

PROPRIEDADE INTELECTUAL NA ERA DA GEOAI: AUTORIA, TITULARIDADE DE MODELOS E DIREITOS SOBRE DADOS DE TREINAMENTO

Hérciles Barbosa Lopes¹, Jesica Fatima Scopel², Dione Locks³, Daniela Asquidamini⁴,
Elciane Regina Zanatta⁵, Jairo Marlon Corrêa⁶

^{1,2,3,4} Discente do Programa de Pós-Graduação em Propriedade Intelectual e Transferência de Tecnologia para a Inovação (PROFNIT) da Universidade Tecnológica Federal do Paraná - UTFPR, Campus Medianeira, Medianeira, Paraná, Brasil, e-mail: hericlesbl@hotmail.com;

^{5,6} Docente da Universidade Tecnológica Federal do Paraná - UTFPR, Campus Medianeira, Medianeira, Paraná, Brasil, e-mail: elcianezanatta@utfpr.edu.br

Resumo: A difusão da GeoAI tensiona o arcabouço de propriedade intelectual, construído sobre autoria humana e obras discretas. Esta revisão narrativa (2015-2026; 154 fontes, 28 citadas) indica convergência quanto à autoria humana, indeterminação sobre titularidade de modelos e produtos, e deslocamento da proteção para contratos, licenças e soluções técnicas pouco conectado aos dados geoespaciais. Hipótese confirmada em grande medida; implicações para a Infraestrutura Nacional de Dados Espaciais (INDE), o PL 2.338/2023 e os Núcleos de Inovação Tecnológica (NITs).

Palavras-chave: inteligência artificial; dados geoespaciais; propriedade intelectual; direitos autorais.

INTELLECTUAL PROPERTY IN THE GEOAI ERA: AUTHORSHIP, MODEL OWNERSHIP AND RIGHTS OVER TRAINING DATA

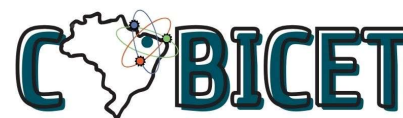
Abstract: The widespread adoption of GeoAI is challenging the intellectual property framework, which was built upon the concepts of human authorship and discrete works. This narrative review (2015–2026; 154 sources, 28 cited) identifies a convergence around the requirement of human authorship, persistent uncertainty regarding the ownership of models and their outputs, and a shift in protection toward contracts, licensing arrangements, and technical safeguards an approach that remains only weakly connected to geospatial data. The proposed hypothesis was largely confirmed, with implications for the Brazilian National Spatial Data Infrastructure (INDE), Bill No. 2,338/2023, and Technology Innovation Centers (NITs).

Keywords: artificial intelligence; geospatial data; intellectual property; copyright.

INTRODUÇÃO

A inteligência artificial (IA) generativa alcançou a ciência da informação geográfica. Em poucos anos, a GeoAI (do inglês *Geospatial Artificial Intelligence*, ou inteligência artificial geoespacial) aplicação de técnicas de IA espacialmente explícitas à descoberta

de conhecimento geográfico (JANOWICZ *et al.*, 2020) evoluiu de classificadores especializados a modelos de fundação multimodais (MAI *et al.*, 2024), modelos de linguagem integrados à geociência (WANG *et al.*, 2024), e a aplicações de alto impacto, como a saúde pública e ambiental (RESCH *et al.*, 2025), discute-se já a produção de mapas por sistemas



generativos, com implicações para a prática e a ética cartográfica, incluindo preocupações com alucinações e direitos autorais (KANG; WANG, 2025).

Esse salto colide com um pressuposto do direito de propriedade intelectual (PI): obras e invenções são criações humanas discretas, com autores identificáveis. Os sistemas generativos produzem ativos em cadeia, dados coletados e processados, modelos treinados, produtos derivados sem sujeito autoral evidente em cada elo (KRETSCHMER; MARGONI; ORUÇ, 2024). Nos Estados Unidos, litígios de grande porte disputam a licitude do uso de obras protegidas no treinamento (SAMUELSON, 2023), e o escritório de direitos autorais reafirma a autoria humana como requisito de proteção das saídas geradas por IA (UNITED STATES COPYRIGHT OFFICE, 2025). Na União Europeia, doutrina e legislação testam os limites de obra, originalidade e das exceções de mineração de textos e dados (HUGENHOLTZ; QUINTAIS, 2021; MARGONI; KRETSCHMER, 2022).

O domínio geoespacial adiciona camadas próprias. Sobre os dados espaciais incidem, de forma sobreposta, o direito autoral cartográfico, direitos sobre bases de dados, contratos e licenças, e a constituição dos mercados de informação geográfica depende de como esses direitos e a interoperabilidade legal e técnica são configurados institucionalmente (ALVAREZ LEÓN, 2024). Parte expressiva desses dados é produzida pelo setor público sob mandatos de abertura, em infraestruturas de dados espaciais, ou por colaboração voluntária, caso em que direitos e obrigações de hospedeiros, contribuidores e usuários se entrelaçam (SCASSA, 2013). Se um modelo é treinado com dados públicos e abertos, o modelo herda essa abertura ou pode ser apropriado privadamente? A pergunta ecoa a que já se fez sobre dados governamentais, em termos de soberania: quem é o dono do mapa? (JOHNSON; SCASSA, 2023).

Para o Brasil, o tema é atual e prático: a Lei nº 9.610/1998 vincula a autoria à pessoa física criadora (BRASIL, 1998b); a Lei nº 9.609/1998 rege a proteção do software (BRASIL, 1998a); a Lei Geral de Proteção de Dados Pessoais alcança dados de localização (BRASIL, 2018); e o Decreto nº 6.666/2008 instituiu a Infraestrutura Nacional de Dados Espaciais (INDE), voltada ao compartilhamento da geoinformação pública (BRASIL, 2008). O Projeto de Lei (PL) nº 2.338/2023, aprovado no Senado em dezembro de 2024 e em tramitação na Câmara, disputa pontos como a mineração de textos e dados e a remuneração de titulares de direitos autorais. Órgãos públicos produtores de geoinformação, universidades e Núcleos de Inovação Tecnológica (NIT) precisam decidir, sob incerteza regulatória, como licenciar

dados, proteger modelos e registrar softwares (CASTRO *et al.*, 2020; CAVALCANTE *et al.*, 2025).

A produção científica sobre essas questões está dispersa em três literaturas que pouco dialogam: a dos juristas na interseção IA e PI; a dos geocientistas que desenvolvem a GeoAI; e a da comunidade de infraestruturas de dados espaciais, dedicada ao licenciamento de geodados. mapeamentos bibliométricos confirmam o crescimento constante embora não acentuado do primeiro grupo, sem conexão com as especificidades geoespaciais (VALENTE; SARTORI; MARIN, 2023); não se identificou revisão que integrasse as três vertentes, lacuna que este artigo começa a preencher.

Nesse contexto, o objetivo geral deste estudo é sintetizar criticamente as evidências e proposições da literatura sobre a gestão da propriedade intelectual em *pipelines* de inteligência artificial geoespacial, dados de treinamento, modelos e produtos gerados no período de 2015 a 2026, com implicações para o marco regulatório brasileiro. Quatro questões norteiam a revisão: (i) como a literatura caracteriza a adequação (ou inadequação) dos institutos vigentes de PI diante de conteúdos gerados por inteligência artificial, e que critérios propõe para distinguir obras assistidas de obras geradas autonomamente?; (ii) o que se estabelece sobre os direitos incidentes e as condições de uso de dados de treinamento, em especial dados geoespaciais públicos e abertos, incluindo exceções de mineração de textos e dados, consentimento e compensação?; (iii) como se resolve, ou permanece irresoluta, a titularidade sobre modelos de GeoAI e sobre seus produtos, e que soluções de licenciamento são propostas para *pipelines* que integram múltiplas fontes?; (iv) que soluções técnico-institucionais e arranjos de governança emergem, e quais suas implicações para o contexto regulatório brasileiro?

Do objetivo geral decorrem cinco objetivos específicos: (i) caracterizar a avaliação da literatura sobre a adequação dos institutos de PI a conteúdos gerados por IA e sobre o requisito de autoria humana; (ii) sistematizar o debate sobre dados de treinamento, com atenção às especificidades dos dados geoespaciais; (iii) mapear as posições sobre a titularidade de modelos e produtos e as soluções de licenciamento para *pipelines* multifonte; (iv) identificar soluções técnicas e institucionais de proveniência, rastreabilidade e gestão de direitos aplicáveis a ecossistemas geoespaciais; e (v) extrair implicações para o Brasil e para a atuação de NITs e gestores públicos de geoinformação.

Sustenta-se, como hipótese de trabalho, que o arcabouço vigente de PI estruturado sobre a autoria humana e obras e invenções discretas não responde adequadamente à cadeia de valor da GeoAI (dados de

treinamento, modelo, produto gerado), produzindo incerteza jurídica que desloca a proteção efetiva dos institutos clássicos de PI para instrumentos contratuais, licenças e soluções técnicas de proveniência e rastreabilidade movimento que a literatura registra de forma crescente, porém fragmentada e pouco conectada às especificidades dos dados geoespaciais. A hipótese é examinada, na discussão, em três componentes: o diagnóstico de inadequação; o deslocamento contratual e técnico; e a desconexão com o domínio geoespacial.

Além desta introdução, o artigo organiza-se em quatro seções: percurso metodológico; resultados, com panorama do corpus e síntese por eixos e discussão, com retorno à hipótese e implicações para o Brasil; e considerações finais. A planilha de seleção de fontes acompanha o artigo como material suplementar, documentando as 154 fontes do corpus preliminar, os critérios de inclusão e exclusão aplicados a cada uma.

MATERIAL E MÉTODOS

Trata-se de revisão narrativa de literatura, desenho apropriado para descrever e discutir o estado da arte

de um tema amplo, em perspectiva teórica e contextual (ROTHER, 2007), com as salvaguardas de transparência recomendadas: questões norteadoras explícitas, critérios de seleção declarados e registro documentado das decisões (GREEN; JOHNSON; ADAMS, 2006; VOSGERAU; ROMANOWSKI, 2014), sem pretensão de exaustividade. Cobre o período de 2015 a 2026, da consolidação do aprendizado profundo ao ciclo recente de litígios e reformas regulatórias, com admissão pontual de clássicos anteriores do domínio geoespacial; prioriza o inglês, admitindo português e espanhol; escopo global, com atenção ao Brasil. O núcleo do corpus restringiu-se a artigos revisados por pares em veículos indexados (Scopus, Web of Science, SciELO, PubMed); *preprints*, capítulos de livro e documentos institucionais foram admitidos apenas como evidência complementar.

A triagem aplicou quatro critérios de inclusão e três de exclusão, conforme apresentado na Figura 1, sobre título, resumo e, quando necessário, texto completo.

Figura 1 — Critérios de elegibilidade aplicados à seleção das fontes do corpus: quatro critérios de inclusão (I1 a I4) e três de exclusão (E1 a E3), utilizados para triar as 154 fontes preliminares a partir de título, resumo e, quando necessário, texto completo.

Critérios de inclusão		Critérios de exclusão	
I1	Interseção IA × PI: abordar substantivamente autoria, titularidade, direitos sobre dados de treinamento ou licenciamento em sistemas de IA, especialmente generativa.	E1	Vagueza metodológica: ausência de descrição transparente de fontes ou procedimentos.
I2	Dimensão geoespacial ou transponibilidade: tratar de dados/sistemas geoespaciais (SIG, sensoriamento remoto, cartografia, IDE) ou oferecer fundamento de PI-IA transponível ao pipeline geoespacial.	E2	Conteúdo não acadêmico: textos sem revisão por pares (admissíveis somente como evidência complementar sinalizada).
I3	Densidade analítica: análise jurídica, técnica, institucional ou empírica com procedimentos transparentes — não menção incidental.	E3	Domínio irrelevante: IA sem dimensão de PI; geotecnologia sem IA nem PI; PI sem conexão com IA ou dados.
I4	Contribuição às questões norteadoras: aportar evidência ou proposição a pelo menos uma das quatro questões da revisão.		

Fonte: Autoria própria (2026).

O corpus preliminar totalizou 154 fontes, classificadas em eixos temáticos com a seguinte nomenclatura: C1 (autoria e criatividade na era da IA), C2 (titularidade de modelos e conteúdos gerados), C3 (dados de treinamento e mineração de textos e dados), C4 (propriedade intelectual e dados geoespaciais), C5 (capacidades da GeoAI), C6 (soluções técnico-institucionais), e, C7 (governança e contexto brasileiro). Dessas, 28 obras foram selecionadas para citação, por densidade analítica, verificabilidade e cobertura equilibrada dos eixos — priorizando periódicos indexados e sinalizando as evidências complementares. A síntese é qualitativa, organizada por eixos e orientada pelas questões norteadoras.

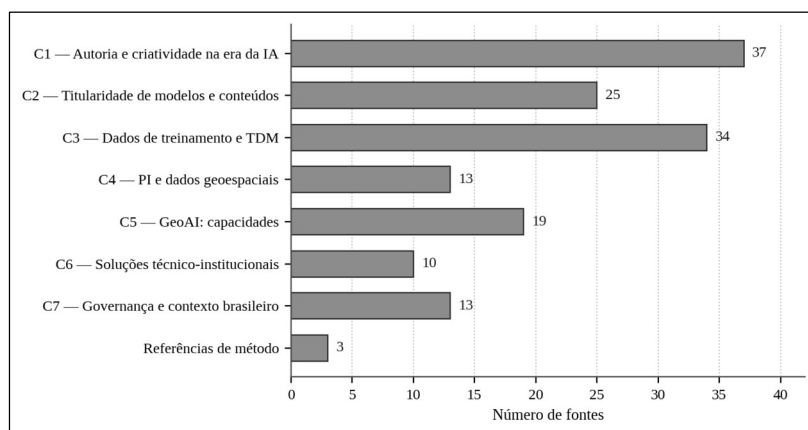
RESULTADOS E DISCUSSÃO

A Figura 2 apresenta a distribuição por eixo temático: os eixos jurídicos da interseção IA × PI (C1 a C3) concentram 96 fontes, contra 42 dos eixos de ponte com o domínio geoespacial (C4 a C6) medida objetiva da desconexão que a hipótese antecipa. Entre as 28 obras citadas nesta síntese predominam as análises jurídico-doutrinárias e comparadas, ao lado de revisões e mapeamentos, editoriais e artigos de visão do campo técnico, estudos de caso, uma proposta técnica com protótipo e um documento institucional. Esse panorama dimensiona a base de evidência com que os cinco objetivos específicos são enfrentados nas subseções seguintes: a adequação dos institutos de PI

e o requisito de autoria humana (objetivo i), no eixo C1; o debate sobre dados de treinamento (objetivo ii), nos eixos C3 e C4; a titularidade de modelos e produtos e as soluções de licenciamento (objetivo iii), no eixo C2; as soluções técnicas e institucionais de

proveniência e gestão de direitos (objetivo iv), nos eixos C5 e C6; e as implicações para o Brasil e para a atuação de NITs e gestores públicos (objetivo v), no eixo C7 e nas subseções que encerram esta seção.

Figura 2 — Distribuição das 154 fontes do corpus preliminar pelos sete eixos temáticos da revisão: C1 (autoria e criatividade na era da IA), C2 (titularidade de modelos e conteúdos gerados), C3 (dados de treinamento e mineração de textos e dados), C4 (propriedade intelectual e dados geoespaciais), C5 (capacidades da GeoAI), C6 (soluções técnico-institucionais) e C7 (governança e contexto brasileiro).



Fonte: Autoria própria (2026).

Os resultados são apresentados a seguir por eixo temático, com leitura transversal orientada pelas quatro questões norteadoras.

Autoria e criatividade na era da inteligência artificial (C1)

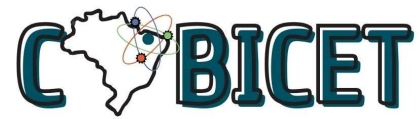
O primeiro eixo responde à questão (i) e apresenta a convergência mais nítida da literatura: a manutenção do requisito de autoria humana, acompanhada de uma zona cinzenta operacional entre a obra assistida e a obra gerada autonomamente. No plano doutrinário europeu, Hugenholtz e Quintais (2021) sustentam que o acervo da União Europeia é suficientemente flexível para acomodar produções assistidas por IA, desde que a saída reflita escolhas criativas humanas em qualquer fase do processo criativo concepção, execução ou redação final; a proteção como obra depende, em um teste de quatro etapas próprio, de que a produção esteja no domínio literário, científico ou artístico, envolva esforço intelectual humano, revele originalidade e se expresse na produção final. No plano institucional norte-americano, o relatório do United States Copyright Office (2025) reafirma que apenas contribuições humanas são protegíveis, que a mera formulação de comandos (*prompts*) não confere, por si, o controle criativo exigido, e que o uso assistivo de ferramentas não descaracteriza a proteção com aferição caso a caso.

A distinção entre assistido e autônomo, contudo, resiste à operacionalização. A partir de nove casos,

Zdravkova (2025) conclui que obras apoiadas por IA podem ser atribuídas à pessoa que as conduziu, ao passo que produções competentemente autônomas não encontram abrigo em nenhuma forma vigente de propriedade intelectual. Na doutrina brasileira, Divino e Magalhães (2020) chegam a resultado convergente por via ontológica: à luz do art. 11 da Lei nº 9.610/1998, a autoria é atributo da pessoa natural, sendo impossível reconhecê-la no sistema o que desloca a discussão para o programador ou criador do sistema, a quem a fonte atribui os direitos quando existentes. Transposta ao domínio geoespacial, a questão ganha materialidade: o mapa derivado de um modelo generativo, com intervenção humana limitada à consulta, tende a situar-se justamente na faixa de proteção duvidosa inquietação que a própria comunidade cartográfica começa a registrar (KANG; WANG, 2025).

Titularidade de modelos e conteúdos gerados (C2)

O segundo eixo responde à questão (iii) e concentra a indeterminação mais consequente. Kretschmer, Margoni e Oruç (2024) percorrem o ciclo de vida dos modelos de aprendizado de máquina coleta de dados, organização dos dados, treinamento do modelo e operação do modelo e mostram que os instrumentos que governam cada elo têm natureza ora legal (exceções e limitações), ora contratual (licenças individuais, públicas ou coletivas), frequentemente combinadas. O achado central é o efeito da incerteza acumulada na entrada: em razão do estatuto jurídico



incerto da coleta e do processamento, somado à vantagem competitiva de não revelar avanços tecnológicos, os insumos dos modelos em operação são, com frequência, desconhecidos de modo que a proteção efetiva se materializa na opacidade e no contrato, mais do que nos institutos clássicos de propriedade intelectual.

No domínio geoespacial, a pergunta pela titularidade precede a IA e já se punha quanto aos próprios dados governamentais. Johnson e Scassa (2023) identificam quatro modelos pelos quais governos lidam com a soberania de seus dados espaciais do controle integral à contratação de terceiros com retenção pública da titularidade, da aquisição sob termos do coletor à renúncia, alertando que a perda de controle compromete o uso da geoinformação no interesse público. O treinamento de modelos privados com dados espaciais públicos radicaliza esse dilema: na ausência de regra específica, a resposta tem sido dada por contratos e termos de uso, e não pelos institutos de propriedade intelectual em linha com o deslocamento diagnosticado no plano geral.

Dados de treinamento e mineração de textos e dados (C3)

O terceiro eixo responde à questão (ii). Samuelson (2023) sintetiza o que está em jogo nos litígios norte-americanos entre eles o movido pela Getty Images contra a Stability AI: se prevalecer a tese dos autores, apenas sistemas treinados com obras em domínio público ou sob licença seriam lícitos, com efeitos em cadeia sobre a pesquisa e a indústria. Em perspectiva comparada, de la Durantaye (2025) percorre a legislação e o debate de política em oito jurisdições (Estados Unidos, Canadá, Reino Unido, União Europeia, Israel, China, Singapura e Japão) e identifica quatro preocupações dominantes: controle dos titulares, compensação, transparência e segurança jurídica. A transparência sobre os conjuntos de treinamento converteu-se em candidata a solução transversal, embora sozinha, não resolva a questão de fundo, como adverte Buick (2025). O contencioso, por fim, ganha corpo em estudos de caso, como o exame do litígio da Getty Images, em que a aparição de marcas-d'água nas saídas geradas se tornou elemento probatório central (KANAYA; SANTOSO, 2025), e em propostas de saída pela via da gestão coletiva de direitos, a exemplo do modelo híbrido de licenciamento coletivo de Shamov (2025).

A leitura geoespacial desse debate revela um desencontro. As exceções de mineração e o mecanismo de *opt-out* foram desenhados para obras textuais e audiovisuais dispersas; os geodados, diferentemente, circulam sob licenças abertas padronizadas com cláusulas de atribuição e, por vezes, de compartilhamento pela mesma licença cujas obrigações não se sabe se sobrevivem à passagem pelo

treinamento. Ademais, camadas de dados colaborativos acrescentam a dimensão do consentimento dos contribuidores, já problematizada por Scassa (2013) antes da era generativa. O corpus não registra estudo que aplique sistematicamente o regime de mineração de textos e dados às licenças de geodados.

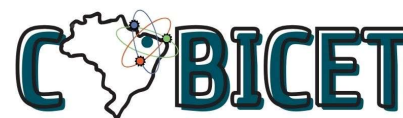
Propriedade intelectual e dados geoespaciais (C4)

O quarto eixo fornece os fundamentos pré-IA sobre os quais o problema contemporâneo se assenta. Scassa (2013) decompõe os papéis de hospedeiro, contribuidor e usuário na informação geográfica voluntária e mostra que direitos e obrigações da propriedade intelectual à responsabilidade e à privacidade variam conforme a posição na cadeia, de modo que a incerteza sobre direitos de terceiros condiciona o reuso. Alvarez León (2024), por sua vez, reconstrói a formação do mercado de informação geográfica nos Estados Unidos e identifica seus fundamentos: o desenvolvimento da interoperabilidade legal e técnica na coleta.

Formulados para o compartilhamento e para o mercado de dados, esses diagnósticos aplicam-se com força redobrada aos *pipelines* de GeoAI, que combinam dezenas de fontes sob termos heterogêneos e produzem artefatos modelos e saídas cuja relação com as licenças de origem não foi regulada. A interoperabilidade legal, condição histórica das infraestruturas de dados espaciais, torna-se condição também da GeoAI juridicamente sustentável.

Capacidades da GeoAI que tensionam a propriedade intelectual (C5)

O quinto eixo caracteriza o objeto técnico. O editorial fundador de Janowicz *et al.* (2020) delimitou a GeoAI como o desenvolvimento de técnicas de inteligência artificial espacialmente explícitas e fixou sua agenda da resposta a perguntas geográficas à reprodutibilidade e à ética. O salto recente está nos modelos de fundação: Mai *et al.* (2024) avaliam modelos existentes em sete tarefas geoespaciais, constatando que modelos de linguagem de propósito geral, em regime *zero-shot* ou *few-shot*, já superam modelos totalmente supervisionados em várias dessas tarefas embora não em todas, e discutem os riscos e desafios do desenvolvimento de modelos de fundação propriamente geoespaciais, treinados sobre grandes volumes de dados multimodais. A revisão sistemática de Wang *et al.* (2024) confirma a rápida incorporação de modelos generativos aos fluxos da geociência, e o editorial de Resch *et al.* (2025) registra a chegada dessa onda a um domínio sensível a saúde pública e ambiental, no qual desafios de governança de dados se somam aos de desempenho. Na cartografia, a produção generativa de mapas deixa de ser hipótese e passa a agenda de pesquisa, com implicações éticas



explícitas entre elas as de direitos autorais (KANG; WANG, 2025).

Para os fins desta revisão, o essencial é que cada capacidade descrita treinar modelos de fundação com dados multifonte, gerar mapas e camadas sintéticas, integrar dados sensíveis produz exatamente os ativos cujo estatuto jurídico os eixos C1 a C3 mostraram ser incerto. A literatura técnica, contudo, trata a propriedade intelectual de forma periférica: aparece, quando muito, como preocupação ética pontual não como eixo de análise, e a agenda dos modelos de fundação geoespaciais discute domínios de aplicação, tipos de dado e papéis do modelo (purificador, conversor, gerador, raciocinador), mas raramente licenças e titularidade.

Soluções técnico-institucionais (C6)

O sexto eixo responde à primeira parte da questão (iv). No plano dos dados, o levantamento de Šarčević *et al.* (2024) organiza taxonomicamente as defesas técnicas disponíveis contra usos indevidos em inteligência artificial generativa das marcas-d'água aos testes de pertencimento a conjuntos de treinamento (*membership inference*), evidenciando um campo em rápida estruturação. No plano da infraestrutura, Zhang *et al.* (2020) propuseram e prototiparam um modelo descentralizado de gestão de direitos digitais para dados espaciais, baseado em contratos inteligentes *Ethereum* e no sistema de arquivos distribuído IPFS, dispensando o terceiro de confiança. Não por acaso, essas técnicas já transbordam para o contencioso: no litígio da Getty Images, a aparição de marcas-d'água nas saídas geradas opera como evidência (KANAYA; SANTOSO, 2025).

O conjunto sugere que parte da resposta ao vácuo normativo é tecnológica e dialoga com ativos tradicionais do domínio geoespacial em especial os metadados de linhagem e proveniência das infraestruturas de dados espaciais, ponto de encaixe natural para as novas técnicas de rastreamento. Essa convergência, porém, ainda não foi sistematizada pela literatura.

Governança e contexto brasileiro (C7)

O sétimo eixo responde à segunda parte da questão (iv). A produção brasileira acompanha o debate internacional sobretudo pela via do direito autoral: Castro *et al.* (2020) examinam os aspectos jurídicos e tecnológicos do uso de ferramentas de IA sob a Lei nº 9.610/1998, e o mapeamento de Valente, Sartori e Marin (2023) confirma o crescimento constante embora não acentuado, e ainda pouco consolidado da produção científica na interseção entre inteligência artificial e direitos autorais. No plano regulatório, Moreira, Silva e Santos (2024) discutem a arquitetura do Projeto de Lei nº 2.338/2023, cuja tramitação mantém em aberto justamente os dispositivos sobre

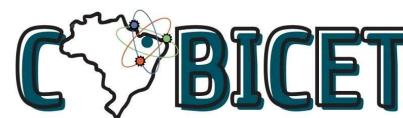
mineração de textos e dados e remuneração de titulares (BRASIL, 2023). No plano institucional, Cavalcante *et al.* (2025) documentam as barreiras enfrentadas por instituições de ciência e tecnologia no registro e na transferência de softwares acadêmicos sob a Lei nº 9.609/1998 retrato do aparato que, na ausência de figura própria, tende a ser mobilizado também para os modelos de IA desenvolvidos em ambiente acadêmico.

O que a pesquisa não registra é a conexão dessas discussões com a geoinformação pública. A INDE, instituída pelo Decreto nº 6.666/2008 com mandato de compartilhamento (BRASIL, 2008), e a incidência da Lei Geral de Proteção de Dados sobre dados de localização (BRASIL, 2018) permanecem à margem do debate de PI e IA no país. Não se localizou estudo brasileiro que enfrente a pergunta central deste artigo o estatuto de modelos treinados com os dados abertos da INDE, o que configura lacuna nacional específica.

A hipótese de trabalho comportava três componentes. O primeiro a inadequação do arcabouço vigente à cadeia de valor da GeoAI encontra amplo suporte. O requisito de autoria humana é mantido pelas instâncias doutrinárias e institucionais (HUGENHOLTZ; QUINTAIS, 2021; UNITED STATES COPYRIGHT OFFICE, 2025; DIVINO; MAGALHÃES, 2020), mas sua aplicação à produção assistida permanece casuística (ZDRAVKOVA, 2025); a opacidade sobre os próprios insumos dos modelos, sustentada pela vantagem competitiva do sigilo, deixa o elo intermediário do *pipeline* sem resposta estável (KRETSCHMER; MARGONI; ORUÇ, 2024); e o regime dos dados de treinamento segue disputado em litígios e reformas (SAMUELSON, 2023; DE LA DURANTAYE, 2025). Nenhum elo do *pipeline* dado, modelo, produto conta com resposta estável.

O segundo componente o deslocamento da proteção efetiva para contratos, licenças e soluções técnicas também se confirma. É a conclusão central da análise do ciclo de vida dos modelos (KRETSCHMER; MARGONI; ORUÇ, 2024), reaparece na aposta regulatória em transparência e licenciamento (BUICK, 2025) e materializa-se nas soluções técnicas de marca-d'água, gestão descentralizada de direitos e verificação de pertencimento (ZHANG *et al.*, 2020; ŠARČEVIĆ *et al.*, 2024; KANAYA; SANTOSO, 2025).

O terceiro componente a desconexão entre esse movimento e as especificidades geoespaciais confirma-se com uma nuance. De um lado, a literatura técnica da GeoAI trata a PI periféricamente (MAI *et al.*, 2024; WANG *et al.*, 2024), e a literatura jurídica ignora as camadas próprias dos geodados licenças abertas padronizadas, direito de bases de dados, mandatos públicos de abertura descritas pelo eixo C4. De outro, existem pontes: a pergunta pela titularidade



dos dados espaciais governamentais (JOHNSON; SCASSA, 2023), a tradição de proteção técnica de geodados (ZHANG *et al.*, 2020) e a preocupação emergente da cartografia com autoria (KANG; WANG, 2025). A hipótese resulta, portanto, confirmada em grande medida: o deslocamento contratual-técnico é real e crescente, e a conexão geoespacial, embora não inexistente, é fragmentária e não sistematizada o espaço exato desta agenda de pesquisa.

Implicações para o Brasil

Três implicações práticas decorrem da síntese. A primeira dirige-se aos produtores públicos de geoinformação: enquanto o marco legal da IA não se estabiliza (BRASIL, 2023; MOREIRA; SILVA; SANTOS, 2024), o instrumento disponível é a própria licença dos dados. Termos de uso da geoinformação pública podem explicitar as condições de utilização em treinamento de modelos inclusive obrigações de atribuição, transparência e não exclusividade, convertendo em cláusula aquilo que a lei ainda não resolve, na linha do deslocamento contratual diagnosticado. A experiência internacional de estruturação de mercados de geoinformação sobre interoperabilidade legal e regimes claros de propriedade intelectual (ALVAREZ LEÓN, 2024) oferece parâmetros; a INDE (BRASIL, 2008), por sua vez, constitui o marco institucional correspondente.

A segunda questão envolve os Núcleos de Inovação Tecnológica (NITs). Os modelos de GeoAI produzidos por universidades e instituições de pesquisa tendem a ser protegidos por meio do regime jurídico do software (BRASIL, 1998a) e do segredo empresarial. A interação entre essas duas formas de proteção, entretanto, já apresenta conflitos conhecidos (CAVALCANTE *et al.*, 2025). Políticas institucionais de PI que tratem expressamente de conjuntos de dados, modelos treinados e saídas geradas inclusive quanto à proveniência dos dados de treinamento reduzem a incerteza nas parcerias e na transferência de tecnologia. A terceira implicação é legislativa: o desenho final dos dispositivos de mineração de textos e dados do PL nº 2.338/2023 deveria considerar o caso dos dados públicos abertos evitando tanto o fechamento da geoinformação pública quanto a apropriação privada silenciosa de modelos nela treinados, riscos simétricos apontados pela literatura de soberania de dados (JOHNSON; SCASSA, 2023).

Três limitações devem ser explicitadas. Primeira, trata-se de revisão narrativa: a despeito das salvaguardas de transparência adotadas, não há pretensão de exaustividade nem de replicabilidade estrita (ROTHER, 2007; GREEN; JOHNSON; ADAMS, 2006). Segunda, o objeto está em movimento litígios em curso e projetos de lei em tramitação podem alterar o quadro descrito; o corte da

síntese é julho de 2026. Terceira, a pesquisa reflete vieses de indexação e de idioma das ferramentas de busca utilizadas, com predominância do inglês e sub-representação possível de literaturas locais.

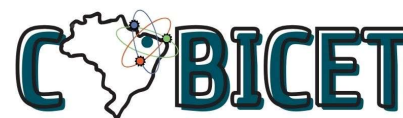
CONCLUSÃO

Este artigo perguntou a quem pertencem, e sob quais condições podem ser protegidos, licenciados e reutilizados, os ativos gerados ao longo do pipeline da GeoAI. Quanto à adequação dos institutos de propriedade intelectual a conteúdos gerados por IA, a revisão permite concluir que a literatura converge na manutenção da autoria humana como requisito de proteção: a produção assistida é examinada caso a caso, e a produção autônoma não encontra abrigo em nenhuma forma vigente de propriedade intelectual. Quanto ao debate sobre dados de treinamento, conclui-se que esse é o front mais disputado do pipeline: as exceções de mineração de textos e dados, o consentimento e a compensação estão no centro de litígios e reformas, sem que suas categorias tenham sido testadas contra as licenças abertas típicas dos geodados. Quanto à titularidade de modelos e produtos, conclui-se que ela permanece juridicamente indeterminada, sendo resolvida, na prática, por contratos e segredo e não pelos institutos clássicos de propriedade intelectual. Quanto às soluções técnicas e institucionais de proveniência, rastreabilidade e gestão de direitos, conclui-se que despontam como resposta complementar, com raízes antigas no domínio geoespacial, ainda que dispersas e não sistematizadas. Quanto às implicações para o Brasil e para a atuação dos NITs e dos gestores públicos de geoinformação, conclui-se que o debate legislativo brasileiro segue aberto e desconectado da geoinformação pública: não se localizou estudo nacional que enfrente o estatuto de modelos treinados com os dados abertos da INDE.

CONSIDERAÇÕES PARA PESQUISAS FUTURAS

Três frentes de pesquisa decorrem diretamente dos resultados deste artigo. Primeira, o exame empírico das licenças praticadas pelas instituições que compõem a INDE, verificando se seus termos de uso autorizam, vedam ou silenciam sobre o treinamento de modelos de inteligência artificial. Segunda, a avaliação comparativa de técnicas de proveniência e marca-d'água aplicadas a produtos cartográficos generativos, com atenção à sua robustez probatória em eventual litígio. Terceira, o acompanhamento do texto final do Projeto de Lei nº 2.338/2023 e de seus efeitos concretos sobre os ecossistemas geoespaciais, em especial os dispositivos sobre mineração de textos e dados e remuneração de titulares.

DECLARAÇÃO DE USO DE INTELIGÊNCIA ARTIFICIAL



Em atenção à transparência, declara-se que a ferramenta de Inteligência Artificial ChatGPT 5 foi empregada como apoio em etapas específicas da elaboração deste trabalho, a saber: a organização e a conferência das referências bibliográficas, apoio na análise de dados, melhoria da aparência das figuras e o apoio à redação e ao refinamento linguístico do texto. As decisões metodológicas, a interpretação dos resultados e a verificação das informações, dos dados e das citações foram conduzidas e revisadas integralmente pelos autores, que assumem total responsabilidade pelo conteúdo científico apresentado.

MATERIAL SUPLEMENTAR

Os dados brutos desta pesquisa estão disponíveis para consulta:

<https://docs.google.com/spreadsheets/d/1plpx2kSiHebRHpICY4eTAgdMkdzqiSaN/edit?usp=sharing&ouid=107582554004621766514&rtopof=true&sd=true>

REFERÊNCIAS

ALVAREZ LEÓN, L. F. Assembling the geographic information market in the United States. In: GLÜCKLER, J.; PANITZ, R. (ed.). **Knowledge and digital technology**. Cham: Springer, 2024. p. 131-152. (Knowledge and Space, v. 19). DOI: 10.1007/978-3-031-39101-9_7.

BRASIL. Lei nº 9.609, de 19 de fevereiro de 1998. Dispõe sobre a proteção da propriedade intelectual de programa de computador, sua comercialização no País, e dá outras providências. **Diário Oficial da União**: Brasília, DF, 20 fev. 1998a.

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

BRASIL. Decreto nº 6.666, de 27 de novembro de 2008. Institui, no âmbito do Poder Executivo federal, a Infraestrutura Nacional de Dados Espaciais — INDE, e dá outras providências. **Diário Oficial da União**: Brasília, DF, 28 nov. 2008.

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. Câmara dos Deputados. **Projeto de Lei nº 2.338, de 2023**. Dispõe sobre o uso da inteligência artificial. Brasília, DF: Câmara dos Deputados, 2023. Em tramitação.

BUICK, A. Copyright and AI training data — transparency to the rescue? **Journal of Intellectual Property Law & Practice**, Oxford, v. 20, n. 3, 2025. DOI: 10.1093/jiplp/jpae102.

CASTRO, C. F. P.; OLIVEIRA, J. A.; ARAÚJO, L. B.; PINHEIRO, L. A. O direito autoral e o uso de

ferramentas de inteligência artificial: aspectos jurídicos e tecnológicos. **Cadernos de Prospecção**, Salvador, v. 13, n. 4, p. 989, 2020. DOI: 10.9771/cp.v13i4.32551.

CAVALCANTE, M. F. A. F. *et al.* Propriedade intelectual e transferência de tecnologia em softwares acadêmicos: barreiras no registro. **Contribuciones a las Ciencias Sociales**, São José dos Pinhais, v. 18, n. 10, p. 1-23, 2025. DOI: 10.55905/revconv.18n.10-369.

DE LA DURANTAYE, K. Control and compensation: a comparative analysis of copyright exceptions for training generative AI. **IIC — International Review of Intellectual Property and Competition Law**, v. 56, p. 737-770, 2025. DOI: 10.1007/s40319-025-01569-6.

DIVINO, S. B. S.; MAGALHÃES, R. A. Propriedade intelectual e direito autoral de produção autônoma da inteligência artificial. **Revista Direitos e Garantias Fundamentais**, Vitória, v. 21, n. 1, 2020. DOI: 10.18759/rdgf.v21i1.1537.

GREEN, B. N.; JOHNSON, C. D.; ADAMS, A. Writing narrative literature reviews for peer-reviewed journals: secrets of the trade. **Journal of Chiropractic Medicine**, v. 5, n. 3, p. 101-117, 2006. DOI: 10.1016/S0899-3467(07)60142-6.

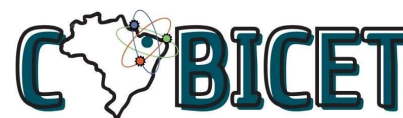
HUGENHOLTZ, P. B.; QUINTAIS, J. P. Copyright and artificial creation: does EU copyright law protect AI-assisted output? **IIC — International Review of Intellectual Property and Competition Law**, v. 52, n. 9, p. 1190-1216, 2021. DOI: 10.1007/s40319-021-01115-0.

JANOWICZ, K.; GAO, S.; McKENZIE, G.; HU, Y.; BHADURI, B. GeoAI: spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. **International Journal of Geographical Information Science**, v. 34, n. 4, p. 625-636, 2020. DOI: 10.1080/13658816.2019.1684500.

JOHNSON, P.; SCASSA, T. Who owns the map? Data sovereignty and government spatial data collection, use, and dissemination. **Transactions in GIS**, v. 27, p. 275-289, 2023. DOI: 10.1111/tgis.13024.

KANAYA, A. L.; SANTOSO, B. Analysis of copyright implementation in Getty Images and Stability AI as a case study of generative AI. **International Journal of Social Science and Human Research**, v. 8, n. 9, p. 7123-7129, 2025. DOI: 10.47191/ijsshr/v8-i9-49.

KANG, Y.; WANG, C. Envisioning generative artificial intelligence in cartography and mapmaking. **arXiv**, 2025. DOI: 10.48550/arXiv.2508.09028. Preprint.



KRETSCHMER, M.; MARGONI, T.; ORUÇ, P. Copyright law and the lifecycle of machine learning models. **IIC — International Review of Intellectual Property and Competition Law**, v. 55, n. 1, p. 110-138, 2024. DOI: 10.1007/s40319-023-01419-3.

MAI, G. *et al.* On the opportunities and challenges of foundation models for GeoAI (vision paper). **ACM Transactions on Spatial Algorithms and Systems**, v. 10, n. 2, art. 11, p. 1-46, 2024. DOI: 10.1145/3653070.

MARGONI, T.; KRETSCHMER, M. A deeper look into the EU text and data mining exceptions: harmonisation, data ownership, and the future of technology. **GRUR International**, v. 71, n. 8, p. 685-701, 2022. DOI: 10.1093/grurint/ikac054.

MOREIRA, A. E. P.; SILVA, T.; SANTOS, T. M. Regulamentação da inteligência artificial: desafios para a proteção dos direitos autorais. **Revista de Direitos Humanos e Desenvolvimento Social**, Campinas, v. 5, 2024. DOI: 10.24220/2675-9160v5a2024e13772.

RESCH, B.; KOLOKOUSSIS, P.; HANNY, D.; BROVELLI, M. A.; KAMEL BOULOS, M. N. The generative revolution: AI foundation models in geospatial health — applications, challenges and future research. **International Journal of Health Geographics**, v. 24, art. 6, 2025. DOI: 10.1186/s12942-025-00391-0.

ROTHER, E. T. Revisão sistemática x revisão narrativa. **Acta Paulista de Enfermagem**, São Paulo, v. 20, n. 2, p. v-vi, 2007. DOI: 10.1590/S0103-21002007000200001.

SAMUELSON, P. Generative AI meets copyright. **Science**, v. 381, n. 6654, p. 158-161, 2023. DOI: 10.1126/science.adi0656.

ŠARČEVIĆ, T.; KARLOWICZ, A.; MAYER, R.; BAEZA-YATES, R.; RAUBER, A. U can't gen this? A survey of intellectual property protection methods for data in generative AI. **arXiv**, 2024. DOI: 10.48550/arXiv.2406.15386. Preprint.

SCASSA, T. Legal issues with volunteered geographic information. **The Canadian Geographer / Le Géographe canadien**, v. 57, n. 1, p. 1-10, 2013. DOI: 10.1111/j.1541-0064.2012.00444.x.

SHAMOV, O. Ai and copyright: from a doctrinal crisis to a hybrid model of collective licensing. **Legal Horizons**, v. 26, n. 3, p. 68-76, 2025. DOI: 10.54477/lh.25192353.2025.3.pp.68-76.

UNITED STATES COPYRIGHT OFFICE. **Copyright and artificial intelligence: part 2 — copyrightability**. Washington, DC: USCO, 2025.

VALENTE, C.; SARTORI, R.; MARIN, J. P. Inteligência artificial e direitos autorais: um

mapeamento da produção científica. **Cadernos de Prospecção**, Salvador, v. 16, n. 4, p. 1137-1150, 2023. DOI: 10.9771/cp.v16i4.50583.

VOSGERAU, D. S. R.; ROMANOWSKI, J. P. Estudos de revisão: implicações conceituais e metodológicas. **Revista Diálogo Educacional**, Curitiba, v. 14, n. 41, p. 165-189, 2014.

WANG, S. *et al.* GPT, large language models (LLMs) and generative artificial intelligence (GAI) models in geospatial science: a systematic review. **International Journal of Digital Earth**, v. 17, n. 1, art. 2353122, 2024. DOI: 10.1080/17538947.2024.2353122.

ZDRAVKOVA, K. Professor: who holds the copyright for AI-assisted and AI-generated contents? **Language and Law / Linguagem e Direito**, Porto, v. 12, n. 1, p. 144-161, 2025. DOI: 10.21747/21833745/lanlaw12_1a7.

ZHANG, Y. *et al.* A decentralized model for spatial data digital rights management. **ISPRS International Journal of Geo-Information**, v. 9, n. 2, art. 84, 2020. DOI: 10.3390/ijgi9020084.