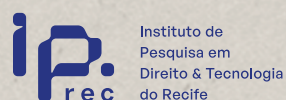


policy paper

IA E DIREITOS AUTORAIS:

EXPROPRIAÇÃO DE DIREITOS E A EXPLORAÇÃO DA CRIAÇÃO



Coordenação:

| André Lucas Fernandes
| Clarissa Mendes

Autores:

| Helton Leyendecker
| André Lucas Fernandes

Revisão:

| André Lucas Fernandes
| Raquel Saraiva
| Rhaiana Valois

INTRODUÇÃO

O uso massivo de obras protegidas por direitos autorais para treinar modelos de inteligência artificial (IA), sem permissão ou compensação aos autores, tornou-se prática disseminada e, hoje, dificilmente controlável. Isso vem consolidando um cenário de desrespeito às bases jurídicas e econômicas que sustentam a produção cultural e informacional: autores, artistas, jornalistas – criadores de conteúdo original passível de proteção autoral no geral –, têm seu trabalho explorado com **finalidades comerciais**, sem qualquer retorno ou reconhecimento.

No contexto nacional, a regulação da IA avança no PL 2338/23. Nessa toada, permitir um afrouxamento normativo que legitime a continuidade desse modelo predatório de apropriação do trabalho criativo significaria, na prática, **institucionalizar** tal cenário, acarretando consequências irreversíveis para o ecossistema cultural e informacional. A institucionalização desse modelo afeta não apenas a proteção autoral dos criadores e a saúde de todo o setor cultural, mas também o próprio ecossistema da Internet aberta na forma que conhecemos.

Por anos, motores de busca indexaram conteúdos e devolveram **tráfego** aos criadores. Essa dinâmica proporcionava **visibilidade** aos criadores através do direcionamento de pessoas aos conteúdos originais, sustentando – ainda que precariamente, já que nunca houve, nessa dinâmica, recompensa patrimonial direta aos criadores – um fluxo de criação, informação, cultura e livre acesso pelo usuário comum, embora com diversos problemas e críticas¹. Hoje, sistemas de IA coletam e recombina conteúdos para fornecer respostas diretas, desestimulando os usuários a consultarem as fontes originais, além de apresentarem alucinações e erros factuais em suas respostas².

O resultado é a perda de **reconhecimento** para quem cria, minando o único e frágil incentivo à produção de conhecimento, informação e cultura na rede e, em última análise, colocando em risco o ciclo, bastante argumentado na literatura, que permitiu a Internet aberta florescer.

Ainda assim, no Brasil, a regulação do uso de conteúdo protegido para treinamento de modelos de IA enfrenta resistência, alimentada pela crença de que impor contrapartidas

1 “Assim, parece no mínimo injusto que as plataformas e seus defensores estejam promovendo uma cultura do ‘gratuito’, enquanto as próprias plataformas lucram com esse conteúdo (e ainda por cima com as informações dos usuários, por meio da venda dos dados gerados sobre esses usuários)” (Tradução Livre) – O’CONNOR, Sean. Creators, Innovators, and Appropriation Mechanisms. George Mason Law Review, Vol. 22, No. 4, 2015. Disponível em: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2647052>. Acesso em: 25/07/2025.

2 **ESTADÃO**. Ferramentas de IA podem falhar em checagens e buscas; entenda. 2025. Disponível em: <<https://www.estadao.com.br/estadao-verifica/ferramentas-de-ia-podem-falhar-em-checagens-e-buscas-entenda/>>. Acesso em: 25/07/2025. Outras pesquisas e notícias chegam a relatar que mesmo sistemas considerados robustos e sucesso como produto fornecido no mercado informacional, tem apresentado uma piora a partir da implantação de sistemas de IA embutidos nestes produtos. Sobre isso, ver: “Why Google’s AI Overviews gets things wrong”. Disponível em: <<https://www.technologyreview.com/2024/05/31/1093019/why-are-googles-ai-overviews-results-so-bad/>>; “Google users are less likely to click on links when an AI summary appears in the results”. Disponível em: <<https://www.pewresearch.org/short-reads/2025/07/22/google-users-are-less-likely-to-click-on-links-when-an-ai-summary-appears-in-the-results/>>.

e limites a tal uso colocaria o país em desvantagem na “corrida” pelo desenvolvimento. É importante levar o argumento a sério, a partir de uma análise técnica, jurídica e socialmente orientada. Na prática, tratar regulação como obstáculo ao progresso é ignorar que **o avanço tecnológico não acontece no vácuo**: ele se apoia em setores estruturados por direitos, obrigações e salvaguardas. Se as normas que sustentam a produção cultural e a circulação de conteúdo nacional são negligenciadas, não há base para inovação duradoura – o que há é a entrega de patrimônio cultural e informacional a poucos agentes predatórios, numa profecia anunciada de canibalização da cultura.

O desenvolvimento e uso de sistemas de IA deve servir ao interesse público, de modo a fortalecer a **centralidade humana**³ e, por consequência, a sustentabilidade dos ecossistemas culturais e informacionais. Não se justifica uma corrida desenvolvimentista que se apoia na exploração do setor autoral nacional e de seu rico patrimônio cultural, enquanto devolve à sociedade resultados que não guardam a mínima proporcionalidade com o valor apropriado. Ainda que guardassem, isso não legitimaria a apropriação desautorizada, sem qualquer compensação ou reconhecimento, e nem a usurpação do controle que os autores deveriam ter sobre suas obras⁴, prerrogativas há muito estabelecidas e indissociáveis do ato criativo⁵ – os direitos morais e patrimoniais que lhes são assegurados⁶.

Cada dia sem diretrizes claras torna mais complexo, no futuro, reduzir as assimetrias econômicas e informacionais geradas por esse cenário de incertezas. Um marco regulatório sólido, além do que já foi exposto, cria condições para que empresas brasileiras se desenvolvam em diálogo com o setor cultural, alinhadas às tendências globais de sustentabilidade⁷. É esse mesmo marco que fortalece o posicionamento do país internacionalmente, dando lastro para pleitear maior proteção ao patrimônio cultural brasileiro, enfrentando a apropriação externa e sinalizando o compromisso do Brasil com um desenvolvimento de IA responsável, sustentável e alinhado ao interesse público.

3 Deriva do princípio da dignidade da pessoa humana (CF, art. 1º, III). Norteia o PL 2338/23, de acordo com sua ementa: “Dispõe sobre o desenvolvimento, o fomento e o uso ético e responsável da inteligência artificial com base na **centralidade da pessoa humana**.”

4 **Art. 24.** São direitos morais do autor:

I - o de reivindicar, a qualquer tempo, a autoria da obra;

II - o de ter seu nome, pseudônimo ou sinal convencional indicado ou anunciado, como sendo o do autor, na utilização de sua obra; [...]

IV - o de assegurar a integridade da obra, opondo-se a quaisquer modificações ou à prática de atos que, de qualquer forma, possam prejudicá-la ou atingi-lo, como autor, em sua reputação ou honra;

V - o de modificar a obra, antes ou depois de utilizada;

VI - o de retirar de circulação a obra ou de suspender qualquer forma de utilização já autorizada, quando a circulação ou utilização implicarem afronta à sua reputação e imagem; – Lei nº 9.610/98 (Lei de Direitos Autorais - LDA). BRASIL.

5 “A identidade pessoal atinge, pelo exercício da liberdade de descoberta e de invenção, a obra: a obra foi feita por **A**; **A** é autor da obra; **A** tem direito a que a obra se publique ou se exponha com alusão à sua pessoa; ninguém pode negar que **A** seja o autor, sem ofensa a esse direito absoluto. Direito de personalidade? Se o é, não se renuncia a êle, não se dispõe dêle, nem êle se transmite. [...] Com o fazer-se a obra, adquire-se o direito autoral de personalidade, o direito de ligar o nome à obra, o direito patrimonial de autor, ligado à disponibilidade do bem imaterial, separado da pessoa, e a propriedade da obra-coisa (pela species nova).” – MIRANDA, Francisco Cavalcanti. Tratado de Direito Privado, T. VII. RT, 2012. p. 213 e 215.

6 **Art. 22.** Pertencem ao autor os direitos morais e patrimoniais sobre a obra que criou. – LDA.

7 GEIGER, Cristophe. DI LAZZARO, Francesca. Sustainability as a Guiding Principle for Copyright Reform? Moral Rights Issues concerning the Use of Generative AI in the Field of Research and Education. Innovation Law and Ethics Observatory (ILEO), 2025. Disponível em: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5191571&download=yes>. Acesso em: 25/07/2025.

PROBLEMA

Dentre as etapas de desenvolvimento de sistemas de IA generativa, a de **treinamento** é crucial na avaliação da questão presente. Nessa etapa, modelos são expostos a conjuntos massivos de dados (datasets), nos quais se incluem, em larga escala, obras protegidas por direitos autorais – textos, imagens, músicas, vídeos – extraídas da Internet ou de bases privadas, não raro sem qualquer autorização ou compensação aos titulares⁸. Esses dados são a matéria-prima que torna o modelo funcional: sem os dados não há sobre o que um modelo operar (seja no treinamento, seja no resultado). Existe, portanto, um **nexo causal** direto entre a **utilização** dessas obras no treinamento da IA, sua transformação em produto e posterior **comercialização**.

Casos emblemáticos já foram documentados, como a utilização, sem licença, de coleções de **imagens** protegidas para treinar modelos de geração visual da Stability AI, que gerou imagens exibindo marcas d'água do Getty Images⁹; ou ainda **acervos musicais** completos incorporados aos datasets de empresas como Suno e Udio para alimentar modelos de geração sonora, prática que as próprias empresas reconheceram ao invocar a doutrina do fair use (uso justo) como justificativa¹⁰.

Também se destaca a apropriação de **conteúdo jornalístico** do The New York Times, cujos textos foram reproduzidos quase literalmente por sistemas como o ChatGPT, segundo alegações do próprio jornal em processo judicial¹¹. Houve ainda, na China, o caso em que uma IA gerou um personagem idêntico ao registrado como Ultraman sem deter os direitos, resultando em decisão que reconheceu a infração autoral¹².

Reiteramos: na forma como a tecnologia é concebida e implementada hoje, é o uso intensivo desses dados que permite ao modelo “aprender” e gerar resultados e, assim, tornar-se comercialmente atrativo. As empresas exploram esses modelos em produtos pagos ou incorporados a soluções corporativas, obtendo **receita**. Mesmo nos casos em que

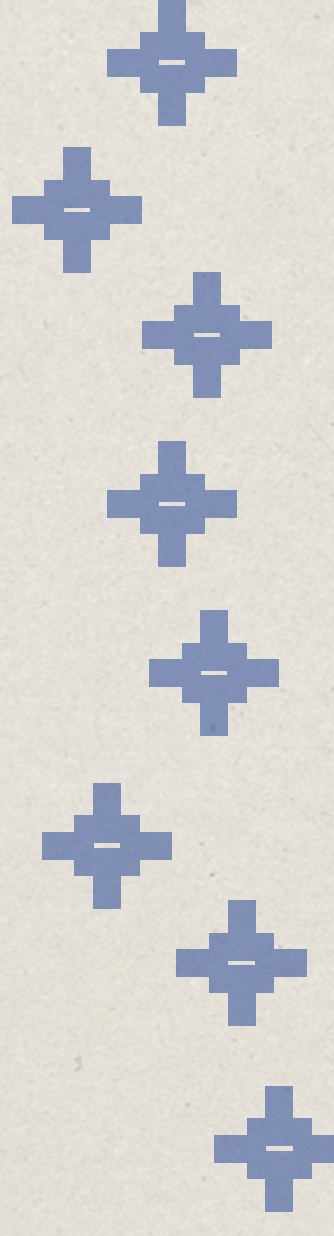
8 **WIRED.** Meta Secretly Trained Its AI on a Notorious Piracy Database, Newly Unredacted Court Docs Reveal. Disponível em: <<https://www.wired.com/story/new-documents-unredacted-meta-copyright-ai-lawsuit/>>. Acesso em: 25/07/2025.

9 **REUTERS.** Getty Images processa Stability AI por uso indevido de imagens. Disponível em: <<https://epocanegocios.globo.com/tecnologia/noticia/2023/02/getty-images-processa-stability-ai-por-uso-indevido-de-imagens.ghtml>>. Acesso em: 23/07/2025.

10 **THE VERGE.** AI music startups say copyright violation is just rock and roll. Disponível em: <<https://www.theverge.com/2024/8/2/24211842/ai-music-riaa-copyright-lawsuit-suno-udio-fair-use>>. Acesso em: 23/07/2025.

11 POPE, Audrey. NYT v. OpenAI: The Times's About-Face. Harvard Law Review, 2024. Disponível em: <https://harvardlawreview.org/blog/2024/04/nyt-v-openai-the-timess-about-face/>. Acesso em: 25/07/2025.

12 **FORBES.** China decide que empresa de IA cometeu violação de direitos autorais. Disponível em: <<https://www.forbes.com/sites/johannacostigan/2024/02/29/china-rules-ai-firm-committed-copyright-infringement/>>. Acesso em: 23/07/2025.



parte das ferramentas é oferecida gratuitamente (a chamada modalidade freemium), isso funciona como estratégia de mercado para ampliar a base de usuários e, consequentemente, aumentar a receita proveniente de assinaturas, licenças ou outros serviços¹³. Em síntese, a exploração do trabalho criativo é elemento indispensável para a existência – e lucro – desses modelos de negócio, restando configurado, nesse contexto, o **uso comercial** dessas obras. E esse é um ponto inelidível numa análise técnica estrita.

Apesar do argumento de que não haveria infração autoral porque “os dados não ficam estocados”, “o que resta do treinamento são tokens matemáticos”, para treinar um modelo é necessário criar e manter localmente cópias de obras protegidas, mesmo que venham a ser descartadas em etapa posterior¹⁴. O próprio U.S. Copyright Office reconhece, em sua mais recente publicação sobre o tema, que o processo de coleta e curadoria de dados para treinamento “claramente implica o direito de **reprodução**”¹⁵, pois envolve a criação de cópias locais de milhões ou bilhões de obras protegidas. Apesar dos indícios de que muitas empresas mantêm essas bases para usos futuros¹⁶, o fato de tais dados poderem ser descartados logo após o treinamento não altera a análise de infração, nem à luz da doutrina do fair use, nem nos termos da legislação brasileira de direitos autorais. Não há, na doutrina e ordenamento positivo do direito autoral, uma regra pré-excludente de juridicidade atrelada à deleção posterior, fora das hipóteses de adequação

13 Esse é um tema com ao menos uma década de debate no campo da economia de dados, plataformizada. EVANS, Peter C.; GAWER, Annabelle. The Rise of the Platform Enterprise: A Global Survey. The Center for Global Enterprise. Disponível em: <https://www.thecge.net/app/uploads/2016/01/PDF-WEB-Platform-Survey_01_12.pdf>. Acesso em: 25/07/2025.

14 SAMUELSON, Pamela; SAG, Mattheu; SPRINGMAN, Cristopher Jon. Comments In Response To The Copyright Office’s Notice Of Inquiry On Artificial Intelligence And Copyright. Publicado em resposta à Chamada para Contribuições sobre IA e Direitos Autorais do Escritório de Direitos Autorais dos Estados Unidos, 2024. p. 7. Disponível em: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4976391>. Acesso em: 25/07/2025.

15 “The steps required to produce a training dataset containing copyrighted works clearly implicate the right of reproduction.” – U.S. COPYRIGHT OFFICE. Copyright and Artificial Intelligence, Part 3: Generative AI Training. 2025. p. 26. Disponível em: <<https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-3-Generative-AI-Training-Report-Pre-Publication-Version.pdf>>. Acesso em: 25/07/2025.

16 “Além disso, relatos públicos indicam que grandes desenvolvedores frequentemente mantêm conjuntos de dados de treinamento para uso em projetos futuros.” (Tradução Livre) Id. Ibid. –p. 27

a um dos tipos de exceções.

A lei de direitos autorais brasileira (LDA – Lei nº 9.610/98) define **reprodução** como a cópia de uma obra “de qualquer forma tangível, incluindo qualquer armazenamento permanente ou temporário por meios eletrônicos”¹⁷. A LDA assevera que a utilização da obra por quaisquer modalidades depende de autorização prévia e expressa do autor, inclusive a reprodução parcial ou integral (art. 29, I). O art. 30 reforça que cabe ao titular decidir sobre a forma e as condições de exploração econômica da sua obra, excetuando situações de uso autorizado. Em suma, ao copiar obras para treinar modelos de IA sem autorização, resta frontalmente ofendido o direito de reprodução, ainda que o armazenamento seja temporário.

A LDA contempla, indiretamente, hipóteses de **exceções** (por vezes equiparáveis, sem total propriedade, à doutrina do fair use) em que a reprodução de obras não constitui ofensa aos direitos autorais, elencadas no art. 46. No direito brasileiro, são regras que apresentam a exceção, impedindo a incidência da norma geral sobre o tema como medida pela qual o ordenamento jurídico restringe a regra patrimonial básica, projetada dos direitos do autor, para que se contemple elementos públicos relevantes – no Brasil, com maior destaque, a proteção e acesso à cultura, acesso ao conhecimento e outros valores de nível constitucional.

Essas exceções abrangem situações como: reprodução de trechos para uso privado e **sem fins lucrativos** (incisos I, d; II; VI); citação para fins de estudo, crítica ou polêmica com **indicação de autoria e origem** (III); ou ainda, a reprodução de pequenos trechos de obras preexistentes, desde que não prejudique a exploração normal da obra reproduzida nem cause prejuízo aos legítimos interesses dos autores (VIII). Nenhuma dessas hipóteses abarca a utilização massiva de obras protegidas para treinamento de IA. Conclui-se que a LDA não admite a prática de coleta e uso indiscriminado de obras protegidas como insumo para modelos de IA que são explorados economicamente.

Não obstante os argumentos de que a lei autoral deveria ser ajustada para atender às novas realidades – questão que vem de muito antes do atual cenário da IA, destaque-se –, a LDA, como qualquer norma, é passível de aperfeiçoamentos para melhor atender aos legítimos interesses de seus titulares e da sociedade. Contudo, além de não ser este o escopo ao tratar de IA e autorais, qualquer proposta de alteração da referida lei deve ter como norte o fortalecimento da proteção autoral a partir do valor constitucional da proteção e acesso à cultura, ao conhecimento, e não o contrário. Ajustar a LDA para legitimar práticas de coleta predatória de dados sob o argumento de que esta é a inescapável “nova realidade” seria não apenas um retrocesso, mas uma inversão completa de sua finalidade, afrontando a própria razão de existir desse diploma normativo, tanto sob o seu tradicional aspecto privatista, como do insuficiente aspecto social.

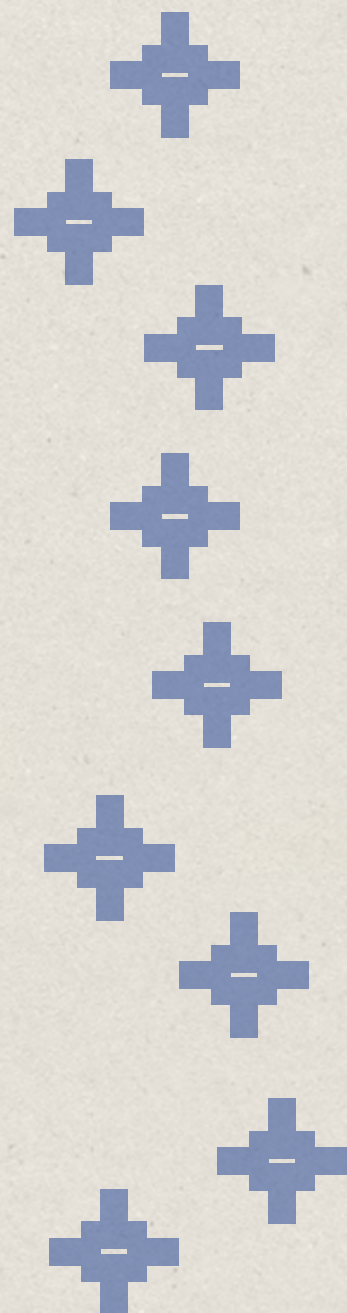
DESAFIOS

Reconhecer que criadores têm, indiscutivelmente, direitos morais e patrimoniais no tocante ao uso de suas obras em treinamentos de IA é o primeiro passo. O grande desafio, no entanto, é efetivar esse reconhecimento, o que demanda enfrentar questões complexas:

a) **Definir objetivamente a forma e extensão da remuneração ou compensação:** Estabelecer como a remuneração será calculada e distribuída é um desafio central. O relatório mais recente do U.S. Copyright Office aponta modelos debatidos internacionalmente, como: **licenciamento voluntário** (negociações diretas, altamente complexas diante do volume massivo de obras e da dificuldade de rastreamento); **licenciamento coletivo** (via entidades de gestão); **licenciamento obrigatório** (em que se paga valores fixados, sem autorização prévia); e **licenciamento coletivo estendido com possibilidade de opt-out**. Cada modelo carrega prós e contras, exigindo um equilíbrio entre viabilidade técnica, segurança jurídica e adaptação ao contexto nacional. Tais opções, embora relevantes, não esgotam as possibilidades de solução.

b) **Exigência de transparência e accountability:** Outro ponto crítico é garantir a transparência por parte dos desenvolvedores sobre suas bases de treinamento, sobretudo diante de práticas de raspagem de dados, que, na maioria das vezes, ocorrem de forma opaca, informal e em escala massiva. Instrumentos de accountability efetivos que permitam auditorias, verificação de conformidade e sanções proporcionais em caso de uso irregular precisam ir além de “diretrizes éticas”. É necessário estabelecer formas reais de efetividade (enforcement) na exigência de cumprimento das normas aplicáveis, assegurando que a coleta e o processamento respeitem os direitos dos titulares e as obrigações legais vigentes.

c) **Redução de assimetrias e estímulo a novas entrantes:** Há ainda o desafio de enfrentar as assimetrias estruturais criadas pela coleta massiva de dados por um número restrito de empresas, que lhes conferiu uma vantagem competitiva relevante. No entanto, esse cenário não significa uma vantagem incompensável: estudos mostram que, entre 2026 e 2032, todo



conteúdo textual humano na Internet terá sido raspado¹⁸, de modo que o próprio setor já alerta sobre riscos de degradação progressiva dos modelos devido a retroalimentação com dados sintéticos (model collapse), reconhecendo que o sucesso de um sistema de IA depende criticamente da **qualidade** dos dados¹⁹. A lógica que hoje dita as tendências de desenvolvimento no mundo é pautada pela sustentabilidade, com uma transição para práticas baseadas na qualidade como fator predominante à quantidade de dados²⁰. Nesse contexto, é chave fomentar um ambiente de dados sustentável, baseado na valorização e uso responsável de **conteúdos originais**, incentivando – inclusive por meio de políticas públicas e estímulos concretos – que novas entrantes adotem as melhores práticas alinhadas ao estado da arte dos novos paradigmas em desenvolvimento tecnológico a nível global.

Por fim, cada dia de inação é um dia em que criadores permanecem desamparados, sem compensação justa pelo uso de suas obras, enquanto assistem à exploração massiva de seu trabalho por modelos que não retribuem financeiramente, em reconhecimento, e tampouco devolvem um produto tecnológico eficiente e de ponta para a sociedade.

O Brasil tem chance de ser pioneiro na construção de um ecossistema que valorize a criação original e incentive inovações responsáveis, mas é necessário agir rápida e assertivamente com o fito de mitigar os riscos a que, atualmente, estamos expostos.

Além disso, o país é um ator fundamental, dado o seu amplo acervo de produções culturais, para propor alternativas aos modelos exploratórios do setor da cultura e do conhecimento, tudo em harmonia ao programa constitucional vigente.



18 VILLALOBOS, Pablo; et al. **Will we run out of data? Limits of LLM scaling based on human-generated data**, ARXIV, 2024. Disponível em: <<https://arxiv.org/abs/2211.04325>>. Acesso em: 25/07/2025.

19 “The success of an AI system is critically dependent on data quality”. **Data Quality Over Quantity: Pitfalls and Guidelines for Process Analytics**. <https://www.sciencedirect.com/science/article/pii/S2405896323013046>

20 “Recent research from major developers suggests that quality may even be a more important consideration than quantity.” U.S. Copyright Office Report Part. 3, Op. Cit. – p. 11

O cenário atual é de instabilidade e apresenta uma série de riscos, como: **(1)** a degradação técnica dos modelos de IA, dependentes de dados sintéticos – vulneráveis ao model collapse; **(2)** o estrangulamento da Internet aberta, com criadores bloqueando acesso a seus conteúdos para evitar raspagem indiscriminada, prejudicando também o acesso à informação por usuários comuns; e **(3)** o desmonte progressivo de setores criativos essenciais, como arte, jornalismo e literatura, desestimulados por um ambiente que ignora seus direitos.

A qualidade dos dados utilizados no treinamento de modelos de IA, e, por conseguinte, da entrega útil de valor (output) por estes modelos, está diretamente vinculada à originalidade de tais dados, em contraste com dados sintéticos, cuja utilização tem se expandido para suprir a escassez de conteúdo original disponível. Dados sintéticos são informações geradas artificialmente a partir de conteúdos já existentes, reproduzindo ou recombinao padrões aprendidos. Tal alternativa passou a ser empregada porque a quantidade massiva de dados exigida pelos modelos atuais pressiona um “repositório global” de conteúdos humanos que está muito próximo de ser totalmente raspado²¹.

Ao fim e ao cabo, insistir em uma narrativa que concebe o avanço dos modelos como dependente de quantidades exorbitantes de dados é um argumento pouco inovador, apesar da tentativa vã de narrar o contrário. E tal fato se deve a uma acomodação conservadora de tipos de modelos que não são aprimorados em sua internalidade, mas apenas em “poder bruto de processamento”. É uma receita falida, exauriente de recursos, reprodutora de modelos sem uma percepção do desafio ecológico do nosso tempo²².

Portanto, utilizar esse raciocínio como argumento para fomentar o desenvolvimento tecnológico é um contrassenso, já que, por essa lógica, estaríamos na iminência de atingir o limite da tecnologia, tendo em vista o risco concreto de esgotamento da base de dados textuais originais ainda não utilizados em treinamento já no ano de 2026, conforme

21 No limite, discute-se o risco de uma “internet morta” (Dead Internet Theory), conceito que descreve um ambiente digital progressivamente dominado por conteúdos gerados por IA, enquanto a participação humana se torna pontual e cada vez menos visível. Trata-se de uma hipótese antes especulativa, mas que, diante do avanço da IA generativa, ganha contornos concretos. Como explica Yoshija Walter em *Artificial Influencers and the Dead Internet Theory* (2024) [Disponível em: <<https://link.springer.com/article/10.1007/s00146-023-01857-0>>]. Acesso em: 25/07/2025], observa-se hoje, em primeira mão, um cenário em que a IA molda percepções e narrativas digitais, confundindo fronteiras entre dados humanos e sintéticos e comprometendo a integridade informacional necessária para interações digitais saudáveis.

22 Yann LeCun, um dos mais renomados cientistas do campo, tem dado declarações diretas sobre este problema, especialmente quanto à IA generativa. Meta’s AI chief says scaling AI won’t make it smarter. Business Insider. Disponível em: <<https://www.businessinsider.com/meta-yann-lecun-scaling-ai-wont-make-it-smarter-2025-4>>. Acesso em: 25/07/2025. Meta’s Yann LeCun predicts ‘new paradigm of AI architectures’ within 5 years. TechCrunch. Disponível em: <<https://techcrunch.com/2025/01/23/metas-yann-lecun-predicts-a-new-ai-architectures-paradigm-within-5-years-and-decade-of-robotics>>. Acesso em: 25/07/2025..

aponta Pablo Villalobos²³.

Nesse contexto, uma visão centrada na quantidade de dados impulsiona a adoção desarrazoada de soluções espúrias, como o uso indiscriminado de dados sintéticos. Esse processo, como já antecipado, contribui para a degradação progressiva dos modelos, fenômeno conhecido como model collapse, em que os sistemas passam a reproduzir cada vez mais conteúdo derivado de si mesmos, alimentando-se de seus próprios subprodutos sintéticos, num movimento de ruminação que deforma e corrompe progressivamente os resultados – com queda de qualidade e perda de diversidade já observada e descrita na literatura técnica²⁴. Além disso, o desempenho sustentável desses modelos depende de um fluxo constante de dados atualizados, já que sistemas de IA não produzem informação nova nem alimentam a web com eventos do mundo real, dependendo integralmente de **dados gerados por humanos em primeira mão** para manter a relevância e a precisão de seus outputs.

No cenário internacional, cresce o reconhecimento de que o modelo atual de extração irrestrita está corroendo o ecossistema da Internet aberta. Medidas como a da Cloudflare, que passou a adotar um sistema baseado em permissão para evitar a exploração de conteúdo sem contrapartida, ilustram essa preocupação²⁵. Iniciativas como o Signals da Creative Commons²⁶ também demonstram um movimento global de defesa da transparência e do controle sobre como os conteúdos são utilizados por sistemas de IA.

Nesse sentido, evidencia-se o risco do estrangulamento progressivo da rede enquanto fonte plural de conhecimento e informação acessível. O tráfego que dava visibilidade a sites e criadores está sendo erodido por mecanismos de busca e plataformas de intermediação que passaram a oferecer respostas sintéticas, diretamente na página de resultados, geradas por sistemas de IA integrados, afetando drasticamente o fluxo de audiência para os conteúdos originais.

23 “Nossa análise sugere que, se o rápido crescimento no tamanho dos conjuntos de dados continuar, os modelos poderão esgotar todo o suprimento de texto público produzido por humanos entre 2026 e 2032, ou até um ou dois anos antes, caso sejam treinados em excesso. A partir desse ponto, a disponibilidade de textos humanos públicos pode se tornar um fator limitante para a ampliação futura dos modelos de linguagem”. (Tradução Livre) – VILLALOBOS, Pablo; et al. Op. Cit. p. 9.

24 “Nossas descobertas sugerem que a quantidade de dados gerados deve ser consideravelmente menor em comparação com os dados originais para evitar o colapso do modelo.” (Tradução Livre) – SEDDIK, Mohamed, et al. How Bad is Training on Synthetic Data? A Statistical Analysis of Language Model Collapse. ARXIV, 2024. Disponível em: <<https://arxiv.org/pdf/2404.05090>>. Acesso em: 25/07/2025.

25 **CLOUDFLARE**. Cloudflare Just Changed How AI Crawlers Scrape the Internet-at-Large; Permission-Based Approach Makes Way for A New Business Model. Disponível em: <<https://www.cloudflare.com/pt-br/press-releases/2025/cloudflare-just-changed-how-ai-crawlers-scrape-the-internet-at-large/>>. Acesso em: 23/07/2025.

26 **CREATIVE COMMONS**. CC Signals: A New Social Contract for the Age of AI. Disponível em: <<https://creativecommons.org/ai-and-the-commons/cc-signals/>>. Acesso em: 23/07/2025.

Cerca de 60% das buscas já são consideradas zero-click²⁷, isto é, concluídas sem que o usuário visite qualquer site. À medida que os criadores sentem os efeitos da supressão da única contraparte²⁸ que recebiam por manter a Internet enquanto ecossistema de acesso ao conhecimento e circulação de informação, cresce o fechamento do acesso a conteúdos antes livres – como já ocorre em larga escala no jornalismo –, numa tentativa de frear a extração predatória de seus dados para alimentar sistemas que contribuem com a não efetivação de seus direitos de personalidade. Assim, é afetado também o usuário comum, que vê restringida a principal característica que fez a Internet florescer: o acesso livre à informação.

Para o setor cultural, o risco é o desmonte de um ecossistema que reconhece e estimula, ainda que com suas limitações, a criação original e os direitos autorais. Mesmo em casos em que a exploração econômica direta de uma obra não gera compensação significativa, o reconhecimento e a tutela jurídica do trabalho criativo humano têm valor estruturante para a cultura nacional.

Ignorar isso desestimula a produção autoral e contribui para um cenário em que o volume de conteúdos sintéticos tende a superar o de obras humanas, limitando o próprio desenvolvimento tecnológico e apagando progressivamente traços da cultura humana, num processo epistemicida²⁹ corrosivo que empobrece exponencialmente o patrimônio cultural.

27 Goodbye Clicks, Hello AI: Zero-Click Search Redefines Marketing. Disponível em: <<https://www.bain.com/insights/goodbye-clicks-hello-ai-zero-click-search-redefines-marketing/>>. Acesso em: 25/07/2025.

28 No caso, o reconhecimento enquanto efetivação de seus direitos morais, já que os patrimoniais, na prática, se diluíram sob a prerrogativa – que não passa isenta de críticas – do acesso ao conhecimento. Sobre isso, Sean O'Connor (Op. Cit., 2015, p. 911): “Nesse debate, duas reivindicações morais conflitantes precisam ser equilibradas: o direito dos criadores de exercer algum controle sobre suas criações, e o interesse público em não restringir conhecimentos importantes apenas a uma minoria privilegiada que possa pagar por eles. [...] Embora as empresas de busca e a iniciativa A2K não expliquem como os criadores deveriam se sustentar, tem-se a impressão de que acreditam que os criadores simplesmente encontrarão cargos assalariados ou apoio governamental”. (Tradução Livre)

29 FERNANDES, André. IA epistemicida e a continuidade do projeto colonial. IP.rec, 2025. Disponível em: <<https://ip.rec.br/blog/ia-epistemicida-e-a-continuidade-do-projeto-colonial/>>; Acesso em: 25/07/2025.

Há, nesse cenário, um argumento no mínimo questionável: promessas grandiloquentes de melhorias sociais tecnosolucionistas³⁰ são impostas; a cultura é depredada, sem opção de oposição; e os resultados entregues vão na contramão do prometido, sendo não só irrisórios como ofensivos frente ao valor da cultura nacional, no mais das vezes produzida a partir da criação popular.

O Brasil corre o risco de sacrificar um de seus setores mais ricos e relevantes – a cultura – em troca de uma suposta corrida tecnológica que, nos moldes postos, já está perdida e profetiza seu esgotamento. Mesmo uma abertura irrestrita não faria o país alcançar a infraestrutura e o poder de coleta das grandes empresas estrangeiras, que historicamente operam nesse modelo predatório. E mesmo essas jurisdições, como demonstra o recente relatório do U.S. Copyright Office, já expressam sérias preocupações sobre a legalidade e a ética dessas práticas.

30 MENDES, Clarissa. Soluções tecnológicas para problemas sociais? O que é tecnosolucionismo. IP.rec, 2025. Disponível em: <<https://ip.rec.br/blog/solucoes-tecnologicas-para-problemas-sociais-o-que-e-tecnosolucionismo/>>; Acesso em: 25/07/2025.

Ao contrário do discurso de que o Brasil “ficará para trás” caso imponha limites ao uso de dados protegidos no treinamento de IA, há uma constatação clara: a corrida que o país pode ganhar é a do pioneirismo em desenvolvimento sustentável. Essa é a tendência global no momento, e o Brasil – com seu histórico de legislação avançada em setores estratégicos – tem condições concretas de se colocar como referência ao estruturar um modelo de inovação que respeite direitos, valorize a cultura e garanta equilíbrio entre tecnologia e interesse público.

De acordo com a UNESCO³¹, a sustentabilidade envolve a realização de um conjunto complexo e dinâmico de objetivos em dimensões humanas, sociais, culturais, econômicas e ambientais. Tecnologias como a inteligência artificial podem contribuir para esses objetivos ou, ao contrário, comprometer sua realização, dependendo de como forem implementadas e reguladas. A avaliação contínua dos impactos da IA sobre essas dimensões é, portanto, condição indispensável para assegurar um desenvolvimento que dialogue com os Objetivos de Desenvolvimento Sustentável (ODS) e não apenas com métricas de crescimento econômico.

Nesse sentido, a regulação do uso de obras no treinamento de IA contribui de forma ativa para a sustentabilidade, pois se articula com todas essas dimensões – humanas, culturais, sociais, econômicas e ambientais. Ao deslocar o foco do volume indiscriminado de dados para a qualidade dos dados utilizados, reduz-se inclusive a necessidade de expansão agressiva de datacenters, mitigando impactos ambientais, ao mesmo tempo em que se promove um modelo de inovação mais contido, responsável e coerente com as diretrizes internacionais de desenvolvimento sustentável. O tema, como se vê, exige uma análise holística complexa – tarefa que, em boa medida, o PL 2338/2023 tenta desempenhar.

Essa visão ampliada de sustentabilidade vem sendo endossada em fóruns internacionais recentes. Na Declaração dos Líderes do G20 (Rio de Janeiro, 2024), a cultura é reconhecida como fator essencial para promover um mundo mais sustentável em todas as suas dimensões, fomentando solidariedade, diálogo e cooperação. O documento também destaca a necessidade de um debate global mais robusto sobre direitos autorais e direitos conexos no ambiente digital, diante dos impactos da IA sobre os detentores de direitos, e conclama o fortalecimento de políticas que garantam condições dignas e remuneração justa aos trabalhadores da cultura e da economia criativa.³²

No mesmo sentido, o BRICS incluiu, expressamente, a cultura como parte da solução para os desafios contemporâneos e vem construindo consenso sobre a necessidade

31 **UNESCO.** Recommendation on the Ethics of Artificial Intelligence. 2023. Disponível em: <<https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence>>. Acesso em 24/07/2025.

32 **G20.** Declaração de Líderes do Rio de Janeiro. 2024. Disponível em: <<https://www.gov.br/planalto/pt-br/acompanhe-o-planalto/g20-rio-de-janeiro-leaders-em-portugues.pdf>>. Acesso em: 24/07/2025.

de mecanismos regulatórios que assegurem a prevalência dos direitos autorais na era da IA. Sob a presidência brasileira, o bloco tem priorizado a criação de um catálogo de melhores práticas e o desenvolvimento de parâmetros internacionais que promovam uma governança inclusiva e responsável da IA, garantindo remuneração justa e equitativa aos criadores e reconhecendo a cultura como vetor de mudança social, ambiental e econômica.³³

Até mesmo jurisdições historicamente permissivas, como os Estados Unidos, vêm revisitando o tema. O próprio Copyright Office, em seu mais recente relatório técnico, reconheceu que a coleta massiva de obras protegidas para treinamento de IA implica níveis de ofensa a direitos autorais e que as práticas atuais das empresas não se encaixam no fair use tradicional. A repercussão política foi imediata, com a demissão da então diretora pelo presidente Trump³⁴ – um gesto interpretado como resposta às pressões de alguns atores e aos interesses estratégicos norte-americanos, inclusive em detrimento da soberania de outros países.

O Brasil, no contexto do PL 2338/23, tem a oportunidade de se posicionar internacionalmente como referência em desenvolvimento verdadeiramente sustentável, defendendo a primazia da cultura e da centralidade humana frente a agendas desenvolvimentistas rasas, desvinculadas de critérios éticos e de responsabilidade social, que, além de não contribuírem significativamente com o avanço da tecnologia, drenam o patrimônio cultural e ambiental, num movimento que afasta a inovação da realidade das necessidades sociais.

33 **BRICS.** Direitos autorais na era digital com IA é prioridade da Cultura do BRICS. 2025. Disponível em: <<https://brics.br/pt-br/noticias/direitos-autorais-na-era-digital-com-ia-e-prioridade-da-cultura-do-brics>>. Acesso em: 24/07/2025.

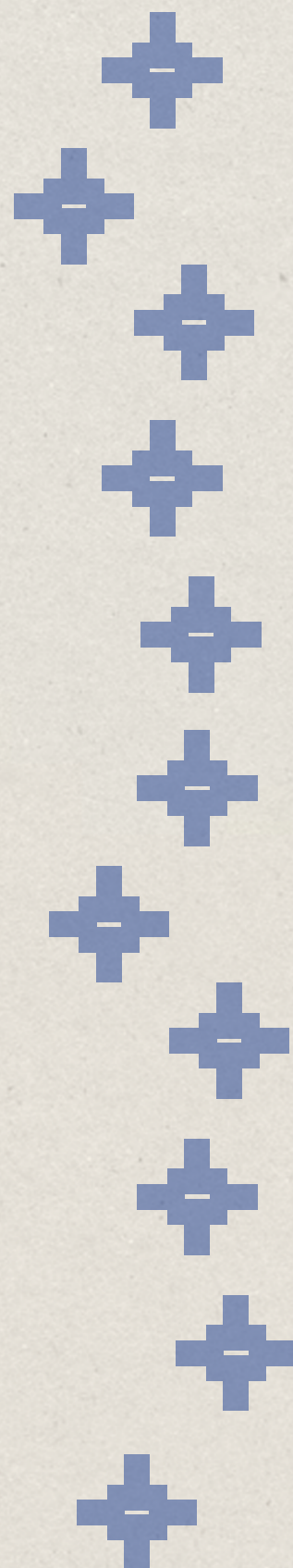
34 **PODER 360.** Trump demite chefe do escritório de Direitos Autorais. 2025. Disponível em: <<https://www.poder360.com.br/poder-tech/trump-demite-chefe-do-escritorio-de-direitos-autorais/>>. Acesso em: 25/07/2025.

SUGESTÕES AO PROJETO DE LEI 2338/23

O PL 2338/23 dedica a Seção IV (arts. 62 a 66) aos Direitos de Autor e Conexos, estabelecendo salvaguardas específicas para o uso de obras protegidas no desenvolvimento de sistemas de IA.

O **art. 62** estabelece que o desenvolvedor de IA que utilizar conteúdo protegido por direitos autorais deve informar publicamente sobre os conteúdos empregados, mediante a publicação de um sumário em sítio eletrônico de fácil acesso, respeitados os segredos comercial e industrial. Essa previsão de **transparência ativa** é um passo importante para reduzir a opacidade típica da coleta de dados em larga escala. Contudo, o dispositivo dependerá de **regulamentação** para especificar o nível de detalhamento exigido no sumário, a frequência de atualização, os mecanismos de contestação pelos titulares e critérios objetivos para invocar segredo comercial.

O **art. 63** cria uma exceção específica para a mineração de textos e dados feita por instituições científicas, educacionais, museus, arquivos públicos e bibliotecas, desde que o acesso seja lícito, não haja finalidade comercial e o uso seja limitado ao necessário, sem prejuízo aos interesses econômicos dos titulares. O texto também veda a exibição ou disseminação das obras utilizadas e determina que as cópias sejam armazenadas apenas pelo tempo estritamente necessário. Trata-se de um alinhamento com exceções internacionais de text and data mining (TDM), mas que exige mecanismos de auditoria e fiscalização para evitar desvio de finalidade e uso indireto em contextos comerciais. É importante ressaltar, entretanto, que a redação proposta no parágrafo quarto pode dar fundamento à prática deletéria conhecida como “copyright trolls”, que instrumentaliza o direito autoral em favor de uma restrição ao acesso à cultura, informação e conhecimento. Além disso, o parágrafo terceiro da minuta sob avaliação da Câmara dos Deputados requer uma hermenêutica sistemática, especialmente com o fito de evitar a criação de um cenário no qual atores com maior poder econômico façam uso de atores menores com vistas a criar uma situação de “exceção” (testa de ferro), para apropriar a exceção proposta numa economia espúria e anticoncorrencial. O texto pode e deve avançar com um novo comando sobre a necessidade de impor a tais processos de mineração e aos modelos treinados a partir deles, albergados pela exceção, um regime de licenciamento aberto e destinação pública-social – seja no desenvolvimento científico nacional, seja na incorporação em novas soluções de



mercado, estas não protegidas pelo manto excepcional.

O **art. 64** reforça a prerrogativa do titular de proibir a utilização de suas obras em contextos não abarcados pelo artigo anterior. Essa previsão concretiza o direito moral do autor, mas o texto não especifica instrumentos práticos para o exercício desse direito, como a existência de um registro centralizado ou mecanismos de notificação padronizados. É conhecido que tal exercício, no Brasil, se desdobra em processos de notificação extrajudicial e, em situações de recalcitrância, com a devida ação judicial. O tema está suficientemente tratado no texto do PL 2338/2023, mas exigirá um amadurecimento em processos regulamentares posteriores, ou em debates mais amplos de reforma da LDA.

O **art. 65** é um dos dispositivos mais centrais, ao determinar que o agente de IA que utilizar conteúdos protegidos em mineração, treinamento ou desenvolvimento deve remunerar os titulares dessas obras. O texto assegura:

(i) a possibilidade de negociação coletiva ou individual, inclusive gratuita ou onerosa,

(ii) a consideração de critérios de razoabilidade e proporcionalidade na fixação de valores, levando em conta porte da empresa e impactos concorrenciais,

(iii) a livre negociação para promover pesquisa e inovação, sem restringir a pactuação entre as partes.

Essa previsão abre espaço para modelos de remuneração como licenciamento coletivo, licenciamento estendido com opt-out ou acordos diretos, mas o PL ainda não detalha como operacionalizar essa remuneração, como rastrear o uso de obras específicas e como calcular valores proporcionais ao uso efetivo. No texto há, ainda, um dilema ou omissão quanto à “cadeia de transmissão” do aspecto patrimonial do direito autoral, de modo que atores com maior poder econômico (especialmente grandes atores estrangeiros), através do processo de cessão, passam a explorar uma remuneração que deveria valorizar a figura do criador originário. É preciso avançar neste debate, privilegiando e garantindo, por meios jurídicos e técnicos, a remuneração da figura do criador na economia da IA, concretizando o fim da norma autoralista.

Por fim, o **art. 66** garante que a utilização de conteúdos de imagem, áudio, voz ou vídeo que retratem ou identifiquem pessoas naturais respeite os direitos da personalidade, conforme o Código Civil e legislação correlata, trazendo para a discussão o alinhamento com proteção de dados e integridade individual. Apesar de importante, tal dispositivo estaria mais bem posicionado nas seções I – “Dos Direitos da Pessoa ou Grupo Afetado por Sistema de IA” e II – “Dos Direitos da Pessoa ou Grupo Afetado por Sistema de IA de Alto Risco”, no Capítulo II do texto do PL.

Em conjunto, os dispositivos representam um esforço de harmonização entre inovação tecnológica e salvaguarda de direitos autorais. Ainda assim, alguns pontos merecem atenção, especialmente no tangente à **regulamentação** específica: a necessidade de detalhar a implementação da transparência prevista no art. 62; a criação de mecanismos efetivos para o exercício do direito de exclusão do art. 64; e o

estabelecimento de diretrizes claras e auditáveis para a remuneração prevista no art. 65. A aplicabilidade desses instrumentos serão determinantes para que o PL cumpra seu objetivo de fomentar o desenvolvimento ético e responsável da IA, com base na centralidade da pessoa humana e na valorização da cultura e do patrimônio informacional brasileiros.

Como medida complementar, seria recomendável incluir no projeto um dispositivo que estimule a formação de bases de dados públicas para treinamento e pesquisa em IA, compostas por obras em domínio público, conteúdos protegidos cujos titulares optem por licenciamento voluntário e materiais disponibilizados sob licenças abertas. Devem constar, também, incentivos para ampliar, sanear e estruturar o acervo disponível, como programas de estímulo ao licenciamento aberto e até mesmo a discussão, em sede própria, sobre a redução progressiva do prazo geral de proteção autoral em determinadas categorias de obras, com o fito de fortalecer o acesso ao patrimônio cultural e cultivar um ambiente de dados próprios sustentável, equilibrado e propício à inovação duradoura.



TABELA DE SUGESTÕES

PL 2338/2023 - ARTS. 62 a 66

	TEXTO ORIGINAL	SUGESTÃO DE EMENDA
Art. 62	O desenvolvedor de IA que utilizar conteúdo protegido por direitos de autor e conexos deverá informar sobre os conteúdos protegidos utilizados nos processos de desenvolvimento dos sistemas de IA, por meio da publicação de sumário em sítio eletrônico de fácil acesso, observados os segredos comercial e industrial, nos termos de regulamento específico. Parágrafo único. Para fins deste Capítulo, o desenvolvimento compreende as etapas de mineração, treinamento, retreinamento, testagem, validação e aplicação de sistemas de IA.	Mantido original
Art. 63	Não constitui ofensa aos direitos de autor e conexos a utilização automatizada de conteúdos protegidos em processos de mineração de textos e dados para os fins de pesquisa e desenvolvimento de sistemas de IA por organizações e instituições científicas, de pesquisa e educacionais, museus, arquivos públicos e bibliotecas, desde que observadas as seguintes condições: I – o acesso tenha se dado de forma lícita; II – não tenha fins comerciais; III – a utilização de conteúdos protegidos por direitos de autor e conexos seja feita na medida necessária para o objetivo a ser alcançado, sem prejuízo dos interesses econômicos dos titulares e sem concorrência com a exploração normal das obras e conteúdos protegidos. § 1º Cópias de conteúdos protegidos por direitos de autor e conexos utilizadas nos sistemas de IA deverão ser armazenadas em condições de segurança, e unicamente pelo tempo necessário para a realização da atividade ou para a finalidade específica de verificação dos resultados. § 2º É vedada a exibição ou a disseminação das obras e conteúdos protegidos por direitos de autor e conexos utilizados no desenvolvimento de sistemas de IA. § 3º Este artigo não se aplica a instituições vinculadas, coligadas ou controladas por entidade com fins lucrativos que forneçam sistemas de IA ou que tenham, entre elas, participação acionária. § 4º Aplica-se o disposto no caput deste artigo à mineração de dados, por entidades públicas ou privadas, no contexto de sistemas de IA para combate a ilícitos civis e criminais, que atentem contra direitos de autor e conexos.	[... Inclusão] IV – os processos de mineração e os modelos treinados a partir deles, albergados pela presente exceção, sejam submetidos a regime de licenciamento aberto e destinação pública ou social, nos termos do regulamento. [... Inclusão] § 3º Este artigo não se aplica a instituições vinculadas, coligadas ou controladas por entidade com fins lucrativos que forneçam sistemas de IA ou que tenham, entre elas, participação acionária, nem a instituições que atuem como interpostas ou representantes de tais entidades, com o objetivo de se beneficiar indevidamente da exceção prevista neste artigo. [... Inclusão] § 4º Aplica-se o disposto no caput deste artigo à mineração de dados, por entidades públicas ou privadas, no contexto de sistemas de IA para combate a ilícitos civis e criminais, que atentem contra direitos de autor e conexos, vedada, entretanto, a utilização da presente disposição como fundamento para práticas abusivas ou de caráter exclusivamente repressivo que restrinjam o acesso legítimo à cultura, informação e conhecimento.

Art. 64	O titular de direitos de autor e conexos poderá proibir a utilização dos conteúdos de sua titularidade no desenvolvimento de sistemas de IA nas hipóteses não contempladas pelo art. 63 desta Lei. Parágrafo único. A proibição do uso de obras e conteúdos protegidos nas bases de dados de um sistema de IA posterior ao processo de treinamento não exime o agente de IA de responder por perdas e danos morais e materiais, nos termos da legislação aplicável.	[...] §2º A forma e os procedimentos pelos quais o titular poderá efetivar a proibição deverão ser especificados em regulamentação posterior.
Art. 65	O agente de IA que utilizar conteúdos protegidos por direitos de autor e conexos em processos de mineração, treinamento ou desenvolvimento de sistemas de IA deve remunerar os titulares desses conteúdos em virtude dessa utilização, devendo-se assegurar: I – que os titulares de direitos de autor e conexos tenham condições efetivas de negociar coletivamente, nos termos do Título VI da Lei nº 9.610, de 19 de fevereiro de 1998 (Lei dos Direitos Autorais), ou diretamente a utilização dos conteúdos dos quais são titulares, podendo fazê-lo de forma gratuita ou onerosa; II – que o cálculo da remuneração a que se refere o caput considere os princípios da razoabilidade e da proporcionalidade e elementos relevantes, tais como o porte do agente de IA e os efeitos concorrenciais dos resultados em relação aos conteúdos originais utilizados; III – a livre negociação na utilização dos conteúdos protegidos, visando à promoção de ambiente de pesquisa e experimentação que possibilite o desenvolvimento de práticas inovadoras, e que não restrinjam a liberdade de pactuação entre as partes envolvidas, nos termos dos arts. 156, 157, 421, 422, 478 e 479 da Lei nº 10.406, de 10 de janeiro de 2002 (Código Civil), e o art. 4º da Lei nº 9.610, de 19 de fevereiro de 1998 (Lei dos Direitos Autorais). § 1º A remuneração a que se refere o caput deste artigo é devida somente: I – aos titulares de direitos de autor e conexos nacionais ou estrangeiros domiciliados no Brasil; II – a pessoas domiciliadas em país que assegure a reciprocidade na proteção, em termos equivalentes a este artigo, aos direitos de autor e conexos de brasileiros, conforme disposto nos arts. 2º, parágrafo único, e 97, § 4º, da Lei nº 9.610, de 19 de fevereiro de 1998 (Lei dos Direitos Autorais), sendo vedada a cobrança nos casos em que a reciprocidade não estiver assegurada. § 2º O titular do direito de remuneração previsto no caput que optar pela negociação e autorização direta, nos termos do inciso I do caput, poderá exercê-lo independentemente de regulamentação posterior.	[... caput, I, II... modificação de redação original] III - a livre negociação na utilização dos conteúdos protegidos, com vistas à promoção de ambiente de pesquisa e experimentação que possibilite o desenvolvimento de práticas inovadoras, respeitados os direitos dos titulares e os princípios da boa-fé e do equilíbrio contratual. [... inclusão de inciso] IV - a adoção de modelos coletivos de licenciamento, inclusive com cláusulas de opt-out, conforme regulamentação específica, de modo a facilitar o cumprimento das obrigações e garantir eficiência na remuneração dos titulares. [... inclusão de parágrafos] § 3º Nos casos de cessão total ou parcial dos direitos patrimoniais, a remuneração prevista no caput deverá contemplar, sempre que possível, o autor originário da obra utilizada, de forma proporcional e razoável, especialmente quando se tratar de cessão imposta por desequilíbrio contratual ou posição dominante de mercado. § 4º A futura autoridade reguladora regulamentará guias acerca dos meios técnicos e jurídicos para rastreamento da utilização de obras protegidas e para cálculo proporcional da remuneração devida, observando critérios de transparência, auditabilidade e participação da sociedade.
Art. 66	A utilização de conteúdos de imagem, áudio, voz ou vídeo que retratem ou identifiquem pessoas naturais pelos sistemas de IA deverá respeitar os direitos da personalidade, na forma prevista na Lei nº 10.406, de 10 de janeiro de 2002 (Código Civil), e na legislação pertinente.	Mantido original. Sugere-se realocar para o TÍTULO II – DOS DIREITOS.

Art. 66-A	[criação de novo artigo]	[criação de novo artigo] Art. 66-A. O Poder Público, através, prioritariamente de iniciativas da autoridade coordenadora do SIA e autoridades setoriais, fomentará a criação e a manutenção de bases de dados públicas destinadas ao treinamento e à pesquisa em inteligência artificial, compostas por: I – obras em domínio público; II - conteúdos protegidos por direito autoral e conexos, cujos titulares tenham optado, expressamente, por licenciamento voluntário; III – materiais disponibilizados sob licenças abertas, nos termos da legislação aplicável. §1º O Poder Público deverá estimular, através da criação de programas e incentivos, a ampliação do acervo disponível nas bases de dados públicas; §2º As políticas de fomento previstas no caput poderão incluir a destinação de recursos públicos, incentivos fiscais, programas de estímulo à abertura de acervos e ações voltadas à digitalização, estruturação, saneamento e interoperabilidade das bases; §3º O Poder Público promoverá processos de escuta e participação pública e multissetorial, para fundamentar estudos e propor medidas para a redução progressiva do prazo geral de proteção autoral em determinadas categorias de obras, a serem especificadas em regulamentação posterior, com vistas a fortalecer o acesso ao patrimônio cultural e à promoção da sustentabilidade informacional no país.
-------------------------	---------------------------------	--

Este policy paper faz parte do projeto “Desafios Emergentes da Inteligência Artificial”, com um financiamento inicial da Google Brasil e, atualmente, sem financiamento específico.



Instituto de
Pesquisa em
Direito & Tecnologia
do Recife

