A UTILIZAÇÃO DE OBRAS PROTEGIDAS POR DIREITOS AUTORAIS PARA TREINAMENTO DE INTELIGÊNCIAS ARTIFICIAIS GENERATIVAS: Desafios jurídicos e possibilidades regulatórias

Aluna: Ariadne Oliveira Felix

Orientadora: Maria Edelvacy Marinho

Apoio: PIVIC Mackenzie

Resumo

Este artigo tem como finalidade analisar se o uso de obras protegidas por direitos autorais no treinamento de Inteligências Artificiais Generativas (IAg) configura violação aos direitos autorais ou pode ser enquadrado nas limitações legais existentes, bem como examinar as responsabilidades dos desenvolvedores e as possibilidades de regulamentação da matéria. A pesquisa adota metodologia qualitativa, fundamentada em revisão bibliográfica, análise legislativa e doutrinária e estudo comparado de modelos nacionais e internacionais de proteção autoral e regulação da inteligência artificial. Os resultados indicam que o atual regime jurídico brasileiro apresenta limitações para disciplinar o uso de obras protegidas no treinamento de IAs, especialmente diante da ausência de critérios específicos sobre autorização, transparência, responsabilização e remuneração dos titulares. Verificou-se, ainda, que as diferenças entre os modelos regulatórios internacionais, como o sistema de *fair use* norte-americano e as disposições europeias sobre mineração de textos e dados, evidenciam a complexidade da matéria. Também foram identificados obstáculos técnicos relacionados à impossibilidade de "desaprender" dados já incorporados aos modelos. Conclui-se que a utilização dessas obras demanda maior segurança jurídica e mecanismos regulatórios específicos, capazes de conciliar o desenvolvimento tecnológico e científico com a proteção dos direitos autorais, mediante critérios de transparência, responsabilização e compensação aos criadores, garantindo equilíbrio entre inovação, interesses econômicos e valorização da produção intelectual.

Palavras-chave: Inteligência Artificial. Direitos autorais. Regulação Normativa.

Abstract

This article aims to analyze whether the use of copyrighted works in the training of Generative Artificial Intelligence (GenAI) systems constitutes copyright infringement or may fall within existing legal limitations, as well as to examine the responsibilities of developers

and the possibilities for regulating the matter. The research adopts a qualitative methodology based on a literature review, legislative and doctrinal analysis, and a comparative study of national and international models of copyright protection and artificial intelligence regulation. The results indicate that the current Brazilian legal framework faces limitations in regulating the use of copyrighted works for AI training, particularly due to the absence of specific criteria regarding authorization, transparency, liability, and remuneration of rights holders. The study also found that differences among international regulatory models, such as the U.S. *fair use* system and European provisions on text and data mining, highlight the complexity of the issue. Technical obstacles related to the impossibility of “unlearning” data already incorporated into AI models were also identified. It is concluded that the use of such works requires greater legal certainty and specific regulatory mechanisms capable of reconciling technological and scientific development with copyright protection through criteria of transparency, accountability, and compensation for creators, thereby ensuring a balance between innovation, economic interests, and the recognition of intellectual production.

Keywords: Artificial intelligence. Copyright. Regulation.

Introdução

A Inteligência Artificial Generativa configura-se como uma das inovações tecnológicas mais relevantes do século XXI, reconfigurando a experiência humana e tornando-se onipresente (Kissinger, 2023). Modelos como o ChatGPT (OpenAI), o MidJourney e o Gemini transformam setores criativos e corporativos, revelando-se capazes de produzir textos, músicas, artes e vídeos com qualidade assemelhada à humana.

Entretanto, a aparente criatividade dessas ferramentas esconde um dilema jurídico muito atual, o uso massivo de obras protegidas por direitos autorais no processo de treinamento. O que antes era o fruto do trabalho humano, agora se converte em dado a ser processado por algoritmos, sem autorização dos titulares. Assim, essa prática evidencia uma assimetria, já conhecida, entre grandes corporações, que lucram com os sistemas, e os artistas, que permanecem invisibilizados. (Giblin; Doctorow, 2024)

Este artigo tem como objetivo geral analisar se o uso de obras protegidas no treinamento de IAs generativas configura violação ao direito autoral, à luz da legislação brasileira, ou pode ser enquadrado como hipótese legítima de limitação ou uso transformativo, bem como discutir as implicações jurídicas para desenvolvedores e autores. Para alcançar esse objetivo primeiro iremos analisar os fundamentos do direito autoral e sua aplicação ao uso de obras no treinamento de IAs generativas; em seguida, examinaremos os principais desafios jurídicos e técnicos dessa utilização, bem como as perspectivas internacionais; por fim, discutiremos possíveis alternativas regulatórias para o Brasil e suas implicações para autores e desenvolvedores.

1. O Direito Autoral como Instrumento de Proteção da Criatividade Humana

O direito autoral constitui um dos pilares fundamentais da propriedade intelectual e desempenha papel essencial na preservação da criatividade humana. Sua função é dupla e complementar: de um lado, protege o vínculo pessoal entre o autor e sua obra, assegurando os direitos morais; de outro, garante ao criador o aproveitamento econômico de sua produção intelectual, garantindo os direitos patrimoniais (Barbosa, 2003). Essa estrutura visa equilibrar o reconhecimento da autoria com a justa remuneração pelo trabalho criativo, promovendo a continuidade e a diversidade da produção cultural.

No plano internacional, a Convenção de Berna (1886), ratificada pelo Brasil, estabelece que qualquer forma de reprodução, adaptação ou transformação de uma obra depende de autorização prévia do autor (Brasil, 1975). Nesse contexto, o conceito de obra protegida é compreendido por “todas as produções do domínio literário, científico e artístico,

qualquer que seja o modo ou a forma de expressão", incluindo, entre outras, obras literárias, musicais, cinematográficas, fotográficas e artísticas (Brasil, 1975).

Nesse contexto, é importante distinguir os dois principais sistemas de proteção autoral: o *droit d'auteur*, de tradição continental, que atribui especial relevância à relação entre o autor e sua criação e à proteção dos direitos morais (Souza; Castro, 2024) e o *copyright*, de tradição anglo-americana, que historicamente enfatiza os direitos de exploração econômica da obra (WIPO, 2026).

Sob essa perspectiva, o uso de obras de autoria humana, para o treinamento de sistemas de IA, sem consentimento, contraria os princípios centrais do sistema internacional de proteção autoral e enfraquece o equilíbrio entre incentivo à criação e respeito à titularidade intelectual.

Desse modo, o direito autoral não se limita à função econômica de assegurar remuneração justa ao criador, ele também se constitui como um instrumento ético e jurídico de proteção à originalidade e à individualidade criativa. Ao impedir a apropriação indevida de expressões intelectuais para fins não autorizados, o sistema autoral preserva o valor simbólico da autoria.

A evolução histórica do direito autoral demonstra uma tensão permanente entre a proteção dos interesses dos autores e titulares e o interesse público no acesso à cultura e ao conhecimento. O próprio sistema autoral busca estabelecer esse equilíbrio, garantindo direitos exclusivos aos criadores, mas também admitindo limitações e exceções destinadas a preservar o acesso, a educação, a pesquisa e a circulação do conhecimento (WIPO, 2026). Na sociedade informacional, essa tensão se intensifica com a desmaterialização das obras e com o enfraquecimento da centralidade do autor nas relações de exploração econômica (Ascensão, 2022). No contexto das IAs generativas, o problema assume nova dimensão, pois obras produzidas por autores humanos podem ser coletadas e convertidas em dados para o treinamento de sistemas desenvolvidos e explorados economicamente por grandes empresas, muitas vezes sem participação ou remuneração dos criadores. Isso demonstra que, embora as novas tecnologias ampliem a circulação do conhecimento, não podem ser usadas de justificativa para uma erosão de bases éticas, que asseguram o valor simbólico da autoria.

No ordenamento jurídico brasileiro, a Lei nº 9.610/1998 (Lei de Direitos Autorais) estabelece que a utilização de obras protegidas depende de autorização prévia e expressa do autor, salvo nas hipóteses de limitações legais previstas no artigo 46. Entre essas exceções, incluem-se os usos para fins de citação, ensino e pesquisa, desde que não haja

finalidade lucrativa. Tal estrutura normativa tem como propósito viabilizar o acesso à cultura e ao conhecimento, sem contudo, comprometer os direitos fundamentais do criador. Essa lógica exclui práticas empresariais, como o treinamento de IAs generativas, cujo objetivo final é, invariavelmente, a comercialização.

O avanço das tecnologias de inteligência artificial generativa (IAg) apresenta novos desafios interpretativos à aplicação do direito autoral. O treinamento de IAs generativas envolve, frequentemente, o uso massivo de obras intelectuais, como textos, músicas e imagens, com finalidade comercial, o que extrapola as hipóteses legais de limitação ao direito do autor.

O grande desafio contemporâneo consiste, portanto, em compatibilizar o regime jurídico autoral, com um ambiente tecnológico marcado pela coleta massiva e automatizada de dados criativos. Essa conciliação exige reflexões normativas e éticas profundas, capazes de garantir que o avanço da inteligência artificial não se dê às custas da erosão da criatividade humana e da substituição do autor pelo algoritmo.

2. A Cultura da Pirataria e a invisibilidade do Autor

Com a popularização da internet, disseminou-se uma falsa percepção de que todo conteúdo disponível em meio digital pertence automaticamente ao domínio público (Santos, 2011). Essa crença equivocada enfraquece o reconhecimento da autoria e contribui para a consolidação de uma cultura de desvalorização dos direitos autorais. Nesse contexto, tornou-se comum a prática de apropriação quase automática de obras, como textos, imagens, músicas e demais criações intelectuais, que circulam livremente nas plataformas digitais sem a devida identificação ou respeito aos seus criadores.

A internet, portanto, não apenas ampliou o acesso à informação, mas também contribuiu para a naturalização do uso indevido de obras protegidas, diante da equivocada percepção de que conteúdos disponíveis na rede seriam automaticamente de uso livre. No mesmo sentido, o Superior Tribunal de Justiça reconhece que a disponibilidade de uma obra na internet não afasta a proteção autoral nem autoriza sua utilização sem consentimento do titular (Brasil, 2020).

A proteção autoral, conforme previsto na Lei n.º 9.610/1998 (Brasil, 1998), pressupõe o consentimento prévio do autor para qualquer forma de exploração econômica de sua obra. Nesse sentido, a ausência de transparência quanto às fontes de dados e critérios de filtragem utilizados no treinamento das IAg, torna o cenário preocupante. O criador humano

torna-se invisível, e sua produção intelectual é reduzida a mera informação estatística e matemática empregada na geração de novos conteúdos (Samuelson, 2023).

Embora as empresas de IAg aleguem finalidade legítima, segundo a doutrina norte-americana do fair use, e algumas organizações, como o jornal The New York Times, tenham adotado mecanismos de bloqueio, para impedir a captação de seus conteúdos específicos na base de dados do Chat GPT (OpenAI, 2024), a apropriação de obras sem autorização afasta-se do conceito jurídico de uso justo. Ademais, ações corretivas posteriores, como exclusões ou compensações, não eliminam os impactos negativos sofridos pelos autores cujas obras foram utilizadas indevidamente.

A doutrina do fair use, uso justo, reconhece que o requisito de originalidade das obras possui caráter relativo, mas ainda assim é essencial para a tutela autoral. Segundo Bittar (2018), as criações derivadas ou transformativas exigem elementos de criatividade e inovação para se caracterizarem como novas expressões intelectuais. No entanto, quando empresas de IA incorporam massivamente obras preexistentes sem adição substancial de originalidade, os outputs resultantes carecem de novidade essencial, aproximando-se de meras reproduções estilísticas.

Como observa Metz (2024), ex-pesquisador da Open AI, os modelos generativos baseados em grandes volumes de dados online tendem a replicar padrões estéticos e linguísticos existentes de dados online, o que contribui para um processo de esvaziamento da autoria. Esse fenômeno não se restringe à violação patrimonial dos direitos dos criadores, mas também gera impactos éticos e culturais profundos. A sistemática substituição do autor humano pela máquina produz uma tendência à homogeneização cultural, na qual a diversidade criativa é gradualmente substituída por reproduções padronizadas de estilos e ideias já existentes.

3. A Utilização do Estilo do Studio Ghibli por Sistemas Generativos de IA

No último ano, em 2025, o caso envolvendo o Studio Ghibli ganhou grande repercussão nos debates sobre inteligência artificial, quando modelos generativos passaram a reproduzir o estilo visual característico do estúdio (De Tilia, 2025). A controvérsia ganhou força em março de 2025, quando a OpenAI disponibilizou novos recursos de geração de imagens no ChatGPT. Usuários passaram a transformar fotografias, memes e acontecimentos públicos em imagens semelhantes às animações do Studio Ghibli, tendo a própria OpenAI registrado uma demanda excepcional pela ferramenta (Zeff, 2025).

Nesse contexto, a reprodução do traço Ghibli por sistemas generativos tornou-se um exemplo emblemático das tensões entre inovação tecnológica e proteção da autoria. A utilização de deep learning (aprendizado profundo), demonstra que esses modelos, para buscar aperfeiçoar seus resultados, devem absorver diretamente elementos característicos das animações, como texturas, paletas de cores e traços específicos, por meio do uso de imagens e vídeos extraídos das obras originais (Kottapalli; Kundargi, 2023). O episódio provocou intenso debate jurídico, especialmente sobre a origem da capacidade do sistema de reproduzir características tão específicas. Embora não exista confirmação pública de que obras do Studio Ghibli tenham sido individualmente utilizadas no treinamento da ferramenta, surgiu o questionamento sobre a possibilidade de imagens e outros conteúdos protegidos integrarem os conjuntos de dados utilizados para desenvolver o modelo.

Apesar do caso não ter sido judicializado, em outubro de 2025, a Content Overseas Distribution Association (CODA), entidade japonesa de proteção e distribuição de conteúdos, encaminhou solicitação formal à OpenAI sobre o Sora 2, modelo de criação de vídeos de IA, afirmando ter identificado conteúdos semelhantes a obras japonesas e questionando a utilização desses materiais no treinamento do sistema. A entidade sustentou que a reprodução de obras protegidas durante o processo de aprendizado poderia configurar violação autoral e também criticou o sistema de opt-out, defendendo a necessidade de autorização prévia para a utilização das obras no sistema japonês (CODA, 2025).

O caso evidencia, portanto, que o debate não se restringe à capacidade da IA de imitar determinado estilo, mas envolve a origem dos dados utilizados no treinamento, a autorização dos titulares, a reprodução de elementos protegidos e a exploração econômica dos resultados gerados. Dessa forma, o fenômeno do “estilo Ghibli” constitui exemplo concreto da dificuldade de aplicação dos conceitos tradicionais de direito autoral às novas formas de criação algorítmica.

4. O Dilema da “Impossibilidade de Desaprender”

Estudos recentes têm demonstrado que a remoção de dados específicos, já incorporados a modelos de inteligência artificial, constitui um desafio técnico de extrema complexidade, quando não se trata de uma tarefa inviável. Pois, uma vez assimilados os padrões de dados durante o treinamento, o modelo de IA não é capaz de simplesmente “desaprender” essas informações sem comprometer a integridade estrutural e o desempenho do sistema. (Cooper et al. 2025). Assim, a exclusão pontual de obras utilizadas de forma irregular mostra-se tecnicamente impraticável, dada a natureza interdependente dos modelos de aprendizado profundo, utilizados em inteligências artificiais generativas.

De modo análogo, não há garantias de que um dado eliminado não venha a ser posteriormente reinserido no sistema por meio de novas interações de usuários Ademais, é possível que o próprio modelo, ao recombinar padrões já internalizados, gere resultados substancialmente semelhantes ao conteúdo anteriormente excluído (Cooper et al. 2025). Esse fenômeno evidencia as limitações técnicas atuais no controle da base de dados de IAg, no que se refere ao processo de “esquecimento” algorítmico.

Todavia, a existência desse desafio técnico não deve ser utilizado como justificativa para a ausência de regulação jurídica. A inviabilidade técnica de desaprender não pode ser utilizada como argumento para legitimar violações de direitos autorais ou outras infrações normativas. Ao contrário, deveria servir como impulso para o desenvolvimento de soluções regulatórias preventivas, com mecanismos como o licenciamento obrigatório e a filtração dos dados desde as etapas iniciais de treinamento e inserção.

Outro aspecto relevante refere-se à participação direta dos usuários na alimentação e evolução dos modelos de IA. Diversos sistemas permitem que conteúdos autorais, incluindo obras completas, trechos de livros ou composições artísticas, sejam inseridos manualmente pelos próprios usuários. Esse processo altera significativamente os resultados gerados e contribui para a expansão não supervisionada da base de dados do modelo.

Nessas circunstâncias, a responsabilidade pode tornar-se difusa, uma vez que tanto desenvolvedores quanto usuários participam da utilização de dados protegidos. Contudo, sob a perspectiva da teoria do risco do empreendimento (Tartuce, 2024), as empresas desenvolvedoras, por explorarem economicamente sistemas de IAs Generativas e deterem maior controle sobre seu funcionamento, devem responder pelos riscos previsíveis decorrentes de sua atividade. No âmbito dos direitos autorais, isso significa que, embora o usuário possa ser responsável pela inserção ou solicitação de determinado conteúdo protegido, a empresa também pode ser responsabilizada quando tiver participação relevante na disponibilização, processamento ou reprodução da obra e não adotar mecanismos adequados para prevenir ou cessar violações. Assim, a responsabilidade deve considerar o grau de controle da empresa, a previsibilidade do dano e sua capacidade de prevenção.

5. Perspectivas Contrárias

Outra perspectiva sustenta que os instrumentos já previstos no ordenamento jurídico podem, em determinadas hipóteses, ser utilizados para solucionar os conflitos decorrentes do treinamento de modelos de IA, sem que seja necessária a criação imediata de um novo regime autoral específico (Morimoto; Prado, 2025; Brasil, 2022). Parte-se da ideia de que a legislação atual, aplicada caso a caso, seria capaz de coibir abusos e que novas regulações

poderiam impor entraves desnecessários à inovação. Contudo, essa confiança nas ferramentas existentes desconsidera que elas foram elaboradas para um contexto anterior, no qual as violações ocorriam de forma localizada, identificável e em menor escala.

Outro argumento contrário à regulação sustenta que, por trabalharem com dados matemáticos e não com cópias literais, as obras utilizadas para treinar modelos de IAs não configurariam reprodução protegida pelo direito autoral (Sobel, 2017), o que afastaria a necessidade de novos marcos legais. Essa visão, porém, reduz a discussão a um aspecto técnico e ignora os efeitos concretos da apropriação massiva de obras protegidas.

Diante de modelos generativos que utilizam conjuntos de dados incomparavelmente maiores e diversos, confiar apenas nessa interpretação restrita amplia o risco de aprofundar a invisibilização autoral e fortalecer assimetrias entre criadores e grandes plataformas tecnológicas. Assim, longe de constituir barreira à inovação, marcos regulatórios atualizados tornam-se instrumentos essenciais para equilibrar interesses, transparência ao processo de treinamento e assegurar uma proteção efetiva aos autores na era da inteligência artificial.

6. Perspectivas Comparadas de Modelos Internacionais

A ausência de uniformidade internacional na regulação da inteligência artificial evidencia as limitações dos sistemas jurídicos tradicionais, ainda estruturados sob o princípio territorialista da lex loci protectionis, segundo o qual a proteção autoral é determinada pela legislação do país em que se pleiteia tutela. Tal lógica revela-se progressivamente tensionada diante da natureza transnacional das tecnologias de IA, que operam simultaneamente em múltiplas jurisdições e processam dados provenientes de diferentes territórios.

A própria Organização Mundial da Propriedade Intelectual (OMPI) reconhece que a IA generativa possui potencial para provocar significativa “destruição criativa” (creative destruction), ao transformar setores e modelos de produção, mas ressalta a necessidade de avaliar seus impactos sobre os criadores humanos e sobre o sistema de propriedade intelectual (OMPI, 2024). Nesse contexto, a organização destaca dificuldades relacionadas à proteção das obras utilizadas no treinamento, transparência, atribuição e compensação dos titulares, evidenciando a necessidade de soluções capazes de conciliar inovação tecnológica e proteção autoral.

Nesse cenário, a fragmentação normativa expõe uma limitação do modelo clássico de tutela autoral, cuja arquitetura foi concebida para relações jurídicas territorialmente delimitadas, e não para fluxos digitais globais e descentralizados (Medina, 2024). Como

consequência, o tratamento jurídico do uso de obras protegidas no treinamento de sistemas de inteligência artificial generativa passou a variar significativamente entre ordenamentos, refletindo distintas concepções acerca do equilíbrio entre proteção da criatividade, estímulo à inovação e liberdade econômica.

A análise comparada dessas experiências revela não apenas divergências normativas, mas também diferentes compreensões sobre o papel do Estado, do mercado e do próprio autor na economia informacional contemporânea.

6.1 Estados Unidos: A Centralidade do Fair Use

No contexto norte-americano, o principal fundamento jurídico invocado pelas empresas de tecnologia é a doutrina do fair use, prevista na seção 107 do Copyright Act. Esse instituto admite determinados usos de obras protegidas sem autorização prévia do titular, desde que considerados transformativos e avaliados segundo critérios como finalidade, natureza da obra, quantidade utilizada e impacto no mercado.

As empresas desenvolvedoras de sistemas de inteligência artificial generativa sustentam que o treinamento de seus modelos não implica reprodução literal das obras originais, mas sim a extração de padrões estatísticos e estilísticos que permitam a geração de novos conteúdos. Assim, defendem que o uso seria transformativo e, portanto, juridicamente admissível.

Contudo, essa interpretação tem sido crescentemente questionada. Observa-se o aumento de litígios envolvendo grandes corporações do setor, como OpenAI e Microsoft, processadas por artistas, escritores e outros profissionais criativos sob a alegação de uso indevido de obras protegidas no treinamento de modelos algorítmicos (CNN Brasil, 2025; Forbes Brasil, 2023).

Tais disputas demonstram que o enquadramento automático do treinamento como uso transformativo não constitui consenso consolidado, especialmente quando se examina o potencial impacto econômico sobre os titulares. Desse modo, o sistema estadunidense revela forte dependência da atuação judicial e da interpretação casuística do fair use, o que pode gerar significativo grau de incerteza jurídica.

Embora os defensores do fair use argumentem que o treinamento de modelos de IAg possui caráter transformativo, tal interpretação ainda carece de consenso jurídico e ético. Mesmo usos considerados transformativos podem reproduzir estruturas criativas reconhecíveis, gerando resultados que preservam semelhanças substanciais com obras

preexistentes, o que intensifica o debate quanto aos limites éticos e jurídicos do treinamento automatizado (Cooper et al., 2025). Essa possibilidade levanta preocupações quanto ao mercado criativo e à redução da visibilidade dos autores humanos, cujas obras serviram de base para o treinamento dos modelos.

6.2 União Europeia: Regulação Preventiva e Transparência

Em contraste com o modelo norte-americano, a União Europeia adota uma abordagem mais restritiva e preventiva. O AI Act (Regulamento n.º 2024/1689) estabeleceu obrigações de transparência quanto às bases de dados utilizadas no treinamento de sistemas de IA, assegurando a rastreabilidade das informações e a observância da legislação de direitos autorais (União Europeia, 2024). O artigo 53 do referido regulamento reforça essa diretriz ao determinar a necessidade de licenciamento prévio para a utilização de obras protegidas, representando, assim, um avanço normativo substancial em relação ao modelo norte-americano no que se refere à proteção dos criadores.

Além disso, vêm sendo implementados mecanismos de licenciamento coletivo estendido (ECL), que asseguram compensação justa aos titulares de direitos (Strowel, 2011). O debate internacional sobre a utilização de obras protegidas por sistemas de IA ultrapassa a dimensão meramente jurídica, da legalidade formal. A questão central não se limita a determinar se há ou não violação de direitos autorais, mas abrange uma reflexão mais ampla acerca de como a reprodução de padrões criativos humanos por máquinas pode impactar o valor, a originalidade e a diversidade das produções culturais.

6.3 Japão: Ampliação das Exceções para Análise de Dados

Para além das abordagens ocidentais, o Japão promoveu, em 2018, reforma relevante em sua Lei de Direitos Autorais, introduzindo o artigo 30-4, que permite a utilização de obras para fins de análise de dados, inclusive com finalidade comercial, desde que não haja fruição expressiva do conteúdo original (Japan, 2018).

A medida foi apresentada como estratégia de fortalecimento da competitividade tecnológica e estímulo à inovação no campo da inteligência artificial. Contudo, a ampliação das exceções pode fragilizar o controle econômico do autor sobre sua obra, sobretudo quando o uso ocorre de forma massiva por grandes empresas privadas. O modelo japonês, portanto, evidencia opção regulatória mais permissiva, orientada prioritariamente ao desenvolvimento tecnológico.

6.4 China: Governança Estatal e Responsabilização dos Provedores

A China, por sua vez, adotou abordagem caracterizada por maior intervenção estatal. Em 2023, a Administração do Ciberespaço da China publicou as Interim Measures for the Management of Generative Artificial Intelligence Services, estabelecendo que os provedores de serviços de IA devem assegurar que os dados utilizados no treinamento estejam em conformidade com a legislação de propriedade intelectual e assumam responsabilidade por eventuais violações (China, 2023).

O modelo chinês combina incentivo estratégico ao desenvolvimento tecnológico com forte controle administrativo, impondo deveres de verificação, supervisão e responsabilização das plataformas. Trata-se de estrutura regulatória que reforça o papel do Estado como agente central na governança da inovação.

6.5 Síntese Comparativa e Reflexões para o Brasil

A análise comparada evidencia que não há solução uniforme para o tratamento jurídico do treinamento de inteligências artificiais generativas. Enquanto os Estados Unidos privilegiam flexibilidade interpretativa e protagonismo judicial, a União Europeia aposta em regulação preventiva e transparência; o Japão adota postura mais permissiva voltada ao estímulo tecnológico; e a China combina inovação com controle estatal robusto.

Diante desse panorama, o Brasil encontra-se em posição estratégica para estruturar modelo próprio, capaz de harmonizar proteção constitucional da autoria, segurança jurídica e incentivo à inovação. A construção de solução equilibrada exige diálogo entre direito autoral, regulação tecnológica e princípios constitucionais, evitando tanto a desregulação excessiva quanto a rigidez normativa que possa comprometer o desenvolvimento científico e econômico nacional.

Essa preocupação também já se manifesta no âmbito institucional brasileiro. O Guia de IA Generativa, elaborado pela Secretaria de Governo Digital e pelo Serviço Federal de Processamento de Dados (SERPRO), reconhece que muitas ferramentas de IAG são treinadas com dados que podem estar protegidos por direitos autorais e propriedade intelectual, ressaltando que a legalidade da coleta e do tratamento desses dados não é necessariamente garantida. O documento recomenda, ainda, que o uso dessas ferramentas seja precedido de uma avaliação rigorosa quanto à origem e à natureza dos dados, em conformidade com a legislação aplicável e com as boas práticas institucionais (Brasil, 2025). Dessa forma, observa-se que a preocupação com a proteção autoral e com os riscos jurídicos decorrentes da utilização de dados no desenvolvimento de sistemas de IA não se restringe ao debate acadêmico ou às experiências regulatórias estrangeiras, encontrando também respaldo nas orientações oficiais do próprio Governo Federal brasileiro.

7. Propostas de Regulamentação para o Brasil

Diante do cenário internacional e dos desafios decorrentes do uso de obras protegidas em treinamentos de inteligência artificial generativa, torna-se urgente que o Brasil avance na formulação de um marco regulatório próprio, capaz de conciliar inovação tecnológica e proteção dos direitos autorais.

Uma primeira medida fundamental seria a atualização da Lei de Direitos Autorais (Lei nº 9.610/1998), de modo a incluir dispositivos específicos relacionados à utilização de obras por sistemas de IA. Essa atualização legislativa deveria estabelecer que qualquer uso de obras protegidas para fins de treinamento ou geração de conteúdo por IAg dependa de autorização ou licenciamento prévio do titular dos direitos, em conformidade com os princípios da Convenção de Berna (Brasil, 1975). Tal adequação garantiria maior segurança jurídica e reforçaria o respeito à propriedade intelectual no contexto digital contemporâneo.

Além disso, propõe-se a criação de mecanismos de compensação financeira inspirados em modelos internacionais, especialmente o europeu. A experiência da União Europeia com o Extended Collective Licensing (ECL) apresenta uma alternativa para lidar com a complexidade dos direitos autorais na era da inteligência artificial. O ECL permite que entidades de gestão coletiva negociem licenças com empresas de tecnologia, estendendo os efeitos desses acordos a todos os titulares de direitos, inclusive àqueles que não participam diretamente da negociação (Strowel, 2011).

Esse modelo oferece duplo benefício, assegura remuneração ampla e proporcional aos autores e facilita o uso legítimo de grandes volumes de obras para fins tecnológicos e científicos, sem violações legais. A adoção de um sistema semelhante no Brasil poderia equilibrar a necessidade de acesso a dados, essencial para o avanço de IAs como um todo, com o reconhecimento econômico e moral dos autores e detentores de direitos.

No contexto nacional, observa-se que já há movimentos legislativos relevantes nesse sentido. O Projeto de Lei nº 2.338/2023, atualmente em tramitação, propõe diretrizes para a governança responsável de sistemas de inteligência artificial e inclui dispositivos voltados à proteção dos direitos autorais e conexos. O texto prevê, entre outras medidas, a remuneração justa e proporcional pela utilização de obras protegidas nos processos de treinamento e desenvolvimento de modelos de IAs.

Em síntese, a regulamentação brasileira deve buscar um modelo híbrido, que promova tanto a inovação e o desenvolvimento tecnológico quanto a valorização da autoria e da diversidade cultural. A adoção de práticas baseadas em licenciamento coletivo,

transparência nos dados de treinamento e responsabilização é um passo essencial para assegurar um futuro digital mais equitativo e juridicamente sólido.

8. Modelos Híbridos de Remuneração e Gestão Coletiva

Uma alternativa promissora para conciliar o avanço tecnológico com a efetiva proteção dos direitos autorais é a adoção de modelos híbridos de remuneração e gestão coletiva, que permitam o uso ético e sustentável de obras no contexto da IA (Lemos, 2023). Esses modelos preveem a constituição de fundos coletivos de redistribuição entre os criadores, financiados por empresas desenvolvedoras de IA. Trata-se de uma solução intermediária que não inviabiliza o progresso tecnológico, ao mesmo tempo em que garante compensação econômica justa aos titulares de direitos.

Juntamente à experiências como o licenciamento coletivo estendido europeu, esse modelo assegura remuneração proporcional e transparente, mesmo quando as obras são utilizadas de forma automatizada e em larga escala. A implementação de mecanismos híbridos promove previsibilidade nas relações entre criadores e empresas, reduz o risco de exploração indevida e contribui para um ecossistema cultural equilibrado e sustentável.

A implementação de modelos híbridos desse tipo promove maior transparência e previsibilidade nas relações entre criadores e empresas de tecnologia. Além disso, garante remuneração proporcional ao uso efetivo das obras e mitiga os impactos econômicos negativos que a utilização massiva de conteúdo autoral pode gerar sobre os criadores.

Outro benefício relevante consiste na redução do risco de exploração indevida de obras e na formação de um ecossistema cultural mais equilibrado e sustentável. Ao redistribuir os valores arrecadados por meio de fundos coletivos de compensação, assegura-se que os autores sejam beneficiados de maneira ampla e equitativa, mesmo em contextos de uso automatizado e em larga escala.

Desse modo, os modelos híbridos de remuneração e gestão coletiva configuram uma via inovadora para harmonizar os interesses da indústria tecnológica e da comunidade criativa. Eles permitem que o desenvolvimento da inteligência artificial ocorra de forma ética e juridicamente segura, preservando os valores centrais do direito autoral, a proteção da criatividade, a justa remuneração e o reconhecimento da autoria, sem frear o potencial inovador das novas tecnologias.

9. Implicações da Utilização de Obras Autorais

9.1 Implicações Éticas

O uso de obras protegidas para treinar IAs gera dilemas éticos complexos, principalmente relacionados à invisibilização do autor. Ao transformar a criatividade humana em dados para processamento algorítmico, há uma redução da autoria a insumo técnico, desconsiderando o vínculo moral entre autor e obra (Brasil, 1998; Barbosa, 2003).

Além disso, a prática de utilizar massivamente conteúdos sem consentimento pode levar à homogeneização cultural, na qual os modelos de IA reproduzem padrões dominantes. Isso impacta negativamente a inovação artística e a pluralidade cultural, pois obras originais e minoritárias tendem a ser menos representadas no conjunto de dados.

Outro ponto ético relevante é a transparência pois, os usuários finais e os criadores, frequentemente, não têm conhecimento de como seus conteúdos são processados, nem de que forma contribuem para o treinamento da IAg. A ausência de transparência impede que autores tomem decisões informadas sobre a utilização de suas obras e limita o controle sobre seu uso econômico e moral, desrespeitando princípios fundamentais do direito autoral.

9.2 Implicações Econômicas

Do ponto de vista econômico, o uso não autorizado de obras cria um desequilíbrio entre empresas de tecnologia e criadores. Corporações que se apropriam de conteúdos autorais sem remuneração direta concentram lucros e reduzem incentivos financeiros à produção cultural, configurando externalidades negativas para o mercado criativo (Lemos, 2023). Essa prática pode afetar editoras, músicos, escritores e outros setores artísticos que dependem da exploração econômica de suas obras. A reprodução de padrões já existentes por IAs diminui a necessidade de investimento em novas produções originais, prejudicando tanto o incentivo à criação quanto a diversidade de oferta cultural.

Sob uma perspectiva mais ampla, também pode-se compreender como uma forma de concorrência desleal. O artigo 195, da Lei nº 9.279/1996, (Brasil, 1996) define como ato de concorrência desleal toda conduta que vise obter vantagem econômica indevida em prejuízo de concorrentes ou da coletividade, inclusive mediante a utilização de meios ilícitos ou contrários aos bons costumes comerciais. Ao utilizarem obras protegidas para treinar modelos de IA sem autorização ou remuneração, as empresas de tecnologia se beneficiam da criatividade humana, sem arcar com os custos de sua produção, o que distorce o princípio da livre concorrência e cria uma assimetria estrutural no mercado digital.

As empresas, contudo, defendem posição contrária. A OpenAI sustenta que o treinamento de seus modelos constitui uso transformativo e analítico, pois os sistemas extraem padrões e relações dos dados para produzir conteúdos novos, enquadrando essa

utilização na doutrina do fair use (OpenAI, 2025). De forma semelhante, a Anthropic argumenta que a utilização de obras protegidas no treinamento de seu modelo Claude possui finalidade transformativa, pois os conteúdos são utilizados para desenvolver uma nova tecnologia, e não para reproduzir as obras originais (Anthropic, 2025). Esses argumentos, entretanto, são controvertidos, especialmente diante da possibilidade de os sistemas gerarem conteúdos que concorram economicamente com os próprios criadores cujas obras foram utilizadas no treinamento.

Essa prática configura, uma apropriação do trabalho criativo. Enquanto os titulares de direitos arcam com os custos de criação, tempo e investimento, as corporações tecnológicas internalizam os benefícios desses esforços sem qualquer contrapartida. Essa inundação de obras automatizadas reduz o preço médio de bens culturais e desincentiva o investimento em originalidade e no artista, provocando uma crise econômica para quem vive de suas artes.

A discussão acerca do uso de obras protegidas por sistemas de inteligência artificial também assume relevância sob a perspectiva do direito concorrencial. O acesso a grandes volumes de dados constitui hoje um ativo estratégico essencial ao desenvolvimento tecnológico, funcionando como barreira de entrada para novos competidores. A Organização para Cooperação e Desenvolvimento Econômico destaca que a concentração de dados pode gerar efeitos anticoncorrenciais e consolidar posições dominantes no mercado (OCDE, 2023). Nesse cenário, a utilização irrestrita de acervos criativos por grandes plataformas tecnológicas, sem mecanismos de remuneração ou licenciamento, pode reforçar assimetrias estruturais e deslocar valor econômico da produção cultural para corporações globais, exigindo reflexão também à luz da Lei nº 12.529/2011, que estrutura o Sistema Brasileiro de Defesa da Concorrência (Brasil, 2011).

9.3 Implicações à Luz da Lei Geral de Proteção de Dados

Para além das repercussões autorais e concorrenciais, a utilização de obras protegidas no treinamento de inteligências artificiais também pode envolver tratamento de dados pessoais, atraindo a incidência da Lei nº 13.709/2018 (LGPD), especialmente quando as obras incorporam elementos identificáveis da personalidade do autor, como imagem, voz ou dados biográficos (Brasil, 2018). Nos termos do art. 5º, X, da LGPD, qualquer operação de coleta, armazenamento ou processamento automatizado caracteriza tratamento de dados, exigindo fundamento jurídico adequado.

Ainda que parte dos conteúdos utilizados esteja disponível publicamente na internet, tal circunstância não afasta a necessidade de observância dos princípios da finalidade,

adequação e necessidade. A Autoridade Nacional de Proteção de Dados tem ressaltado que o tratamento deve ser compatível com a finalidade informada ao titular e respeitar os direitos fundamentais de liberdade e privacidade (ANPD, 2023).

Conclusão

Em última análise, a interseção entre o desenvolvimento acelerado das inteligências artificiais generativas e a proteção do direito autoral não representa um conflito insolúvel, mas exige a atualização do sistema jurídico. O desafio central consiste em equilibrar inovação tecnológica e proteção da criatividade humana. Entretanto, não basta estabelecer mecanismos para impedir novas utilizações indevidas, considerando que inúmeras informações já foram utilizadas no treinamento de sistemas existentes. Surge, portanto, o questionamento sobre como reparar os titulares por usos já realizados, especialmente diante da dificuldade de identificar todas as obras utilizadas e de "desaprender" os dados incorporados aos modelos. Nesse contexto, mecanismos de licenciamento coletivo e fundos de compensação financiados pelas empresas de IA podem representar alternativas para assegurar remuneração aos autores. Assim, conclui-se que a transparência, responsabilização e reparação tornam-se essenciais para que o desenvolvimento tecnológico ocorra sem transformar a criatividade humana em recurso econômico gratuito.

Referências:

ANPD – AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS. **Guia orientativo sobre tratamento de dados pessoais pelo poder público**. Brasília, DF: ANPD, 2023. Disponível em: https://www.gov.br/anpd. Acesso em: 10 fev. 2026.

ASCENSÃO, José de Oliveira. **Estudos de José de Oliveira Ascensão sobre direito autoral & sociedade informacional**. Curitiba: IODA, 2022.

BARBOSA, Denis Borges. **Uma introdução à propriedade intelectual**. Rio de Janeiro: Lumen Juris, 2003.

BITTAR, Carlos Alberto. **Direito de autor**. São Paulo: Saraiva, 2018.

BRASIL. **Decreto nº 75.699, de 6 de maio de 1975**. Promulga a Convenção de Berna para a Proteção das Obras Literárias e Artísticas, de 9 de setembro de 1886, revista em Paris, a 24 de julho de 1971. Diário Oficial da União: Brasília, DF, 9 maio 1975.

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.

BRASIL. **Lei nº 9.610, de 19 de fevereiro de 1998**. Altera, atualiza e consolida a legislação sobre direitos autorais. Diário Oficial da União: Brasília, DF, 20 fev. 1998.

BRASIL. **Lei nº 12.529, de 30 de novembro de 2011**. Estrutura o Sistema Brasileiro de Defesa da Concorrência. Diário Oficial da União: Brasília, DF, 1 dez. 2011.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados Pessoais (LGPD). Diário Oficial da União: Brasília, DF, 15 ago. 2018.

BRASIL. Senado Federal. **Relatório final da Comissão de Juristas responsável por subsidiar a elaboração de substitutivo sobre inteligência artificial no Brasil**. Brasília, DF: Senado Federal, 2022.

BRASIL. Ministério da Gestão e da Inovação em Serviços Públicos. Secretaria de Governo Digital; SERVIÇO FEDERAL DE PROCESSAMENTO DE DADOS (SERPRO). **Guia IA Generativa**. Brasília, DF, 2025. Disponível em: https://www.gov.br/governodigital/pt-br/infraestrutura-nacional-de-dados/inteligencia-artificial-1/publicacoes/guia-ia-generativa. Acesso em: 19 ago. 2026.

BRASIL. Superior Tribunal de Justiça. **Direito autoral deve ser respeitado mesmo que foto esteja disponível na internet**. Brasília, DF, 9 mar. 2020. Disponível em: https://www.stj.jus.br/sites/portalp/Paginas/Comunicacao/Noticias/Direito-autoral-deve-ser-respeitado-mesmo-que-foto-esteja-disponivel-na-internet.aspx. Acesso em: 19 ago. 2026.

CHINA. Cyberspace Administration of China. **Interim Measures for the Management of Generative Artificial Intelligence Services**. Beijing, 2023. Disponível em: http://www.cac.gov.cn. Acesso em: 10 fev. 2026.

CNN BRASIL. **Microsoft é processada por autores pelo uso de livros em treinamento de IA**. São Paulo, 25 jun. 2025. Disponível em: https://www.cnnbrasil.com.br/economia/negocios/microsoft-e-processada-por-autores-pelo-uso-de-livros-em-treinamento-de-ia/. Acesso em: 26 set. 2025.

CODA – CONTENT OVERSEAS DISTRIBUTION ASSOCIATION. **CODA Issues Written Request to OpenAI Regarding Sora 2**. 28 out. 2025. Disponível em: https://coda-cj.jp/en/news/817/. Acesso em: 19 jun. 2026.

COOPER, A. Feder; CHOQUETTE-CHOO, Christopher A. et al. **Machine Unlearning Doesn't Do What You Think: Lessons for Generative AI Policy, Research, and Practice**. 2025. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5060253. Acesso em: 23 fev. 2025.

COOPER, et al. **Artificial Intelligence and Copyright Law: Transformative Use and Creative Outputs**. 2025.

DE TILIA, Caroline. **Plágio de Studio Ghibli é sinal de alerta: até que ponto plataformas podem copiar conteúdos disponíveis online**. Forbes Brasil, 3 abr. 2025. Disponível em: https://forbes.com.br/forbes-tech/2025/04/plagio-de-studio-ghibli-e-sinal-de-alerta-ate-que-ponto-plataformas-podem-copiar-conteudos-disponiveis-online/. Acesso em: 11 ago. 2026.

FORBES BRASIL. **Processo diz que OpenAI violou direitos autorais nos EUA para treinar ChatGPT**. 29 jun. 2023. Disponível em: https://forbes.com.br/forbes-tech/2023/06/processo-diz-que-openai-violou-direitos-autorais-nos-eua-para-treinar-chatgpt/. Acesso em: 7 set. 2025.

GIBLIN, Rebecca; DOCTOROW, Cory. **Capitalismo anticompetição: como os oligopólios corporativos estão asfixiando os mercados criativos**. Tradução de Carol Colffield. 1. ed. Rio de Janeiro: Alta Books, 2024.

HANS, Abhimanyu; WEN, Yuxin; JAIN, Neel; KIRCHENBAUER, John; KAZEMI, Hamid; SINGHANIA, Prajwal; SINGH, Siddharth; SOMEPELLI, Gowthami; GEIPING, Jonas;

BHATELE, Abhinav; GOLDSTEIN, Tom. **Be like a Goldfish, Don't Memorize! Mitigating Memorization in Generative LLMs**. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 37., 2024. Proceedings [...]. 2024. p. 24022-24045. Disponível em: https://proceedings.neurips.cc/paper_files/paper/2024/file/2ad2dffba5079687651226ac8752df97-Paper-Conference.pdf. Acesso em: 23 nov. 2026.

JAPAN. **Copyright Act of Japan (Act No. 48 of 1970, as amended in 2018)**. Disponível em: https://www.japaneselawtranslation.go.jp. Acesso em: 10 fev. 2026.

KISSINGER, Henry A.; SCHMIDT, Eric; HUTTENLOCHER, Daniel. **A era da IA**. Tradução de Vanessa Schreiner. Rio de Janeiro: Alta Books, 2023.

LEMOS, Ronaldo. **A batalha da IA será travada no direito autoral**. Folha de S. Paulo, São Paulo, 12 jun. 2023. Disponível em: https://www1.folha.uol.com.br/colunas/ronaldolemos/2023/06/batalha-da-ia-sera-travada-no-direito-autoral.shtml. Acesso em: 18 ago. 2025.

MEDINA, Janny Carrasco. **A inteligência artificial versus direitos autorais: a impunibilidade do uso não autorizado, específica referência ao Brasil**. Revista da Faculdade de Direito da UFMG, Belo Horizonte, n. 84, p. 243-263, jan./jun. 2024. Disponível em: https://revista.direito.ufmg.br/index.php/revista/article/view/2721/2115. Acesso em: 3 out. 2024.

METZ, Cade. **Former OpenAI Researcher Says the Company Broke Copyright Law**. The New York Times, Nova York, 23 out. 2024. Disponível em: https://www.nytimes.com/2024/10/23/technology/openai-copyright-law.html. Acesso em: 18 ago. 2025.

MORIMOTO, Lucas Baltasar da Silva; PRADO, Luis Fernando. **Fundamentos técnico-jurídicos para o treinamento de IA e a necessária revisão do PL 2.338/2023**. In: CANTARINI, Paola (org.). **Parecer jurídico sobre o Projeto de Lei n.º 2.338/2023 – Instituto EthikAI**. São Paulo: Instituto EthikAI, 2025.

OCDE – ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. **Competition and Artificial Intelligence**. Paris: OECD Publishing, 2023. Disponível em: https://www.oecd.org/competition/. Acesso em: 10 fev. 2026.

OPENAI. **OpenAI and journalism**. 8 jan. 2024. Disponível em: https://openai.com/index/openai-and-journalism/. Acesso em: 8 out. 2025.

ORGANIZAÇÃO MUNDIAL DA PROPRIEDADE INTELECTUAL (OMPI). **Generative AI**. Genebra: WIPO, 2024. Disponível em: https://www.wipo.int/publications/en/details.jsp?id=4750. Acesso em: 19 ago. 2026.

SAMUELSON, Pamela. **Generative AI meets copyright**. Science, v. 381, n. 6654, p. 158-161, 14 jul. 2023. Disponível em: https://siliconflatirons.org/wp-content/uploads/2023/08/Generative-AI-Meets-Copyright_Pam-Samuelson.pdf. Acesso em: 23 out. 2025.

SANTOS, José Carlos Sales dos. **Direitos autorais na internet: o caso do Google Books**. 2011. Disponível em: https://periodicos.ufba.br/index.php/revistaici/article/view/4084. Acesso em: 18 ago. 2025.

SOBEL, Benjamin. **Artificial Intelligence's Fair Use Crisis**. Columbia Journal of Law & the Arts, v. 41, n. 1, p. 45-88, 2017. Disponível em: https://www.bensobel.org/files/articles/41.1_Sobel-FINAL.pdf. Acesso em: 24 nov. 2025.

SOUZA, Gabriela Biscotto de; CASTRO, João Francisco Chacarosque de. **Os direitos morais de integridade e atribuição no Brasil e nos EUA**. Humanidades em Diálogo, v. 13, p. 200-213, 2024. Disponível em: https://doi.org/10.11606/issn.1982-7547.hd.2024.192317. Acesso em: 19 ago. 2026.

STROWEL, Alain. **The European "Extended Collective Licensing" Model**. Columbia Journal of Law & the Arts, v. 34, p. 665, 2011.

TARTUCE, Flávio. **Manual de Direito Civil: volume único**. 14. ed. São Paulo: Método, 2024.

UNIÃO EUROPEIA. **Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024**. Estabelece regras harmonizadas em matéria de inteligência artificial. Jornal Oficial da União Europeia, 12 jul. 2024. Disponível em: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/por. Acesso em: 27 nov. 2024.

WORLD INTELLECTUAL PROPERTY ORGANIZATION (WIPO). **Copyright**. Geneva: WIPO, 2026. Disponível em: https://www.wipo.int/en/web/copyright. Acesso em: 23 nov. 2026.

ZEFF, Maxwell. **OpenAI's viral Studio Ghibli moment highlights AI copyright concerns**. TechCrunch, 26 mar. 2025. Disponível em: https://techcrunch.com/2025/03/26/openais-viral-studio-ghibli-moment-highlights-ai-copyright-concerns/. Acesso em: 19 ago. 2026.

Contatos: ariadne.felix@mackenzista.com.br e maria.marinho@mackenzie.br