

“O DIREITO AUTORAL ESTÁ NO PROMPT” – a quem pertencem os textos gerados por ChatGPT e similares?

> **JORGE MACHADO:** Professor do Programa de Mudança Social e Participação Política, Universidade de São Paulo. Contato: machado@usp.br

> **ÍTALO DO NASCIMENTO SOUZA:** Doutorando do Programa de Mudança Social e Participação Política, Universidade de São Paulo. Contato: italoengenharia@usp.br

Resumo

O surgimento do ChatGPT e outras ferramentas similares de Inteligência Artificial (IA) tem gerado muita polêmica pela facilidade de produzir textos de razoável qualidade. Sistemas de inteligência artificial generativa, como ChatGPT, são capazes de gerar ensaios, artigos, resumos, poemas, textos publicitários, discursos, postagens para blogs e redes sociais, resolver equações matemáticas, produzir códigos de programação, entre outras tarefas. Esse cenário aponta para mudanças significativas no paradigmas da produção do conhecimento e da necessidade de adequações regulatórias. O objetivo desse texto é, a partir de uma breve descrição de como essas tecnologias de IA generativa funcionam, analisar e discutir o lugar do autor na produção desse tipo de conhecimento e os possíveis impactos que isso pode ter na ideia de produção autoral. Esse trabalho se insere na linhagem analítica da moldagem social da tecnologia, utilizando os métodos de revisão bibliográfica e de análise crítica

Palavras-chave: IA generativa, ChatGPT, produção do conhecimento, direito autoral

Summary

The emergence of ChatGPT and other similar Artificial Intelligence (AI) tools has generated a lot of controversy due to the ease of producing texts of reasonable quality. Generative artificial intelligence systems, such as ChatGPT, are capable of generating essays, articles, summaries, poems, advertising texts, speeches, posts for blogs and social networks, solving mathematical equations, producing programming codes, among other tasks. This scenario points to significant changes in the paradigms of knowledge production and the need for regulatory adjustments. The objective of this text is, based on a brief description of how these generative AI technologies work, to analyze and discuss the author's place in the production of this type of knowledge and the possible impacts that this may have on the idea of authorial production. This work is part of the analytical lineage of the social shaping of technology, using the methods of bibliographic review and critical analysis.

Keywords: Generative AI, ChatGPT, knowledge production, copyright

Resumen

La aparición de ChatGPT y otras herramientas similares de Inteligencia Artificial (IA) ha generado mucha controversia debido a la facilidad para producir textos de calidad razonable. Los sistemas de inteligencia artificial generativa, como ChatGPT, son capaces de generar ensayos, artículos, resúmenes, poemas, textos publicitarios, discursos, posts para blogs y redes sociales, resolver ecuaciones matemáticas, producir códigos de programación,

entre otras tareas. Este escenario apunta a cambios significativos en los paradigmas de producción de conocimiento y la necesidad de ajustes regulatorios. El objetivo de este texto es, a partir de una breve descripción de cómo funcionan estas tecnologías de IA generativa, analizar y discutir el lugar del autor en la producción de este tipo de conocimiento y los posibles impactos que esto puede tener en la idea de autor. producción. Este trabajo se inscribe en la línea analítica de la conformación social de la tecnología, utilizando los métodos de revisión bibliográfica y análisis crítico.

Palabras clave: IA generativa, ChatGPT, producción de conocimiento, derechos de autor

Résumé

L'émergence de ChatGPT et d'autres outils d'intelligence artificielle (IA) similaires a généré de nombreuses controverses en raison de la facilité de produire des textes de qualité raisonnable. Les systèmes d'intelligence artificielle générative, tels que ChatGPT, sont capables de générer des essais, des articles, des résumés, des poèmes, des textes publicitaires, des discours, des articles pour blogs et réseaux sociaux, résoudre des équations mathématiques, produire des codes de programmation, entre autres tâches. Ce scénario met en évidence des changements significatifs dans les paradigmes de production de connaissances et la nécessité d'ajustements réglementaires. L'objectif de ce texte est, à partir d'une brève description du fonctionnement de ces technologies d'IA générative, d'analyser et de discuter de la place de l'auteur dans la production de ce type de connaissances et des impacts possibles que cela peut avoir sur l'idée d'auteur. production. Ce travail s'inscrit dans la lignée analytique du façonnage social de la technologie, utilisant les méthodes de revue bibliographique et d'analyse critique.

Mots-clés: IA générative, ChatGPT, production de connaissances, droit d'auteur

Introdução

O surgimento do ChatGPT e outras ferramentas similares de Inteligência Artificial (IA) tem gerado muita polêmica pela facilidade de produzir textos de razoável qualidade. Sistemas de inteligência artificial generativa, como ChatGPT, são capazes de gerar qualquer forma de textos escritos e imagens. É possível produzir ensaios, artigos, resumos, poemas, textos publicitários, discursos, postagens para blogs e redes sociais, resolver equações matemáticas, produzir códigos de programação, entre outras tarefas. Tais textos têm povoado os trabalhos escolares, ambiente acadêmico, ambiente corporativos, redes sociais, etc. Apesar do impacto disso, banir ou ignorar ferramentas de IA já não está mais em questão dada a rápida aderência social e econômica que apresentam. Esse cenário aponta para gerar mudanças significativas no paradigmas da produção do conhecimento e da necessidade de adequações regulatórias. O objetivo desse texto é, a partir de uma breve descrição de como essas tecnologias de IA generativa funcionam, analisar e discutir o lugar do autor na produção desse tipo de conhecimento e os possíveis impactos que isso pode ter na ideia de produção autoral. Esse trabalho se insere na linhagem analítica da moldagem social da tecnologia (Williams, 2015), utilizando os métodos de revisão bibliográfica e de análise crítica.

Fundamentação teórica

Há cerca de 20 anos, em *Code: Version 2.0* — versão revisada de um influente livro publicado originalmente em 1999 —, Lawrence Lessig afirmava que a sociedade em rede é efetivada através de intermediários da comunicação humana. Em poucas palavras: quem controla o código, tem o poder. Assim, uma forma de evitar concentração de poder, seria promover programas, algoritmos, protocolos e padrões abertos, transparentes e plenamente auditáveis (Lessig, 2006, Machado, 2016). O que se viu ao longo das últimas duas décadas foi a formação de concentração de poder em corporações do campo da informação, como Google, Facebook, Amazon, Microsoft, Apple (“the big five”) (Parra et al., 2018). Estas empresas controlam plataformas, produtos e serviços, controlando softwares, protocolos e mesmo infraestrutura de informação.

As resistências à promoção da informação aberta sempre foram muito grandes, em especial pois afeta o interesse de grandes conglomerados da mídia e diversas áreas da chamada “indústria criativa”. A chamada Cultura Livre (Lessig, 2004) baseada no software livre (Stallman, 1999), no código aberto (Raymond, 2000) e “cultura hacker” (Himanen, 2001), no contexto de intensa digitalização da cultura, da arte e da criação na primeira década dos anos 2000, proporcionou um enfrentamento a isso, promovendo uma quebra de paradigma é um processo de

regulamentação através dos modelos livres de licenciamento livre de conteúdos, que quebraram com a lógica monolítica dos copyrights.

As leis de “propriedade intelectual” praticamente se tornaram inaplicáveis à lógica da internet. Vale ressaltar que a Internet - antes chamada de Arpanet e nascida de um projeto militar -, foi criada basicamente com três objetivos práticos: 1) o compartilhamento de processamento de informação, 2) o compartilhamento de banco de dados e conteúdos e o 3) compartilhamento de banda de transmissão. Portanto, o objetivo por trás da internet é ser uma imensa rede de compartilhamento, baseada num conceito de segurança e resiliência - para resistir até a um ataque nuclear.

É possível fazer um paralelo do surgimento da Internet com o da IA generativa. Ambas guardam características disruptivas sobre a forma como se produz e se distribui conhecimento.

A internet trouxe a facilidade de reproduzir e distribuir qualquer coisa que seja transformável em bits. Antes disso, conhecimento e cultura estavam dentro de uma lógica onde a necessidade de acesso físico gerava uma dinâmica de escassez controlada por indústrias. Uma vez que obras intelectuais precisavam ser fixadas em meios físicos - como livros, discos, cassetes, livros - envolviam uma complexa logística de produção, armazenamento e distribuição. Essa lógica mostrou-se ineficiente e obsoleta frente ao mundo digital.

Nesse cenário, o direito autoral, baseado na propriedade intelectual - abstração jurídica criada para mediar os interesses entre autores, consumidores e produtores e regular o interesse de promover a cultura, a ciência e as artes (Machado, 2008, 2013)¹ - foi se transformando ao longo do tempo, até se adequar a esse novo contexto. O modelo de negócio que veio a prosperar em meio à pirataria digital foi o de licenciamento de conteúdo, baseado principalmente no *streaming*. Assim surgiram plataformas de comercialização e de gerenciamento de direito autoral como Netflix, Spotify, Deezer, YouTube Music, Kindle Unlimited e outros. De modo geral, foi mantido o monopólio que confere ao detentor dos direitos autorais o poder de excluir do acesso todos aqueles que não podem pagar.

Benkler em *The Wealth of Networks* (2006), afirma que os benefícios e transformações positivas advindas da digitalização dependem de mudanças do arcabouço jurídico, que incentivem à produção de conhecimento comum (*informational commons*), disponível ao uso e reuso criativo da sociedade.

Como bem aclaram Simon & Vieira (2007: 13). No modelo convencional de propriedade intelectual, o conhecimento passa a existir quando há um “proprietário” reconhecido. É como se, para todos os casos, não houvesse informação ou conhecimento preexistentes à criação. O resultado dessa concepção é que quanto maior é a “propriedade intelectual”, mais restrito ficam os *commons*. Mais correto seria dizer que os “bens intelectuais” deixam os commons para se tornar “propriedade” de alguém.

A IA generativa coloca agora um desafio tão grande como a internet colocou para o monólito dos direitos autorais/propriedade intelectual. Desta vez os bens intelectuais deixam de ser propriedade de alguém para retornar ao commons. A solução das grandes plataformas que faturam gerenciando direitos autorais e vendendo acesso também passam a ser ameaçadas pela IA generativa, que cria sem copiar. A IA generativa tem o potencial de “violar o direito autoral” sem deixar qualquer prova, uma vez que uma base de conhecimento proprietária pode ser literalmente dissolvida em padrões e relações de conjuntos de dados.

Do “copia e cola” para o “cria e cola”

O “copia e cola” sempre representou um desafio àqueles profissionais que, por ofício, avaliam a origem, qualidade e a integridade do texto autorais. O digital trouxe o copia e cola e a Internet o espalhou pelo mundo. O copia e cola não é necessariamente ruim, pois deu origem também a remixes e caracterizou o salto criativo (Lessig, 2006) na primeira década de expansão da Internet, sendo essenciais para o surgimento das licenças “creative commons” (CC). Copiar e colar faz sempre parte da experiência do digital proporcionada pela Internet, uma rede global cuja dinâmica de funcionamento se baseia na replicação da informação em alta velocidade e a um custo banal ao usuário.

A “cultura” do copia e cola criou até sites específicos para esse propósito para facilitar a cópia de trabalhos inteiros. Isso resultou numa abundância de materiais disponibilizados para cópia e modificação, desde materiais técnicos, acadêmicos passando por código de programação até imagens e vídeos. Assim, a Internet potencializou o acesso à informação e estimulou a criatividade da mesma forma que gerou uma multidão de vagabundos digitais tentando enganar outros, principalmente professores(as).

As ferramentas antiplágio, que vieram a se integrar a esse ecossistema para deter o uso predatório do “copia e cola”, mostram-se agora ineficazes na oposição da IA generativa.

Essa nova tecnologia de IA trouxe o “cria e cola” no lugar do “copia e cola”. O ChatGPT, da empresa OpenAI, foi o primeiro sistema de IA efetivo de criação disponível à massa de usuários da Internet. Seu sucesso fez com que seus concorrentes se adiantassem no lançamento de seus produtos, como o Google, como o Bard, a Meta, com o Llama. Ademais, boa parte do código do ChatGPT está aberto, o que permitiu o desenvolvimento de mais de uma dezena de sistemas similares. A rapidez e o impacto alcançados indicam estarmos apenas no início de uma grande jornada na evolução das aplicações de IA generativa.

A Inteligência Artificial já estava entre nós

Podemos definir as tecnologias de inteligência artificial como aquelas que são capazes de substituir o humano no processo de tomada de decisão. Elas existem há muito tempo e podemos identificá-las desde uma porta automática que decide quando deve abrir e permanecer aberta (IA simples), até aplicações mais complexas, que demandam maior capacidade computacional e capazes de tomar uma sequência mais ampla de decisões (IA complexa).

A IA já nos acompanha, portanto, há muito tempo. Ela está, por exemplo, nos *bots* de atendimento automático, nos algoritmos das redes sociais e computação em nuvem ou em artefatos físicos, carregados com sensores, desde dispositivos mais simples até veículos autônomos. A novidade do ChatGPT, foi a de apresentar um “oráculo” potente, capaz de resolver equações matemáticas, elaborar textos, códigos de programação e até dar conselhos. O nome “Chat GPT” significa *Generative Pre-Trained Transformer*, ou “transformador pré-treinado, que é um algoritmo por trás do LLM (Large Language Model), em português “Modelo de Linguagem Grande”, que realiza a tarefa de gerar linguagem. Para fazer isso são utilizados modelos sofisticados de tecnologias de aprendizagem profunda. Esse tipo de ferramenta funciona através de redes neurais que analisam grandes volumes de dados, criando conexões entre conteúdos das bases de conhecimento. Essas bases, mesmo sendo autorais, acabam sendo “diluídas” pela forma como funciona o modelo. Isso permite criar algo novo a partir da combinação dessas tecnologias.

A IA poderia ser desenhada para trabalhar com blocos de textos e gerar textos respeitando o conhecimento autoral? Provavelmente sim. Mas essa não é sua “natureza”, sua lógica de funcionamento. Gerar mosaicos autorais significaria uma limitação tão grande que eliminaria seu maior potencial: o de criar novas conexões, muitas jamais feitas, sobre o conhecimento humano disponível em dado assunto.

É por essa razão que nos deparamos com algo completamente novo que ameaça diversos setores altamente autorais, como escritores, acadêmicos, roteiristas, artistas e outros profissionais. Não apenas gera algo novo, como o faz com uma eficiência incomparável. É compreensível que haja protestos. Esse cenário projeta uma concorrência perigosa de humanos com sistemas inteligentes, a obsolescência de profissões e ameaças a empregos.

No entanto, um novo tipo de *autor* surge na capacidade de interagir com as ferramentas de IA generativa através do *prompt*. Da capacidade intelectual e perspicácia de verificar e explorar suas respostas e agregar a inteligência e a criatividade humana na produção de conteúdos com base em sistema de inteligência artificial generativa. Surge aí, portanto, um novo autor, produtor de conteúdos através do *prompt*.

O direito autoral está no *prompt*?

Não é possível proteger direitos autorais com aquilo que é gerado por sistemas de IA. O autor não está na resposta obtida – que resulta da aplicação de um modelo sobre uma base de conhecimento. *O sistema de IA é apenas uma ferramenta geradora.*

Sistemas como o ChatGPT não criam algo do zero. Possuem em suas bases de conhecimento obras protegidas por direitos autorais, como artigos científicos, livros, literatura não acadêmica, imagens, etc. Cabem aqui as perguntas. *É a OpenAI quem detém os direitos autorais dos produtos de seu sistema? Ou seria o usuário? Não há titular de direitos autorais?*

A OpenIA não reivindica – e nem poderia – o direito autoral do que é produzido por seu sistema (OpenIA, 2023). Não há qualquer remuneração ao autor. Mas tampouco pode haver acusação de plágio, pois esse tipo de sistema foi desenhado para gerar, não para copiar. Inclusive pode gerar muitas respostas com graus diferentes de similaridade.

A princípio, a informação produzida por um sistema como ChatGPT está no escopo de uma informação de domínio público. De acordo com a Wikipédia inglesa:

O domínio público consiste em todo o trabalho criativo ao qual não se aplicam direitos exclusivos de propriedade intelectual. Esses direitos podem ter expirado, sido perdidos, expressamente renunciados ou podem ser inaplicáveis. Como ninguém detém os direitos exclusivos, qualquer pessoa pode usar ou fazer referência legal a essas obras sem permissão. (Wikipedia, 2023)

O *input* de texto do *prompt* é provido pelo usuário do sistema. Esse *input* pode ser altamente detalhado e complexo, podendo ser caracterizado como uma criação autoral, sendo passível de proteção de direitos autorais. No entanto, para o sistema de inteligência artificial, o texto digitado no *prompt* é visto como apenas um comando, podendo gerar respostas similares com a de outros usuários.

Por exemplo, quando inserimos o *prompt* no ChatGPT: “Explique o que é um ensaio científico”, obtemos a seguinte resposta:

“Um ensaio científico é um tipo de documento acadêmico que tem como objetivo discutir, analisar e apresentar informações de maneira clara, lógica e fundamentada em evidências dentro do contexto de uma investigação científica. Ele é uma forma de comunicação amplamente usada na academia para compartilhar conhecimento e resultados de pesquisas.”

Essa é uma resposta simples e generalizada, que qualquer um pode obter ao inserir um *prompt* sucinto e equivalente a esse do exemplo. Ou seja, esse é um texto gerado mais amplamente pelo algoritmo e conhecimento coletivo da Web do que da pessoa que entrou de fato com o *prompt*.

Agora quando inserimos no *prompt* no ChatGPT: “Explique o que é um ensaio científico. É importante partir do pressuposto que gerar um ensaio científico é uma forma de criar textos científicos de forma mais livre nas criações mas sem deixar de ter embasamento científico.”, obtemos a seguinte resposta:

“Um ensaio científico, na forma mais ampla, é uma abordagem de escrita que permite que os autores explorem tópicos científicos de maneira mais livre e criativa, ao mesmo tempo em que mantêm um embasamento científico sólido. Os ensaios científicos são uma forma de expressão acadêmica que combina a liberdade da escrita criativa com a necessidade de aderir a princípios científicos.”

Essa é uma resposta mais personalizada segundo o viés adotado na escrita do *prompt*. Ela foge do formato padrão porque adapta também a visão do usuário que escreveu o *prompt* na geração do resultado. É como se a resposta criada passasse por uma customização, ou refinamento para seguir uma ideia mais próxima do que o usuário deseja ter como resposta.

As respostas geradas pelo usuário em seu “direito de uso” podem ser consideradas de sua autoria? A resposta não é clara, pois as leis não foram desenhadas considerando a existência desses sistemas. No entanto, pode ser identificado claramente conteúdo autoral na informação do *prompt* do chat. Tanto os textos inseridos como as respostas geradas na interação com o sistema podem permitir a identificação de direito autoral da pessoa que interage com o chat por meio do *prompt*. Como numa pintura, o artista não tem qualquer direito sobre a tinta ou as cores utilizadas, mas sim sobre a composição que gerou. **Por analogia, o direito autoral está no *prompt*.**

Assim, a criação do usuário surge da interação com o sistema de inteligência artificial, ao elaborar perguntas ou fazer requisições de tarefas com precisão, explorando as respostas geradas. Isso inclui algum grau de conhecimento prévio e verificação de dados. Esse conjunto de ações lhe conferem a criatividade autoral. Não se trata de algo banal, como fazer simples perguntas ao sistema, mas sim de interagir com o sistema para criar algo presumivelmente novo.

Para seguir com as analogias, a interação do autor com o sistema é como aquela que tem o músico com o instrumento musical. Este pode simplesmente tocar composições publicamente conhecidas, como fazer músicas novas, dependendo do músico. A criatividade estaria na combinação da inteligência humana e o sistema de IA. Para complicar ainda mais, podemos destacar as adaptações, traduções e outras transformações de obras originais, que são consideradas no Brasil e em outros países, como criações intelectuais novas, onde se estabelece uma nova camada de direito sobre a obra original.

Como resolver o problema do direito autoral?

Quando se insere no *prompt* no ChatGPT a pergunta “Como resolver o problema do direito autoral com a criação de textos por IA generativa como o ChatGPT?”. Temos a seguinte resposta:

“O problema dos direitos autorais com a criação de textos por IA generativa, como o ChatGPT, é uma questão complexa e ainda em evolução. A resolução desse problema envolve a consideração de vários aspectos legais, éticos e tecnológicos. Aqui estão algumas abordagens que podem ajudar a lidar com o problema:

1. Revisão da legislação de direitos autorais: A legislação de direitos autorais existente precisa ser revisada e adaptada para incluir disposições específicas para criações geradas por IA. Isso pode envolver a definição de critérios para determinar a autoria em casos de obras geradas por máquinas e a revisão das limitações e exceções dos direitos autorais.

2. Autenticidade e identificação: É importante desenvolver métodos para autenticar a autoria de textos gerados por IA. Isso pode incluir marcas d'água digitais, assinaturas criptografadas ou outros mecanismos que tornem claro que o texto foi gerado por uma IA específica.

3. Licenças e contratos: Criar licenças e contratos específicos para textos gerados por IA pode ser uma abordagem eficaz. Os termos dessas licenças podem ser usados para determinar como o texto pode ser usado e quem detém os direitos.

4. Registro de direitos autorais: Estabelecer um sistema de registro de direitos autorais adaptado para criações de IA pode ajudar a esclarecer a propriedade e os direitos sobre os textos. Os criadores de IA podem registrar suas criações em um banco de dados centralizado.

5. Acordos entre criadores de IA e criadores humanos: Promover a colaboração e a negociação entre criadores de IA e criadores humanos pode levar a acordos mutuamente benéficos sobre a autoria e os direitos autorais.

6. Transparência na criação de IA: Promover a transparência na criação de IA generativa, de modo que seja possível rastrear e entender como a IA chegou a determinadas conclusões, pode ajudar a determinar a autoria e a responsabilidade.

7. Educação e conscientização: A conscientização sobre questões de direitos autorais e ética em relação a IA generativa deve ser ampliada entre os desenvolvedores, criadores de conteúdo e o público em geral.

8. Respeito pelos direitos autorais existentes: Enquanto as leis e regulamentos estão sendo adaptados, é fundamental que as criações de IA respeitem os direitos autorais existentes e não violem os direitos de terceiros.

É importante notar que a resolução desse problema é complexa e pode exigir a colaboração entre legisladores, especialistas em tecnologia, criadores de conteúdo e a sociedade em geral. À medida que a tecnologia de IA continua a evoluir, a regulamentação e a ética em torno dela também precisarão se adaptar para refletir as mudanças no cenário da criação de conteúdo.”

Como podemos perceber, esse texto gerado já apresenta uma série de alternativas para lidar com a questão. Esse texto foi tanto gerado a partir de conexões feitas pelo *Transformer* consumindo informações do conhecimento coletivo da Web, como também do *Fine tuning*, que é uma técnica algorítmica desempenhada para afinar um modelo computacional de forma a gerar um resultado mais próximo do esperado, que foi nesse caso desempenhada pelos próprios desenvolvedores por trás do ChatGPT - que provavelmente assim o fizeram para lidar com questões mais críticas, como essa do direito autoral aqui discutida.

Não apenas os desenvolvedores do ChatGPT podem realizar um *Fine tuning*, mas também desenvolvedores no geral que saibam construir ou modificar um LLM. A execução de um *Fine tuning* se assemelha a de criação de um *prompt*, só que exige um conhecimento computacional maior, mas a ideia principal de afinar uma estrutura para gerar saídas próximas do desejado se equivale a elaborar *prompt* personalizados para gerar resultados mais próximos do esperado.

É importante destacar esse ponto do *Fine tuning* pois o próprio ChatGPT e afins não utilizam exatamente os conhecimentos coletados da Web, eles fazem uma curadoria desses conhecimentos, condicionando os resultados para gerar algo que possa seguir os princípios que estabelecem em seus modelos de negócios ou na própria apresentação das suas soluções.

Nesse sentido, se formos tratar de “donos” de alguma coisa, textos simples de *prompt* sem muita personalização inseridos no ChatGPT por exemplo pertencem muito mais aos donos dessa ferramenta do que aos escritores iniciais do textos consumidos para alimentar o modelo, pois o Fine Tuning desse LLM foi o responsável pela curadoria desses textos.

No entanto, textos bastante customizados inseridos em um *prompt* têm como maior responsável quem o personalizou na inserção do LLM que gerou o resultado final. Isso ocorre pois o viés utilizado na personalização desse prompt foi bem mais responsável pela geração do resultado final do que os fundamentos seguidos pelo modelo base. Ou seja, a “curadoria” foi feita por quem criou o texto inserido no prompt, levando a conclusão de que o resultado final, caso venha ter um dono, seja essa mesma pessoa. Essa discussão traz como implicação, por exemplo, que o agente por trás das informações acaba sendo mais dono delas do que as próprias fontes de onde elas vieram.

E nesse ponto temos então uma questão que está ligada às fontes de informação utilizadas e sua devida referência. Como o ChatGPT não traz referências das fontes que usa em suas coletas, como estas podem ser referenciadas por quem as utiliza por meio da ferramenta? Ou no caso do utilizador se referenciar como autor, qual seria o grau de customização de tais textos inseridos no prompt? E como comparar o texto gerado com as fontes originais?

Conclusões

Este texto buscou discutir e analisar, a partir de modo como sistemas de IA generativa funcionam para produzir informação e conhecimento, os possíveis impactos na autoria das obras.

Assim como as tecnologias digitais - em especial, a Internet - trouxe o “copia e cola”, a IA generativa trouxe “cria e copia”. Tal como a Internet veio trazer desafios e quebrar muitos paradigmas, a IA generativa parece estar desempenhando um papel disruptivo semelhante, transformando radicalmente a forma como lidamos com informação e conhecimento.

A autoria de obras intelectuais, da forma monolítica como tem sido tratada, está sendo ameaçada pelo modo de funcionamento dos sistemas de IA generativa. Baseada em análise de padrões e relações de conjuntos de dados, tais sistemas têm o potencial de dissolver fontes autorais para criar novos conhecimentos a partir de complexas inferências lógicas. Tais sistemas, ainda de desenvolvimento, apresentam ainda erros e imprecisões, mas a sua eficiência e capacidade de geração de conteúdos inéditos abarcando muitas áreas do conhecimento é impressionante.

Podemos dizer que a lógica da criação, baseada na apropriação do conhecimento comum e prévio para gerar novos conhecimentos protegidos por direitos autorais (portanto de acesso e uso restrito) dá lugar a outra lógica inversa. A partir de artigos, livros, imagens outros tipos de dados e códigos), a IA Generativa transforma conhecimento autoral protegido o em conhecimento comum e disponível ao uso livre. Esses sistemas de IA podem ser vistos como “máquinas geradoras de commons” para o pânico de autores.

No entanto, esse trabalho mostra que o direito autoral - identificado com a labor criativa e a inspiração, pode estar no *prompt*. Poderá o *autor* ser reconhecido em sua interação qualificada com os sistemas de IA, onde sua inteligência e criatividade são determinantes para a geração de conhecimento através de IA generativa. Obviamente que não se trata de qualquer inserção no prompt. Por isso, o autor “pode estar” e não “está”. *Saber gerar informação e conhecimento de qualidade*, com base em conhecimentos preexistentes e através da interação com sistemas de IA poderá ser o ativo mais importante dos autores e autoras do futuro.

Esse cenário de quebra de paradigma da produção do conhecimento demanda adaptação de nossos sistemas jurídicos, que ainda não estão preparados para enfrentá-lo.

¹ O objetivo dos copyrights não é limitar o acesso ao conhecimento, à cultura e às criações humanas, senão discipliná-lo, de modo a balizar decisões judiciais sobre os conflitos morais e econômicos com respeito às formas de expressão, usos e disseminação das ideias.

Bibliografia

Benkler, Yochai (2011). *The Penguin and the Leviathan: How Cooperation Triumphs over Self-Interest*. New York: Crown Business.

_____. (2006). *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. New Haven: Yale University Press. Online: <http://www.congo-education.net/wealth-of-networks/> (visitado em 02 de outubro de 2023)

Lessig, Lawrence (2006). *Code 2.0*. NY: Basic Books, 2006. Online: <http://codev2.cc/download+remix/Lessig-Codev2.pdf> (visitado em 02 de outubro de 2023)

_____. (2004). *Cultura Livre. Como a Grande Mídia usa a Tecnologia e a Lei para Barrar a Criação Cultural e Controlar a Criatividade*. New York: Tramas, 2004. Online: <https://www.ufmg.br/proex/cpinfo/educacao/docs/10d.pdf> (visitado em 02 de outubro de 2023)

Himanen, Pekka (2001). *A Ética dos Hackers e o Espírito da Era da Informação*. Rio de Janeiro: Campus.

Machado, Jorge (2016) *Políticas Públicas de informação para o Conhecimento Aberto*, tese para a obtenção de Livre-docente na Escola de Artes, Ciências e Humanidades da USP. (visitado em 02 de outubro de 2023). Online: <http://each.usp.br/machado/docs/tese-livre-docencia-Jorge-Machado.pdf> (visitado em 02 de outubro de 2023)

_____. Open or “half-open” Access?: Re-thinking Open Access Initiative (OAI) Policies. *Observatório OBS**, vol. 7 n. 1 (2013) ISSN 1646-5954. Online: <http://obs.obercom.pt/index.php/obs/article/view/594/575> (visitado em 02 de outubro de 2023)

_____. (2008). “Desconstruindo Propriedade Intelectual”. *Observatório OBS** Vol 2, No 1 (2008) (visitado em 02 de outubro de 2023)

_____. O Ciberespaço como Arquitetura da Liberdade - Tentativas de Territorialização e Controle na Rede, in Alves, Giovanni e Martinez, Vinício (orgs.) *Dialética do Ciberespaço - Trabalho, Tecnologia e Política no Capitalismo Global*. Bauru: Ed. Práxis. Online: http://each.usp.br/machado/bm/articles/ciberespaço_territorializacao_jorgemachado.htm (visitado em 02 de outubro de 2023)

OpenIA (2023) Will OpenAI claim copyright over what outputs I generate with the API?, OpenAI website, *Policy*, <https://help.openai.com/en/articles/5008634-will-openai-claim-copyright-over-what-outputs-i-generate-with-the-api> (visitado em 02 de outubro de 2023)

PARRA, H.; AMIEL, T.; Cruz, L; MACHADO, J. Infraestruturas, Economia e Política Informacional: o Caso do Google Apps For Education. *Mediações - Revista de Ciências Sociais*. v. 23, n. 1 (2018). DOI: 10.5433/2176-6665.2018.1v23n1p63 ISSN: 1414-0543 (visitado em 02 de outubro de 2023)

Raymond, Eric (2000). *The Cathedral and the Bazaar*. Online <http://www.catb.org/esr/writings/cathedral-bazaar/cathedral-bazaar/index.html#catbmain> (visitado em 02 de outubro de 2023)

Simon, Imre; Vieira, Miguel Said (2007). *A Propriedade Intelectual diante da Emergência da Produção Social*. Paper. 19 de maio. Online: <http://www.ime.usp.br/is/papir/pips-19mar07.pdf> (visitado em 02 de outubro de 2023)

Stallman, Richard (1999). The GNU Operating System and the Free Software Movement. In DiBonna, Ockman & Stone, *Open Sources: Voices from the Open Source Revolution*, O'Reilly Media. Online: <http://www.oreilly.com/openbook/opensources/book/stallman.htm> (visitado em 02 de outubro de 2023)

Wikipedia, 2023. *Public domain* https://en.wikipedia.org/wiki/Public_domain

Williams, Raymond. (2015). *Televisão: tecnologia e forma cultural*. São Paulo: Boitempo.