



Tecnologia Disputa nos tribunais

# Justiça dos EUA pode definir o futuro da inteligência artificial e do jornalismo

— Ação judicial movida pelo 'The New York Times' acusa a OpenAI e a Microsoft de violarem direitos autorais; especialistas dizem que a lei sobre o tema tem muitas brechas

WILL OREMUS  
ELANE IZADI  
THE WASHINGTON POST

Se um veículo de comunicação copiasse um monte de artigos do *New York Times* e os publicasse em seu site, isso provavelmente seria visto como uma violação flagrante de direitos autorais. Mas o que acontece quando uma empresa de tecnologia copia esses mesmos artigos, os combina com inúmeros outros trabalhos copiados e os usa para treinar um chatbot de inteligência artificial (IA) capaz de conversar sobre praticamente qualquer assunto?

Essa é a questão legal no centro do processo que o *Times* moveu contra a OpenAI e a Microsoft em um tribunal federal, alegando que as empresas usaram ilegalmente "milhões" de artigos protegidos por direitos autorais para ajudar a desenvolver os modelos de IA por trás de ferramentas como o ChatGPT e o Bing. É o mais recente, e alguns acreditam ser o mais forte, de uma série de processos que acusam empresas de IA de violação de propriedade intelectual.

Juntos, os casos têm o potencial de abalar os alicerces do setor de IA generativa, dizem alguns especialistas jurídicos — mas também podem fracassar. Isso porque as empresas de tecnologia provavelmente se apoiaram fortemente em um conceito jurídico que lhes serviu bem no passado: a doutrina conhecida como "uso justo" (ou fair use).

Em termos gerais, a lei de direitos autorais nos EUA faz distinção entre copiar literalmente o trabalho de outra pessoa — o que geralmente é ilegal — e "remixar" ou fazer um uso novo e criativo. No caso de sistemas de IA, o que gera confusão é que eles parecem estar fazendo as duas coisas, segundo James Grimmelmann, professor de direito digital e da informação da Universidade de Cornell.

A IA generativa representa "essa grande transformação tecnológica que pode criar uma versão remixada de qualquer coisa", diz Grimmelmann. "O desafio é que esses modelos também podem memorizar descaradamente obras nas quais foram treinados e, muitas vezes, produzir

cópias quase exatas", o que, segundo ele, é "tradicionalmente o cerne do que a lei de direitos autorais proíbe".

Desde os primeiros videocassetes, que podiam ser usados para gravar programas de TV e filmes, até o Google Books, que digitalizou milhões de livros, as empresas americanas convenceram os tribunais de que suas ferramentas tecnológicas equivaliam ao uso justo de obras. A OpenAI e a Microsoft já estão montando uma defesa semelhante.

"Acreditamos que o treinamento de modelos de IA se qualifica como um uso justo, estando diretamente alinhado com os precedentes estabelecidos que reconhecem que o uso de materiais protegidos por direitos autorais por inovadores de tecnologia de forma transformadora é totalmente consistente com a lei de direitos autorais", escreveu a OpenAI em um pedido ao Escritório de Direitos Autorais dos EUA em novembro.

**COMO FUNCIONA.** Em geral, os sistemas de IA são "treinados" em conjuntos de dados gigantescos que incluem grandes quantidades de material publicado, grande parte dele protegido por direitos autorais. Por meio desse treinamento, eles passam a reconhecer padrões na disposição de palavras e pixels, que podem ser usados para montar texto e imagens.

Alguns entusiastas de IA veem esse processo como uma forma de aprendizado, não muito diferente de um estudante de arte que devora livros sobre Monet ou de um violado em notícias que lê o *Times* de capa a capa para desenvolver seu próprio conhecimento. Mas os críticos veem um processo mais cotidiano em ação sob o capô desses modelos: é uma forma de cópia, e uma cópia não autorizada.

Há duas vertentes principais no caso do *New York Times* contra a OpenAI e a Microsoft. Em primeiro lugar, como em outros processos recentes de direitos autorais e IA, o jornal argumenta que seus direitos foram violados quando seus artigos foram "raspados" — ou digitalizados e copiados — para inclusão nos conjuntos de dados gigantes nos quais o GPT-4 e outros modelos de IA foram treinados. Isso às vezes



Processo movido pelo 'Times' é o mais recente de uma série de ações abertas contra empresas de IA

**"O desafio é que esses modelos podem memorizar obras nas quais foram treinados e produzir cópias quase exatas"**

James Grimmelmann  
Universidade de Cornell

**"Acreditamos que o treinamento de modelos de IA se qualifica como um uso justo"**

Resposta da OpenAI ao  
Escritório de Direitos Autorais

jornal alega que as ferramentas também violaram seus direitos autorais com sua "produção".

**JUSTIÇA.** Até o momento, os juízes têm sido cautelosos com o argumento de que o treinamento de um modelo de IA em obras protegidas por direitos autorais equivale a uma violação em si, diz Jason Bloom, sócio do escritório de advocacia Haynes and Boone e presidente do grupo de litígio de propriedade intelectual.

"Técnicamente, fazer isso pode ser uma violação de direitos autorais, mas é mais provável que seja considerado fair use (uso justo, em tradução livre), porque você não está exibindo publicamente o trabalho", diz Bloom. O fair use também pode ser aplicado quando a cópia é feita para uma finalidade diferente da simples reprodução da obra original, como para crítica-la ou usá-la para fins educacionais ou de pesquisa. Foi assim que o Google defendeu o Google Books.

O dispositivo gerou uma ação judicial em 2005 por parte da Authors Guild, que o considerou uma "violação descarada da lei de direitos autorais". Mas o Google argumentou que, como exibia apenas "trechos" em resposta às pesquisas, não estava prejudicando o mercado de livros, mas fornecendo um serviço diferente. Em 2015, um tribunal federal

de apelação concordou com o Google.

Esse precedente deve favorecer a OpenAI, a Microsoft e outras empresas de tecnologia, afirma Eric Goldman, professor da Faculdade de Direito da Universidade de Santa Clara e codiretor do Instituto de Direito para a Alta Tecnologia. "Vou adotar essa posição, com base em precedentes", diz.

Uma vitória do *Times* poderia ter consequências importantes para o setor de notícias, que está em crise desde que a internet começou a suplantiar os jornais e revistas há quase 20 anos. Desde então, a receita de publicidade dos jornais tem sofrido um declínio constante, o número de jornalistas em atividade caiu drasticamente e centenas de comunidades em todo o país não têm mais jornais locais.

Blake Reid, professor associado de Direito do Colorado, observou que os gigantes da tecnologia podem ser sérios menos simpáticos hoje para muitos juízes e juristas do que há uma década, como o caso do Google Books. "Há um motivo pelo qual se ouve falar muito em inovação, código aberto e startups", disse ele. "Há uma corrida para definir quem é o Davi e quem é o Golias aqui." ●

ESTE CONTEÚDO FOI PRODUZIDO COM O AUXÍLIO DE FERRAMENTAS DE INTELIGÊNCIA ARTIFICIAL E REVISADO POR NOSSA EQUIPE EDITORIAL.