008).

A teoria também foi objeto de críticas, inclusive as tradicionalmente associadas ao segredo de negócios como propriedade, já que não se confere um direito absoluto de uso exclusivo, até porque, caso houver uma descoberta independente da informação, e não uma apropriação indevida, não há remédios (Friedman; Landes;Posner, 1991). Sua importância resta, nesse sentido, na recondução aos fundamentos axiológicos do segredo de negócios e à sua aproximação com o sistema de patentes. Deve-se notar, entretanto, que essas não se confundem. Além da diferença quanto à mediação da publicidade, o escopo das patentes é limitado legislativamente, excluindo-se ideias, métodos abstratos, entre outros, enquanto o escopo do segredo de negócios é tendencialmente infinito, abrangendo tudo aquilo que tenha valor comercial. É por isso, inclusive, que assume renovada importância nos tempos atuais, em que muito do que tem valor econômico tem elevado caráter de abstração, nem sempre cumprindo com os requisitos de patenteabilidade ou proteção pelos demais ramos da propriedade intelectual. É, nesse sentido, "*perfectly suited to the evolutionary (progression of old ideas) and revolutionary (creation of new ideas) nature of innovation*" (Almeling, 2012).

Feito tal escorço pelas notas basilares do segredo de negócios, passa-se a sua proteção nos diplomas internacionais. Quanto a esses, destaca-se, pela aplicabilidade ao Brasil, o art. 39 do Acordo sobre Aspectos dos Direitos de Propriedade Intelectual relacionados ao Comércio, que estabelece que:

> 1. Ao assegurar proteção efetiva contra competição desleal, como disposto no ARTIGO 10bis da Convenção de Paris (1967), os *Membros protegerão informação confidencial de acordo com o parágrafo 2 abaixo*, e informação submetida a Governos ou a Agências Governamentais, de acordo com o parágrafo 3 abaixo.
> 2. Pessoas físicas e jurídicas terão a possibilidade de *evitar que informação legalmente sob seu controle seja divulgada, adquirida ou usada por terceiros, sem seu consentimento, de maneira contrária a práticas comerciais honestas, desde que tal informação*:
> a) seja *secreta*, no sentido de que não seja conhecida em geral nem facilmente acessível a pessoas de círculos que normalmente lidam com o tipo de informação em questão, seja como um todo, seja na configuração e montagem específicas de seus componentes;
> b) tenha *valor comercial* por ser secreta; e
> c) tenha sido objeto de *precauções razoáveis*, nas circunstâncias, pela pessoa legalmente em controle da informação, para mantê-la secreta (Brasil, 1994, grifo nosso).

Nota-se, portanto, a sedimentação legislativa dos requisitos de segredo, valor comercial e precauções razoáveis. Ainda, é de se destacar que a restrição à divulgação, aquisição ou uso das informações, sem o consentimento do titular, é restrita àquelas maneiras contrárias às práticas comerciais honestas, pelo que se visualiza a manifestação da restrição apenas à apropriação indevida, encorajando inovações incrementais e balanceando a proteção ao comerciante com o interesse geral na circulação das informações (Fekete, 2017). A remissão à honestidade das práticas comerciais também estabelece uma ligação direta com o direito concorrencial que, aliás, foi a porta de entrada do segredo de negócios no Brasil.

A Lei 9279/96 estabelece, em seu art. 195, que comete crime de concorrência desleal quem:

> XI - divulga, explora ou utiliza-se, sem autorização, de conhecimentos, informações ou dados confidenciais, utilizáveis na indústria, comércio ou prestação de serviços, excluídos aqueles que sejam de conhecimento público ou que sejam evidentes para um técnico no assunto, a

> que teve acesso mediante relação contratual ou empregatícia, mesmo após o término do contrato;
> XII - divulga, explora ou utiliza-se, sem autorização, de conhecimentos ou informações a que se refere o inciso anterior, obtidos por meios ilícitos ou a que teve acesso mediante fraude; [...] (BRASIL, 1996).

Percebe-se que é vedado, no sistema brasileiro, duas situações distintas, ambas reconduzíveis ao segredo de negócios (Fekete, 2017). No inc. XI, cobre-se o segredo clássico, sobre o qual incide uma obrigação de confidencialidade que, devendo ser mantida, não o é (como no caso das relações contratuais e empregatícias, sobre as quais também se prevê a divulgação de informações confidenciais como justa causa para rescisão do contrato de trabalho no art. 482, g, CLT). Protege-se, portanto, a expectativa criada na condução contratual da empresa, inclusive nas transferências voluntárias do seu conjunto de conhecimentos e experiências, que não esteja no domínio público, o que também é chamado de *know-how*, e cujos contratos devem ser registrados no Instituto Nacional da Propriedade Industrial (art. 12, Lei 9279/96). Já no inc. XII, protege-se o acesso ao segredo não por uma violação a regras de confidência, mas por meios ilícitos, como a fraude (Fekete, 2017).

Em síntese, pode-se diferenciar ambas as modalidades pela licitude do acesso, presente na primeira, mas não na segunda (Barbosa, 2003). Nota-se, ademais, que a divulgação, exploração ou utilização das informações, em quaisquer dos casos, é ilícita, trazendo consequências criminais e cíveis (estas por influxo da cláusula geral de responsabilidade civil do art. 186, CC).

Por fim, há de se notar que, na ocasião de um processo judicial seja para responsabilização cível ou criminal daquele que divulga ilicitamente um segredo de negócios, seja para obter tutela preventiva para impedir a divulgação, ou em qualquer outro processo em que divulguem as informações para defesa dos interesses das partes, "deverá o juiz determinar que o processo prossiga em segredo de justiça, vedado o uso de tais informações também à outra parte para outras finalidades" (art. 206, Lei 9279/96). Novamente, privilegia-se o âmbito de confidencialidade da informação secreta de indústria ou de comércio.

Analisado o segredo de negócios no plano internacional e nacional, resta entender suas intersecções gerais com a proteção de dados pessoais e, mais especificamente, no plano das decisões automatizadas, sobre o qual se sobressai o fenômeno da opacidade algorítmica.

3. Segredo de negócios, proteção de dados pessoais e opacidade algorítmica

Uma simples conferência do texto literal da Lei Geral de Proteção de Dados (LGPD) leva à conclusão de que se ressalva o segredo de negócios (ou, nos termos adotados pela lei, "segredos comercial e industrial") nove vezes (Brasil, 2018).

O segredo é ressalvado no princípio da transparência (art. 6º, VI); no direito de acesso a informações sobre o tratamento de dados pessoais (art. 9º, II); na solicitação ao controlador pela autoridade nacional de relatório de impacto à proteção de dados pessoais quando o tratamento se fundar no legítimo interesse (art. 10, § 3º); no direito de portabilidade dos dados a outros fornecedores (art. 18, V); no direito de confirmação da existência do tratamento e no acesso aos dados, por meio de declaração completa (art. 19, II); na obtenção de cópia integral dos dados quando o tratamento tiver origem no consentimento ou em contrato (art. 19, § 3º); no fornecimento de informações sobre os critérios e os procedimentos usados para decisões automatizadas (art. 20, § 1º); na elaboração de relatório de impacto à proteção de dados pessoais a ser determinada pela autoridade nacional (art. 38, caput); na indicação das medidas adotadas para a segurança dos dados (art. 48, § 1º, III).

Em todos esses casos, estabelece-se que devem ser "observados os segredos comercial e industrial" (Brasil, 2018). Ademais, compete diretamente à Agência Nacional de Proteção de Dados "zelar pela preservação do segredo empresarial e do sigilo das informações, nos termos da lei" (art. 55-J, § 5º). Nota-se, nesse sentido, que, em diversas ocasiões, o legislador estabelece que deverá ocorrer uma ponderação entre o direito do titular de ter a si garantidos sejam seus dados, informações sobre eles ou sobre o tratamento,

e a confidencialidade das informações de valor comercial que são protegidas sob o rótulo de segredo de negócios (Fujimoto, 2023).

Por um lado, tal ponderação é de extrema relevância: tanto a defesa da livre concorrência, ao qual remete o segredo de negócios, quanto a proteção de dados pessoais são bens jurídicos relevantes com proteção no plano constitucional, aquela no art. 170, IV, CF/88, esta no art. 5º, LXXIX. Ademais, retome-se os fundamentos de proteção do segredo de negócios, que reconduzem ao próprio fomento do desenvolvimento científico-tecnológico, tutelados no art. 218 e seguintes da CF/88 como parte integrante do mercado interno, e que devem ser viabilizados.

Por outro lado, isso cria um cenário em que o segredo de negócios pode ser invocado pelas empresas que realizam operações de tratamento de dados pessoais para limitar a publicidade das informações sobre tais tratamentos, mesmo que com a finalidade legítima de manter sua vantagem competitiva, contribuindo para a opacidade dessas (Fujimoto, 2023). A isso, remete-se uma das facetas do fenômeno de opacidade algorítmica.

As decisões automatizadas, entendidas, de maneira geral, como aquelas mediadas por sistemas algorítmicos de tomada de decisão, assumiram crescente importância no mundo hodierno: definem o que acessamos como informação, o que divulgamos de nós e quem a isso tem acesso, o que temos oportunidade de comprar por meio da obtenção de crédito, o que efetivamente compramos, além de outras áreas específicas, como se vamos ser contratados, como vai ser nosso tratamento de saúde ou até mesmo em qual medida teremos contato com forças policiais (O'Neil, 2021).

A amplitude de escopo é acompanhada de uma amplitude de poder. Na medida que os algoritmos influenciam virtualmente todas as facetas da vida contemporânea, quem os define, o que objetivam e como atingem seus objetivos, quando e por que refletem poder sobre, igualmente, todas os aspectos da modernidade (Pasquale, 2015). Nesse sentido, mecanismos que ocultam a resposta a tais perguntas e introduzem uma opacidade sobre o poder mediado pelos algoritmos geram problemas de legitimidade e prestação de contas ao público (Pasquale e Brevini, 2020).

Dentre tais mecanismos, destacam-se formas de opacidade técnicas e legais. Quanto à opacidade técnica, deriva do próprio funcionamento dos algoritmos. Na medida em que modelam a realidade a partir de mecanismos nem sempre explicáveis para os seres humanos, como quando consideram um número de dimensões e variáveis que nos é impossível acompanhar, podem se tornar caixas-pretas (Yang *et al.*, 2023). Em outras palavras, suas entradas e saídas (*inputs* e *outputs*) são conhecidas, mas não o modo como se relacionam, como chega-se a estas a partir daquelas.

À opacidade técnica, une-se a opacidade legal, advinda justamente da proteção aos segredos de negócio das empresas que desenvolvem ou aplicam os sistemas de tomada de decisão automatizadas (Pasquale, 2015). Quanto ao segredo, além do que já se notou, observa-se que, muitas vezes, os algoritmos não são patenteáveis, seja por limitações legais seja por serem considerados, por vezes, de abstração tamanha que se assemelham a métodos ou processos excluídos do âmbito de patenteabilidade. Nesses casos, assume relevo renovado o segredo de negócios como proteção de seu valor comercial (Almeling, 2012).

Por fim, a própria transparência sobre o sistema pode ser instrumentalizada em nome da opacidade, no que se chama de ofuscação (Pasquale, 2015). Boas explicações podem ser entendidas como aquelas que são contrafactuais, ou seja, elucidam porque um evento ocorreu ao invés de outros e são seletivas, focando em algumas causas do evento ao contrário de todas as possíveis (Miller, 2019). Nesse sentido, a transparência total não é uma boa estratégia para se explicar o funcionamento de um modelo, dificultando a compreensão pelos sujeitos, especialmente aqueles não-especialistas, e podendo se tornar uma estratégia de ofuscação (Mostowy, 2020). Isso porque os seres humanos não conseguem levar em consideração tantos fatores ao mesmo tempo para tomar uma decisão como a IA, devido à chamada "maldição da dimensionalidade" e, portanto, não veem sentido nessas decisões se tiverem disponíveis todos os fatores que foram considerados, mas não uma explicação (Mostowy, 2020).

Delineada a opacidade algorítmica, em suas diversas modalidades, há de se antecipar um possível argumento: de que a opaci-

dade não conflitaria com a proteção de dados pessoais que, afinal, também subtrai do público em geral dados. Nesse sentido, a única diferença entre a opacidade legal e a proteção de dados pessoais seria que aquela subtrai do domínio geral dados de valor comercial e esta dados que fazem referência a pessoas identificadas ou identificáveis.

O argumento não deve subsistir na medida em que se reconheça que, no que tange à opacidade algorítmica, o que se cria não é uma restrição fundamentada à publicidade, como na tutela da proteção de dados pessoais, mas uma transparência dos comportamentos oposta a uma opacidade sobre o funcionamento dos mecanismos que analisam e decidem sobre os titulares de tais comportamentos (Pasquale, 2015).

Nesse sentido, chegou-se a propor o conceito de um capitalismo de vigilância, que, a partir da ubiquidade da tecnologia digital e da datificação da experiência humana, utilizaria tais dados "como dados comportamentais para o aperfeiçoamento do controle de outros sobre nós" (Zuboff, 2020, p. 115), por meio da predição automatizada de comportamentos futuros. Também por isso sintetizou-se que "*we do not live in a peaceable kingdom of walled gardens; the contemporary world more closely resembles a one-way mirror*" (Pasquale, 2015, p. 9).

Por isso, justifica-se uma análise mais particularizada da intersecção entre as decisões automatizadas e a opacidade algorítmica, em especial aquela gerada pelo uso estratégico do segredo de negócios, principalmente quando as decisões se mostrem danosas aos sujeitos das quais tratem, como quando os discriminam.

4. Decisões automatizadas discriminatórias e opacidade algorítmica

a.Tratamento automatizado e envolvimento humano

O tratamento automatizado de dados pessoais, pode envolver diferentes graus de envolvimento humano e, portanto, diferentes

esferas de autonomia conferidas aos sistemas de processamento de dados pessoais (Waldman, 2019). A conjugação do binômio envolvimento humano e automação é bem explicitada no Regulamento Geral de Proteção de Dados (GDPR), já que este traz obrigações regulatórias e direitos ao titular que diferem conforme a tomada de decisão seja automatizada *lato sensu*, seja exclusivamente baseada no tratamento automatizado ou, ainda, seja uma decisão de definição de perfis (*profiling*) (Article 29 Data Protection Working Party, 2018).

A definição de perfis é definida legalmente no GDPR como "qualquer forma de tratamento automatizado de dados pessoais que consista em utilizar esses dados pessoais para avaliar certos aspectos pessoais de uma pessoa singular [...]" (União Europeia, 2016). O artigo procede com um rol de aspectos pessoais, como os relacionados ao desempenho profissional, situação econômica, saúde, preferências pessoais, comportamento, entre outros. Independente dos aspectos específicos envolvidos no perfilamento, percebe-se que a definição se centra em (I) um tratamento automatizado; (II) que envolva dados pessoais; (III) que analise aspectos subjetivos de uma pessoa natural. A partir desses elementos, infere-se que o tratamento automatizado é um gênero no qual se inclui a definição de perfis, logo, há tratamentos automatizados de dados pessoais que não a envolvem (Article 29 Data Protection Working Party, 2018).

Uma segunda categoria adotada no GDPR são as decisões tomadas exclusivamente com base no tratamento automatizado, referidas no art. 22 (União Europeia, 2016). Também nesse artigo se qualifica um subconjunto de tratamento automatizado em que se exclui o envolvimento humano, podendo-se defini-lo como "a habilidade de tomar decisões por meio tecnológico sem envolvimento humano" (Article 29 Data Protection Working Party, 2018)..

Extrai-se, portanto, que no gênero tratamento automatizado, há como espécies decisões automatizadas nele exclusivamente baseadas e decisões de definição de perfis, que podem ou não se sobrepor. A distinção é relevante porque há previsões que se aplicam de maneira geral a todas as categorias (como hipóteses legais

que justificam as operações de tratamentos gerais, os princípios gerais, entre outras), mas outras dependem das categorias em que se inclui o tratamento (Article 29 Data Protection Working Party, 2018).

Sobre as decisões exclusivamente baseadas no tratamento automatizado de dados pessoais, incluindo ou não a definição de perfis, e que produzam efeitos na esfera jurídica do titular ou o afetem significativamente, incide o art. 22, GDPR, que prevê uma proibição geral a tais decisões sem envolvimento humano (União Europeia, 2016). Estabelece-se, assim, o direito do titular a não estar sujeito a decisões automatizadas sem envolvimento humano.

A proibição, como regra, admite exceções, previstas no art. 22.2, que autoriza tais decisões se forem necessárias para a celebração ou execução de contrato, forem autorizadas pelo direito da União ou do Estado-Membro ao qual o responsável pelo tratamento estiver sujeito, ou forem baseadas no consentimento explícito do titular.

No caso das decisões que se englobem em uma das exceções, o controlador deve ainda "aplica[r] medidas adequadas para salvaguardar os direitos e liberdades e legítimos interesses do titular dos dados [...]" (União Europeia, 2016), incluindo o direito de obter intervenção humana, manifestar seu ponto de vista e contestar a decisão.

A integração das previsões dos arts. 13, 14 e 15, que se referem à definição de perfis, e do art. 22, que se refere às decisões automatizadas sem envolvimento humano, foi objeto de intenso debate doutrinário, especialmente quanto à existência de um direito à explicação de decisões automatizadas sem envolvimento humano, mas que não envolvessem definição de perfis, o que será retomado (Wachter; Mittelstadt; Floridi, 2017; Selbst; Powles, 2018).

Mais que detalhar os contornos da interpretação sistemática do GDPR, até mesmo porque a LGPD não reproduz a estrutura normativa do GDPR nesse ponto, cabe notar que o envolvimento humano foi um fator ponderado pelo legislador comunitário europeu como assumindo tamanha relevância que foram estabelecidos direitos diferentes ao titular submetido a decisões exclusivamente baseadas no tratamento automatizado.

Nesse sentido, humano envolvido na tomada de decisões deve ter a autoridade e a competência para alterar a decisão tomada a partir do processamento automatizado de dados pessoais, devendo ainda considerar todos os dados relevantes naquele contexto, possivelmente extrapolando aqueles considerados pelo sistema automatizado (Article 29 Data Protection Working Party, 2018). O nível concreto de envolvimento deve ser documentado no relatório de impacto à proteção de dados pessoais, para se evitar que o controlador forje um envolvimento humano quando na verdade o que há é mera aprovação das decisões, sem qualquer influência no resultado, pelo humano, que agiria como mero "carimbador" (*rubber stamping*) (Wagner, 2019).

Passando à LGPD, o enquadramento regulatório erigido pelo art. 20 parte de uma premissa diferente daquela adotada na União Europeia. Enquanto lá se parte da proibição das decisões baseadas exclusivamente no tratamento automatizado, o art. 20 implicitamente as autoriza na medida em que atribui ao titular dos dados pessoais o direito de solicitar sua revisão (Almada, 2019). O direito à explicação é mais diretamente previsto no § 1º, que impõe ao controlador o dever de fornecer informações sobre os critérios e procedimentos utilizados para a decisão automatizada. Faz-se uma "*avaliação funcionalizada da opacidade*", na medida em que se ressalvam os segredos comerciais e industriais (Marrafon; Medon, 2019). Mesmo no âmbito destes, estabelece-se a possibilidade de uma auditoria pela ANPD para que se verifiquem aspectos discriminatórios (art. 20, § 2º, LGPD).

Observa-se, portanto, que a lógica decisional e os fatores ponderados no âmbito de uma decisão automatizada, particularmente naquelas que não recebem envolvimento humano, assumem destacada importância no contexto da discriminação algorítmica (Sá e Lima, 2020) , o que remete às etapas das operações de tratamento às quais a discriminação pode remeter e o que, inclusive, foi reconhecido legalmente na medida em que se estabelece competência regulatória específica para aferir a existência de fatores discriminatórios (Frazão, 2021).

b. Decisões automatizadas e discriminação algorítmica

Em paper seminal, Selbst e Barocas (2016) apresentaram uma classificação dos momentos nos quais a discriminação algorítmica pode surgir. Segundo os autores, esta pode se originar (I) da escolha da variável de destino e dos rótulos de classe; (II) do enviesamento advindo do banco de dados de treinamento; (III) da definição dos atributos do modelo; (IV) do uso de proxies. Adiciona-se, ainda, o enviesamento intencional dos algoritmos (Croll, 2012).

I. Enviesamento advindo da escolha da variável de destino e dos rótulos de classe

Quanto ao primeiro momento de origem da discriminação algorítmica, a variável de destino (target variable) é o resultado de interesse de um modelo, como a classificação de e-mails em spam ou não-spam (Selbst e Barocas, 2016). Já os rótulos de classe dividem os possíveis valores da variável de destino em categorias excludentes (Selbst; Barocas, 2016).

Entretanto, nem sempre o processo de formalização matemática da variável de destino e dos rótulos de classe é tão objetivo quanto a separação dos e-mails em spam ou não, categorias binárias, relativamente evidentes e estanques. Como os cientistas da computação devem traduzir problemas da realidade em termos mensuráveis, a criação de classes pode envolver maior arbitrariedade dos desenvolvedores (Selbst; Barocas, 2016).

Exemplo disso é um algoritmo de contratação cuja variável de destino fosse "ser um bom trabalhador". Tal variável demandaria a tradução de uma questão abstrata em algo quantitativamente dimensionável, como chegar cedo ao trabalho, ficar mais tempo no mesmo emprego ou realizar um maior número de vendas a cada mês. A escolha não é neutra. Por exemplo, caso se escolha relacionar "ser um bom trabalhador" a chegar mais cedo ao trabalho, é possível que pessoas de baixa renda que, por isso, demorem mais tempo para se deslocar ao trabalho, sejam desproporcionalmente prejudicadas.

II. Enviesamento advindo do banco de dados de treinamento

Quanto ao segundo momento, do enviesamento advindo do banco de dados de treinamento, os modelos algorítmicos dependem de duas suposições, que formam o paradigma das variáveis aleatórias independentes e identicamente distribuídas: (I) que as características da população dos dados de treinamento continuem as mesmas quando o modelo for aplicado, no futuro, a outros dados e (II) que os dados de treinamento sejam representativos da população tomada em sua totalidade (Calders; Žliobaitė, 2013).

Caso alguma das suposições não seja atendida durante as etapas de treinamento de um modelo de IA, esse poderá ter desempenho discriminatório. Respectivamente aos pressupostos acima arrolados, pode haver um banco de dados cujos rótulos de classe estejam permeados por algum viés, ou um conjunto de dados de treinamento que sub- ou super-representa certos grupos sociais (Calders; Žliobaitė, 2013).

Exemplo do primeiro seria o modelo mencionado anteriormente, que associava perfis de candidatos a "ser um bom trabalhador". Com base no mesmo perfil, dois avaliadores humanos diferentes podem divergir quanto ao rótulo que deve ser atribuído. Suponha-se que a mesma empresa quer medir "ser um bom trabalhador", usando para treinamento dados de avaliações semestrais de desempenho, que serão associadas a rótulos como "bom trabalhador" ou "trabalhador ruim". Se, historicamente, trabalhadoras receberem avaliações piores que seus pares do sexo masculino, o algoritmo que for treinado com tais dados apresentará comportamento discriminatório em relação às mulheres que posteriormente ingressarem na empresa.

Um caso que ilustra bem a reprodução de vieses presentes nos dados de treinamento é, historicamente, o do Hospital St. George, no Reino Unido. Na década de 70, no Hospital St. George, foi desenvolvido um modelo de IA para orientar a contratação futura de residentes, com o treinamento do modelo sendo feito com dados históricos sobre os candidatos aceitos no processo seletivo. No entanto, os dados de treinamento mostravam decisões que desfa-

voreciam sistematicamente aplicações de estudantes negros e de mulheres com as mesmas credenciais dos colegas. Nesse contexto, o algoritmo inferiu que tais grupos populacionais seriam menos apropriados para as vagas, reproduzindo um viés histórico (Lowru; Macpherson, 1988).

Já se os dados de treinamento não representarem adequadamente a população em questão, a segunda suposição do paradigma das variáveis aleatórias independentes e identicamente distribuídas será desrespeitada (Calders; Žliobaitė, 2013).

Inicialmente, é de se notar que certos grupos podem estar sub-representados por estarem localizados nas margens do sistema de Big Data. Nesta situação, enquadram-se pessoas que, por seus hábitos de vida, sua condição econômica ou localização geográfica, geram menos pontos de dados diariamente e que, por isso, estão menos representadas nos bancos de treinamento (Lerman, 2013). Isso pode limitar o acesso de certos grupos a determinadas situações, gerando um ciclo vicioso no qual, por serem excluídas digitalmente, as pessoas são sub-representadas nos bancos de treinamento e, por serem sub-representadas, o acesso delas a direitos e oportunidades, inclusive de acesso digital, é progressivamente reduzido (Lerman, 2013).

Por outro lado, a super-representação de grupos nos dados de treinamento também pode gerar vieses. Exemplo disso são algoritmos de policiamento preditivo, como o Predpol, em que se observa que áreas de baixa renda apresentam maior número de ocorrências policiais para crimes como posse de pequenas quantidades de drogas. (Ferguson, 2017) Ao serem alimentados os registros policiais no modelo, há a super-representação de pessoas de baixa renda, o que, por sua vez, correlaciona-se com raça. Assim, o algoritmo direciona ainda maior atenção policial a tais locais, o que gera mais registros policiais, em um ciclo vicioso que configura uma profecia autorrealizável (Ferguson, 2017).

III. Enviesamento advindo da definição dos atributos do modelo

Outro momento do qual pode derivar a discriminação algorítmica é a seleção dos atributos de um modelo, já que é impossí-

vel levar em conta todos os fatores envolvidos em um problema, de forma granular, ao desenvolver um modelo (Calders; Žliobaitė, 2013). Logo, a formalização matemática implica em uma redução da complexidade dos fenômenos como se apresentam na realidade (Carusi, 2008). Outra razão prática para que nem todos os aspectos sejam considerados é que, apesar de potencialmente levar a uma maior acurácia, tal consideração demandaria custos muitas vezes tidos como excessivos e impeditivos à consecução da atividade (Selbst; Barocas, 2016).

Por isso, selecionam-se os atributos que a análise algorítmica levará em conta ao relacionar um *input* a um *output*. Um exemplo disso seria um algoritmo de contratação que atribuísse peso significativo à universidade de origem do candidato. Uma vez que minorias têm acesso dificultado às universidades de renome, tais grupos seriam discriminados, mesmo que tenham as mesmas qualidades de outros candidatos, já que o modelo falhará em considerar adequadamente outros aspectos granulares além das credenciais acadêmicas.

IV. Enviesamento advindo do uso de proxies

Quanto ao uso de *proxies*, nota-se que, na estatística, "*uma variável proxy é aquela que se apresenta no lugar da real variável de interesse, a qual pode não estar disponível, ser muito cara ou muito demorada de medir*" (Bruce, 2019). Em outras palavras, uma variável *proxy* é aquela que está correlacionada à variável desejada. Por exemplo, para medir renda, pode-se utilizar como variável *proxy* o CEP, já que, em bairros onde o metro quadrado é mais caro, em geral moram pessoas de maior renda. Para definir se algo é uma variável independente ou correlacionada a outra, deve-se observar se cada atributo contribui separadamente para a tomada de decisão do modelo (Calders; Žliobaitė, 2013). Caso isso não ocorra, pode ser difícil identificar qual variável contribui mais para o modelo, por exemplo, o CEP ou a renda.

Mesmo que os desenvolvedores de *software* não tenham nenhuma intenção prejudicial ao utilizar uma variável, e sim desejem maior acurácia, pode haver casos nos quais estar em uma classe protegida esteja encoberto por outros dados. Logo, a análise des-

ses dados pode ser discriminatória, como seria o caso de precificações diferenciadas por CEP (Selbst; Barocas, 2016).

V. Enviesamento intencional

O enviesamento intencional pode ocorrer de várias formas. Entre elas, destaca-se a seleção consciente de um banco de dados de treinamento que esteja enviesado, por exemplo, pela sub-representação de uma população, como nos algoritmos de contratação supracitados (Selbst; Barocas, 2016). Pode-se alegar que tais decisões de contratação que formaram o banco de dados foram imparciais e que permitiriam inferir regras válidas e universais, o que mascararia a discriminação. Ainda, pode-se usar *proxies* complexos para contornar as leis de proteção aos grupos minoritários, em um tipo de discriminação difícil de identificar, já que, à primeira vista, não envolve características protegidas (Selbst; Barocas, 2016). Por fim, visando fins discriminatórios, pode-se inferir de dados não-sensíveis informações como orientação política ou sexual, que são considerados sensíveis para grande parte das legislações de proteção de dados pessoais (Fico; Assunção, 2022).

c. "Desconhecidos desconhecidos", o segredo de negócios e a discriminação algorítmica

Epistemologicamente, a ideia de que há lacunas no conhecimento sobre o que é ou não conhecido para um determinado agente, remete à ignorância socrática (Proctor; Schiebinger, 2008). Expandindo-se a noção de conhecimento ou ignorância para uma matriz de dois sujeitos, é possível extrair quatro quadrantes, (I) conhecido para si, conhecido para terceiros; (II) conhecido para si, desconhecido para terceiros; (III) desconhecido para si, conhecido para terceiros; (IV) desconhecido para si, desconhecido para terceiros (Shenton, 2007). Respectivamente, seriam áreas de transparência epistemológica; de obscurecimento, em que se esconde uma informação conhecida; pontos cegos no conhecimento de um sujeito; de verdadeira ignorância e desconhecimento. Esquematicamente, poder-se-ia representá-las da seguinte forma:

	Conhecido para si	Desconhecido para si
Conhecido para terceiros	(I) Transparência	(III) Pontos cegos
Desconhecido para terceiros	(II) Obscurecimento	(IV) Verdadeira ignorância

Fonte: Do autor.

Transpondo tal enquadramento epistemológico para a discriminação algorítmica, e estabelecendo-se uma relação com a opacidade estratégica que pode ser gerada pela proteção ao segredo de negócios dos agentes econômicos, é possível perceber que tal proteção pode tanto dificultar a percepção de um sistema de tomada de decisões automatizadas é discriminatório quando o agente que o desenvolve ou aplica conhece tal situação quanto quando tal informação está na zona de verdadeira ignorância. No primeiro caso, o que se tem é o risco de que o agente ofusque a informação de que o sistema é discriminatório do conhecimento geral ou de terceiros específicos (como reguladores) (Mostowy, 2020). Já no segundo, há uma verdadeira incerteza, um desconhecido desconhecido (*unknown unknown*) quanto à discriminação algorítmica (Lakkaraju *et al*., 2017).

Em ambos os casos, percebe-se que funcionalmente prejudica-se a obtenção de reparações de danos eventualmente causados. Mais do que isso, dificulta-se o cumprimento da função preventiva, para que o sistema seja alterado ou retirado de circulação para não causar danos a indivíduos ou parcelas da população que possam ser por ele discriminados (Pasquale, 2015). Resta inquirir, portanto, quais mecanismos jurídicos podem responder ao obscurecimento ou à ignorância para aumentar a transparência do fenômeno discriminatório para o agente e para terceiros, como reguladores, julgadores ou até mesmo vítimas que sejam por ele lesadas, "observado o segredo comercial e industrial".

d. Soluções no Direito Comparado

I. União Europeia

Na União Europeia, as soluções giram em torno do direito à explicação. Apesar de parte da doutrina reconhecer que não há direito à explicação na General Data Protection Regulation (GDPR), legislação de proteção de dados europeia, já que este não está previsto expressamente como ocorria na Diretiva de Proteção de Dados (Wachter; Mittelstadt; Floridi, 2017), a doutrina majoritária reconhece que tal direito deriva do direito de acesso a informações úteis (Selbst; Powles, 2018) e é pressuposto para exercer o direito de revisão das decisões automatizadas e foi este também o entendimento do Tribunal de Justiça da União Europeia (2025) no caso C-203/22.Em decisão de fevereiro de 2025, em caso que tratava da recusa do controlador de fornecer informações sobre decisão automatizada no âmbito da concessão de crédito, o Tribunal considerou que o art. 15(1)(h) do GDPR era instrumental para o exercício do direito de revisão e que não seria possível contestar uma decisão sem explicações inteligíveis e acessíveis sobre o procedimento e os princípios utilizados para tomá-la. Em síntese:

> Com efeito, se as pessoas afetadas por uma decisão automatizada, incluindo a definição de perfis, não estivessem em condições de compreender as razões que conduziram a essa decisão antes de exprimirem o seu ponto de vista ou de a contestarem, esses direitos não poderiam, por esse facto, cumprir plenamente a sua finalidade de proteger essas pessoas contra os riscos específicos para os seus direitos e liberdades decorrentes do tratamento automatizado dos seus dados pessoais. (Tribunal de Justiça da União Europeia, 2025.

No referido caso, também se discutiu em que medida o segredo comercial seria suscetível de restringir o direito de acesso que fora reconhecido. Notou-se, preliminarmente, que o direito à proteção de dados pessoais não assume caráter absoluto, devendo ser equilibrado com outros direitos fundamentais conforme a proporcionalidade (Tribunal de Justiça da União Europeia, 2025). Dentre tais direitos de terceiros, destaca-se o segredo comercial e a propriedade intelectual.

Conclui-se, a esse respeito, que, sempre que possível, deve-se optar por formas de comunicações e prestação de informações ao titular de dados pessoais que não viole o segredo comercial. Por outro lado, "*essas considerações não deverão resultar na recusa de prestação de todas as informações ao titular dos dados*" (Tribunal de Justiça da União Europeia, 2025). A análise será casuística e caberá à autoridade de controle de proteção de dados pessoais ou ao tribunal competente, que pode solicitar que as informações lhe sejam apresentadas para ponderar devidamente os interesses no caso concreto.

II. Estados Unidos da América

Também nos Estados Unidos, país no qual não há legislação federal transversal de proteção de dados pessoais, os tribunais já foram chamados a responder sobre a intersecção entre o segredo de negócios e a proteção de dados pessoais e, especificamente, sobre a discriminação algorítmica.

Um dos casos concretos em que isso fica claro centrava-se no algoritmo Compas. Tal ferramenta de IA balanceia diversos fatores criminógenos e o histórico criminal para promover duas escalas de risco, a General Recidivism Risk Scale (GRRS) e a Violent Recidivism Risk Scale (VRRS), que visam auxiliar na tomada de decisão de agências judiciais e correcionais (Brennan, Dieterich; Ehret, 2009).

Entretanto, como revelou investigação da ProPublica, com base no exame das pontuações de risco de reincidência de 10.000 réus criminais na Flórida e de suas verdadeiras taxas de reincidência, o algoritmo Compas é enviesado. Réus negros apresentaram predições de reincidência mais altas que as reais, sendo classificados como de alto risco mesmo sem reincidir (falsos positivos) em taxas quase duas vezes superiores às de réus brancos (45% x 23%) (Larson *et al*., 2016). Enquanto isso, réus brancos tinham quase o dobro de chance de serem considerados de baixo risco, mesmo reincidindo (falsos negativos), do que réus negros (48% x 28%) (Larson *et al*, 2016).

Após tal investigação, um réu negro chamado Eric L. Loomis, que tinha sido sujeito à análise algorítmica pelo Compas durante a persecução penal, recorreu à Suprema Corte do Wisconsin, com a

alegação de que a utilização de um algoritmo proprietário para auxiliar o juiz a sentenciar violava o devido processo legal e a garantia do contraditório, já que não era possível contestá-lo sem saber exatamente como o modelo de IA funcionava. A Suprema Corte de Wisconsin (2016) negou o recurso uma vez que considerou que ambas as partes tinham acesso ao relatório final do Compas e tinham acesso aos pontos de dados iniciais (*inputs*) que foram usados pela IA, sendo que o réu poderia contestar qualquer dado incorreto. Ademais, o uso dos escores de risco era acompanhado de outros fatores independentes a serem ponderados pelo juiz, não sendo determinativo na decisão de liberdade comercial. Em outro caso no qual o algoritmo Compas foi utilizado, discutiu-se particularmente o segredo de negócios como óbice à prestação de informações sobre a lógica decisional. Em Rayner v. New York State Department of Corrections, o requerente questionava a negativa de um pedido de informações baseado na Freedom of Information Law, sendo que a negativa havia se baseado nos interesses de negócio da empresa proprietária do Compas. A Suprema Corte de Albany (2023) considerou que o conteúdo do algoritmo e informações sobre seu funcionamento eram segredos de negócio bona fide, e que divulgá-los afetaria a competitividade da empresa, já que: "*if petitioner's FOIL application were granted and equivant's technology opened to the public, it would be a simple matter for one of the State's information technology vendors to supply DOCCS with the functionality of COMPAS-NY at a much lower cost*". Assim, considerou-se justificada a negativa.

Com base nestes casos, a doutrina norte-americana dedica-se a estabelecer salvaguardas procedimentais, com o chamado devido processo tecnológico (*technological due process*), para que haja a notificação dos indivíduos que serão sujeitos de processos de decisão automatizados (Citron, 2007). Esses sujeitos devem ter acesso ao processo, ter suas posições levadas em consideração e os registros de tais processos devem ser armazenados para possibilitar auditorias posteriores (Citron; Pasquale, 2014). Entretanto, verifica-se que a mediação doutrinária nos casos concretos resta prejudicada.

Se, por um lado, a solução conformada na União Europeia remete a ponderações nos casos concretos, a experiência norte-ameri-

cana, país no qual casos que lidavam com a discriminação algorítmica já se colocaram, leva-nos a um segundo impasse: qual o papel efetivo das autoridades regulatórias frente a decisões automatizadas discriminatórias na qual releva-se a opacidade algorítmica?

Isto se torna ainda mais relevante frente à opção pelo legislador brasileiro de conferir à Agência Nacional de Proteção de Dados competência para realizar auditoria para verificação de aspectos discriminatórios em tratamento automatizado de dados pessoais nos casos em que as informações sobre o tratamento não sejam prestadas sob alegação de segredo de negócios.

5. Uma proposta de solução: regulação responsiva e o segredo de negócios na proteção de dados pessoais

O art. 20, LGPD, estabelece, além do direito de explicação, um mecanismo original de prestação de contas, a partir de uma avaliação funcionalizada da opacidade em que se pondera o segredo de negócios e a necessidade de se verificar aspectos discriminatórios de decisões automatizadas (Marrafon; Medon, 2019; Frazão, 2020).

Para melhor entender suas conformações, com a determinação de auditorias pela Agência Nacional de Proteção de Dados, faz-se necessário ir além de uma concepção da prestação de contas como uma virtude que traz legitimidade ao processo, para concebê-la como um arranjo relacional, um mecanismo em que alguém presta contas sobre algo a outrem, de alguma maneira específica, orientado por um por quê como diretriz (Bovens, 2010). Algumas perguntas centrais são, nesse sentido: Quem presta contas? A quem? Sobre o que? Como? Por quê? Resta analisar as respostas a tais perguntas face ao contexto específico do art. 20, § 2º, LGPD.

a. Quem presta contas?

De uma interpretação sistemática da LGPD, estabelece-se que as decisões tomadas unicamente com base em tratamento automatizado de dados pessoais são uma dentre várias formas de tra-

tamento de dados pessoais, logo, realizadas por agentes de tratamento. Destes, destacam-se os controladores, que determinam os meios e os fins a que se dirigem o tratamento, e os operadores, que processam a operação face às determinações centrais erigidas por aqueles.

Assim, pela literalidade do art. 20, § 1º, é o controlador que deverá fornecer, sempre que solicitadas, informações a respeito dos critérios e dos procedimentos utilizados para a decisão automatizada. Se não as prestar, alegando segredo comercial ou industrial, incide o art. 20, § 2º, que trata da auditoria para verificação de aspectos discriminatórios. Logo, é o controlador que se alberga sob a alegação de segredo de negócios que presta contas (Frazão, 2020; Bueno, 2023).

A essa conclusão, também se chega axiologicamente. A simples previsão de um direito à explicação poderia incorrer em uma falácia da transparência, culminando em inefetividade, seja porque os titulares de dados pessoais, a quem caberia solicitar as informações, têm pouco tempo, recursos e expertise para das explicações fazer sentido quanto pelo fato de que as explicações em si podem ser pouco claras ou significativas, na prática, em especial quando presente o segredo de negócios (Edwards; Veale, 2017). Por isso, ao lado de tal direito, complementando-o quando for frustrado pelo segredo comercial ou industrial, prevê-se uma competência de auditoria de aspectos discriminatórios.

Reconhecer tal complementariedade leva à constatação de que não basta que aquele que preste contas responda às indagações do titular, já que isto sequer é absoluto, podendo ser excepcionado pelo segredo de negócios, mas é necessário que, em todo caso, saiba responder sobre os critérios e os procedimentos utilizados para a decisão automatizada (Bueno, 2023), ou seja, é preciso que seja capaz de prestar contas sobre eles, seja ao titular seja à autoridade regulatória.

Esta prestação de contas sob uma concepção epistêmica justifica-se em um contexto em que há vários possíveis arranjos de tratamentos automatizados que podem garantir a mesma acurácia, mas nem sempre com as mesmas garantias individuais ou propriedades coletivas (como vieses) (Black; Raghavan; Barocas, 2022).

Nesse cenário, não basta ao controlador alegar boa-fé ou ignorância técnica: exige-se que tenha estruturado internamente as condições para responder a contento, seja ao titular, seja à autoridade. Trata-se de uma responsabilidade epistêmica que vincula o dever de saber ao dever de justificar sobre o arranjo que escolheu.

b. A quem presta contas?

O controlador, quanto à prestação de contas passiva, primeiro será instado a fornecer, a partir da solicitação do titular, informações a respeito dos critérios e dos procedimentos utilizados para a decisão automatizada. Não as fornecendo sob ressalva de segredo comercial ou industrial, incide a auditoria para verificação de aspectos discriminatórios (Bueno, 2023).

Assim, em um primeiro momento, o controlador presta contas ao titular e, subsidiariamente, à Agência Nacional de Proteção de Dados, como interlocutor estratégico da estrutura prestação de contas erigida pelo legislador ordinário. Essa deve detectar comportamentos indesejáveis, desenvolver ferramentas para responder a tais comportamentos, aplicar tais ferramentas (enforcement), avaliar os resultados e modificar sua estratégia, caso necessário (Baldwin; Black, 2008).

Para isso, revela-se útil o paradigma da regulação responsiva, em que se identifica, a partir de interações sucessivas, o paradigma atitudinal do agente regulado: se é um líder corporativo em compliance, com colaboração aumentada com o regulador, se é um agente com intenção de cooperar, mas baixa competência no domínio, se é um agente que coopera relutantemente, ou se é recalcitrante (Ayres; Braithwaite, 1992). Um comportamento responsivo é de máxima importância já que, mesmo diante do segredo, a LGPD delega à ANPD a função de auditar os sistemas de tratamento automatizado, o que exige um modelo regulatório não puramente punitivo (Fujimoto, 2023).

No paradigma responsivo, presume-se inicialmente a virtude do agente, ou seja, a máxima cooperação, passa-se às interações, escalando para ferramentas mais coercitivas caso se verifique, em uma estratégia de reciprocidade, que o agente tem relutância ou

não intenciona colaborar com o enforcement (Ayres; Braithwaite, 1992). Para que tal modelo funcione, é preciso que se esteja diante de um setor em que seja possível uma interação repetida, o que é o caso de uma atividade de fiscalização de aspectos discriminatórios legalmente autorizada (Gunningham, 2010).

É também necessário que o regulador capte sinais informacionais de com qual grupo efetivamente está lidando a partir das interações, o que é prejudicado por opacidades estruturais por parte dos regulados e que pode ocorrer nas referidas auditorias, já que se está diante de um cenário em que já se alegou o segredo de negócios. Nesse contexto, o raciocínio do regulador será o de "*dadas as circunstâncias que tenho diante de mim, qual técnica de enforcement é mais provável de ter um efeito adequado, enquanto se mantém a confiança dos regulados?*" (Gunningham, 2010).

c. Sobre o que e como se presta contas?

A prestação de contas no contexto do art. 20 da LGPD não se limita ao resultado da decisão automatizada, mas incide, destacadamente, sobre os próprios critérios e procedimentos utilizados para sua formação. Em outras palavras, presta-se contas não apenas sobre o que se decidiu, mas sobre como se decidiu. Essa dimensão processual do dever de prestação de contas assume centralidade especial diante da crescente sofisticação técnica dos sistemas algorítmicos, nos quais os vieses e discriminações frequentemente decorrem do próprio design, como já explicitado.

Dessa constatação emerge a necessidade de conceber sistemas de decisão automatizada a partir de uma lógica de contestabilidade desde o design (Alfrink *et al.*, 2023). Isso significa estruturar os sistemas de forma que terceiros, sejam titulares, autoridades ou entidades independentes, possam escrutiná-los de maneira significativa, viabilizando a audibilidade dos sistemas (Almada, 2019).

Materialmente, isto deve-se estruturar por ferramentas que promovam o escrutínio do sistema, seja por diversos métodos de explicabilidade, como os métodos globais, que explicam a lógica de funcionamento do modelo como um todo, seja intrinsecamente,

no qual o modelo de IA é desenvolvido para ser explicável por si só, ou extrinsecamente, em que se desenvolve um algoritmo separado do modelo de IA auxilia na explicação da decisão após ela ser feita (Alfrink *et al.*, 2023).

Outras ferramentas dignas de notas dizem respeito à documentação do processo de desenvolvimento do sistema de tratamento automatizado. Nesses casos, o desenvolvedor deve estabelecer um sistema de governança dos dados que forem usados para treinar e validar o modelo, de modo que sejam submetidos a testes que analisem se os dados são representativos, se podem levar a enviesamentos ou se há lacunas indesejáveis. Exemplo de modelos para a documentação desses procedimentos de governança podem ser extraídos de ferramentas já existentes como os *Model Cards*, propostos pelo Google (Mitchell, 2019), ou as *AI Factsheets*, propostas pela IBM (2023).

Finalmente, há de se questionar sobre o nível das informações sobre as quais se prestará contas: se focadas no plano individual do titular a quem foi negada a explicação ou no plano global ou coletivo. Não se discrimina um indivíduo por si, mas pelo que representa, ou, em termos mais técnicos, por possuir uma característica típica de uma coletividade, que é, pela associação estrutural à qual remete, legalmente protegida (Selbst; Barocas, 2016). Como a discriminação algorítmica assume dimensão coletiva, também deve fazê-lo a prestação de contas sobre os aspectos a ela relacionados.

d. Síntese: Por que se presta contas?

Presta-se contas, em última instância, para assegurar que o poder inscrito nas decisões automatizadas possa ser legitimamente controlado por aqueles sobre os quais incide (Bovens, 2010), o que é especialmente relevante naqueles casos em que o segredo de negócios se afigure de maneira estrutural como estratégia para negar visibilidade à lógica decisional. Assim, evita-se a reprodução de desigualdades estruturais por meio de sistemas automatizados que escapam ao controle democrático.

A regulação responsiva, nesse cenário, apresenta-se como modelo refinado de *enforcement* escalável, em que a confiança é

proporcional à cooperação epistêmica (Ayres; Braithwaite, 1992). A pirâmide de fiscalização permite diferenciar entre aqueles que colaboram para a construção de sistemas auditáveis e transparentes, e aqueles que instrumentalizam o segredo para preservar assimetrias informacionais.

Retomando o quadro de conhecimento relacional apresentado anteriormente, o objetivo de maior transparência, para que por meio dela se possa expor o poder algorítmico, demanda a fiscalização ativa das zonas de ignorância e obscurecimento por meio da criação de condições institucionais que a contestação e, quando necessário, a correção pública dos sistemas automatizados que se demonstrem discriminatórios. Assim, reverte-se o processo em que o comportamento dos indivíduos se torna transparente e os sistemas automatizados opacos, trazendo esses últimos ao escrutínio público (Frazão, 2020; Bueno, 2023; Fujimoto, 2023).

e. Limitações do modelo regulatório proposto

Não se pode afirmar, entretanto, que o modelo de regulação responsiva proposto seja suficiente para enfrentar, de forma plena, os riscos gerados pela opacidade algorítmica escudada sob o segredo de negócios. Em primeiro lugar, trata-se de um mecanismo discricionário, que depende da atuação concreta da autoridade regulatória caso a caso, geralmente em resposta a uma negativa específica de fornecimento de informações pelo controlador (Sá; Lima, 2020). Isso significa que a auditoria prevista no art. 20, § 2º, da LGPD atua *ex post* e de modo pontual, e não como forma de acompanhamento contínuo e sistemático do funcionamento dos sistemas automatizados potencialmente discriminatórios (Sá; Lima, 2020).

Em segundo lugar, mesmo quando acionada, essa auditoria opera dentro dos limites institucionais e cognitivos da autoridade regulatória, que não detém, por si só, a capacidade de antecipar a totalidade dos possíveis usos e significações sociais que a informação produzida por tais sistemas pode assumir. A capacidade de identificar discriminações, sobretudo aquelas difusas, estrutural-

mente arraigadas ou manifestas apenas por correlações indiretas, não se esgota no exame técnico-regulatório, mas depende de uma pluralidade de interpretações, como os da sociedade civil, academia, imprensa especializada, comunidades afetadas e órgãos de defesa coletiva (Jain *et al*, 2024). Dessa maneira, ao condicionar o escrutínio do sistema à atuação de um único agente institucional, ainda que responsivo, restringe-se a esfera pública da contestação, e, com ela, a possibilidade de desfechos que desvelem discriminações por sistemas automatizados (Rogal, 2020; Sunstein, 2020).

Por tais razões, além do modelo responsivo quanto à prestação de contas nos casos em que se reconheça presente o segredo de negócios, é preciso instituir mecanismos complementares que ampliem o acesso público à informação relevante, dentro de balizas proporcionais, e favoreçam o controle social e institucional difuso dos sistemas automatizados. Entre tais mecanismos, podem-se considerar modelos de *transparency by default* em certos setores, como já se faz, por exemplo, na Administração Pública, como com a Lei de Acesso à Informação (Brasil, 2011), garantias processuais a práticas de *whistleblowing* (Rogal, 2020) e de auditorias independentes (Groves *et al*., 2024; Longpre *et al*., 2024). A legitimidade do uso de sistemas opacos, em suma, não pode repousar exclusivamente na confiança institucional na autoridade reguladora, mas deve ser confrontada, continuamente, pela pluralidade epistêmica de uma sociedade democrática.

Conclusão

O artigo demonstrou que a proteção jurídica ao segredo de negócios, embora fundamental para o incentivo à inovação e à competitividade, pode resultar em opacidade incompatível com os direitos assegurados pela legislação de proteção de dados pessoais, sobretudo diante de decisões automatizadas com potencial discriminatório. A análise evidenciou os múltiplos pontos de fricção entre esses regimes normativos, ao mesmo tempo em que identificou a necessidade de mediações institucionais e regulatórias que não enfraqueçam nem a proteção da livre iniciativa nem os direitos fun-

damentais. A comparação entre abordagens europeias e norte-americanas evidenciou diferentes estratégias de ponderação entre transparência e segredo comercial, sendo a regulação responsiva proposta como caminho promissor no contexto brasileiro. No entanto, seu êxito depende da articulação com mecanismos complementares capazes de garantir uma prestação de contas efetiva e a contestação pública por múltiplos agentes. Em síntese, o escrutínio sobre sistemas opacos não pode depender unicamente de autoridades reguladoras, exigindo-se uma infraestrutura institucional que favoreça a pluralidade epistêmica e o controle social contínuo sobre o poder algorítmico.

Referências

ALFRINK, Kars *et al*. Contestable AI by design: Towards a framework. **Minds and Machines**, v. 33, n. 4, p. 613-639, 2023.

ALMADA, Marco. Human intervention in automated decision-making: Toward the construction of contestable systems. *In*: **Proceedings of the Seventeenth International Conference on artificial intelligence and law**. 2019. p. 2-11.

ALMADA, Marco. Revisão humana de decisões automatizadas. **Pós-Debate**, v. 7, 2019.

ALMELING, David S. Seven reasons why trade secrets are increasingly important. **Berkeley Technology Law Journal**, p. 1091-1117, 2012.

ARTICLE 29 DATA PROTECTION WORKING PARTY. **Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679**. 2018. Disponível em: https://ec.europa.eu/newsroom/article29/items/612053. Acesso em: 02 mai. 2025.

AYRES, Ian; BRAITHWAITE, John. Responsive regulation: Transcending the deregulation debate. **Oxford University Press**, USA, 1992.

BALDWIN, Robert; BLACK, Julia. Really responsive regulation. **The Modern Law Review**, v. 71, n. 1, p. 59-94, 2008.

BARBOSA, Denis Borges. **Uma introdução à propriedade intelectual.** 2a ed. 2003.

BAROCAS, Solon; SELBST, Andrew D. Big data's disparate impact. **Calif. L. Rev.**, v. 104, p. 671, 2016.

BLACK, Emily; RAGHAVAN, Manish; BAROCAS, Solon. Model multiplicity: Opportunities, concerns, and solutions. *In*: **Proceedings of the 2022 ACM conference on fairness, accountability, and transparency**. 2022. p. 850-863.

BOVENS, Mark. Two Concepts of Accountability: Accountability as a Virtue and as a Mechanism. **West European Politics**, v. 33, n. 5, p. 946-967, 2010.

BRASIL. Lei nº 9.279, de 14 de maio de 1996. Regula direitos e obrigações relativos à propriedade industrial. **Diário Oficial da União:** seção 1, Brasília, DF, 15 maio 1996. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l9279.htm.

BRASIL. **Decreto nº 1.355, de 30 de dezembro de 1994**. Promulga a Ata Final que incorpora os resultados da Rodada Uruguai de Negociações Comerciais Multilaterais do GATT. Brasília, DF: Presidência da República, 1994. Disponível em: https://www.planalto.gov.br/ccivil_03/decreto/Antigos/D1355.htm Acesso em: 2 maio 2025.

BRASIL. **Lei nº 12.527, de 18 de novembro de 2011**. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências. Brasília, DF: Presidência da República, 2011. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm. Acesso em: 2 maio 2025.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 2 maio 2025.

BRENNAN, Tim; DIETERICH, William; EHRET, Beate. **Evaluating the predictive validity of the COMPAS risk and needs assessment system**. Criminal Justice and behavior, v. 36, n. 1, p. 21-40, 2009.

BREVINI, Benedetta; PASQUALE, Frank. Revisiting the Black Box Society by rethinking the political economy of big data. **Big Data & Society**, v. 7, n. 2, 2020. DOI:10.1177/2053951720935146.

BRUCE, Peter. **Estatística prática para cientistas de dados**. 1º ed. Rio de Janeiro: Alta Books, 2019, p. 29.

BUENO, Rafael Carvalho. **Lei Geral de Proteção de Dados, segredo de negócio e decisões automatizadas: a suficiência da análise contrafactual para verificação de aspectos discriminatórios**. Dissertação de Mestrado. Universidade Federal de Santa Catarina, 2023.

CALDERS, Toon; ŽLIOBAITĖ, Indrė. **Why unbiased computational processes can lead to discriminative decision procedures**. In: CUSTERS, Bart; CALDERS, Toon; SCHERMER, Bart; ZARSKY, Tal. **Discrimination and privacy in the information society**. 1. ed. Berlin: Springer, 2013. p. 43-57.

CARUSI, Annamaria. Data as representation: Beyond anonymity in e-research ethics. **International Journal of Internet Research Ethics**, v. 1, n. 1, p. 37-65, 2008.

CITRON, Danielle Keats; PASQUALE, Frank. The scored society: Due process for automated predictions. **Wash. L. Rev.**, v. 89, p. 1, 2014.

CITRON, Danielle Keats. Technological due process. **Wash. UL Rev**., v. 85, p. 1249, 2007.

FERGUSON, Andrew Guthrie. **The rise of Big Data policing:** surveillance, race and the future of law enforcement. 1º ed. Nova York: New York University, 2017, pp. 9-14.

FICO, Bernardo; ASSUNÇÃO, Isadora. **A case for regulation: impacts of artificial intelligence on the LGBTQIA+ community**. Legal Tech Center, 2022. Disponível em: https://legaltechcenter.net/files/sites/159/2022/06/Assuncao_Fico_A_Case_for_Regulation_Impacts_of_AI_on_the_LGBTQIA_Community.pdf. Acesso em: 02 mai. 2025.

FRAZÃO, Ana; OLIVA, Milena Donato; TEPEDINO, Gustavo. **Lei Geral de Proteção de Dados Pessoais e suas repercussões no direito brasileiro.** Thomson Reuters Brasil, 2019.

FRAZÃO, Ana. **Transparência de algoritmos x segredo de empresa.** Jota, 2021. Disponível em: https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/transparencia-de-algoritmos-x-segredo-de-empresa. Acesso em: 09 abr. 2025.

FRIEDMAN, David D.; LANDES, William M.; POSNER, Richard A. Some economics of trade secret law. **Journal of Economic Perspectives**, v. 5, n. 1, p. 61-72, 1991.

FUJIMOTO, Milton Yasuo. **Segredo de negócios, proteção de dados pessoais e inteligência artificial-os desafios do diálogo**. Dissertação de Mestrado. Universidade de São Paulo, 2023.

GARANTE PRIVACY. **Registro dei provvedimenti.** n. 741 del 27 novembre 2024. Disponível em: https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/10077129. Acesso em: 02 maio 2025.

GROVES, Lara *et al.* Auditing work: Exploring the New York city algorithmic bias audit regime. *In*: **Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency.** 2024. p. 1107-1120.

GUNNINGHAM, Neil. **Enforcement and compliance strategies**. 2010. In: BALDWIN, Robert; CAVE, Martin; LODGE, Martin (eds.). **The Oxford Handbook of Regulation**. New York: Oxford University Press, 2010.

JAIN, Shomik *et al.* Algorithmic pluralism: A structural approach to equal opportunity. In: **Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency**. 2024. p. 197-206.

LAKKARAJU, Himabindu *et al.* Identifying unknown unknowns in the open world: Representations and policies for guided exploration. In: **Proceedings of the AAAI Conference on Artificial Intelligence**. 2017.

LARSON, Jeff *et al.* **How we analyzed the COMPAS Recidivism Algorithm**. ProPublica, 2016. Disponível em: https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm. Acesso em: 04 mar. 2025.

LEMLEY, Mark A. How generative AI turns copyright upside down. Science & Technology Law Review, v. XXV, 2024.

LEMLEY, Mark A. The surprising virtues of treating trade secrets as IP rights. **Stanford Law Review,** v. 61, p. 311, 2008.

LERMAN, Jonas. Big data and its exclusions. **Stan. L. Rev. Online**, v. 66, p. 55, 2013.

LONGPRE, Shayne *et al.* **A safe harbor for ai evaluation and red teaming**. arXiv preprint arXiv:2403.04893, 2024.

MARRAFON, Marco Aurélio; MEDON, Filipe. **Importância da revisão humana das decisões automatizadas na Lei Geral de Proteção de Dados**. Conjur, 2019. Disponível em: https://www.conjur.com.br/2019-set-09/constituicao-poder-importancia-revisao-humana-decisoes-automatizadas-lgpd/. Acesso em: 09 abr. 2025.

MILLER, Tim. Explanation in artificial intelligence: Insights from the social sciences. **Artificial intelligence**, v. 267, p. 1-38, 2019.

MITCHELL, Margaret *et al.* Model cards for model reporting. In: **Proceedings of the conference on fairness, accountability, and transparency.** 2019. p. 220-229.

MOSTOWY, Walter A. Explaining Opaque AI Decisions, Legally. **Berkeley Technology Law Journal**, v. 35, n. 4, p. 1291-1330, 2020.

O'NEIL, Cathy. **Algoritmos de destruição em massa**. Santo André: Editora Rua do Sabão, 2021.

ORGANIZAÇÃO MUNDIAL DA PROPRIEDADE INTELECTUAL. **Guide to Trade Secrets and Innovation**. Genebra: WIPO, 2024.

PASQUALE, Frank. **The black box society: The secret algorithms that control money and information**. Harvard University Press, 2015.

PROCTOR, Robert N.; SCHIEBINGER, Londa. **Agnotology: The making and unmaking of ignorance.** São Francisco: Stanford University Press, 2008.

RISCH, Michael. Why do we have trade secrets. **Marq. Intell. Prop. L. Rev**., v. 11, p. 1, 2007.

ROGAL, Lauren. Secrets, lies, and lessons from the Theranos scandal. **Hastings** LJ, v. 72, p. 1663, 2020.

SÁ, Maria de Fátima Freire de; LIMA, Taisa Maria Macena de. Inteligência artificial e Lei Geral de Proteção de Dados Pessoais: o direito à explicação nas decisões automatizadas. **Revista Brasileira de Direito Civil**, v. 26, n. 04, p. 227-227, 2020.

SCHILLER, A. Arthur. Trade secrets and the Roman Law; the actio servi corrupti. **Colum. L. Rev.**, v. 30, p. 837, 1930.

SELBST, Andrew; POWLES, Julia. "Meaningful information" and the right to explanation. *In*: **conference on fairness, accountability and transparency**. PMLR, 2018. p. 48-48.

SHENTON, Andrew K. Viewing information needs through a Johari Window. **Reference Services Review**, v. 35, n. 3, p. 487-496, 2007.

SUNSTEIN, Cass R. **Too much information: understanding what you don't want to know**. MIT Press, 2020.

SUPREMA CORTE DE ALBANY. **Matter of Rayner v. New York State Dept. of Correction & Community Supervision**, 2023. Disponível em: https://law.justia.com/cases/new-york/other-courts/2023/2023-ny-slip-op-23293.html. Acesso em: 02 mai. 2025.

SUPREMA CORTE DE WISCONSIN. **No. 2015AP157–CR**. Sentença. State v. Loomis. Wisconsin, 13 de julho de 2016. Disponível em: https://cases.justia.com/wisconsin/supreme-court/2016-2015ap000157-cr.pdf?ts=1468415026. Acesso em: 08 abr. 2025.

TRIBUNAL DE JUSTIÇA DA UNIÃO EUROPEIA. PROCESSO C-203/22, **CK contra Magistrat der Stadt Wien**. 27 fev. 2025. Disponível em: https://curia.europa.eu/juris/document/document.jsf;jsessionid=CFDA0D70CD29C4B4A2820B36474754F7?text=&docid=295841&pageIndex=0&doclang=PT&mode=lst&dir=&occ=first&part=1&cid=3335197. Acesso em: 10 abr. 2025.

UNIÃO EUROPEIA. **Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho, de 27 de abril de 2016**. Relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados – GDPR). 2016. Disponível em: https://eur-lex.europa.eu/eli/reg/2016/679/oj. Acesso em: 02 mai. 2025.

USING AI Factsheets for AI Governance. IBM Cloud Pak for Data. 2023. Disponível em: https://dataplatform.cloud.ibm.com/docs/content/wsj/analyze-data/factsheets-model-inventory.html?context=cpdaas. Acesso em: 18 de nov. de 2024.

VARELA-PEZZANO, Eduardo. On slaves, guilds and the origins of trade secrets. **Vniversitas**, n. 121, p. 217-232, 2010.

WACHTER, Sandra; MITTELSTADT, Brent; FLORIDI, Luciano. Why a right to explanation of automated decision-making does not exist in the general data protection regulation. **International Data Privacy Law,** v. 7, n. 2, p. 76-99, 2017.

WAGNER, Ben. Liable, but not in control? Ensuring meaningful human agency in automated decision-making systems. **Policy & Internet,** v. 11, n. 1, p. 104-122, 2019.

WALDMAN, Ari Ezra. Power, process, and automated decision-making. **Fordham L. Rev**., v. 88, p. 613, 2019.

YANG, Wenli *et al.* Survey on explainable AI: From approaches, limitations and applications aspects. **Human-Centric Intelligent Systems**, v. 3, n. 3, p. 161-188, 2023.

ZUBOFF, Shoshana. **A era do capitalismo de vigilância**. Rio de Janeiro: Editora Intrínseca, 2020.

3º lugar

A identificação da criança como pressuposto jurídico-operacional para a aplicação proporcional da LGPD: limites técnicos e diretrizes regulatórias

Nome Giuseppe Grando

Resumo

O presente artigo analisa um paradoxo central na proteção de dados contemporânea: a massiva e crescente presença de crianças e adolescentes em plataformas digitais ocorre sob um manto de "invisibilidade programática", que neutraliza a eficácia das salvaguardas previstas na Lei Geral de Proteção de Dados Pessoais (LGPD). Partindo da premissa de que a aplicação do princípio do melhor interesse da criança depende fundamentalmente do reconhecimento de sua condição etária pelo agente de tratamento, este trabalho argumenta que a verificação de idade (*age assurance*) não constitui uma barreira, mas a pré-condição técnico-jurídica indispensável para a "calibragem" proporcional das medidas protetivas. Para tanto, por meio de análise jurídico-dogmática e documental, investigam-se as consequências sistêmicas dessa lacuna operacional, mapeia-se o panorama de soluções tecnológicas e seus dilemas, e utiliza-se o estudo de caso do TikTok para demonstrar a convergência regulatória global. Como contribuição principal, são delineadas propostas para a Agência Nacional de Proteção de Dados (ANPD), notadamente a criação de um *framework* de risco, e para os agentes de tratamento, baseadas em *age-appropriate de-*

sign. Conclui-se que a superação deste desafio exige uma mudança de paradigma, de uma "internet por padrão adulta" para uma que, desde sua concepção, reconheça e se adapte às diferentes infâncias, materializando um imperativo ético e legal.

Palavras-chave: *Proteção de dados; crianças e adolescentes; verificação de idade; melhor interesse; LGPD.*

1. Introdução

A imersão de crianças e adolescentes no ecossistema digital é um fenômeno consolidado e crescente. Dados da pesquisa TIC Kids Online Brasil de 2024 revelam que 93% da população brasileira entre 9 e 17 anos é usuária de internet, o que corresponde a 24,5 milhões de jovens (Comitê Gestor da Internet no Brasil, 2025). Dentro desse universo, plataformas de compartilhamento de vídeos curtos ganharam protagonismo, com o TikTok sendo apontado como a principal rede social por mais de um terço das crianças na faixa de 9 a 12 anos (Henriques; Ribeiro, 2024).

Essa massiva adesão infantojuvenil expõe um paradoxo regulatório central: embora os Termos de Serviço da plataforma restrinjam seu uso a maiores de 13 anos, a realidade empírica, corroborada por dados da própria empresa, que admite a remoção de milhões de contas infantis no Brasil anualmente (Brasil, 2024b), demonstra uma presença sistêmica e desprotegida deste público. A Agência Nacional de Proteção de Dados (ANPD) já concluiu que os mecanismos de verificação etária da plataforma são frágeis e insuficientes, permitindo que o tratamento de dados de milhões de crianças ocorra à margem da regulação (Brasil, 2023b).

Esse cenário evidencia uma lacuna operacional que desafia a aplicação do arcabouço protetivo da Lei Geral de Proteção de Dados Pessoais (LGPD). Em seu Enunciado nº 1/2023, a ANPD estabeleceu que o tratamento de dados de crianças e adolescentes pode se fundamentar em diversas hipóteses legais dos artigos 7º e 11 da lei, para além do consentimento parental, contanto que seja "*obser-*

vado e prevalecente o seu melhor interesse, a ser avaliado no caso concreto" (Brasil, 2023a, p. 1).

Tal flexibilidade, contudo, longe de diminuir as obrigações dos agentes de tratamento, eleva a complexidade da conformidade, pois a aplicação efetiva do princípio do melhor interesse depende de um pressuposto fático e lógico: saber quando se está lidando com uma criança ou adolescente.

Conforme reconhece a própria ANPD, a falta de um mecanismo eficaz na verificação de idade, além de infringir o melhor interesse da criança, demonstra a displicência do agente de tratamento com os deveres de segurança e prestação de contas (Brasil, 2024b). Esse vácuo cria um "efeito em cascata" que afeta todo o ecossistema: agentes de tratamento vivem em constante insegurança jurídica; pais e responsáveis perdem a capacidade de mediar a vida digital de seus filhos; e as próprias crianças são submetidas a uma "adultificação" forçada ou a uma exclusão digital desproporcional.

Diante desse cenário de insegurança jurídica e vulnerabilidade, emerge a questão de pesquisa que orienta este trabalho: como garantir que o tratamento de dados de crianças e adolescentes, nos termos da LGPD, observe o seu melhor interesse e as salvaguardas adequadas, se os agentes de tratamento não possuem meios eficazes para identificar a condição etária do titular nos ambientes digitais?

Para responder a essa questão, o presente trabalho adota a utilização do método dedutivo, de forma que a investigação se desenvolve por meio de uma análise jurídico-dogmática, que parte do arcabouço normativo da LGPD e das manifestações da ANPD, e de uma análise documental de normas internacionais, decisões de outras autoridades e do panorama de soluções técnicas disponíveis. O percurso lógico busca, assim, identificar as contradições operacionais entre a norma e sua aplicabilidade em contextos digitais, a fim de fundamentar as propostas apresentadas

Para ilustrar o problema estrutural em um contexto prático e global, o caso do TikTok é analisado sob o escrutínio regulatório no Brasil e no mundo. A partir desse diagnóstico, são delineadas propostas concretas para uma regulação proporcional centrada na

criança, com diretrizes para a ANPD e recomendações operacionais para os agentes de tratamento. Por fim, a conclusão reafirma a tese central, consolidando a identificação etária como pré-condição técnica e jurídica para a proteção efetiva e proporcional de crianças e adolescentes na era digital.

O presente artigo argumenta, portanto, que a verificação de idade (*age assurance*) não deve ser interpretada como uma barreira ao acesso, mas como um instrumento jurídico-operacional indispensável para a calibragem da proteção de dados. A identificação etária é a pré-condição para modular o ambiente digital e aplicar as salvaguardas da LGPD de forma escalonada e proporcional ao risco, superando a inércia setorial que, muitas vezes, reflete uma "*reluctance on the part of service providers to take on the responsibilities they would have to children once their age is known*" (5rights Foundation, 2021, p. 7).

2. Identificar para proteger: a condição etária como chave de calibragem regulatória

Uma primeira e mais literal leitura do arcabouço normativo da LGPD poderia conduzir à interpretação de que o tratamento de dados de crianças estaria invariavelmente condicionado à obtenção do consentimento específico e em destaque de um dos pais ou do responsável legal, conforme preceitua o art. 14, § 1º, da referida lei.

Contudo, essa visão centrada unicamente no consentimento parental se mostra insuficiente para endereçar a complexidade das interações no ambiente digital, uma percepção já consolidada pela própria ANPD. Em um movimento que denota maturidade e sofisticação interpretativa, a ANPD, por meio do Enunciado CD/ANPD nº 1/2023, estabeleceu que "o tratamento de dados pessoais de crianças e adolescentes poderá ser realizado com base nas hipóteses legais previstas no art. 7º ou no art. 11 da [...] LGPD, desde que observado e prevalecente o seu melhor interesse" (Brasil, 2023a, p. 1).

A publicação desse enunciado promoveu um fundamental deslocamento do centro de gravidade hermenêutico: a legalidade do tratamento de dados infantojuvenis não mais reside exclusivamen-

te em uma única base legal, mas na observância de um "requisito adicional: a observância e a prevalência do princípio do melhor interesse da criança ou adolescente" (Brasil, 2024a, p. 9). Este princípio, longe de ser um conceito etéreo, possui um denso arcabouço jurídico e axiológico, com raízes na Convenção sobre os Direitos da Criança da ONU.

Como detalhado pelo Comitê dos Direitos da Criança e ecoado em publicações nacionais, o melhor interesse deve ser compreendido em uma tripla dimensão: (I) como um direito substantivo de ter seus interesses considerados de forma primária em qualquer ponderação; (II) como um princípio jurídico fundamental e interpretativo, que orienta o aplicador da lei a sempre escolher a interpretação mais protetiva; e (III) como uma regra de processo, que demanda uma avaliação concreta dos possíveis impactos de qualquer ação que afete este público (Instituto Alana; Ministério Público do Estado de São Paulo, 2022).

Dessa forma, a LGPD, ao positivar o melhor interesse como cláusula geral norteadora do tratamento de dados de crianças e adolescentes, impõe aos agentes de tratamento um dever que transcende a simples verificação formal de uma hipótese legal. Exige-se uma análise material dos riscos e benefícios de suas operações, tratando os dados deste público com um grau de cuidado elevado, análogo ao dispensado aos dados sensíveis, dada a sua condição peculiar de desenvolvimento e vulnerabilidade intrínseca (Instituto Alana; Ministério Público do Estado de São Paulo, 2022).

O grande desafio que se impõe, e que constitui o cerne deste trabalho, reside na questão subsequente: como pode um agente de tratamento cumprir o dever de avaliar e garantir o "melhor interesse" se, na prática, não possui os meios para saber que o titular dos dados é, de fato, uma criança ou adolescente?

A dimensão procedimental do melhor interesse impõe ao agente de tratamento o ônus de avaliar os impactos de suas operações, o que pressupõe uma diligência ativa e prévia. Contudo, essa obrigação se torna inexequível em um cenário de "cegueira etária", no qual a idade do titular é desconhecida, exatamente conforme apontado pela própria ANPD, "a ausência de mecanismos adequa-

dos de verificação de idade compromete a privacidade e a segurança dos titulares de dados" (Brasil, 2024b, p. 20).

Ignorar a condição etária do usuário significa, na prática, aplicar um tratamento de dados padronizado, concebido para adultos, a um público com vulnerabilidades e necessidades de proteção acentuadas. Tal abordagem falha em respeitar o princípio das capacidades evolutivas da criança (e*volving capacities of the child*), que reconhece que os riscos e as oportunidades no ambiente digital se alteram drasticamente conforme a idade e o estágio de desenvolvimento do indivíduo (United Nations, 2021).

A identificação etária, portanto, funciona como um gatilho técnico-jurídico, uma chave que ativa e permite a calibragem de um conjunto de salvaguardas previstas na LGPD. A primeira e mais evidente é a obrigação de transparência qualificada, disposta no art. 14, § 6º, da LGPD, que exige o fornecimento de informações de maneira simples, clara e acessível, adequada ao entendimento da criança (Brasil, 2018), ou seja, sem saber a idade do interlocutor, é impossível adequar a linguagem.

Adicionalmente, o reconhecimento da idade é condição para a aplicação efetiva dos princípios da segurança e da privacidade desde a concepção e por padrão (*security and privacy by design and by default*). Um tratamento de dados desenhado para o melhor interesse de um adolescente de 16 anos, por exemplo, deve ter configurações de privacidade mais restritivas por padrão e medidas de segurança mais robustas do que as aplicadas a um usuário adulto, uma vez que se deve considerar o "nível de certeza de forma proporcional e razoável aos riscos a direitos e liberdades das crianças e adolescentes no caso concreto" (Brasil, 2023b, p. 11).

Por fim, a identificação da idade é o que permite ao controlador realizar a escolha consciente e defensável da hipótese legal mais adequada. Ao se valer do legítimo interesse para tratar dados de um público que sabe ser infantil, por exemplo, o agente de tratamento ativa para si o dever de realizar e documentar um criterioso teste de balanceamento, que demonstre inequivocamente a prevalência do melhor interesse da criança sobre os seus interesses comerciais (Brasil, 2024a).

Sem o conhecimento da idade, essa ponderação essencial sequer é iniciada, resultando em um tratamento de dados que, mesmo que aparente conformidade formal, carece de legitimidade material. A verificação de idade, nesse sentido, não é um fim em si mesma, mas o meio pelo qual a proteção deixa de ser abstrata e se torna operacional.

3. O desafio do invisível: ausência de verificação etária e suas consequências sistêmicas

A gênese do desafio sistêmico que impede a aplicação efetiva da LGPD na proteção de crianças e adolescentes reside na prática de mercado dominante para verificação etária: a autodeclaração. Frequentemente referida como uma mera verificação de "marcar a caixa" (*'tick box' age assurance*), essa abordagem está diretamente associada ao fracasso generalizado em se estabelecer a idade real das crianças no ambiente online (5Rights Foundation, 2021).

Tal método transfere toda a responsabilidade da veracidade da informação para a própria criança (Brasil, 2024b), um ator que, por razões de transgressão ou aspiração, sabidamente pode fornecer dados inverídicos (5Rights Foundation, 2021), então, embora seja uma solução de simples implementação, a autodeclaração como mecanismo isolado só é considerada adequada para produtos e serviços que apresentam baixo risco para o público infantojuvenil (5Rights Foundation, 2021).

A fragilidade da autodeclaração como método isolado é reconhecida pela própria ANPD. Em sua análise sobre o TikTok, a ANPD concluiu que o uso do *Age Gate* se torna "problemático quando utilizado como única barreira à utilização da plataforma por crianças e adolescentes, especialmente considerando o público do TikTok e o fato de que elevado número desse grupo de vulneráveis tem conseguido contornar essa barreira" (Brasil, 2024b, p. 3).

A ANPD ainda aprofunda a crítica ao notar que a forma como o método é desenhado "torna-o incapaz de desencorajar ou de de-

tectar a inserção de datas de nascimento inverídicas" (Brasil, 2024b, p. 3), uma falha que viola as noções de privacidade desde a concepção (*privacy by design*) e expõe crianças a riscos desnecessários.

O resultado dessa prática generalizada é a criação de um estado de "invisibilidade programática", onde as plataformas, ao adotarem um mecanismo sabidamente falho, podem alegar uma conveniente ignorância sobre a real faixa etária de seus usuários. Essa postura, descrita como "não olhe, não veja" (*don't look don't see*), desincentiva o desenvolvimento de serviços e produtos adequados às crianças (5Rights Foundation, 2021). A suposta dificuldade técnica em se realizar a verificação etária torna-se questionável, como aponta a 5Rights Foundation:

> Many companies have a detailed picture of their users' interests, location, relationships, family status, income, sexuality and so on. Understanding users (profiling) and tailoring user journeys (personalisation) are the bread and butter of the tech sector, so it is perplexing that companies claim it is difficult, impossible or intrusive to identify children by age (5Rights Foundation, 2021, p. 4).

Essa aparente dificuldade, portanto, pode ser mais bem compreendida como uma relutância em assumir as responsabilidades inerentes à proteção de dados infantojuvenis (5Rights Foundation, 2021). É essa invisibilidade deliberada que gera um efeito em cascata, cujas consequências se espalham por todo o ecossistema digital, afetando titulares, controladores, pais e a própria autoridade reguladora.

Os primeiros e mais vulneráveis a sofrerem com a ausência de verificação etária são as próprias crianças e adolescentes. Sem que as plataformas digitais reconheçam suas idades e, consequentemente, suas capacidades evolutivas distintas (United Nations, 2021), elas são submetidas a um falso dilema: ou são tratadas por padrão como adultas ("adultificação"), ou são sumariamente excluídas ("exclusão").

Na primeira hipótese, a da "adultificação", elas são expostas a riscos para os quais não possuem maturidade para lidar, como a

exploração comercial por meio de publicidade comportamental e técnicas de *design* persuasivo, conhecidas como "*nudges*", que podem impactar negativamente seu desenvolvimento. O modelo de negócio de muitas plataformas, baseado na coleta massiva de dados para análises preditivas, agrava essa exposição (Instituto Alana; Ministério Público do Estado de São Paulo, 2022).

A inadequação dos ambientes digitais é um problema sistêmico. Uma análise comparativa de redes sociais populares no Brasil, apresentada na pesquisa TIC Kids Online Brasil, constatou que as plataformas se mostraram, "no mínimo, despreparadas para atender às necessidades especiais de crianças e adolescentes, seja por falhas de *design*, seja pela ausência de termos adequados ou de mecanismos eficazes de controle e denúncia" (Henriques; Ribeiro, 2024, p. 94).

Na segunda hipótese do dilema, a da "exclusão", a reação das plataformas à pressão regulatória pode levar à adoção de barreiras excessivamente restritivas, que acabam por privar as crianças de experiências digitais que são seu direito. Conforme a 5Rights Foundation:

> Age assurance should not be used to freeze out children from areas of the digital world which they have a right to enjoy, as a way of companies avoiding their responsibilities to make a service age appropriate (5Rights Foundations, 2021, p. 49-50).

O resultado é um ambiente digital que não foi desenhado para a infância e que, por omissão ou por ação desproporcional, falha em garantir os direitos fundamentais desse público.

Para os agentes de tratamento, a "invisibilidade programática" dos usuários infantojuvenis se traduz em um estado de permanente insegurança jurídica e alto risco operacional. Como estabelecido no capítulo anterior, a aplicação correta da LGPD para esse público depende do conhecimento de sua idade para a ativação de salvaguardas e para a escolha da base legal adequada.

Ao operar sem essa informação essencial, as empresas atuam em uma zona cinzenta de conformidade, incapazes de demonstrar

proativamente que suas operações respeitam o melhor interesse da criança. Essa falha em implementar medidas técnicas e organizacionais apropriadas para garantir a proteção de dados desde a concepção é uma violação direta da LGPD (Henriques; Ribeiro, 2024).

Este risco não é meramente teórico. Decisões de autoridades de proteção de dados ao redor do mundo materializam as consequências financeiras e reputacionais. A Data Protection Commission (DPC) da Irlanda, por exemplo, ao analisar o tratamento de dados de crianças pelo TikTok, apontou os "severos riscos possíveis" aos direitos e liberdades dos usuários infantis e concluiu pela falha da empresa em implementar as medidas técnicas e organizacionais adequadas para garantir a conformidade com o General Data Protection Regulation (GDPR) (Ireland, 2023).

De forma semelhante, o European Data Protection Board (EDPB), em sua decisão vinculante sobre o caso Instagram, reforçou que as empresas devem dar a devida consideração aos fatores agravantes ou atenuantes (European Data Protection Board, 2022, p. 60), onde a vulnerabilidade dos titulares é um fator central. A inação ou a adoção de medidas frágeis, portanto, coloca os controladores em uma posição de constante vulnerabilidade perante os órgãos de fiscalização.

Não obstante, o “efeito cascata” também recairá nos pais e responsáveis das crianças e adolescentes, afinal, é dever primário deles a orientação e cuidado desses menores. No entanto, a eficácia de sua atuação no ambiente digital é diretamente minada pela "invisibilidade programática" de seus filhos perante as plataformas.

Ao não reconhecerem a presença de uma criança, os serviços deixam de oferecer as ferramentas adequadas e necessárias para uma mediação parental efetiva, deixando os responsáveis "à deriva" em um ecossistema complexo e, por vezes, hostil.

Essa dificuldade é agravada por um contexto em que a orientação parental sobre o uso da internet já é um desafio, com pesquisas indicando que uma parcela significativa de crianças e adolescentes navega com pouca ou nenhuma supervisão direta (Instituto Alana; Ministério Público do Estado de São Paulo, 2022).

A falta de reconhecimento etário pelas plataformas esvazia o papel dos pais de mediadores conscientes da privacidade dos filhos. Sem configurações de controle parental que sejam ativadas e adaptadas para a faixa etária específica da criança, torna-se extremamente difícil para os responsáveis gerenciarem a exposição a conteúdos, o tempo de tela e as interações com outros usuários.

Conforme aponta o Comentário Geral nº 25 da ONU, o dever de apoiar e orientar os filhos no ambiente digital pressupõe que os pais e cuidadores estejam cientes dos riscos e das estratégias disponíveis para proteger as crianças (United Nations, 2021). Contudo, quando o próprio design da plataforma não oferece essas estratégias por padrão, a responsabilidade é desproporcionalmente alocada sobre os ombros das famílias, que muitas vezes não possuem o conhecimento técnico ou os recursos para suprir essa falha estrutural.

Por fim, a ausência sistêmica de verificação etária eficaz gera um impacto direto sobre a própria ANPD, criando uma espécie de impasse regulatório. A ANPD, em sua atuação, tem demonstrado sofisticação ao construir um arcabouço interpretativo flexível, que busca equilibrar proteção e inovação, como visto na permissão do uso de diversas bases legais para o tratamento de dados infantojuvenis (Brasil, 2023a). Entretanto, a eficácia desse arcabouço é frustrada na prática quando a indústria falha em implementar o requisito mais basilar de todos: a capacidade de distinguir um usuário criança de um adulto.

Essa falha obriga a ANPD a um esforço de fiscalização reativo e exaustivo para provar violações que decorrem de uma omissão primária dos agentes de tratamento. A própria existência das extensas Notas Técnicas sobre o caso TikTok (Brasil, 2023b; Brasil, 2024b) ilustra o volume de trabalho necessário para regular uma questão que deveria ser um padrão de mercado.

Frequentemente, as empresas se defendem alegando a ausência de uma orientação regulatória clara sobre qual método de verificação adotar, como observado na decisão da autoridade irlandesa sobre o mesmo caso (Ireland, 2023). Contudo, essa alegação cria um ciclo vicioso: as empresas não inovam por falta de

uma regra prescritiva, e a ausência de um padrão mínimo de diligência por parte delas impede que a regulação avance para temas mais complexos. Essa inércia sistêmica não apenas compromete a proteção dos titulares mais vulneráveis, mas também enfraquece a capacidade da ANPD de garantir o cumprimento efetivo da LGPD em sua plenitude.

4. Entre técnicas e dilemas: panorama de soluções de verificação etária

A superação do estado de "invisibilidade programática" de crianças e adolescentes no ambiente digital, detalhado no capítulo anterior, depende da capacidade técnica e jurídica dos agentes de tratamento em implementar mecanismos eficazes de identificação etária.

Existe um crescente e diversificado mercado de soluções de *age assurance* (garantia de idade), que abrange desde métodos de simples declaração até complexos sistemas de inferência e verificação biométrica. O conceito de *age assurance* é amplo e compreende diferentes ambições: algumas ferramentas buscam a verificação precisa da idade (*age verification*), outras se propõem a estimar uma faixa etária (*age estimation*), enquanto algumas são desenhadas para a identificação de um indivíduo específico (*identification*) (5Rights Foundation, 2021).

Contudo, a escolha e a implementação dessas tecnologias não representam uma "bala de prata". Muitas vezes, as abordagens são combinadas ou sequenciadas ao longo da jornada do usuário, sendo aplicadas em diferentes momentos a depender da interação, como em um site de comércio eletrônico que apenas exige a verificação no momento da compra de um produto restrito (5Rights Foundation, 2021).

O debate sobre qual ferramenta adotar é perpassado por um dilema fundamental: a tensão entre a eficácia da verificação e a privacidade do titular. Uma preocupação central é que, para verificar ou estimar a idade, é necessário processar dados, inclusive

de adultos, o que pode infringir direitos e contrariar o princípio da minimização de dados (UNICEF, 2021a).

Nesse contexto, devemos analisar criticamente o panorama de soluções de verificação etária, partindo das mais simples e problemáticas para as mais robustas e complexas, com base na proporcionalidade, minimização e não discriminação, demonstrando que a solução não reside em uma única técnica, mas na aplicação ponderada e contextual de diferentes métodos, conforme o nível de risco de cada atividade de tratamento.

O ecossistema de garantia de idade oferece um leque de opções que variam em método, fonte de dados e nível de certeza. A análise a seguir explora as principais abordagens, seus mecanismos operacionais e os dilemas jurídicos e éticos que cada uma suscita.

A abordagem mais comum é a autodeclaração, na qual o usuário simplesmente informa dados como nome, e-mail e data de nascimento. Apesar de sua simplicidade, esse método é considerado insuficiente como prova de idade por órgãos reguladores europeus, que entendem ser um requisito implícito do GDPR que as empresas empreendam esforços razoáveis para verificar a idade de seus usuários. A principal falha da autodeclaração reside no fato de que o tratamento de dados se torna ilícito caso uma criança forneça o consentimento sem ter idade suficiente para fazê-lo em nome próprio (UNICEF, 2021a).

Para adicionar uma camada de robustez, a autodeclaração pode ser combinada com a estimativa por inferência, que utiliza modelos de inteligência artificial para analisar o comportamento do usuário e verificar se a idade declarada é consistente com sua atividade na plataforma. A atividade pode ser derivada de informações contextuais, como o tipo de conteúdo com o qual o usuário interage, sua localização, frequência de uso, a idade de suas conexões ou até mesmo sua proficiência em determinadas tarefas ou o uso da linguagem.Por exemplo, se uma criança declara ter 16 anos mas interage com conteúdo tipicamente associados a um público mais jovem, verificações adicionais podem ser requeridas (5Rights Foundation, 2021).

A análise biométrica representa um avanço em termos de exatidão, utilizando características físicas ou comportamentais para estimar a idade ou identificar um usuário, que pode envolver o uso de impressões digitais, escaneamento de íris e voz, ou dados comportamentais como a dinâmica de digitação (5Rights Foundation, 2021).

No entanto, o uso de biometria acarreta dilemas severos, conforme pondera a UNICEF:

> Because biometric data is based on data generated from the unique characteristics of humans, it can be used to track and profile people across their lives, which carries unknown in the long term. Without strict safeguards in place and strong legal frameworks (such as the GDPR which bans profiling of children), biometric IDs can be used to facilitate discrimination, profiling and mass surveillance. (UNICEF, 2021, p. 19).

Outra técnica de alta confiabilidade é a verificação por documentos, ou "*hard identifiers*", que confronta os dados do usuário com bases de dados oficiais ou privadas, como registros escolares, médicos, de crédito, ou aqueles mantidos por bancos e operadoras de telefonia (5Rights Foundation, 2021).

Nessas situações, o nível de confiança pode variar de baixo, como uma imagem enviada, até alto, como a leitura de um *chip NFC* em um smartphone para validar o documento. O grande dilema, contudo, é a alta intrusão na privacidade, em decorrência do acesso a dados que não seriam necessários no momento, e o risco de exclusão, uma vez que muitas crianças ao redor do mundo não possuem documentos oficiais com foto (UNICEF, 2021a).

Por fim, a verificação por terceiros surge como um modelo que busca equilibrar privacidade e segurança, no qual uma entidade confiável atesta a idade do usuário para um serviço sem necessariamente compartilhar dados adicionais (5Rights Foundation, 2021). Esses sistemas podem se basear em fontes de dados centralizadas, como grandes agregadores de crédito, o que gera preocupações de segurança, como ilustrado pelo vazamento de dados da Equifax em 2017 (UNICEF, 2021a).

Ainda sobre verificações de terceiros, surge a opção de utilizar fontes descentralizadas, que oferecem maior resiliência a ataques por não concentrarem todos os dados em um único local, ou mesmo tecnologias de *blockchain*, que dão ao usuário maior controle sobre seus dados (UNICEF, 2021a).

A prova de que a superação da "invisibilidade programática" é tecnicamente viável já pode ser observada em práticas de mercado proativas, que se anteciparam às exigências regulatórias. A plataforma Yubo, por exemplo, tornou-se a primeira grande rede social a implementar a verificação de idade para 100% de seus usuários, movida pela crença de que a segurança e a confiança são o alicerce para conexões online significativas (Yubo, [2024?]).

Em parceria com a provedora de identidade digital Yoti, a plataforma implementou um sistema de baixo atrito, no qual o usuário captura uma imagem ao vivo no próprio aplicativo, que é então analisada por uma tecnologia de estimativa facial para confirmar a correspondência com a data de nascimento fornecida. O método demonstra alta eficácia, com cerca de 87% dos usuários conseguindo verificar a idade na primeira tentativa, e gerou uma percepção positiva na comunidade, com 79% dos usuários afirmando se sentirem mais seguros na Yubo em comparação a outras plataformas (Yubo, [2024?]).

A credibilidade técnica da solução é reforçada pela transparência do parceiro tecnológico, que publica abertamente seus dados de acurácia e tem sua metodologia validada por certificação independente (Yoti, 2023).

Em um outro espectro de risco, a plataforma Roblox ilustra a aplicação de uma abordagem de alta garantia, demonstrando na prática o princípio da proporcionalidade. Para que os usuários possam "*access innovative social capabilities and age-appropriate content such as experiences for people 17 and up*" (Roblox, [2023?]), a plataforma exige um método de verificação mais robusto, pois, neste caso, o usuário deve ter no mínimo 13 anos e apresentar um documento de identidade com foto emitido pelo governo, como passaporte ou carteira de motorista, ou seja, esse processo envolve o escaneamento do documento e a captura de uma *selfie*, que é

utilizada para garantir, por meio de comparação biométrica, que o usuário é a mesma pessoa presente na identificação oficial (Roblox, [2023?]).

Os casos de Yubo e Roblox, portanto, servem como provas de conceito de que diferentes modelos de *age assurance* podem e devem ser calibrados de acordo com as funcionalidades e os riscos específicos de cada ambiente digital, transformando a proteção de dados infantojuvenis em uma realidade operacional.

Tendo todo esse panorama em vista, fica a evidência que não há uma "solução universal", de forma que a opção dentre todas as ferramentas *age assurance* não pode ser feita de maneira arbitrária, mas deve, conforme destaca Cruz (2024), ser guiada pela proporcionalidade, a finalidade e pelo risco que o conteúdo ou serviço apresenta à criança ou ao adolescente.

A ANPD corrobora essa visão ao afirmar que as práticas de verificação devem observar os riscos envolvidos no tratamento de dados, de modo a se alcançar um "nível de certeza de forma proporcional e razoável aos riscos a direitos e liberdades das crianças e adolescentes no caso concreto" (Brasil, 2023b, p. 11). Isso significa que, em vez de um mecanismo único, o ideal é a implementação de múltiplas verificações, ajustadas à experiência que se deseja personalizar para o usuário (Cruz, 2024).

Para avaliar o nível de risco, a 5Rights Foundation (2021) propõe um *framework* "4C", baseado em quatro categorias de riscos que as crianças enfrentam no ambiente digital, sendo eles: a) Conteúdo; Risco de exposição a material danoso ou inadequado para a idade, como pornografia, violência ou desinformação, b) Contato; Risco de interação com atores maliciosos, que pode levar à exploração ou assédio, especialmente em serviços que permitem mensagens privadas, c) Conduta; Risco de o serviço facilitar ou encorajar comportamentos prejudiciais, como *bullying* ou *trolling*, d) Contrato; Risco de exposição a relações comerciais inapropriadas, como publicidade direcionada agressiva ou mecânicas de jogos que se assemelham a apostas (5Rights Foundation, 2021).

A intensidade desses riscos deve, portanto, ditar a necessidade, e a profundidade, do método de verificação exigido, contudo,

tendo em mente que essa necessidade nem sempre é linear, pois não é uma regra que as crianças mais novas são as que necessitam do maior nível de garantia de idade, uma vez que as mais velhas, ao receberem seus primeiros celulares, computadores e acesso irrestrito a serviços desenhados para adultos – que é, justamente, o ambiente digital –, podem precisar de mais suporte e proteção (5Rights Foundation, 2021).

A natureza do serviço também é crucial, pois certos conteúdos, como apostas e consumo de bebida alcoólica, podem ter um impacto desproporcional em adolescentes, logo, a calibragem da proteção exige uma análise multifacetada, que considere não apenas o risco inerente ao serviço, mas também a vulnerabilidade específica de cada faixa etária (5Rights Foundation, 2021).

Dessa forma, fica evidente que o caminho para a efetiva garantia de idade no ambiente digital não reside na adoção de um método único e infalível, mas na construção de uma arquitetura de confiança baseada na proporcionalidade, de maneira que, aos agentes de tratamento, fica imposta a tarefa de calibrar seus mecanismos de verificação de acordo com a natureza de seus serviços, produtos e ambientes. Com isso, para relacionar esse nosso estudo teórico com a situação fática atual, utilizar-se-á o emblemático caso do TikTok, já estudado pela ANPD, a fim de observar como as autoridades regulatórias, no Brasil e no mundo, têm confrontado este desafio estrutural.

5. O caso tiktok e a resposta regulatória: brasil e mundo

A discussão sobre as fragilidades dos mecanismos de verificação etária, bem como os dilemas inerentes às soluções técnicas, deixa o plano teórico e encontra sua mais contundente expressão prática na análise do caso TikTok.

A plataforma, por sua imensa popularidade global entre crianças e adolescentes, tornou-se centro de atenção das autoridades de proteção de dados do mundo inteiro, de maneira que a forma que esses reguladores, no Brasil e no mundo, têm respondido aos

desafios impostos por este serviço evidencia não apenas as falhas estruturais da plataforma, mas também uma crescente convergência internacional sobre a necessidade de salvaguardas robustas para o público infantojuvenil.

No Brasil, a ANPD tem se debruçado sobre o caso TikTok de forma consistente, produzindo análises técnicas que revelam uma profunda preocupação com a forma como a plataforma lida com os dados de seus usuários mais jovens.

Um dos pontos centrais da crítica da autoridade é a funcionalidade que permite o uso do aplicativo sem a necessidade de criação de uma conta, o chamado "*feed sem cadastro*" (Brasil, 2024b). Segundo a ANPD, essa possibilidade, somada ao fato de que muitas crianças conseguem burlar os mecanismos de verificação de idade no cadastro, demonstra as fragilidades do sistema e a urgência de melhorias.

A autoridade brasileira concluiu que, embora o TikTok afirme não ter a intenção de tratar dados de titulares menores de 13 anos, a combinação do "*feed* sem cadastro" com a baixa efetividade do *age gate* resulta no tratamento fático de dados deste público, que ocorre sem a adoção das garantias e salvaguardas adequadas (Brasil, 2024b), sendo categórica ao afirmar que a manutenção dessa funcionalidade de acesso irrestrito, "sem qualquer mecanismo adequado de verificação de idade, compromete a proteção de dados de crianças e adolescentes" (Brasil, 2024b, p. 20).

A análise da autoridade avança ao apontar que a prática do "*feed* sem cadastro" foi uma "decisão de negócio" da empresa, visando ampliar sua base de usuários e permitir o tratamento de dados em larga escala, inclusive de crianças (Brasil, 2024b), sendo considerada como "incompatível com o ordenamento jurídico vigente e representa um risco à proteção dos titulares de dados, especialmente no que tange às crianças" (Brasil, 2024b, p. 20), o que levou a autoridade a considerar a instauração de processo administrativo sancionador e a necessidade de medidas preventivas para suspender imediatamente o funcionamento do "*feed* sem cadastro" (Brasil, 2024b).

Além disso, a ANPD já havia apontado, em análise anterior, que o mecanismo de verificação de idade do TikTok é "frágil" (Brasil,

2023b), ponderando sobre o dilema entre a praticidade de um cadastro simplificado e os severos riscos associados à permissão de que milhões de crianças ingressem na plataforma e tenham seus dados tratados, inclusive dados potencialmente sensíveis, a depender do conteúdo que compartilham (Brasil, 2023b).

A conclusão é que o controlador deveria reavaliar seus mecanismos para que o filtro seja mais efetivo, sopesando se uma ferramenta de verificação mais robusta, ainda que trate mais dados no momento inicial, não seria mais adequada para mitigar os riscos de um tratamento extenso e desprotegido de dados infantis (Brasil, 2023b).

A narrativa pública da empresa, contudo, contrasta com as conclusões técnicas da autoridade reguladora. Em um painel de alto nível sobre segurança digital durante o G20 no Brasil, a Chefe Global de Política de Segurança e Bem-Estar Juvenil do TikTok, Emma Leiken, ao ser questionada sobre as medidas adotadas pela plataforma, afirmou que a empresa leva "extremamente a sério" a regra que proíbe o uso do serviço por menores de 13 anos (G20 Brasil, 2024).

Segundo a executiva, a primeira barreira é o *Age Gate*, no qual o usuário deve inserir sua data de nascimento para criar uma conta, sendo o acesso negado caso a idade informada seja inferior ao limite. Ela admite que, embora a maioria das pessoas seja honesta, pode existir tentativas de burlar e, por isso, a empresa adota outras medidas para detectar e suspender contas suspeitas, como o uso de moderadores humanos e automatizados que buscam por sinais de que o usuário seja menor de idade. Como prova da seriedade dessas medidas, a executiva destacou a remoção de 20 milhões de contas em 2023 por suspeita de pertencerem a usuários abaixo da idade permitida (G20 Brasil, 2024).

Esse contraste entre o discurso corporativo e a análise regulatória revela a dissonância central no debate sobre a proteção de dados infantojuvenis. Enquanto a empresa apresenta suas medidas como um esforço diligente, a ANPD, ao analisar esse mesmo sistema, classifica-o como "frágil" e "incapaz de desencorajar ou de detectar a inserção de datas de nascimento inverídicas" (Brasil, 2024b, p. 3).

A remoção de 20 milhões de contas, apresentada pela plataforma como um indicador de sucesso, pode ser interpretada, sob a ótica regulatória, como a prova empírica da falha massiva do age gate em barrar a entrada deste público em primeiro lugar. É evidente, portanto, que a autodeclaração, mesmo quando complementada por moderação reativa de conteúdo, é considerada pelos órgãos de fiscalização como uma abordagem insuficiente para endereçar os riscos de uma plataforma de alto alcance como o TikTok, reforçando a necessidade de mecanismos mais robustos e proativos.

Comprovando isso, a preocupação da autoridade brasileira não é um caso isolado, muito pelo contrário, ela reflete, em verdade, um movimento global de escrutínio sobre as práticas do TikTok, com diversas autoridades de proteção de dados chegando a conclusões semelhantes. No âmbito da União Europeia, por exemplo, a Data Protection Commission (DPC) da Irlanda, atuando como autoridade líder, conduziu uma investigação aprofundada sobre a plataforma.

Uma de suas conclusões mais relevantes foi a de que as configurações "públicas por padrão" para contas de crianças criavam riscos significativos. Conforme a decisão da DPC Ireland:

> The public-by-default settings create a risk of unauthorised access to Child Users' personal data as inadvertently or advertently disclosed in video content or via comments. This could take the form of bad actors using the TikTok website to access the personal data of Child Users in a manner that cannot be moderated by TTL. Any such access to that data as a result of utilising the website in this manner would be unauthorised access (Ireland, 2023, p. 21).

Essa falha em garantir a segurança apropriada dos dados, incluindo a proteção contra processamento não autorizado, foi vista como uma violação do princípio da integridade e confidencialidade (Ireland, 2023). A autoridade irlandesa também criticou a falta de transparência da plataforma, apontando que as referências vagas a "terceiros" e a "todos" como potenciais receptores de dados não eram claras, concisas ou inteligíveis para os usuários (Ireland, 2023).

Esse entendimento está alinhado com os princípios estabeleci-

dos pelo European Data Protection Board (EDPB), que reforça a expectativa de que empresas cuja atividade principal é o tratamento de dados devem ter medidas suficientes para proteger esses dados (EDPB, 2022). O caso ilustra a interconexão entre os princípios da legalidade, lealdade e transparência, que devem ser respeitados por qualquer controlador (Ireland, 2023).

Conforme relatado pela própria ANPD (Brasil, 2023b), ações de outras autoridades internacionais corroboram essa tendência, como a Information Commissioner's Office (ICO) do Reino Unido, que aplicou uma multa de 12,7 milhões de libras ao TikTok por falhar em realizar verificações adequadas para identificar e remover crianças menores de idade da rede. De forma similar, a autoridade italiana de proteção de dados também considerou que o mecanismo de autocertificação da plataforma era insuficiente para limitar o acesso de menores de 13 anos (Brasil, 2023b).

O caso TikTok, portanto, materializa a tese central deste trabalho. As ações coordenadas das autoridades de proteção de dados, no Brasil e no mundo, demonstram que a era da "invisibilidade programática" está sendo ativamente combatida, ou seja, que os reguladores não mais aceitam a autodeclaração como medida suficiente para plataformas de alto risco que tratam dados de crianças em larga escala, demonstrando que a pressão regulatória global força os agentes de tratamento a confrontarem o desafio da verificação etária, o que nos leva diretamente à necessidade de debater as propostas de encaminhamento e as diretrizes para uma regulação efetiva e proporcional, tema do próximo capítulo.

6. Propostas para uma regulação proporcional e centrada na criança

A convergência regulatória global analisada no capítulo anterior, que culminou em sanções e determinações contra gigantes da tecnologia como o TikTok, sinaliza o fim da era da “invisibilidade programática”, destacando que a tolerância com a alegação de ignorância sobre a presença de crianças em plataformas digitais está se esgotando.

Diante deste novo cenário de exigibilidade, torna-se imperativo superar o diagnóstico e avançar para a construção de uma arquitetura de solução. Este capítulo surge com a finalidade de delinear propostas concretas e factíveis, direcionadas tanto ao regulador quanto aos regulados, com o objetivo de criar um ecossistema digital mais seguro e protetivo para crianças e adolescentes.

Incialmente, detalha-se que a atuação da ANPD é o ponto de alavancagem para transformar a proteção de dados infantojuvenis em uma realidade operacional no Brasil e, em primeiro lugar, é fundamental a elaboração de um documento, equivalente aos "Guias Orientativos", que estabeleça um *framework* para a verificação de idade baseado em risco.

A própria ANPD já indicou que as práticas de verificação devem observar os riscos envolvidos no tratamento de dados (Brasil, 2023b). A criação de um *framework* oficial daria segurança jurídica ao mercado, que seria moldado tanto pela regulação governamental quanto pela inovação (5Rights Foundation, 2021). Tal regulação deve consolidar a premissa de que "o método de verificação de idade escolhido deve ser proporcional à finalidade e/ou ao risco que o conteúdo ou serviço pode oferecer à criança ou adolescente" (Cruz, 2024, p. 8).

Para a definição dos níveis de risco, a ANPD poderia se inspirar em modelos internacionais consolidados, como o framework "4C" proposto pela 5Rights Foundation (2021), que é ancorado em quatro riscos principais: a) Conteúdo (*Content*): risco de exposição a material danoso ou inapropriado, como violência, desinformação ou conteúdo que promova comportamentos não saudáveis; b) Contato (*Contact*): risco de interação com atores maliciosos, que pode levar à exploração ou assédio, especialmente em serviços que permitem mensagens privadas ou expõem a localização; Conduta (*Conduct*): risco de o serviço facilitar ou encorajar comportamentos prejudiciais, como bullying ou a perda de controle sobre o legado digital, e d) Contrato (*Contract*): risco de exposição a relações comerciais inapropriadas, como publicidade direcionada, custos ocultos ou apropriação indevida de dados pessoais (5Rights Foundation, 2021).

A partir da análise desses riscos, a ANPD pode orientar os agentes de tratamento a adotarem diferentes "camadas" de verificação

(Cruz, 2024), ou seja, um serviço de baixo risco, como uma *newsletter*, exigiria pouca verificação, enquanto um serviço de alto risco, como uma plataforma de jogos com microtransações, demandaria um "um nível de certeza de forma proporcional e razoável aos riscos a direitos e liberdades das crianças e adolescentes no caso concreto" (Brasil, 2023b, p. 11), justificando o uso de métodos mais robustos.

Adicionalmente, poderia reforçar o dever de cuidado que recai sobre as empresas, exigindo uma postura proativa na identificação e mitigação de riscos, uma vez que o ordenamento jurídico brasileiro deixa clara a responsabilidade das empresas fornecedoras de produtos e serviços no ambiente digital frente às crianças (Instituto Alana; Ministério Público do Estado de São Paulo, 2022), tornando fundamental que as empresas realizem avaliações para entender quem são os usuários afetados por seus serviços, incluindo crianças que não deveriam estar utilizando a plataforma.

Paralelamente à atuação regulatória, os agentes de tratamento devem adotar, desde já, um conjunto de boas práticas para garantir a proteção de dados infantojuvenis. Tais práticas devem ser incorporadas ao *design* dos produtos e serviços, em uma abordagem de *Privacy by Design*.

A primeira e mais fundamental recomendação é a segmentação da experiência do usuário, também conhecida como design apropriado à idade (*age-appropriate design*), ou seja, a partir da identificação da faixa etária, as plataformas devem calibrar o serviço para mitigar riscos, garantindo que configurações de privacidade e segurança sejam, naturalmente, mais restritivas para usuários mais jovens, e não uma opção a ser ativada, como desativar por padrão o recebimento de mensagens de estranhos, limitar a coleta de dados para perfis jovens e restringir ou vedar a publicidade comportamental dirigida a esse público (Instituto Alana; Ministério Público do Estado de São Paulo, 2022).

Para que essa segmentação funcione de fato, os responsáveis pelo tratamento devem adotar, desde a concepção de seus produtos, os princípios de *safety-by-design*, *privacy by design* e *security by design* (UNICEF, 2021a). Nesse sentido, elaborar Relatórios de Im-

pacto à Proteção de Dados (RIPD) específicos para os direitos da criança é fundamental: eles permitem identificar, analisar e mitigar riscos antes que se tornem realidade (Instituto Alana; Ministério Público do Estado de São Paulo, 2022).

A 5Rights Foundation (2021) recomenda um processo de avaliação de risco em oito etapas, que serve como um guia prático para essa diligência:

> 1. Know your customer - who is it that you are impacting (in this case a child or children).
> 2. Map impact - interrogate the impact of your service, including the impact on underage children who should not be using it.
> 3. Gather evidence - this will be collected on a risk register that should be created through three lenses: risk, rights and safety-by-design.
> 4. Consult - in and outside your organisation. Solutions may come from surprising places including children themselves.
> 5. Assess, analyse and appraise - what you discover may be surprising or obvious and different risks are likely to require different mitigation strategies.
> 6. Recommend - this is your plan of what to do.
> 7. Publish and report - transparency gives confidence to users and regulators. It also provides learning for others and sets a bar for your organisation.
> 8. Monitor and review - digital products and services are rarely static. Small changes can have big impacts and constant vigilance and iteration is necessary (5Rights Foundation, 2021, p. 21-22).

Por fim, é crucial que os agentes de tratamento se abstenham da exploração econômica de crianças e adolescentes a partir de seus dados pessoais, garantindo ambientes digitais livres de publicidade segmentada ou comportamental dirigida a esse público (Instituto Alana; Ministério Público do Estado de São Paulo, 2022).

A responsabilidade de proteger as crianças deve ser deslocada delas para as empresas e governos, que devem priorizar o melhor interesse desse público em todas as suas práticas de coleta e processamento de dados (UNICEF, 2021b). A adoção dessas medidas

não é apenas uma questão de conformidade, mas um imperativo ético e um dever de cuidado para com os usuários mais vulneráveis do ecossistema digital.

7. Conclusão

O percurso argumentativo deste trabalho partiu da constatação de um paradoxo regulatório fundamental na era digital: a crescente e onipresente participação de crianças e adolescentes em ambientes online ocorre, majoritariamente, sob um manto de invisibilidade que neutraliza as salvaguardas legais destinadas a protegê-los.

Demonstrou-se que, embora a ANPD tenha construído um arcabouço interpretativo sofisticado, que flexibiliza o uso de bases legais para além do consentimento parental, a efetividade dessa abordagem depende de um pressuposto lógico e operacional que a prática de mercado tem sistematicamente ignorado: a capacidade de identificar a condição etária do titular dos dados.

A investigação das consequências dessa falha revelou um "efeito em cascata" destrutivo, que se espalha por todo o ecossistema digital. As crianças, ao não serem reconhecidas, são submetidas a um falso dilema entre a "adultificação" forçada e a exclusão desproporcional. Os agentes de tratamento, por sua vez, operam em um estado de permanente insegurança jurídica, incapazes de demonstrar a conformidade material de suas operações com o princípio do melhor interesse. Os pais e responsáveis perdem sua capacidade de mediação, e a própria ANPD se vê diante de um impasse regulatório, obrigada a um esforço de fiscalização reativo para um problema que deveria ser mitigado na origem.

A análise do panorama de soluções técnicas de *age assurance* demonstrou que, embora não exista uma "bala de prata", há um arsenal de ferramentas disponíveis, cujos dilemas de privacidade, segurança e inclusão podem e devem ser ponderados. Por fim, o estudo de caso do TikTok serviu como prova empírica da convergência regulatória global, que aponta para o fim da tolerância com a inércia dos controladores de dados.

Diante do exposto, este artigo reafirma sua tese central: a verificação de idade não deve ser interpretada como uma barreira ao acesso ou uma mera formalidade burocrática, mas como a pré-condição técnica e jurídica indispensável para a proteção efetiva de dados de crianças e adolescentes, como um ato de "identificar para proteger" que ativa o arcabouço normativo da LGPD, permitindo a "calibragem" das salvaguardas de forma proporcional e transformando o princípio do "melhor interesse" de um conceito abstrato em uma obrigação operacionalizável.

A solução não está na busca por um método único e universalmente aplicável, mas na construção de uma arquitetura de confiança, onde o rigor dos mecanismos de verificação escala de acordo com os riscos inerentes a cada atividade de tratamento.

Nesse sentido, a proposta de um *framework* de risco a ser desenvolvido e orientado pela ANPD, conforme detalhado no capítulo anterior, surge como o encaminhamento mais lúcido e eficaz para superar o impasse atual, uma vez que a criação de diretrizes claras, que estabeleçam critérios objetivos para a avaliação dos riscos de conteúdo, contato, conduta e contrato, oferece um caminho de segurança jurídica para os agentes de tratamento e de proteção efetiva para os titulares.

Tal regulação não apenas incentivaria a inovação responsável no mercado de *age assurance*, mas também capacitaria os controladores a adotarem, de forma defensável e transparente, as "camadas" de verificação mais adequadas para seus serviços, desde a simples autodeclaração em contextos de baixo risco até métodos mais robustos onde os direitos e liberdades infantojuvenis estejam sob maior ameaça. Esta é a contribuição mais significativa que este trabalho busca oferecer, ou seja, um modelo para que a regulação atue, simultaneamente, de maneira protetiva e viável.

Em última análise, o desafio que se impõe transcende a mera conformidade com a LGPD, uma vez que se trata de uma necessária mudança de paradigma na forma como concebemos e construímos o ambiente digital. É preciso migrar de uma "internet adulta por padrão", na qual as crianças são tratadas como “visitantes inesperados” ou como um “problema”, para uma "internet que reconhece e

se adapta às diferentes infâncias", incorporando os princípios de *privacy by design* e *safety by design* em sua essência.

A implementação de mecanismos de verificação etária calibrados e proporcionais não é, portanto, apenas uma obrigação legal, mas sim um imperativo ético e um passo civilizatório fundamental para a construção de um ecossistema digital que, desde sua concepção, seja verdadeiramente seguro, empoderador e digno dos seus usuários mais jovens. É a materialização do dever de cuidado que a sociedade, o Estado e o mercado devem, em conjunto e com prioridade absoluta, a todas as crianças e adolescentes.

Referências

5RIGHTS FOUNDATION. **But how do they know it's a child?** Age assurance in the digital world. [*S. l.*]: 5Rights Foundation, 2021. Disponível em: https://5rightsfoundation.com/resource/but-how-do-they-know-its-a-child/. Acesso em: 27 maio 2025.

BRASIL. Autoridade Nacional de Proteção de Dados. **Enunciado CD/ANPD nº 1**, de 22 de maio de 2023. Diário Oficial da União: seção 1, Brasília, DF, ed. 98, p. 129, 24 maio 2023. 2023a. Disponível em: https://www.in.gov.br/en/web/dou/-/enunciado-cd/anpd-n-1-de-22-de-maio-de-2023-485306934. Acesso em: 18 de maio de 2025.

BRASIL. Autoridade Nacional de Proteção de Dados. **Guia orientativo:** hipóteses legais de tratamento de dados pessoais – legítimo interesse. Brasília, DF: ANPD, 2024a. Disponível em: https://www.gov.br/anpd/pt-br/assuntos/noticias/anpd-lanca-guia-orientativo-sobre-legitimo-interesse. Acesso em: 27 maio 2025.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados Pessoais (LGPD). Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709compilado.htm. Acesso em: 18 maio 2025.

BRASIL. Autoridade Nacional de Proteção de Dados. **Nota Técnica nº 6/2023/CGF/ANPD**. [*S. l.*]: ANPD, 2023b. Disponível em: https://www.gov.br/anpd/pt-br/assuntos/noticias/anpd-abre-processo-sancionador-e-emite-determinacoes-ao-tiktok. Acesso em: 28 maio 2025.

BRASIL. Autoridade Nacional de Proteção de Dados. **Nota Técnica nº 50/2024/FIS/CGF/ANPD**. [*S. l.*]: ANPD, 2024b. Disponível em: https://www.gov.br/anpd/pt-br/assuntos/noticias/anpd-abre-processo-sancionador-e-emite-determinacoes-ao-tiktok. Acesso em: 28 maio 2025.

COMITÊ GESTOR DA INTERNET NO BRASIL. **Pesquisa sobre o uso da Internet por crianças e adolescentes no Brasil: TIC Kids Online Brasil 2024**. São Paulo: CGI.br, 2025. Disponível em: https://cetic.br/pt/publicacao/pesquisa-sobre-o-uso-da-internet-por-criancas-e-adolescentes-no-brasil-tic-kids-online-brasil-2024/. Acesso em: 29 maio 2025.

CRUZ, Sinuhe. **Relatório de Inteligência nº 5** – Proibido para menores de 18 anos: verificação de idade e proteção de dados de crianças e adolescentes. São Paulo: Data Privacy Brasil, 2024. Disponível em: https://clubedata.com.br/painel/biblioteca/exibir/1120/relatorio-de-inteligencia-5-proibido-para-menores-de-18-anos-verificacao-de-idade-e-protecao-de-dados-de-criancas-e-adolescentes. Acesso em: 28 maio 2025.

EUROPEAN DATA PROTECTION BOARD. **Binding Decision 2/2022 on the dispute arisen on the draft decision of the Irish Supervisory Authority regarding Meta Platforms Ireland Limited (Instagram) under Article 65(1)(a) GDPR**. [*S. l.*]: EDPB, 2022. Disponível em: https://www.edpb.europa.eu/our-work-tools/our-documents/binding-decision-board-art-65/binding-decision-22022-dispute-arisen_en. Acesso em: 18 maio 2025.

G20 BRASIL. **Diálogo G20 – Painéis temáticos**. São Paulo: G20 Brasil, 2024. 01 vídeo (209 min). Disponível em: https://www.youtube.com/live/_L3P-gwyrUgs. Acesso em: 01 de junho de 2025.

HENRIQUES, Isabella; RIBEIRO, Emanuella. Proteção de crianças e adolescentes na Internet: análise comparativa de redes sociais. In: COMITÊ GESTOR DA INTERNET NO BRASIL (Org.). **Pesquisa sobre o uso da Internet por crianças e adolescentes no Brasil:** TIC Kids Online Brasil 2024. São Paulo: CGI.br, 2025. p. 89-98. Disponível em: https://cetic.br/pt/publicacao/pesquisa-sobre-o-uso-da-internet-por-criancas-e-adolescentes-no-brasil-tic-kids-online-brasil-2024/. Acesso em: 29 maio 2025.

INSTITUTO ALANA; MINISTÉRIO PÚBLICO DO ESTADO DE SÃO PAULO. **Comentário Geral nº 25 sobre os direitos das crianças em relação ao ambiente digital** – Versão comentada. [*S. l.*]: Instituto Alana, 2022. Disponível em: https://criancaeconsumo.org.br/biblioteca/comentario-geral-no-25-comentado/. Acesso em: 01 de junho de 2025.

IRELAND. Data Protection Commission. **Final decision in case IN-21-9-1**: TikTok Technology Limited. Dublin: DPC, 2023. Disponível em: https://www.edpb.europa.eu/system/files/2023-09/final_decision_tiktok_in-21-9-1_-_redacted_8_september_2023.pdf. Acesso em: 29 de maio de 2025.

ROBLOX. **Age ID Verification**. Roblox Support, [2023?]. Disponível em: https://en.help.roblox.com/hc/en-us/articles/4407282410644-Age-ID-Verification. Acesso em: 01 de junho de 2025.

UNITED NATIONS. **General comment No. 25 (2021) on children's rights in relation to the digital environment.** Geneva: United Nations, 2021. Disponível em: https://www.right-to-education.org/resource/general-comment-no-25-2021-children-s-rights-relation-digital-environment. Acesso em: 26 maio 2025.

UNITED NATIONS CHILDREN'S FUND. **Digital age assurance tools and children's rights online across the globe:** a discussion paper. [*S. l.*]: UNICEF Office of Global Insight and Policy, 2021a. Disponível em: https://c-fam.org/wp-content/uploads/Digital-Age-Assurance-Tools-and-Childrens-Rights--Online-across-the-Globe.pdf. Acesso em: 01 de junho de 2025.

UNITED NATIONS CHILDREN'S FUND. **The case for better governance of children's data:** a manifesto. New York: Office of Global Insight and Policy, 2021b. Disponível em: https://www.unicef.org/innocenti/media/1036/file/UNICEF%20Global%20Insight%20Data%20Governance%20Summary.pdf. Acesso em: 29 maio 2025.

YUBO. **How Yubo Pioneered 100% Age Verification to Set a New Standard for Trust & Safety on Social Media** . Yubo Blog, [2024?]. Disponível em: https://www.yubo.live/blog/how-yubo-pioneered-100-percent-age-verification. Acesso em: 01 de junho de 2025.

YOTI. **On Facial Age Estimation, improvements and updates**. Yoti Blog, 17 jan. 2023. Disponível em: https://www.yoti.com/blog/on-facial-age-estimation-improvements-and-updates/. Acesso em: 01 de junho de 2025.

www.gov.br/anpd

ANPD

Agência
Nacional de
Proteção de Dados