
Inteligência Artificial Generativa e Direitos Autorais: Desmascarando Mitos e Revelando Verdades

Pontos-chave

- O Brasil pode ser um líder global na inovação em IA, mas isso exige um debate mais profundo sobre a questão dos direitos autorais, uma vez que alguns dos debates têm gerado ruído e mesmo desestimulado esse avanço.
- Os direitos autorais protegem expressões originais fixadas em obras, mas não as ideias subjacentes.
- O Brasil já possui todos os instrumentos regulatórios necessários para proteger os direitos autorais na era da inteligência artificial, não necessitando de novas regras.

O surgimento da inteligência artificial generativa (IA) representa uma grande oportunidade para o Brasil conquistar uma posição competitiva e participar dessa nova fase da economia. Para navegar com sucesso nesse cenário, contudo, é essencial compreender o debate atual sobre o uso de conteúdo protegido por direitos autorais no treinamento de modelos de IA, indo além das simplificações que frequentemente dominam essas discussões no Brasil. Vamos analisar algumas concepções equivocadas comuns, além de abordar formas mais precisas de entender esse debate e suas consequências sociais de longo alcance.

1. “Modelos de IA generativa contêm cópias de seus dados de treinamento” – Mito ou verdade?

Mito. Os Sistemas de IA generativa [não armazenam](#) cópias compactadas dos dados nos quais foram treinados. Eles utilizam técnicas matemáticas para aprender padrões e conceitos como parâmetros numéricos. Por exemplo, ao serem treinados com dados de texto, ajustam parâmetros para refletir probabilidades de combinações de palavras, permitindo-lhes gerar respostas coerentes.

A exposição a certo conteúdo durante a fase de treinamento pode influenciar a saída gerada depois, pois o resultado é uma probabilidade estatística. Por exemplo, se um modelo foi exposto a milhares de imagens de gatos durante o treinamento, ele irá aprender com as características de um modelo de ‘gato’ e, assim, gerar uma imagem do gato com precisão,

quando for solicitado.

Assim, como alguém que leu muitos livros sobre um assunto e, depois, escreve um livro a partir da sua própria visão, os sistemas de IA generativa não copiam, mas entendem padrões e são capazes de gerar, a partir destes padrões, resultados de conteúdo original.

2. “Conhecimento, fatos, ideias e informações são de livre circulação e não podem ser protegidos por direitos autorais” – Mito ou verdade?

Verdade. Direitos humanos e marcos legais – incluindo a Declaração Universal dos Direitos Humanos, a Convenção Americana sobre Direitos Humanos e outras declarações internacionais de direitos humanos, assim como a Constituição Federal do Brasil preveem o direito de qualquer pessoa à liberdade de expressão.

É fundamental compreender que fatos, ideias, conhecimento e informações não podem ser protegidos pelos direitos autorais, sendo essencial a manutenção desses princípios fundamentais em meio aos debates sobre o desenvolvimento da IA no Brasil.

[Lei de Direitos Autorais \(Lei nº 9.610/98\)](#)

*Art. 8º Não são objeto de proteção como direitos autorais de que trata esta Lei:
I - as idéias, procedimentos normativos, sistemas, métodos, projetos ou conceitos matemáticos como tais (...)
VII - o aproveitamento industrial ou comercial das idéias contidas nas obras.*

3. “A lei de direitos autorais protege dados – fim da história” – Mito ou verdade?

Mito. A lei de direitos autorais protege expressões originais fixadas em um meio tangível ou intangível, mas não as ideias, fatos ou informações subjacentes. Ou seja, não se pode explorar o trabalho protegido por direitos autorais de outra pessoa sem a sua permissão, mas se pode aprender com ele. Fazer essa distinção é crucial para evitar a sobrecarga da proteção de direitos autorais, mantendo a liberdade de expressão e informação.

A Lei de Direitos Autorais (Lei nº 9.610/98) reconhece essa distinção. O equilíbrio delicado

das exceções de mineração de texto e dados não deve ser enfraquecido para evitar interpretações contrárias ao espírito da lei e consequências não intencionais para os direitos fundamentais.

4. “Todos os governos concordam que os detentores de direitos autorais devem poder dizer não ao treinamento de modelos de IA” – Mito ou verdade?

Mito. Entre as principais regulações na corrida da IA, apenas a União Europeia concedeu aos detentores de direitos o direito legal de optar por não participar da mineração de texto e dados para treinamento. Países como Estados Unidos, Japão, Singapura, Coreia do Sul, Malásia, Israel e Taiwan têm exceções que promovem a inovação e acessibilidade de dados sem essa exclusão.

O Brasil não deve seguir o modelo europeu, muito menos sugerir a implementação de um modelo ainda mais restritivo.

5. “Os detentores de direitos autorais não têm como impedir que seus dados sejam incluídos em conjuntos de treinamento” – Mito ou verdade?

Mito. Os detentores de direitos autorais podem usar o [protocolo robots.txt](#) para impedir que web crawlers capturem seu conteúdo. Embora alguns enfrentem dificuldades técnicas quanto à granularidade do protocolo, as indústrias de tecnologia e criativa podem trabalhar juntas para desenvolver soluções mais direcionadas.

Grandes empresas de tecnologia [já oferecem](#) ferramentas sofisticadas para que detentores de direitos autorais excluam seus dados dos conjuntos de treinamento. Encontrar soluções técnicas adequadas é do interesse de ambos os setores.

Além disso, o uso do opt-out pelos detentores de direitos autorais não deve impedir as atividades de TDM (“text and data mining”) permitidas por lei. Na verdade, os detentores de direitos não podem se opor à TDM em todos os casos, como quando feita para fins de pesquisa ou acessibilidade. Abordar esses desafios de forma intersetorial é vital para garantir a conformidade legal e promover a inovação.

6. “Os detentores de direitos autorais apenas querem que as empresas de IA adquiram licenças de seu conteúdo” – Mito ou verdade?

Mito.

Embora alguns detentores de direitos autorais ofereçam licenças para usar suas obras no desenvolvimento de IA, muitos outros não estão preparados para isso. [Dados recentes](#) mostram que a maioria dos sites e detentores de direitos autorais não bloqueiam o acesso aos seus dados para treinamento de IA, indicando que não veem necessidade de optar por sair. [Pesquisas com criadores de conteúdo revelam](#) que suas opiniões sobre o uso de suas obras para o treinamento de IA são mais complexas e variadas do que frequentemente se sugere – as preferências dos criadores são, na verdade, muito mais refinadas e detalhadas.

Desacelerar a inovação da IA generativa por meio de licenciamento complicado de dados de treinamento pode impactar negativamente os setores de mídia e criativo, que são dos primeiros a se beneficiar dessa inovação. Muitas empresas de mídia já estão explorando o uso da IA para criação de conteúdo. Equilibrar os interesses do setor criativo com os avanços tecnológicos e os direitos fundamentais será essencial para fomentar essa inovação.

Conclusão

Navegar na interseção entre IA generativa e direitos autorais requer a reconciliação de interesses concorrentes, devendo-se manter liberdades fundamentais e princípios chave. É crucial esclarecer equívocos, assegurar marcos regulatórios seguros e fomentar a colaboração entre os diversos setores para garantir um desenvolvimento de IA equitativo e sustentável no Brasil.

Assim, o Brasil já possui todas as ferramentas regulatórias necessárias para proteger os direitos autorais na era da IA. As partes interessadas só precisam trabalhar juntas para aproveitar ao máximo o potencial da IA generativa.

Este texto é uma adaptação do original europeu para a realidade do Brasil.
<https://project-disco.org/european-union/generative-ai-and-copyright-busting-prevalent-myths-revealing-truths/>

Category

1. Material educativo

Tags

1. inteligência artificial

Date

18/10/2024

Date Created
02/07/2024