

Treinamento de Inteligências Artificiais Generativas com ativos distintivos da propriedade industrial: limites jurídicos e lacunas regulatórias no Brasil

Training Generative Artificial Intelligence using industrial property assets: legal limits and regulatory gaps in Brazil

Entrenamiento de la Inteligencia Artificial Generativa con activos distintivos de la propiedad industrial: límites jurídicos y lagunas normativas en Brasil

DOI: 10.54901/educa.v9-659

Eloy Orlando Paiva Lima¹

Flávio Ribeiro da Costa²

Isabel Cristina de Souza Bertoldo³

José Carlos Cardoso de Oliveira Junior⁴

José Luiz de Paula Neto⁵

Renata Medeiros de Souza Miranda⁶

Fernando Melo da Silva⁷

RESUMO: O avanço da Inteligência Artificial Generativa (IAG) intensificou o debate sobre o uso de ativos protegidos durante o treinamento de grandes modelos de linguagem, concentrando-se, contudo, nos direitos autorais. Este estudo analisa a lacuna regulatória relativa à utilização de ativos distintivos da propriedade industrial, especialmente marcas, marcas de certificação e indicações geográficas, no treinamento desses sistemas. Adotou-se abordagem qualitativa, exploratória e descritiva, fundamentada em pesquisa bibliográfica, documental e em Direito Comparado, com análise da legislação brasileira, do Projeto de Lei nº 2.338/2023 e dos modelos regulatórios dos Estados Unidos, União Europeia e China. Os resultados evidenciam que o ordenamento jurídico brasileiro protege esses ativos em sua exploração econômica tradicional, mas permanece insuficiente para disciplinar sua utilização como insumos informacionais na economia baseada em dados. Conclui-se pela necessidade de um marco regulatório que incorpore transparência, rastreabilidade, governança dos dados e proteção dos ativos distintivos, conciliando inovação tecnológica, segurança jurídica e desenvolvimento econômico.

¹ Mestrando em Propriedade Intelectual e Transferência de Tecnologia para a Inovação, Universidade do Estado de Minas Gerais, Frutal, Minas Gerais, Brasil. E-mail: eloy@paivalima.adv.br

² Mestrando em Propriedade Intelectual e Transferência de Tecnologia para a Inovação, Universidade do Estado de Minas Gerais, Frutal, Minas Gerais, Brasil. E-mail: flavio.251100071@discente.uemg.br

³ Mestranda em Propriedade Intelectual e Transferência de Tecnologia para a Inovação, Universidade do Estado de Minas Gerais, Frutal, Minas Gerais, Brasil. E-mail: isabel.261100002@discente.uemg.br

⁴ Mestrando em Propriedade Intelectual e Transferência de Tecnologia para a Inovação, Universidade do Estado de Minas Gerais, Frutal, Minas Gerais, Brasil. E-mail: jose.261100003@discente.uemg.br

⁵ Mestrando em Propriedade Intelectual e Transferência de Tecnologia para a Inovação, Universidade do Estado de Minas Gerais, Frutal, Minas Gerais, Brasil. E-mail: jose.261100012@discente.uemg.br

⁶ Mestranda em Propriedade Intelectual e Transferência de Tecnologia para a Inovação, Universidade do Estado de Minas Gerais, Frutal, Minas Gerais, Brasil. E-mail: renata.261100013@discente.uemg.br

⁷ Doutor em Direito pela Pontifícia Universidade Católica de São Paulo (PUC-São Paulo), Mestre pela Universidade Estadual Paulista (UNESP), Professor e Vice Coordenador do Programa de Pós-Graduação em Propriedade Intelectual e Transferência de Tecnologia para a Inovação (PROFNIT), da Universidade do Estado de Minas Gerais (UEMG), Frutal, Minas Gerais, Brasil. E-mail: fernando.silva@uemg.br

Palavras-chave: Inteligência Artificial Generativa; propriedade industrial; marcas; indicações geográficas; regulação.

1 INTRODUÇÃO

O avanço da Inteligência Artificial Generativa (IAG) constitui uma das mais expressivas transformações tecnológicas do século XXI, promovendo profundas mudanças na produção, circulação e utilização do conhecimento. Modelos baseados em arquiteturas de aprendizado profundo (*deep learning*), como ChatGPT, Claude, Gemini e Midjourney, dependem do processamento de grandes volumes de dados para identificar padrões estatísticos e gerar conteúdos originais em diferentes formatos, incluindo textos, imagens, códigos computacionais e áudios. O treinamento desses modelos ocorre por meio da utilização de extensos *datasets*, frequentemente constituídos por informações obtidas mediante técnicas de *web scraping* e *Text and Data Mining* (TDM), que podem incorporar obras intelectuais, marcas, indicações geográficas, logotipos e outros ativos protegidos pelo sistema de propriedade intelectual (Volpe, 2025).

A expansão dessas tecnologias intensificou a tensão entre inovação tecnológica e proteção jurídica da propriedade intelectual. A literatura especializada concentra-se, predominantemente, na utilização de obras protegidas por direitos autorais durante o treinamento de sistemas de Inteligência Artificial. Em contrapartida, permanece incipiente a discussão acerca da utilização de ativos distintivos da propriedade industrial, tais como marcas, marcas coletivas, marcas de certificação, indicações geográficas, nomes empresariais, *trade dress* e outros sinais distintivos empregados para identificar a origem, a qualidade e a reputação de produtos e serviços. Essa lacuna torna-se especialmente relevante diante da crescente utilização de conteúdos disponíveis na internet como insumos para o desenvolvimento de modelos comerciais de Inteligência Artificial (Alves, 2024).

No cenário internacional, o tema passou a ocupar posição central nas discussões jurídicas em razão do aumento expressivo de litígios envolvendo empresas desenvolvedoras de Inteligência Artificial. Demandas judiciais propostas contra empresas como OpenAI, Anthropic e Stability AI, além de controvérsias envolvendo grupos editoriais e agências de imagens, demonstram que o treinamento de modelos de IA deixou de representar apenas uma questão tecnológica para assumir significativa dimensão econômica, concorrencial e jurídica. Os principais debates concentram-se nos limites da utilização de conteúdos protegidos, na

transparência dos *datasets*, na responsabilização dos desenvolvedores e na eventual remuneração dos titulares dos direitos intelectuais (Volpe, 2025).

No Brasil, entretanto, observa-se a inexistência de disciplina normativa específica voltada à utilização de ativos protegidos durante o treinamento de sistemas de Inteligência Artificial Generativa. A Constituição Federal de 1988, a Lei nº 9.610/1998 (Lei de Direitos Autorais) e a Lei nº 9.279/1996 (Lei da Propriedade Industrial) foram concebidas em contexto tecnológico anterior ao desenvolvimento dos grandes modelos de linguagem (*Large Language Models* – LLMs), circunstância que dificulta a aplicação de seus institutos às novas formas de coleta, mineração e processamento automatizado de dados. Embora os projetos legislativos em tramitação busquem estabelecer um marco regulatório para a Inteligência Artificial, ainda não disciplinam, de forma específica, a utilização de ativos distintivos da propriedade industrial durante o treinamento desses sistemas (Brasil, 2023; Ferreira; Chaves; Sousa, 2025).

Essa omissão normativa revela-se particularmente sensível quando analisada sob a perspectiva da propriedade industrial. Marcas, indicações geográficas e marcas de certificação extrapolam a função de simples identificação de produtos e serviços, desempenhando papel estratégico na diferenciação concorrencial, na proteção da reputação empresarial, na certificação de padrões de qualidade e na promoção do desenvolvimento econômico e regional, em consonância com a proteção conferida pela Constituição Federal e pela Lei da Propriedade Industrial (Brasil, 1988; Brasil, 1996). No caso das indicações geográficas, esses ativos também preservam a identidade territorial, valorizam conhecimentos tradicionais e fortalecem a competitividade de cadeias produtivas locais, aspectos reconhecidos como relevantes para a inovação e o desenvolvimento sustentável (Wipo, 2020; Wipo, 2024). Nesse contexto, a utilização dessas informações como insumos para o treinamento de modelos de Inteligência Artificial Generativa suscita questionamentos acerca da transparência dos *datasets*, da rastreabilidade das fontes utilizadas, do eventual aproveitamento econômico da reputação construída pelos titulares e da suficiência dos mecanismos jurídicos atualmente disponíveis para a tutela desses ativos na economia baseada em dados (Van Dijck, 2017; Volpe, 2025).

Embora o debate jurídico internacional tenha avançado de forma significativa em relação à utilização de obras protegidas por direitos autorais no treinamento de sistemas de Inteligência Artificial, observa-se que a análise dos impactos sobre os ativos distintivos da propriedade industrial ainda permanece incipiente. Essa assimetria evidencia uma lacuna doutrinária e regulatória que ultrapassa a proteção autoral e alcança institutos essenciais para a livre concorrência, a inovação tecnológica e o desenvolvimento econômico, especialmente diante da crescente relevância dos fluxos informacionais e da governança dos dados na

sociedade contemporânea (Braman, 2006; Braman, 2011; Volpe, 2025). Assim, torna-se necessário investigar se o ordenamento jurídico brasileiro dispõe de instrumentos adequados para disciplinar a utilização desses ativos distintivos nos processos de treinamento de modelos de Inteligência Artificial.

Diante desse cenário, o presente artigo tem por objetivo analisar as lacunas da regulação brasileira relativas ao treinamento de sistemas de Inteligência Artificial Generativa com ativos distintivos da propriedade industrial, especialmente marcas, indicações geográficas e marcas de certificação. Busca-se examinar a suficiência do atual arcabouço normativo, à luz da Constituição Federal, da Lei da Propriedade Industrial, do Projeto de Lei nº 2.338/2023 e das experiências regulatórias internacionais, identificando os principais desafios jurídicos decorrentes da utilização desses ativos em processos de aprendizagem algorítmica e propondo diretrizes que promovam o equilíbrio entre inovação tecnológica, segurança jurídica, proteção da propriedade industrial e desenvolvimento econômico (Brasil, 1988; Brasil, 1996; Brasil, 2023; European Union, 2024)

2 METODOLOGIA

A presente pesquisa caracteriza-se como qualitativa, de natureza exploratória e descritiva, desenvolvida por meio de pesquisa bibliográfica, documental e de Direito Comparado, adotando o método dedutivo como abordagem. A pesquisa qualitativa busca compreender fenômenos complexos a partir da interpretação de seus significados e contextos, sendo especialmente adequada para a investigação de questões jurídicas e sociais relacionadas às transformações tecnológicas (Sampieri; Collado; Lucio, 2013; Silva; Menezes, 2001).

A pesquisa bibliográfica fundamentou-se na análise de livros, artigos científicos nacionais e internacionais e documentos especializados nas áreas de Inteligência Artificial, Direito Autoral, Propriedade Industrial, Ciência da Informação, Direito Digital e Inovação Tecnológica. Foram examinados estudos relacionados ao treinamento de Grandes Modelos de Linguagem (Large Language Models – LLMs), às técnicas de Text and Data Mining (TDM) e aos desafios regulatórios decorrentes da utilização de ativos protegidos em sistemas de Inteligência Artificial, com base, entre outros, em Brown *et al.* (2020), OpenAI (2023), Volpe (2025), WIPO (2020; 2024) e Ferreira, Chaves e Sousa (2025).

A pesquisa documental compreendeu a análise da Constituição da República Federativa do Brasil de 1988, da Lei nº 9.279/1996 (Lei da Propriedade Industrial), da Lei nº 9.610/1998 (Lei de Direitos Autorais), do Projeto de Lei nº 2.338/2023, bem como de documentos

normativos e regulatórios nacionais e internacionais relativos à Inteligência Artificial e à propriedade intelectual. Conforme Sá-Silva, Almeida e Guindani (2009), a pesquisa documental possibilita a análise sistemática de fontes primárias, permitindo compreender a evolução normativa e institucional do objeto investigado.

Como procedimento complementar, empregou-se o Direito Comparado mediante o exame das soluções regulatórias adotadas pelos Estados Unidos, pela União Europeia e pela China, com ênfase na doutrina do fair use, na Diretiva (UE) 2019/790, no Regulamento Europeu de Inteligência Artificial (AI Act) e nas Medidas Provisórias da Administração do Ciberespaço da China para os Serviços de Inteligência Artificial Generativa. A análise comparativa teve por finalidade identificar tendências regulatórias capazes de subsidiar reflexões sobre o aperfeiçoamento do ordenamento jurídico brasileiro.

Por fim, o método dedutivo permitiu partir da análise do sistema constitucional e infraconstitucional de proteção da propriedade intelectual, especialmente da propriedade industrial, para identificar lacunas normativas relacionadas à utilização de ativos distintivos, como marcas, marcas de certificação, indicações geográficas e outros sinais distintivos, no treinamento de sistemas de Inteligência Artificial Generativa. A interpretação dos documentos e da literatura especializada foi orientada pela técnica de análise de conteúdo proposta por Bardin (2011), permitindo organizar, categorizar e interpretar criticamente as informações obtidas. A partir dessa análise, foram elaboradas diretrizes voltadas ao aperfeiçoamento da regulamentação brasileira, buscando conciliar inovação tecnológica, segurança jurídica e proteção dos direitos de propriedade industrial.

3 O TREINAMENTO DE INTELIGÊNCIAS ARTIFICIAIS GENERATIVAS E A UTILIZAÇÃO DE ATIVOS PROTEGIDOS

Os sistemas contemporâneos de Inteligência Artificial Generativa (IAG) fundamentam-se em técnicas de *Machine Learning* e *Deep Learning*, capazes de identificar padrões estatísticos em grandes conjuntos de dados para produzir conteúdos originais em linguagem natural, imagens, áudios, vídeos e códigos computacionais. Diferentemente dos sistemas tradicionais, cuja programação depende de comandos previamente definidos, esses modelos desenvolvem sua capacidade de resposta a partir do treinamento sobre extensas bases de dados (*datasets*), constituídas por bilhões de palavras, imagens e demais informações extraídas de diferentes fontes digitais (Volpe, 2025).

O treinamento representa a etapa central da construção dos Grandes Modelos de Linguagem (*Large Language Models* – LLMs). Nessa fase, algoritmos processam sucessivos volumes de dados, identificando relações probabilísticas entre palavras, imagens e outros elementos informacionais, o que lhes permite desenvolver capacidades de inferência, classificação e geração de novos conteúdos. Conforme observa Volpe (2025), a qualidade das respostas produzidas pelos modelos depende diretamente da diversidade, da representatividade e da qualidade dos dados empregados em seu treinamento.

A constituição desses *datasets* ocorre por diferentes mecanismos, destacando-se o licenciamento de bases de dados, a utilização de conteúdos em domínio público, a geração de dados sintéticos e, principalmente, técnicas automatizadas de coleta, como o *web scraping* e o *Text and Data Mining* (TDM). Enquanto o *web scraping* consiste na extração automatizada de informações disponibilizadas em páginas eletrônicas, o TDM corresponde ao processamento computacional desses conteúdos para identificação de padrões utilizados no treinamento dos modelos. Embora essenciais ao desenvolvimento da Inteligência Artificial, essas técnicas suscitam importantes questionamentos jurídicos quanto à utilização de conteúdos protegidos por direitos de propriedade intelectual sem autorização de seus titulares (Volpe, 2025).

A literatura especializada demonstra que parcela significativa dos conteúdos utilizados nos *datasets* é composta por livros, artigos científicos, fotografias, imagens, bases jornalísticas, sítios eletrônicos e outros materiais protegidos pelo ordenamento jurídico. Em razão disso, o debate jurídico concentrou-se, inicialmente, na compatibilidade entre o treinamento dos modelos e a proteção conferida pelo direito autoral, especialmente quanto à possibilidade de utilização dessas obras sem autorização prévia dos respectivos titulares (Alves, 2024).

Todavia, a problemática ultrapassa o campo dos direitos autorais. Durante o processo de coleta automatizada, os sistemas de Inteligência Artificial também incorporam ativos distintivos da propriedade industrial, como marcas, marcas coletivas, marcas de certificação, indicações geográficas (indicações de procedência e denominações de origem), nomes empresariais, nomes de domínio, logotipos, elementos de identidade visual e trade dress. Esses ativos exercem funções econômicas relacionadas à identificação da origem empresarial, à diferenciação concorrencial, à garantia da qualidade e à consolidação da reputação comercial, constituindo instrumentos essenciais para a proteção da concorrência, da inovação e do mercado (Brasil, 1988; Brasil, 1996; Wipo, 2024).

A tutela jurídica desses ativos encontra fundamento no artigo 5º, inciso XXIX, da Constituição Federal, que assegura proteção às criações industriais, às marcas e a outros signos distintivos em atenção ao interesse social e ao desenvolvimento tecnológico e econômico do

País (Brasil, 1988). No plano infraconstitucional, essa proteção é disciplinada principalmente pela Lei nº 9.279/1996 (Lei da Propriedade Industrial), que regula o registro de marcas, a proteção das indicações geográficas, dos desenhos industriais e a repressão aos atos de concorrência desleal (Brasil, 1996). Entretanto, embora esse sistema normativo discipline a exploração econômica e a defesa desses ativos nas relações de mercado, não contempla regras específicas para sua utilização como insumos informacionais no treinamento de sistemas de Inteligência Artificial Generativa, realidade que evidencia novos desafios para a proteção da propriedade intelectual na economia baseada em dados (Wipo, 2020; Volpe, 2025).

Apesar da relevância econômica e jurídica desses ativos, a produção científica nacional ainda lhes dedica atenção limitada. A maior parte das discussões concentra-se na utilização de obras protegidas por direitos autorais durante o treinamento de modelos de Inteligência Artificial, enquanto os impactos relacionados às marcas, indicações geográficas, marcas de certificação e demais sinais distintivos permanecem insuficientemente explorados. Nesse sentido, Ferreira, Chaves e Sousa (2025) observam que o debate regulatório brasileiro, especialmente no âmbito do Projeto de Lei nº 2.338/2023, permanece voltado à disciplina geral dos sistemas de Inteligência Artificial, sem enfrentar de forma específica a utilização dos ativos distintivos da propriedade industrial.

Essa realidade evidencia uma mudança de paradigma. Os ativos distintivos deixam de desempenhar exclusivamente sua função tradicional de identificação mercadológica para integrarem também a infraestrutura informacional utilizada no treinamento de modelos de Inteligência Artificial Generativa. Em uma economia orientada por dados, informações relacionadas à origem, reputação e qualidade de produtos passam a assumir valor estratégico, ampliando sua relevância para além das relações tradicionais de mercado (Braman, 2006; Braman, 2011; Van Dijck, 2017).

Quanto maior a disponibilidade desses dados, maior tende a ser a capacidade dos sistemas de reconhecer padrões visuais, associar produtos às suas origens geográficas, identificar reputações empresariais e reproduzir relações econômicas presentes no mercado. Embora esse fenômeno represente significativo avanço tecnológico, evidencia igualmente a necessidade de definir limites jurídicos para a utilização desses ativos na economia baseada em dados, de modo a compatibilizar inovação, transparência, segurança jurídica e proteção da propriedade industrial (Wipo, 2024; Volpe, 2025).

4 OS ATIVOS DISTINTIVOS DA PROPRIEDADE INDUSTRIAL E OS DESAFIOS DO TREINAMENTO DE INTELIGÊNCIAS ARTIFICIAIS

A evolução da Inteligência Artificial Generativa ampliou significativamente o debate acerca da proteção da propriedade intelectual. Até o momento, as discussões jurídicas concentram-se predominantemente na utilização de obras protegidas por direitos autorais durante o treinamento de modelos de Inteligência Artificial, especialmente em razão da coleta automatizada de grandes volumes de dados para a formação de datasets empregados pelos Grandes Modelos de Linguagem (Large Language Models – LLMs) (Brown *et al.*, 2020; OpenAI, 2023; Volpe, 2025). Entretanto, esses conjuntos de dados não são compostos exclusivamente por textos, livros ou imagens protegidos pela Lei nº 9.610/1998. As técnicas automatizadas de coleta e mineração de dados também incorporam ativos distintivos da propriedade industrial, como marcas, marcas coletivas, marcas de certificação, indicações geográficas, nomes empresariais, logotipos, elementos de identidade visual e trade dress, cuja utilização durante o treinamento desses sistemas ainda carece de disciplina jurídica específica no ordenamento brasileiro (Brasil, 1996; Wipo, 2024).

Esses ativos desempenham funções que extrapolam a simples identificação de produtos e serviços. Além de distinguir a origem empresarial, asseguram padrões de qualidade, agregam reputação econômica e contribuem para a promoção da concorrência leal e da inovação nos mercados (Brasil, 1988; Brasil, 1996; Wipo, 2020). No caso das indicações geográficas, sua relevância é ainda mais ampla, pois vinculam determinados produtos às características naturais e humanas de um território específico, preservando conhecimentos tradicionais, fortalecendo identidades territoriais e promovendo o desenvolvimento econômico sustentável das regiões produtoras (Wipo, 2024).

Embora os regimes jurídicos aplicáveis aos direitos autorais e à propriedade industrial possuam fundamentos distintos, essa diferenciação não é reproduzida durante o treinamento dos modelos de Inteligência Artificial. Os algoritmos processam grandes volumes de informações disponíveis em ambientes digitais sem distinguir, na etapa de coleta de dados, conteúdos protegidos pela Lei de Direitos Autorais daqueles tutelados pela Lei da Propriedade Industrial. Como observa Volpe (2025), a composição dos *datasets* utilizados no treinamento permanece, em grande medida, protegida por segredo comercial, dificultando a identificação das fontes efetivamente empregadas pelos desenvolvedores dos modelos. Essa limitação compromete a transparência e dificulta o exercício dos direitos pelos respectivos titulares.

A ausência de transparência torna-se especialmente relevante porque os debates jurídicos atuais concentram-se quase exclusivamente na utilização de obras protegidas por direitos autorais. Em contraste, permanecem escassos os estudos voltados aos efeitos do treinamento de Inteligência Artificial sobre marcas, indicações geográficas, marcas de certificação e demais ativos distintivos da propriedade industrial. Consequentemente, a discussão jurídica permanece centrada na Lei nº 9.610/1998, relegando a segundo plano a aplicação da Lei nº 9.279/1996 às novas formas de utilização desses ativos na economia digital.

Sob a perspectiva econômica, a utilização desses sinais distintivos durante o treinamento dos modelos produz efeitos que extrapolam a simples reprodução de informações. Sistemas treinados com milhões de imagens de produtos, embalagens, logotipos e referências geográficas desenvolvem capacidade para reconhecer padrões visuais, estabelecer associações entre marcas e categorias de produtos, identificar vínculos entre territórios e reputação econômica e reproduzir essas correlações em suas respostas. Ainda que esse processo não implique, necessariamente, reprodução literal dos sinais protegidos, evidencia que o valor informacional e reputacional desses ativos passa a integrar o processo de aprendizagem algorítmica.

Essa realidade torna-se particularmente sensível no caso das indicações geográficas. Por sua natureza coletiva, esses ativos protegem não apenas a identificação de determinado produto, mas também a reputação construída por comunidades produtoras ao longo do tempo. Cafés especiais, vinhos, queijos, cachaças e diversos outros produtos brasileiros utilizam as indicações geográficas como instrumento de agregação de valor, diferenciação competitiva e desenvolvimento regional. A incorporação sistemática dessas informações aos *datasets* utilizados para treinamento de modelos de Inteligência Artificial suscita questionamentos acerca do eventual aproveitamento econômico indireto dessa reputação coletiva, especialmente diante da inexistência de mecanismos específicos de transparência, autorização ou compensação.

Situação semelhante ocorre em relação às marcas de certificação. Esses sinais distintivos atestam que produtos ou serviços observam determinados padrões técnicos, qualitativos, ambientais ou de origem, desempenhando relevante função de confiança no mercado. Quando incorporadas aos *datasets*, essas informações passam a integrar a base informacional utilizada pelos algoritmos para reconhecimento de padrões mercadológicos, sem que o ordenamento jurídico brasileiro estabeleça critérios específicos para disciplinar essa forma de utilização.

Importa destacar que a utilização de ativos distintivos durante o treinamento de modelos de Inteligência Artificial não configura, por si só, violação automática da propriedade industrial.

O problema jurídico reside justamente na inexistência de parâmetros legais capazes de definir os limites dessa utilização, estabelecer deveres de transparência e disciplinar eventuais formas de exploração econômica desses ativos. Enquanto o treinamento de modelos com obras protegidas por direitos autorais já constitui objeto de intenso debate doutrinário e jurisprudencial, a utilização de ativos distintivos permanece praticamente ausente da produção científica e das iniciativas regulatórias brasileiras.

Nesse contexto, verifica-se uma mudança de paradigma na função econômica dos ativos distintivos da propriedade industrial. Tradicionalmente destinados à identificação da origem empresarial e à diferenciação concorrencial, esses ativos passaram também a integrar a infraestrutura informacional utilizada no treinamento de modelos de Inteligência Artificial Generativa. Essa transformação evidencia que a proteção jurídica desses institutos não pode mais restringir-se às relações tradicionais de mercado, exigindo uma abordagem regulatória compatível com a economia baseada em dados, fundamentada em princípios de transparência, rastreabilidade, segurança jurídica e equilíbrio entre inovação tecnológica e proteção da propriedade industrial.

O ordenamento jurídico brasileiro dispõe de um sistema consolidado de proteção da propriedade intelectual, abrangendo direitos autorais, propriedade industrial, programas de computador, cultivares e outros ativos imateriais. Contudo, esse conjunto normativo foi concebido em contexto tecnológico anterior ao surgimento da Inteligência Artificial Generativa (IAG), razão pela qual não contempla regras específicas destinadas à utilização desses ativos durante o treinamento de modelos computacionais.

A Constituição da República de 1988 reconhece a propriedade intelectual como direito fundamental ao assegurar aos autores o direito exclusivo de utilização, publicação e reprodução de suas obras (art. 5º, XXVII), bem como a proteção das criações industriais, das marcas e de outros signos distintivos, considerando o interesse social e o desenvolvimento tecnológico e econômico do País (art. 5º, XXIX). O texto constitucional evidencia a necessidade de conciliar a tutela dos direitos dos titulares com a promoção da inovação tecnológica. Todavia, não estabelece parâmetros específicos para disciplinar a utilização de ativos protegidos em processos automatizados de aprendizagem algorítmica, realidade inexistente à época de sua promulgação.

No plano infraconstitucional, a Lei nº 9.610/1998 assegura ao autor o direito exclusivo de utilizar, fruir e dispor de sua obra. Diferentemente do sistema norte-americano, a legislação brasileira não adota uma cláusula geral de exceção equivalente ao *fair use*, restringindo as limitações aos direitos autorais às hipóteses expressamente previstas em lei. Em consequência,

parcela significativa da doutrina sustenta que a utilização de obras protegidas para treinamento de modelos de Inteligência Artificial depende, como regra, da autorização do titular, ressalvadas as exceções legais (Alves, 2024; Volpe, 2025).

Situação distinta verifica-se em relação aos ativos disciplinados pela Lei nº 9.279/1996 (Lei da Propriedade Industrial). Embora esse diploma regule a proteção das marcas, indicações geográficas, patentes, desenhos industriais e reprima os atos de concorrência desleal, inexistente qualquer previsão acerca da utilização desses ativos na constituição de *datasets* destinados ao treinamento de modelos de Inteligência Artificial Generativa. A legislação protege a exploração econômica desses sinais nas relações de mercado, mas permanece silente quanto ao seu emprego como insumos informacionais em sistemas de aprendizagem de máquina.

Essa omissão normativa produz significativa insegurança jurídica. Modelos de Inteligência Artificial podem ser treinados com milhões de imagens contendo marcas registradas, logotipos, embalagens (*trade dress*), marcas de certificação e indicações geográficas sem que exista disciplina legal capaz de definir se essa utilização configura mero processamento técnico de informações, uso legítimo de dados disponíveis publicamente ou modalidade de exploração econômica sujeita à autorização dos respectivos titulares.

A ausência de parâmetros legais é agravada pela inexistência de deveres específicos de transparência quanto aos *datasets* utilizados no treinamento. Conforme observa Volpe (2025), a composição dessas bases costuma ser protegida por segredo comercial, dificultando a identificação das fontes empregadas e impedindo que titulares de direitos verifiquem se seus ativos foram incorporados ao processo de aprendizagem dos modelos de Inteligência Artificial. Essa opacidade compromete tanto a proteção dos direitos de propriedade industrial quanto a previsibilidade jurídica necessária ao desenvolvimento tecnológico.

No âmbito legislativo, o Projeto de Lei nº 2.338/2023 constitui a principal iniciativa brasileira destinada à regulamentação da Inteligência Artificial. A proposta estabelece princípios relacionados à transparência, à supervisão humana, à responsabilização, à gestão de riscos e à governança dos sistemas de IA, buscando harmonizar o desenvolvimento tecnológico com a proteção dos direitos fundamentais (Brasil, 2023). Todavia, embora represente um avanço na construção de um marco regulatório nacional, o projeto não disciplina de forma específica a utilização de ativos distintivos da propriedade industrial, como marcas, indicações geográficas e marcas de certificação, durante o treinamento de modelos de Inteligência Artificial Generativa. Conforme observam Ferreira, Chaves e Sousa (2025), o debate legislativo brasileiro permanece concentrado na regulamentação geral dos sistemas de IA, dedicando

reduzida atenção às implicações decorrentes do uso de ativos protegidos pela propriedade industrial.

Essa constatação demonstra que o principal desafio regulatório brasileiro não decorre da inexistência de proteção jurídica da propriedade intelectual, mas da insuficiência de normas voltadas à disciplina de sua aplicação na economia orientada por dados. Embora a utilização de obras protegidas por direitos autorais já constitua objeto de amplo debate doutrinário e legislativo, a disciplina relativa às marcas, indicações geográficas, marcas de certificação e demais sinais distintivos ainda permanece incipiente, apesar de sua relevância para o "desenvolvimento tecnológico e econômico do País" (Brasil, 1988, art. 5º, XXIX).

Sob essa perspectiva, a lacuna normativa produz impactos tanto para os titulares dos direitos quanto para os desenvolvedores de tecnologias. Os primeiros carecem de mecanismos que assegurem transparência, rastreabilidade e controle sobre o emprego de seus ativos em processos de treinamento de modelos computacionais. Os segundos, por sua vez, enfrentam elevado grau de insegurança jurídica diante da inexistência de critérios objetivos que delimitem a licitude da coleta, do processamento e do uso dessas informações para fins de desenvolvimento de sistemas de Inteligência Artificial Generativa. Como destaca Braman (2011), a informação passou a ocupar posição estratégica nas estruturas de poder e governança, o que reforça a necessidade de aperfeiçoar mecanismos regulatórios capazes de conciliar inovação, segurança jurídica e proteção da propriedade intelectual na economia digital.

Dessa forma, conclui-se que o atual sistema jurídico brasileiro oferece proteção adequada aos ativos distintivos da propriedade industrial em suas formas tradicionais de exploração econômica, mas permanece insuficiente para disciplinar sua utilização na infraestrutura informacional que sustenta a Inteligência Artificial Generativa. Essa lacuna evidencia a necessidade de construção de um modelo regulatório capaz de compatibilizar inovação tecnológica, segurança jurídica e proteção da propriedade industrial, tema que será analisado à luz das experiências internacionais no capítulo seguinte.

5 DIREITO COMPARADO E OS DESAFIOS INTERNACIONAIS NA UTILIZAÇÃO DE ATIVOS DISTINTIVOS DA PROPRIEDADE INDUSTRIAL PARA O TREINAMENTO DE INTELIGÊNCIAS ARTIFICIAIS GENERATIVAS

A expansão da Inteligência Artificial Generativa transformou a proteção da propriedade intelectual em um dos principais desafios regulatórios da economia digital. A natureza transnacional dos modelos de Inteligência Artificial faz com que os *datasets* utilizados durante

seu treinamento sejam compostos por informações provenientes de múltiplas jurisdições, tornando insuficiente uma análise restrita ao direito interno de cada país. Nesse contexto, diferentes ordenamentos jurídicos passaram a desenvolver soluções destinadas a compatibilizar inovação tecnológica, segurança jurídica e proteção dos direitos de propriedade intelectual.

Nos Estados Unidos, o principal eixo das discussões concentra-se na aplicação da doutrina do *fair use*, prevista na Seção 107 do *Copyright Act*. Diferentemente do sistema brasileiro, o ordenamento norte-americano admite, em determinadas circunstâncias, a utilização de obras protegidas sem autorização prévia do titular, desde que observados critérios relacionados à finalidade do uso, à natureza da obra, à extensão da utilização e aos seus impactos econômicos sobre o mercado da criação original. Essa interpretação tem sido amplamente invocada pelas empresas desenvolvedoras de Inteligência Artificial como fundamento jurídico para a utilização de grandes volumes de conteúdos protegidos durante o treinamento de seus modelos (Alves, 2024).

Entretanto, a aplicação do *fair use* ao treinamento de modelos de Inteligência Artificial permanece objeto de intensa controvérsia. Litígios envolvendo empresas como OpenAI, Microsoft, Anthropic, Stability AI, além de ações propostas por *The New York Times*, Getty Images e Authors Guild, demonstram que a discussão ultrapassa a utilização de obras protegidas e alcança temas relacionados à transparência dos *datasets*, à responsabilidade dos desenvolvedores e à eventual remuneração dos titulares dos direitos intelectuais (Volpe, 2025).

Entre esses litígios, destaca-se o caso envolvendo a Anthropic, desenvolvedora do sistema Claude, acusada de utilizar milhões de livros provenientes de acervos digitais não autorizados durante o treinamento de seus modelos. Independentemente do desfecho definitivo da controvérsia, o caso evidencia uma mudança de paradigma na regulação da Inteligência Artificial, demonstrando que a origem dos dados empregados no treinamento passou a representar elemento central para a responsabilização jurídica das empresas desenvolvedoras.

Na União Europeia, o enfrentamento do problema ocorreu por meio de uma estratégia regulatória distinta. A Diretiva (UE) 2019/790, relativa ao Mercado Único Digital, disciplinou hipóteses específicas de *Text and Data Mining* (TDM), enquanto o Regulamento Europeu de Inteligência Artificial (*AI Act*) passou a incorporar princípios voltados à transparência, à gestão de riscos e à governança dos sistemas de IA. Embora essas normas não proibam a mineração automatizada de dados, estabelecem mecanismos destinados a compatibilizar o desenvolvimento tecnológico com a proteção dos direitos dos titulares, conferindo maior previsibilidade jurídica às atividades de treinamento dos modelos.

Apesar desses avanços, observa-se que tanto a experiência norte-americana quanto a europeia permanecem fortemente concentradas na tutela dos direitos autorais. A utilização de ativos distintivos da propriedade industrial como marcas, indicações geográficas, marcas de certificação e outros sinais distintivos ainda recebe tratamento normativo incipiente, revelando que a expansão da Inteligência Artificial produz impactos que extrapolam os limites tradicionalmente analisados pela doutrina.

Essa constatação assume especial relevância para países cuja competitividade econômica depende da valorização de produtos vinculados à origem geográfica e à certificação de qualidade. No Brasil, diversos setores produtivos utilizam marcas, indicações geográficas e marcas de certificação como instrumentos de diferenciação competitiva, agregação de valor e fortalecimento das economias regionais. A eventual incorporação sistemática dessas informações aos *datasets* empregados no treinamento de modelos de Inteligência Artificial poderá produzir efeitos sobre a reputação desses ativos, sem que exista, até o momento, disciplina normativa específica para essa modalidade de utilização.

A análise comparada demonstra, portanto, que as principais experiências regulatórias internacionais reconhecem a necessidade de estabelecer limites jurídicos para o treinamento de sistemas de Inteligência Artificial, especialmente quanto à utilização de obras protegidas por direitos autorais. Contudo, também evidencia que a proteção dos ativos distintivos da propriedade industrial permanece em segundo plano, mesmo diante de sua crescente relevância econômica na sociedade orientada por dados.

Nesse contexto, o Direito Comparado constitui importante instrumento para o aperfeiçoamento do ordenamento jurídico brasileiro, ao evidenciar diferentes estratégias regulatórias destinadas a enfrentar os desafios decorrentes do desenvolvimento e do treinamento de sistemas de Inteligência Artificial (European Union, 2024; China, 2023). Contudo, mais do que transpor soluções estrangeiras, impõe-se a construção de um marco normativo compatível com as especificidades do sistema brasileiro de propriedade industrial e com os princípios constitucionais que orientam a proteção da inovação. Nesse sentido, a Constituição Federal assegura proteção à propriedade industrial "tendo em vista o interesse social e o desenvolvimento tecnológico e econômico do País" (Brasil, 1988, art. 5º, XXIX). Essa diretriz revela que marcas, indicações geográficas e marcas de certificação não desempenham apenas a função de identificar a origem e diferenciar produtos e serviços, mas também assumem relevância estratégica na economia baseada em dados, integrando a infraestrutura informacional utilizada no desenvolvimento de modelos de Inteligência Artificial Generativa (Braman, 2011; Wipo, 2024). Assim, a futura disciplina jurídica deverá conciliar

inovação tecnológica, livre concorrência, segurança jurídica e proteção dos ativos distintivos, em consonância com as diretrizes regulatórias atualmente debatidas no Brasil e no cenário internacional (Ferreira; Chaves; Sousa, 2025; European Union, 2024).

6 DIRETRIZES PARA UMA REGULAÇÃO BRASILEIRA DO TREINAMENTO DE INTELIGÊNCIAS ARTIFICIAIS GENERATIVAS COM ATIVOS DISTINTIVOS DA PROPRIEDADE INDUSTRIAL

A crescente difusão da Inteligência Artificial Generativa evidencia que os mecanismos tradicionais de tutela da propriedade intelectual já não são suficientes para responder aos desafios impostos pela economia orientada por dados. O processamento de grandes volumes de informações para o treinamento de modelos computacionais redefiniu a forma de utilização dos ativos protegidos no ambiente digital, conferindo-lhes também a função de insumos informacionais para o desenvolvimento tecnológico. Esse novo cenário exige a construção de instrumentos regulatórios capazes de conciliar inovação, segurança jurídica e proteção dos direitos dos titulares, em conformidade com os princípios constitucionais da propriedade intelectual e com as discussões internacionais sobre governança de dados e Inteligência Artificial (Brasil, 1988; Braman, 2011).

A experiência internacional demonstra que essa preocupação deixou de ser apenas teórica. Nos Estados Unidos, as ações judiciais propostas contra empresas como OpenAI, Anthropic e Stability AI revelam que a utilização de conteúdos protegidos durante o treinamento dos modelos passou a representar um dos principais conflitos jurídicos da economia digital. Na União Europeia, o Regulamento Europeu de Inteligência Artificial (AI Act) e a Diretiva sobre o Mercado Único Digital estabeleceram mecanismos voltados à transparência, gestão de riscos e mineração de dados, procurando compatibilizar inovação tecnológica com proteção dos direitos intelectuais.

A China, por sua vez, adotou uma abordagem mais intervencionista. As Interim Measures for the Management of Generative Artificial Intelligence Services, publicadas em 2023, estabelecem obrigações relacionadas à legalidade das bases de treinamento, à rastreabilidade dos conteúdos utilizados, à prevenção de violações de direitos de propriedade intelectual e à adoção de mecanismos de governança dos sistemas de Inteligência Artificial. Além disso, impõem aos provedores deveres de avaliação de segurança, remoção de conteúdos ilícitos e responsabilização pelos riscos decorrentes da utilização de informações protegidas (China, 2023). Embora esse modelo reflita características próprias do ordenamento jurídico

chinês, demonstra que a regulação da Inteligência Artificial pode ultrapassar a tutela dos direitos autorais, incorporando instrumentos preventivos voltados à governança dos dados e à proteção da propriedade intelectual.

As experiências analisadas evidenciam que a futura disciplina normativa brasileira não deve restringir-se aos direitos autorais. A utilização de marcas, indicações geográficas, marcas de certificação e outros sinais distintivos em processos de treinamento de modelos de Inteligência Artificial também demanda tratamento jurídico específico, considerando sua relevância econômica e sua crescente inserção na infraestrutura informacional da economia digital (Ferreira; Chaves; Sousa, 2025).

Entre as diretrizes identificadas, destaca-se a necessidade de instituir mecanismos mínimos de transparência quanto aos conjuntos de dados empregados no treinamento dos modelos. Não se trata de exigir a divulgação integral dos datasets, frequentemente protegidos por segredo empresarial, mas de estabelecer documentação técnica que identifique categorias de dados, fontes de obtenção, critérios de seleção e procedimentos de exclusão de conteúdos protegidos, fortalecendo a segurança jurídica e reduzindo potenciais conflitos entre desenvolvedores e titulares de direitos (European Union, 2024).

Outra medida relevante consiste na implementação de mecanismos de rastreabilidade capazes de permitir a identificação da origem dos dados utilizados no treinamento dos sistemas. A adoção de parâmetros dessa natureza favorece tanto a fiscalização pelos titulares dos direitos quanto a previsibilidade jurídica necessária ao desenvolvimento tecnológico, aproximando-se das diretrizes internacionais de governança da Inteligência Artificial (European Union, 2024; China, 2023).

Também merece reflexão a adoção de mecanismos de oposição (opt-out), por meio dos quais os titulares possam manifestar previamente sua discordância quanto à utilização de determinados ativos distintivos em aplicações comerciais específicas. Tal mecanismo preserva o desenvolvimento tecnológico sem afastar a autonomia dos titulares sobre bens imateriais dotados de expressivo valor econômico e reputacional (Wipo, 2024).

Outro aspecto relevante refere-se à definição de critérios para eventual compensação quando o emprego desses ativos gerar aproveitamento econômico pelos desenvolvedores dos modelos. A discussão assume especial importância em relação às indicações geográficas e às marcas de certificação, cuja reputação resulta de investimentos coletivos realizados por produtores, cooperativas e entidades gestoras ao longo do tempo, exigindo soluções compatíveis com a natureza coletiva desses direitos (Wipo, 2020; Wipo, 2024).

A análise também evidencia a conveniência de distinguir diferentes modalidades de utilização dos ativos protegidos. O processamento de dados para fins exclusivamente científicos apresenta natureza distinta daquele voltado ao desenvolvimento de aplicações comerciais capazes de explorar economicamente a reputação associada às marcas, às indicações geográficas e às marcas de certificação. A adoção de critérios proporcionais pode contribuir para compatibilizar inovação tecnológica, livre concorrência e proteção da propriedade intelectual (Brasil, 1988; European Union, 2024).

No caso das indicações geográficas, essa discussão assume dimensão estratégica. Diferentemente das marcas individuais, esses sinais distintivos representam instrumentos coletivos de valorização territorial, agregação de valor e preservação de conhecimentos tradicionais. Assim, a incorporação sistemática dessas informações aos processos de treinamento de modelos de Inteligência Artificial pode produzir impactos sobre cadeias produtivas regionais, justificando mecanismos específicos de transparência, rastreabilidade e governança.

Por fim, recomenda-se o fortalecimento da coordenação institucional entre o Instituto Nacional da Propriedade Industrial (INPI), a Autoridade Nacional de Proteção de Dados (ANPD) e os futuros órgãos responsáveis pela supervisão da Inteligência Artificial, com a participação das entidades gestoras das indicações geográficas e das marcas de certificação. Como destaca a Organização Mundial da Propriedade Intelectual, "generative AI is reshaping the intellectual property landscape" (Wipo, 2024), evidenciando a necessidade de atuação coordenada entre instituições responsáveis pela proteção dos direitos de propriedade intelectual e pela governança dos dados. Nesse contexto, a integração desses órgãos poderá favorecer a elaboração de diretrizes técnicas, padrões de rastreabilidade e mecanismos de resolução de conflitos compatíveis com os desafios da economia digital (Brasil, 2023; Ferreira; Chaves; Sousa, 2025).

Em síntese, a análise comparada demonstra que a regulação da Inteligência Artificial evolui para modelos baseados na transparência, na governança dos dados e na responsabilização dos agentes envolvidos. No Brasil, entretanto, permanece uma lacuna normativa quanto ao tratamento dos ativos distintivos da propriedade industrial empregados no treinamento de modelos de Inteligência Artificial Generativa. O reconhecimento desses sinais como elementos estratégicos da infraestrutura informacional da economia digital constitui passo essencial para a construção de um marco regulatório capaz de harmonizar inovação tecnológica, proteção da concorrência, desenvolvimento regional e segurança jurídica (Braman, 2011).

7 RESULTADOS E DISCUSSÃO

A análise da literatura especializada, da legislação brasileira e das experiências regulatórias internacionais permitiu identificar que o debate jurídico sobre o treinamento de sistemas de Inteligência Artificial Generativa permanece concentrado, predominantemente, na tutela dos direitos autorais. Embora essa discussão tenha avançado nos últimos anos, impulsionada pelo aumento de controvérsias envolvendo empresas desenvolvedoras de Inteligência Artificial, verificou-se que os impactos relacionados à utilização de marcas, indicações geográficas, marcas de certificação e outros sinais distintivos ainda recebem atenção limitada, tanto na produção científica quanto nas iniciativas legislativas (Volpe, 2025; Wipo, 2024).

Os resultados demonstram que os sistemas contemporâneos de Inteligência Artificial são treinados a partir do processamento automatizado de grandes volumes de informações provenientes de diferentes fontes e submetidas a distintos regimes de proteção jurídica (Brown *et al.*, 2020; OpenAI, 2023). Enquanto as obras protegidas pela Lei nº 9.610/1998 ocupam posição central nas discussões regulatórias, os sinais distintivos disciplinados pela Lei nº 9.279/1996 permanecem inseridos em um cenário de reduzida previsibilidade normativa, especialmente quanto à sua utilização como insumos informacionais (Brasil, 1996; Wipo, 2024).

No contexto brasileiro, constatou-se que o ordenamento jurídico dispõe de mecanismos consolidados para a tutela da propriedade industrial, disciplinando marcas, indicações geográficas, desenhos industriais e práticas de concorrência desleal. Contudo, tais instrumentos foram concebidos para regular formas tradicionais de exploração econômica e não contemplam, de maneira específica, o emprego desses ativos no treinamento de modelos de Inteligência Artificial Generativa (Brasil, 1996; Ferreira; Chaves; Sousa, 2025).

Os achados da pesquisa indicam que a principal lacuna regulatória não decorre da ausência de proteção jurídica, mas da inexistência de parâmetros voltados à transparência dos datasets, à rastreabilidade das informações empregadas e à definição de critérios para o uso de sinais distintivos na economia orientada por dados. Essa constatação evidencia a necessidade de aperfeiçoar os mecanismos de governança da informação e de compatibilizar o regime da propriedade intelectual com os novos processos de desenvolvimento tecnológico (Braman, 2011; Van Dijck, 2017; Wipo, 2024).

Outro resultado relevante refere-se à assimetria entre a velocidade da evolução tecnológica e a capacidade de adaptação do ordenamento jurídico. Enquanto os sistemas de

Inteligência Artificial ampliam continuamente sua capacidade de reconhecer padrões, estabelecer associações e processar informações de elevado valor econômico, a legislação ainda permanece estruturada segundo uma lógica anterior à economia baseada em dados, o que amplia a insegurança jurídica para titulares de direitos e desenvolvedores de tecnologias (Ferreira; Chaves; Sousa, 2025; Wipo, 2024).

A análise comparativa internacional apresentada na Figura 01 constitui uma contribuição original deste estudo ao sistematizar, de forma integrada e visual, os diferentes níveis de desenvolvimento regulatório relacionados ao treinamento de sistemas de Inteligência Artificial e à utilização de conteúdos protegidos, com especial atenção aos ativos distintivos da propriedade industrial. Ao confrontar as abordagens adotadas pelos Estados Unidos, União Europeia, China e Brasil, o quadro permite identificar, de maneira objetiva, convergências, divergências e, sobretudo, lacunas regulatórias ainda existentes no ordenamento jurídico brasileiro. Além de sua função analítica, a representação gráfica favorece a visualização, compreensão e memorização dos resultados, transformando informações jurídicas complexas e dispersas em uma síntese comparativa de fácil interpretação.

Figura 1 – Comparativo internacional da regulação do treinamento de Inteligência Artificial e utilização de ativos distintivos da propriedade indústria

Tema	 Estados Unidos	 União Europeia	 China	 Brasil
 Transparência dos datasets	Sem obrigação geral. Divulgação pode decorrer de decisões judiciais.	AI Act e Diretiva 2019/790 estabelecem deveres de transparência e documentação.	Exige registro das fontes, rastreabilidade e governança de dados.	Sem regra específica. PL 2.338/2023 aborda transparência de forma parcial.
 Direitos autorais	Aplicação do fair use (Seção 107), ainda objeto de litígios.	Exceções de Text and Data Mining (TDM), com possibilidade de opt-out.	Uso deve observar os direitos autorais e a legalidade dos dados.	Lei nº 9.610/1998 não prevê exceção específica para treinamento de IA.
 Ativos distintivos (marcas, indicações geográficas, marcas de certificação)	Sem disciplina específica; proteção ocorre principalmente por marca, concorrência desleal e publicidade enganosa.	Marcas e indicações geográficas possuem proteção própria, mas sem regra específica para treinamento de IA.	Proteção indireta contra dados que violem marcas e outros direitos de propriedade intelectual.	Lei nº 9.279/1996 protege marcas e indicações geográficas, mas não disciplina seu uso no treinamento de IA.
 Responsabilidade dos desenvolvedores de IA	Definida principalmente pela jurisprudência e pelas normas gerais de responsabilidade.	AI Act estabelece governança, documentação, mitigação de riscos e sanções administrativas.	Responsabilidade administrativa e civil, com mecanismos de controle e sanções.	Regulação ainda em construção; ausência de disciplina específica para ativos distintivos.
 Situação atual das discussões	Litígios em andamento delimitam os usos de conteúdos protegidos.	Regulação em implementação, com foco em transparência e governança.	Regulação em vigor desde 2023, com abordagem preventiva e forte controle dos dados.	Discussão legislativa em evolução, com lacuna quanto aos ativos distintivos.

Fonte: Elaborado pelos autores com base na legislação, em documentos oficiais e na literatura especificada (2024-2026)

Os Estados Unidos adotam um modelo regulatório predominantemente baseado na interpretação judicial e na aplicação de institutos jurídicos já consolidados, especialmente a doutrina do fair use, prevista na Seção 107 do Copyright Act. Esse instituto tem sido invocado como fundamento para justificar determinadas formas de utilização de obras protegidas durante o treinamento de modelos de Inteligência Artificial. Entretanto, o crescimento dos litígios envolvendo empresas desenvolvedoras evidencia que persistem incertezas quanto aos limites jurídicos da coleta e do processamento de grandes volumes de dados, sobretudo quando empregados em aplicações comerciais (Volpe, 2025).

Além da controvérsia sobre direitos autorais, o debate regulatório norte-americano reflete a tensão entre a necessidade de ampliar mecanismos de responsabilização das grandes empresas de tecnologia e a preocupação em preservar a competitividade do país no desenvolvimento da Inteligência Artificial. Esse cenário demonstra que a construção de um

marco regulatório envolve não apenas aspectos jurídicos, mas também interesses econômicos, tecnológicos e estratégicos relacionados à liderança na economia digital.

Em sentido diverso, a União Europeia estruturou um modelo regulatório baseado na transparência, na governança e na gestão de riscos. A Diretiva (UE) 2019/790 e o Regulamento (UE) 2024/1689 (AI Act) estabelecem deveres específicos de documentação, avaliação de riscos e transparência, buscando compatibilizar o desenvolvimento da Inteligência Artificial com a proteção dos direitos fundamentais e da propriedade intelectual (European Union, 2019; European Union, 2024).

A China, por sua vez, adotou um modelo caracterizado por maior intervenção estatal. As Interim Measures for the Management of Generative Artificial Intelligence Services impõem aos desenvolvedores obrigações relacionadas à segurança dos dados, à supervisão administrativa, à responsabilização dos provedores e à governança dos sistemas de Inteligência Artificial, evidenciando uma estratégia regulatória voltada ao controle preventivo dos riscos tecnológicos (China, 2023).

No Brasil, observa-se um estágio intermediário de construção normativa. Embora o ordenamento jurídico disponha de um sistema consolidado de proteção da propriedade intelectual, especialmente por meio das Leis nº 9.279/1996 e nº 9.610/1998, ainda não há disciplina específica para a utilização de ativos distintivos da propriedade industrial no treinamento de modelos de Inteligência Artificial Generativa. O Projeto de Lei nº 2.338/2023 representa importante avanço ao estabelecer princípios relacionados à transparência, à governança e à responsabilização, mas permanece sem disciplinar de forma detalhada a utilização de marcas, indicações geográficas e marcas de certificação como elementos integrantes dos datasets (Brasil, 1996; Brasil, 1998; Brasil, 2023; Ferreira; Chaves; Sousa, 2025).

A análise comparativa evidencia um ponto convergente entre as diferentes experiências regulatórias: a crescente preocupação com mecanismos de governança aplicáveis aos dados utilizados no treinamento dos sistemas de Inteligência Artificial. Todavia, também demonstra que a proteção dos ativos distintivos da propriedade industrial permanece em posição secundária quando comparada ao debate sobre direitos autorais, revelando uma lacuna regulatória comum às principais jurisdições analisadas.

Sob a perspectiva econômica, os resultados indicam que marcas, indicações geográficas e marcas de certificação passaram a desempenhar dupla função na economia digital. Além de identificar a origem empresarial, assegurar padrões de qualidade e diferenciar produtos e serviços, esses sinais distintivos passaram a integrar a infraestrutura informacional utilizada no

desenvolvimento de modelos de Inteligência Artificial. Essa mudança amplia sua relevância jurídica e econômica, especialmente no caso das indicações geográficas, cuja reputação resulta de investimentos coletivos e da valorização de territórios e conhecimentos tradicionais (Braman, 2011; Wipo, 2024).

Por fim, verificou-se que a ausência de parâmetros específicos amplia a insegurança jurídica tanto para os titulares de direitos quanto para os desenvolvedores de tecnologias. Enquanto os primeiros necessitam de mecanismos que assegurem transparência, rastreabilidade e controle sobre o uso de seus ativos, os segundos demandam critérios objetivos que permitam avaliar a licitude do emprego dessas informações no treinamento de modelos de Inteligência Artificial. Esse cenário reforça a necessidade de aperfeiçoar a governança dos dados e de construir um marco regulatório capaz de equilibrar inovação tecnológica, segurança jurídica e proteção da propriedade industrial.

8 CONCLUSÃO

A presente pesquisa demonstra que o debate jurídico envolvendo Inteligência Artificial e propriedade intelectual não pode permanecer restrito à tradicional discussão sobre reprodução de obras autorais. A evolução dos modelos de Inteligência Artificial Generativa revela uma transformação mais profunda: ativos distintivos da propriedade industrial, historicamente concebidos como instrumentos de identificação mercadológica, passaram também a assumir a condição de ativos informacionais utilizados na construção do conhecimento algorítmico. Marcas, indicações geográficas e marcas de certificação deixam, portanto, de representar apenas elementos de diferenciação no mercado e passam a integrar a infraestrutura de dados que sustenta a formação dos modelos computacionais contemporâneos.

A principal contribuição deste estudo consiste em revelar uma lacuna regulatória ainda pouco explorada pela doutrina brasileira: embora o ordenamento jurídico nacional estabeleça mecanismos consolidados de proteção desses ativos, especialmente por meio da Lei nº 9.279/1996, permanece ausente uma disciplina específica capaz de enfrentar sua utilização no treinamento de Inteligências Artificiais Generativas. O desafio jurídico contemporâneo não está apenas em impedir a reprodução indevida desses sinais distintivos, mas em compreender como sua reputação, valor econômico e identidade territorial podem ser incorporados por sistemas capazes de aprender, sintetizar e reproduzir padrões derivados de grandes volumes de dados.

Nesse novo cenário, a propriedade industrial ingressa em uma nova fronteira: a economia cognitiva dos dados. A proteção jurídica deixa de envolver exclusivamente o uso

comercial tradicional dos ativos e passa a alcançar sua participação como insumo estratégico no desenvolvimento tecnológico. Essa mudança exige uma releitura dos instrumentos tradicionais de tutela, incorporando mecanismos de transparência, rastreabilidade, governança dos dados e equilíbrio entre inovação e proteção dos titulares.

Dessa forma, o Brasil não deve limitar-se à reprodução automática dos modelos regulatórios estrangeiros. Embora as experiências dos Estados Unidos, União Europeia e China ofereçam importantes referências, a realidade brasileira exige uma abordagem própria, capaz de considerar a relevância econômica e cultural de seus ativos distintivos, especialmente as indicações geográficas e as marcas de certificação associadas a produtos territoriais, cadeias produtivas tradicionais e estratégias de desenvolvimento regional.

Mais do que regular a Inteligência Artificial, o desafio brasileiro consiste em definir como seus ativos intelectuais serão reconhecidos, protegidos e valorizados na nova economia baseada em dados. A construção de um marco regulatório equilibrado permitirá que o país participe da revolução tecnológica sem transformar sua biodiversidade, sua diversidade territorial, seus conhecimentos produtivos e sua reputação coletiva em simples recursos informacionais disponíveis para exploração algorítmica sem critérios jurídicos previamente estabelecidos.

AGRADECIMENTOS

Agradecemos ao Programa de Bolsas Institucionais de Pós-Graduação (ProBPG/UEMG) pelo apoio financeiro concedido para o desenvolvimento deste trabalho, por meio da bolsa de mestrado vinculada ao Programa de Pós-Graduação em Propriedade Intelectual e Transferência de Tecnologia para a Inovação (PROFNIT), na UEMG – Unidade Frutal. O apoio recebido foi fundamental para a realização desta pesquisa e para o desenvolvimento das atividades acadêmicas ao longo da formação.

REFERÊNCIAS

ARAÚJO, Carlos Alberto Ávila. **O que é Ciência da Informação?** Belo Horizonte: KMA, 2018. Disponível em: <https://teste.eci.ufmg.br/wp-content/uploads/2024/03/O-QUE-E-CIENCIA-DA-INFORMACAO.pdf>. Acesso em: 10 mar. 2025.

BARDIN, Laurence. **Análise de conteúdo**. São Paulo: Edições 70, 2011.

BRAMAN, Sandra. **Change of state: information, policy, and power**. Cambridge: MIT Press, 2006.

BRAMAN, Sandra. **A economia representacional e o regime global da política de informação.** In: MACIEL, Maria Lucia; ALBAGLI, Sarita (org.). *Informação, conhecimento e poder: mudança tecnológica e inovação social*. Rio de Janeiro: Garamond, 2011. p. 41-66. Disponível em:

https://www.livroaberto.ibict.br/bitstream/123456789/1062/2/informacao_conhecimento_e_poder.pdf. Acesso em: 1 mar. 2025.

BRASIL. **Constituição da República Federativa do Brasil de 1988.** Brasília, DF: Presidência da República, 1988. Disponível em:

https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 3 ago. 2026.

BRASIL. Lei nº 9.279, de 14 de maio de 1996. **Regula direitos e obrigações relativos à propriedade industrial.** Brasília, DF: Presidência da República, 1996. Disponível em:

https://www.planalto.gov.br/ccivil_03/leis/l9279.htm. Acesso em: 3 ago. 2026.

BRASIL. Lei nº 9.610, de 19 de fevereiro de 1998. **Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências.** Brasília, DF: Presidência da República, 1998. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l9610.htm. Acesso em: 3 ago. 2026.

BRASIL. Senado Federal. Projeto de Lei nº 2.338, de 2023. **Dispõe sobre o uso da Inteligência Artificial.** Brasília, DF: Senado Federal, 2023. Disponível em:

<https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>. Acesso em: 3 ago. 2026.

BROWN, Tom B. *et al.* **Language models are few-shot learners.** *Advances in Neural Information Processing Systems*, v. 33, 2020. Disponível em:

<https://arxiv.org/abs/2005.14165>. Acesso em: 3 ago. 2026.

CAMPOS, Rafael Saraiva. **Desmistificando a inteligência artificial: uma breve introdução conceitual ao aprendizado de máquina.** *Aoristo: International Journal of Phenomenology, Hermeneutics and Metaphysics*, Paraná, v. 3, n. 1, p. 106-123, 2020. Disponível em: <https://e-revista.unioeste.br/index.php/aoristo/article/view/24880>. Acesso em: 11 mar. 2025.

CHINA. **Interim Measures for the Management of Generative Artificial Intelligence Services.** Cyberspace Administration of China (CAC), 2023. Disponível em:

<https://www.cac.gov.cn/>. Acesso em: 3 ago. 2026.

EUROPEAN UNION. **Directive (EU) 2019/790 of the European Parliament and of the Council on copyright and related rights in the Digital Single Market.** *Official Journal of the European Union*, 2019. Disponível em: <https://eur-lex.europa.eu/eli/dir/2019/790/oj>. Acesso em: 3 ago. 2026.

EUROPEAN UNION. **Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending certain Union legislative acts (Artificial Intelligence Act).** *Official Journal of the European Union*, 2024. Disponível em: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>. Acesso em: 3 ago. 2026.

FERREIRA, Vanessa Claudia Alves; CHAVES, Italo Teixeira; SOUSA, Marckson Roberto Ferreira de. **Inteligência artificial no Brasil: observações a partir do Projeto de Lei nº 2.338/2023. Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 18, n. 2, p. 1-31, jul./dez. 2025.

FERREIRA, Vanessa Claudia Alves; SOUZA, Marckson Roberto Ferreira de. **Tomada de decisões algorítmicas: uma análise na perspectiva da Ciência da Informação considerando a Lei Geral de Proteção de Dados. Pesquisa Brasileira em Ciência da Informação e Biblioteconomia**, v. 18, n. 4, p. 111-119, 2023. Disponível em: <https://pbcib.com/index.php/pbcib/article/view/61890>. Acesso em: 10 mar. 2025.

GONZALES DE GOMES, Maria Nélida. **Novos cenários políticos para a informação. Ciência da Informação**, Brasília, v. 31, n. 1, p. 27-40, jan./abr. 2002. Disponível em: <https://revista.ibict.br/ciinf/article/view/975>. Acesso em: 10 mar. 2025.

MOTA, Denyson; MARTINS, Gracy Kelli; SAMPAIO, Denise Braga. **Vieses nas inteligências artificiais: um estudo sobre a geração de imagens a partir de comandos de raça/etnia e gênero. In: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO**, 22., 2022, Porto Alegre. *Anais [...]*. Porto Alegre: ANCIB, 2022. p. 1-16. Disponível em: <https://enancib.ancib.org/index.php/enancib/xxxiiienancib/paper/view/1852>. Acesso em: 10 mar. 2025.

OPENAI. **GPT-4 technical report**. San Francisco: OpenAI, 2023. Disponível em: <https://arxiv.org/abs/2303.08774>. Acesso em: 3 ago. 2026.

PINHEIRO, Marta Macedo Kerr. **Política de informação e estado informacional. In: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO**, 9., 2009, João Pessoa. *Anais [...]*. João Pessoa: ANCIB, 2009. p. 1-14. Disponível em: <https://brapci.inf.br/v/172862>. Acesso em: 14 mar. 2025.

RUSSELL, Stuart; NORVIG, Peter. **Inteligência artificial**. Rio de Janeiro: Elsevier, 2013.

SÁ-SILVA, Jackson Ronie; ALMEIDA, Cristóvão Domingos de; GUINDANI, Joel Felipe. **Pesquisa documental: pistas teóricas e metodológicas. Revista Brasileira de História & Ciências Sociais**, v. 1, n. 1, p. 1-15, 2009. Disponível em: <https://periodicos.furg.br/rbhcs/article/view/10351>. Acesso em: 10 mar. 2025.

SAMPIERI, Roberto Hernández; COLLADO, Carlos Fernández; LUCIO, María del Pilar Baptista. **Metodologia de pesquisa**. 5. ed. Porto Alegre: Penso, 2013.

SILVA, Edna Lúcia da; MENEZES, Estera Muszkat. **Metodologia da pesquisa e elaboração de dissertação**. 3. ed. Florianópolis: Laboratório de Ensino a Distância da UFSC, 2001.

TRINDADE, Alessandra Stefane Cândido Elias da; OLIVEIRA, Henry Poncio Cruz de. **Inteligência artificial (IA) generativa e competência em informação: habilidades informacionais necessárias ao uso de ferramentas de IA generativa em demandas informacionais de natureza acadêmica-científica. Perspectivas em Ciência da Informação**, Belo Horizonte, v. 29, e47485, 2024. Disponível em: <https://periodicos.ufmg.br/index.php/pci/article/view/47485>. Acesso em: 10 mar. 2025.

VAN DIJCK, José. **Confiemos nos dados?** As implicações da datificação para o monitoramento social. *Matrizes*, São Paulo, v. 11, n. 1, p. 39-59, 2017. Disponível em: <https://revistas.usp.br/matrizes/article/view/131620>. Acesso em: 10 mar. 2025.

VOLPE, M. C. S. **Direito autoral e aprendizagem de IA generativa: análise comparada Brasil e Estados Unidos no uso de obras protegidas para treinamento de IA.** *Revista Direito, Cidadania e Sociedade*, [S. l.], 2025. Disponível em: <https://ojs.revistadcs.com/index.php/revista/article/view/3880>. Acesso em: 3 ago. 2026.

WORLD INTELLECTUAL PROPERTY ORGANIZATION (WIPO). **Artificial intelligence and intellectual property policy.** Geneva: WIPO, 2020. Disponível em: https://www.wipo.int/about-ip/en/artificial_intelligence/. Acesso em: 3 ago. 2026.

WORLD INTELLECTUAL PROPERTY ORGANIZATION (WIPO). **Generative artificial intelligence: navigating intellectual property.** Geneva: WIPO, 2024. Disponível em: <https://www.wipo.int/>. Acesso em: 3 ago. 2026.